MODELLING TEMPORAL INFORMATION IN A TWO-DIMENSIONAL SPACE

A VISUALIZATION PERSPECTIVE

Yi Qiang

UNIVERSITEIT GENT

Modelling Temporal Information in a Two-Dimensional Space

A Visualization Perspective

Yi Qiang

The research reported in this dissertation was conducted at the CartoGIS research unit, Department of Geography, Faculty of Sciences, Ghent University, and funded by the Research Foundation – Flanders (FWO).

Modelling Temporal Information in a Two-Dimensional Space

A Visualization Perspective

Dissertation submitted in accordance with the requirements for the degree of Doctor of Sciences: Geography

Modelleren van Temporele Informatie in een Twee-Dimensionale Ruimte

Een Visualisatieperspectief

Proefschrift aangeboden tot het behalen van de graad van Doctor in de Wetenschappen: Geografie

by/door Yi Qiang

Supervisors

Prof. dr. Nico Van de Weghe
Department of Geography, Ghent University

Prof. dr. Philippe De Maeyer
Department of Geography, Ghent University

Members of the Reading Committee

Dr. Gennady Andrienko
Fraunhofer Institute IAIS - Intelligent Analysis and Information Systems

Prof. dr. Guy De Tré
Department of Telecommunications and Information Processing,
Ghent University

Prof. dr. Frank Witlox
Department of Geography, Ghent University

Remaining members of the Examination Committee

Prof. dr. ir. Alain De Wulf (Chairman)
Department of Geography, Ghent University

Dr. Tijs Neutens
Department of Geography, Ghent University

Prof. dr. Birger Stichelbaut
Department of Archaeology, Ghent University

Prof. dr. Martin Valcke
Department of Educational Studies, Ghent University

PREFACE

If my life is viewed as a timeline, the PhD period in Ghent must be one of the most meaningful intervals. At the first arrival in Ghent, two suitcases were all I had. Now, I have a lovely wife, a coming baby, a PhD degree and abundant spiritual wealth. Compared with China, Belgium is a small country. But this fertile land has granted me so far the most plentiful harvest.

Pursuing a PhD was a long and arduous journey, in which I have been through many challenges, frustrations and depressions. Luckily, I have never walked alone. The support, help and encouragement from many people have enabled me to keep moving forward to this last step. Without them, I could not have gone this far. Exhausting all their names will not be possible within the few pages of the preface. Instead, I would take this opportunity to express my appreciation to some of the most important ones for the accomplishment of my PhD.

To start with, I would like to sincerely thank my supervisor Prof. dr. Nico Van de Weghe, who selected me and offered me the opportunity of this PhD position. Working with Nico was a pleasant and benefiting experience. His scientific vision, experience, enthusiasm and rigorousness have greatly influenced me. I am grateful to him for being always generous with careful and patient advices, offering me great freedom and support for my research, helping me in building connections to extend my research, and even concerning my career and life after the graduation. All of these are priceless treasures from which I will benefit for my whole life. I would also thank my second supervisor, Prof. dr. Philippe De Maeyer who has helped me in many aspects during my work in this department. Furthermore, I am indebted to Prof. dr. Guy De Tré, Prof. dr. Martin Valcke and Prof. dr. Frank Witlox for their fruitful discussions, patient guidance and constructive comments. I am greatly honored and grateful that Dr. Gennady Andrienko has taken time off from his busy schedule to read my PhD, attend my defense, and give my work the most authoritative comments in the area.

I would like to take this opportunity to express my gratitude to all my colleagues and former colleagues in the Department of Geography. I would like to thank Dr. Matthias Delafontaine, Dr. Tijs Neutens and Dr. Wim Kellens for giving advices to my research and publications. I would like to thank my officemates Mathias Versichele, Seyed Hossein Chavoshi and Bart De Wulf for their frequent helps and accompanying during my PhD. I would thank our secretary Helga Vermeulen, who has liberated me from numerous and trivial administrative matters. Besides the already named, my thanks go to the other colleagues, who have made my time in this department enjoyable.

In addition to the academic support above, I have received evenly important support from my friends and family. I appreciate all my friends in Ghent who have cured my homesick and shared a lot of memorable moments with me. Thanks also go to Prof. dr. Xingzhi Qiu and Mr.

Zongqiang Yu, who have treated me as one of their family. I would like to take this opportunity to express my special gratitude to my wife Shanshan Chen, who has accompanied me through the past four years. I deeply appreciate your understanding, thoughtful care, and faithful support, which were indispensible for me to accomplish the PhD. Meeting you and marrying you was the most beautiful thing in Ghent. I believe our upgraded life (being father and mother) will lead to even greater happiness and success. The last but not least, my deepest thanks go to my parents Mr. Ning Qiang and Mrs. Jie Hu, who have reared, educated and supported me with unconditional giving and selfless love. Without you, I would never have the achievement today.

Ghent, June, 2012

Contents

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Background and Motivation

Time is a pervasive concept involved in all aspects of human activities. Human beings have a long history of finding an agreed answer to the question: "*what is time*". The disputes over centuries have not made the answer clearer, in contrast, even more sophisticated. Different views of time arise in daily life and scientific research due to different viewpoints of phenomena, different ways of thought and different levels of abstraction. Although Einstein's theories of relativity can well explain most phenomena that are related to time, seldom people really use these equations in their everyday coping with the world. In spite of the differences, it seems that there exists some sort of common language about time that enables people to manage, communicate and mentally reason about temporal knowledge.

Since several decades ago when computers started to penetrate into people's work and lives, the handling of time appeared to be a key issue in many computer applications, such as planning (Coddington 2002; Aigner, Miksch et al. 2005), automatic control (Moon et al., 1992), workflow management (Chakraborti et al., 2007) and human language processing (Iwanska, 1996; Leith and Cunningham, 1997). In addressing these time problems, computer scientists realized the importance of formalizing common sense knowledge about time in a way that inference and reasoning about time can be performed by computers, which has led to a new research topic of temporal reasoning. On the one hand, many logic frameworks have been developed to encode temporal semantics in human knowledge into formal languages that can be reasoned about by computer programs (Galton, 1991; McCarthy and Hayes, 1968; McDermott, 1982). These logic frameworks aim to be expressive enough to tackle the complexity of temporal semantics and also take account of the computational efficiency. On the other hand, a large portion of contributions have been made in reasoning about time primitives. The milestone in this area is the interval algebra proposed by Allen (Allen, 1983). This algebra assumes that every event is referenced to a time interval bounded by a start and end point, and defines thirteen atomic relations between time intervals. Allen's theory of time intervals has been widely accepted in artificial intelligence and information sciences for time representation, and triggered a variety of research in interval reasoning (Dechter et al., 1991; Freksa, 1992; Galton, 1990; Van Beek, 1989; Vilain et al., 1990) and even spatial reasoning (Cohn et al., 1997; Freksa, 1990).

Allen's interval algebra applies to an ideal situation in which the temporal extent of an event is crisp, well defined and perfectly known. However, in some cases, pinning an event to a crisp time interval is rather difficult or even impossible. This difficulty can be caused by the vagueness of the event. For example, it is difficult to decide when the Industrial Revolution started and finished. Though some historians like to use the invention of the steam engine to mark its beginning, it is unnatural to understand that the revolution suddenly started at the date when the steam engine was invented. Therefore, crisp time intervals are no longer adequate to represent the temporal references of such events. To solve this kind of problem, time intervals of vague events can be modelled as fuzzy sets, which use a real number between zero and one to quantify the truthiness of whether time points in the time line are occupied by the event. Using this approach, the truthiness of a temporal relation holding between two time intervals is also expressed as a number between zero and one. A number of contributions have investigated the reasoning mechanism of such fuzzy time intervals and have proved their usefulness in inference systems (Nagypál and Motik, 2003; Ohlbach, 2004; Schockaert and De Cock, 2008; Schockaert et al., 2008). In some other cases, events are not vague, but they cannot be pinned to a crisp time interval due to the lack of information. It may be known that an event started within a certain range and ended within a certain range. But information of the exact start time and ending time is not available. This often happens in information collection activities in which data referenced to discrete time stamps, such as historical studies based on literature or evidence in specific years or remote sensing studies based on photographs at specific moments. From information at a specific time stamp, one can determine the existence of a certain event or process. However, the status of the event or process between two time stamps is unknown. With these discrete snapshots, the time interval of the event or process is thus uncertain. If there is no prior knowledge or assumption about the distribution of the start and end point, modelling the interval with fuzzy sets may cause extra overhead and unnecessary complexity. In such cases, a half-way approach, i.e. rough set theory, is excellently suited to interval modelling and reasoning. Unlike the gradual truthiness of fuzzy time intervals, the rough set approach categorises time points in the time line into three crisp sets, which are ‘definitely not in', ‘possibly in', and ‘definitely in' the interval. Although this problem is not rare, it is surprising that much less attention has been paid to the rough set approach, compared with that to the fuzzy set approach. Due to the lack of in-depth research on the reasoning principles and semantics of such rough time intervals, the methods of handling and analysing the uncertainties in them are hard to develop.

Maps are used to display the locations and extent of entities in space. People have also been searching for a map that can visually represent the locations and extent of entities in time. Since Joseph Priestley (Priesley, 1765) has drawn a chart of timelines to represent lifetimes of empires, the timeline became a conventional visual representation of time intervals. In this representation, a time interval is depicted as a linear segment parallel to the time axis. The positions of the two extremes with respect to the time axis indicate the start point and end point of the interval. When applied to represent project schedules, the timeline is called the Gantt chart. Since this representation is in line with humans' intuition of the time linearity, it has been widely applied to illustrate the time intervals of events and processes (Aigner et al., 2005; Chasin et al., 2011; Plaisant et al., 1998). Timelines are mostly used for illustration, but are seldom applied into analytical tasks, particularly exploratory data analysis that greatly relies on data visualization. This situation may be explained by the inherent shortcomings of timelines: it is built on a one-dimensional space and the arrangement of timelines in the second dimension has no temporal meaning. Therefore, the visual structure of intervals may vary, according to different sorting rules applied to the second dimension. This polymorphism prohibits the existence of a universal visual approach to detect patterns (e.g. clusters, density trends) among time intervals. Also, the poor efficiency in space use makes timelines unsuitable for the visualization of a large amount of intervals. Let us imagine that, if thousands of time intervals stored in an archaeological inventory are represented as timelines, obtaining an overview of these intervals is already a hard task, not to mention pattern detection and visual analysis. These problems can become more serious when timelines are used to represent vague and uncertain time intervals with complex inner structures. Note, that an explicit visual representation of time intervals with account of the vagueness and uncertainties is a particular challenge in the emerging research area of visual analytics (Andrienko et al., 2010).

The difficulties of visualising time intervals exist in many information systems that deal with interval-based data, for instance, geographical information systems (GIS), which more and more consider time as one of the most important components. In the last decades, considerable amount of work has been done in the representation of time in GIS (Goodchild et al., 2007; Langran, 1993; Peuquet, 2002) and the ability of visualizing spatio-temporal data becomes a common feature of up-to-date GIS (Andrienko et al., 2010; Andrienko et al., 2003; Dykes et al., 2005). The majority of efforts have been devoted to visualise the dynamics of spatial data at time stamps such as sequences of snapshots (Shi and Winter, 2010; Worboys and Duckham, 2006), spatial time series (Andrienko, G. and Andrienko, N. , 2004; Tominski et al., 2005), and trajectories of travelling objects (Demsar and Virrantaus, 2010; Hurter et al., 2009; Kapler and Wright, 2005). Not many approaches are available to visualise geospatial data referenced to time intervals, such as events with durations and objects with lifetimes. For exploratory spatio-temporal analysis, an integrated display of the data distributions in space and time is very important. For instances, analysts may be interested in whether a special pattern in time implies a special pattern in space and vice versa, or whether the data from different regions in space have different distributions in time and vice versa. A number of approaches have been developed to solve this problem, including cartographical metaphors (Kraak and Ormeling, 2003), space-time cube (Gatalsky et al., 2004; Huisman et al., 2009) and linked space-time views (Kraak and van de Vlag, 2007). Since these approaches are built on the linear representation of time, they show limited power in tackling large spatial datasets with reference to time intervals of different length.

Nowadays, visual and dynamic query tools become a common functionality in GIS and other analytical systems. These kinds of query tools have significantly enhanced human-computer interactivity and therefore play a crucial role in exploratory data analysis (Andrienko, G and Andrienko, N, 2006). Some of them can be used for temporal queries as well. The most well-known one is the dynamic sliding bars, which can be manipulated to retrieve the data display at certain time stamps or during certain intervals. However, the sliding bars are rather limited in the types of queries that they can formulate. For example, it is hard to specify the temporal query of selecting the time intervals that are during 6:00am and 8:00am and longer than 30 minutes. Moreover, it is separate from the data display, meaning that users cannot directly perceive how the query range covers the dataset. Another frequently used query technique is brushing (Monmonier, 1989), which allows users to specify queries by clicking on or sketching zones in the graphical data display. With this technique, users can directly and flexibly select the data in the display for views they find interesting.

Brushing has been widely applied in data displays in an appropriate arrangement or a coordinated space, such as scatter plots and geographical maps. However, its application in temporal querying is not prominent. If temporal data would be arranged or plotted in a more effective way, the brushing technique could better support the data exploration from the temporal perspective.

The temporal scale may cause serious issues in data analysis. Analogous to the well-known modifiable areal unit problem in spatial analysis, the way of aggregating temporal data may also significantly affect analysis results. Sometimes patterns or relationships that are detected in a certain scale cannot be detected in other scales. Even in the same scale, different partitions of aggregation intervals may result in different patterns being revealed. On the other hand, a question can be answered in different scales. For example, the answers to the question of when there are a lot of traffic jams in Belgium may include ‘between 7:00am and 9:00am’, ‘in the weeks it snows’ and ‘in the months of school semester’. All these answers make sense because they may guide people to take actions in corresponding scales. Therefore, an appropriate choice of the aggregation of temporal data should take account of the characteristics of phenomena under study, the level of questions being asked, and the scale of actions to be taken. This choice is not easy, particularly in the phase of exploratory analysis when there is not much known about the data and when the goal of the analysis is not accurately specified. In addition to specifying an appropriate scale for analysis, the hierarchy of phenomena in different scales can also be important in certain analytical tasks. Analysts may be interested in how long-term patterns are composed or influenced by short-term patterns within them. As a result, multi-scale analysis is of critical importance for analysing temporal data. Due to the complexity of this issue, the solution would require a lot of human intelligence to be involved. An explicit visualization of temporal data in multiple scales can effectively combine the insight of humans and processing ability of computers to solve this issue. While a number of approaches have been developed for this purpose, none of them can display temporal data aggregated in all different intervals within an integrated visualization.

To address the aforementioned issues, a two-dimensional representation of time intervals can be considered. This idea can be at least dated back to Ligozat’s work (1994, 1997). In his work, an interval is represented as a point in a two-dimensional space, where the horizontal and vertical axes respectively correspond to the start and end points. Ligozat has shown how this representation can help to characterize different classes of temporal relations (e.g. convex, pointizable and ORD-Horn relations), and how it can help to understand tractability questions about temporal relations. Later, Kulpa has claimed that using the midpoint and radius to project an interval to a point is more interesting for temporal reasoning research. He has comprehensively studied the use of this representation in characterizing classes of temporal relations, temporal relations operations and interval arithmetic (Kulpa, 1997; 2001; 2006). Ligozat and Kulpa have shown that this representation has allowed more direct comprehension of abstract temporal concepts and provided an interesting viewpoint to the temporal reasoning. Their work focuses on crisp time intervals. The usefulness of this representation in modelling and reasoning imperfect time intervals (i.e. fuzzy and rough time intervals) has not been investigated yet. More recently, Van de Weghe (2007) made the first trial of using this representation for information visualization, and shown that this representation allows a convenient observation of the chronological structure of archaeological records. He has named this representation the Triangular Model, which is used in this thesis as well. His work remains at a conceptual level and the demonstration limited to static diagrams with a small dataset. The TM not only provides absolute coordinates of time intervals, but also brings in the concept of interval space, in which the sets of intervals in certain temporal relations can be represented as zones. These features can overcome the inherent difficulties of the conventional linear representation (e.g. timelines and line charts), and therefore, can bring added value to the visualization and analysis of temporal data. Beyond the above-mentioned work, this thesis focuses on the use of the TM in reasoning about imperfect time intervals, analysing of time intervals and linear data, and supporting a GIS for exploratory analysis of interval-based geospatial data.

1.2 Research Objectives

The previous section elaborated several problems in temporal reasoning and analysis, and presumed the potential of the Triangular Model in solving these problems. This section presents the general research questions (RQ) that this thesis intends to address and the corresponding chapters in which they are addressed.

RQ 1: How can the Triangular Model facilitate the reasoning research of imperfect time intervals?

This question is addressed in Chapter 3, which investigates the use of the Triangular Model in representing and reasoning about rough time intervals and fuzzy time intervals. First, the definitions and application scopes of these two types of intervals are discussed. While fuzzy set theory is powerful in handling vagueness and uncertainties in time intervals, under some circumstances, rough set theory is a more budget and effective approach. A clear distinction of the circumstances in which one of the two approaches is more suitable than the other

can be very important for the handling of time intervals. Second, the graphic representation of rough time intervals in the Triangular Model is presented. It studies how the temporal relations between rough time intervals can be expressed by the topological relations of the corresponding geometric objects. It shows how the graphic representation offered by the Triangular Model allows more direct perception of temporal relations of rough time intervals and facilitates the finding of new properties and structures in these relations. Also, the methods of deducting the underlying semantics from the graphic configuration are generalised. Furthermore, it examines whether the Triangular Model can be useful in the representation of fuzzy time intervals with continuous structures. At the end, the potential applications of this fundamental study are envisaged.

RQ 2: Can the Triangular Model bring added value to the analysis of time intervals?

This question focuses on the use of the Triangular Model in the analysis of time intervals. This question is addressed in Chapter 4, 5, with different emphases on crisp time intervals and rough time intervals respectively. Chapter 4 shows how the visual structure of crisp time intervals in the Triangular Model allows the pattern detection and comparison within large amounts of intervals. Moreover, an innovative temporal query mechanism is introduced, which relies on the manipulation of geometries in the visualization of intervals. Also, a software prototype that implements the Triangular Model in a geographical information system (GIS) is introduced. This prototype is applied in a concrete scenario to demonstrate how the visualization and temporal queries supported by the Triangular Model can benefit a GIS for exploratory spatio-temporal analysis. Chapter 5 describes how the Triangular Model can be used to analyse rough time intervals. It shows how the polygonal structure in the Triangular Model can be applied in visual analysis of rough time intervals. A probabilistic framework is proposed to tackle the uncertainty issues that arise in the temporal queries with rough time intervals. An updated version of the prototype implementation is introduced, incorporating functionalities based on the probabilistic framework. The usefulness of this approach is demonstrated in an archaeological use case that involves a spatio-temporal dataset with rough time intervals.

RQ 3: How can the Triangular Model be used for the analysis of linear data?

Chapter 6 describes how linear data can be represented and analysed via the Triangular Model. Linear data refers to data sequences in a one-dimensional linear space, such as time series, traffic speed along a road, or signal strength along a cable. The conventional representation of linear data is the line chart, in which the linear data at a certain scale is represented as a curving line. However, the representation of the linear data in multiple scales is not easy in line charts. In the Triangular Model, every interval corresponds to a unique position in the interval space. It is thus possible to assign the value of every sub-interval of a linear dataset to the corresponding position. Using this approach, the values of all sub-intervals of a linear data can fill a 2D continuous field, offering an overview of the linear data in all different scales. This chapter investigates whether this representation brings added value over the line charts in the analysis of linear data. It also demonstrates how multiple linear data can be combined in the coordinate interval space through spatio-temporal analysis methods. Several real-world datasets are used to demonstrate its usefulness in meaningful analytical tasks.

RQ 3: Can the Triangular Model be understood and used by novice users?

To address the fourth research question, Chapter 7 describes an empirical study that aims to evaluate the understandability and usability of the Triangular Model, using the conventional linear model (i.e. timelines) as a reference. This study focused on the use of the Triangular Model in representing crisp time intervals and relations. It was carried out through an online test including a series of questions about interval properties and relations. The test was carried out within a group of 260 participants with no prior knowledge about the Triangular Model. Before the test, these participants followed a training process about both the Triangular Model and the linear model. In the test, every question was asked in both the Triangular Model and the linear model to generate pairs of comparable parameters. After the test, the parameters generated from the two representations were analysed, in order to evaluate whether the Triangular Model is more advantageous over the linear representation for visual observation of intervals.

This result may be considered as the preliminary evidence of the feasibility of promoting the Triangular Model to a wider range of non-expert users.

1.3 Thesis Outline

The remainder of this thesis consists of seven chapters. It starts with a review of the milestones and important contributions in the areas that this thesis is related to (Chapter 2). Chapters 3 to 7 are the substantial part of the thesis, comprising five academic articles that have been published in international peer-reviewed journals, submitted or in preparation for submission. The details of these articles are given in Table 1-1. In order to allow someone to read these papers smoothly and independently from each other, there are some inevitable overlaps in the individual chapters with regard to the literature reviews and the description of the basic concept of the Triangular Model. The users should also bear in mind that a few differences in styles and terminologies reflect the preferences of particular journals. At the end of the thesis (Chapter 8), the general contribution of this thesis is discussed and the avenues of future research are proposed. In Figure 1-1, an outline is presented to illustrate the overall structure of this thesis.

Table 1-1: The academic articles included in this thesis

Chap.	Title	Authors	Outlet	Status
3	Modelling Imperfect Time Intervals in a Two-Dimensional Space	**Yi Qiang** Matthias Delafontaine Katrin Asmussen Birger Stichelbaut Guy De Tré Philippe De Maeyer Nico Van de Weghe	Control & Cybernetics Year: 2010 Volume: 39 Issue: 4 Page: 983-1010	Published
4	Interactive Analysis of Time Intervals in a Two-Dimensional Space	**Yi Qiang** Matthias Delafontaine Mathias Versichele Philippe De Maeyer Nico Van de Weghe	Information Visualization Year: 2012	Published
5	Analysing Imperfect Temporal Information in GIS Using the Triangular Model	**Yi Qiang** Matthias Delafontaine Tijs Neutens Birger Stichelbaut Guy De Tré	The Cartographic Journal Year: 2012	In press

		Philippe De Maeyer Nico Van de Weghe		
6	Multi-scale Analysis of Linear Data in a Two-Dimensional Space	**Yi Qiang** Steven Logghe Philippe De Maeyer Nico Van de Weghe	Information Visualization	In preparation
7	Representing Time Intervals in a Two-Dimensional Space: an Exploratory Empirical Study	**Yi Qiang** Martin Valcke, Philippe De Maeyer Nico Van de Weghe	Spatial Cognition and Computation	Submitted in May 2012

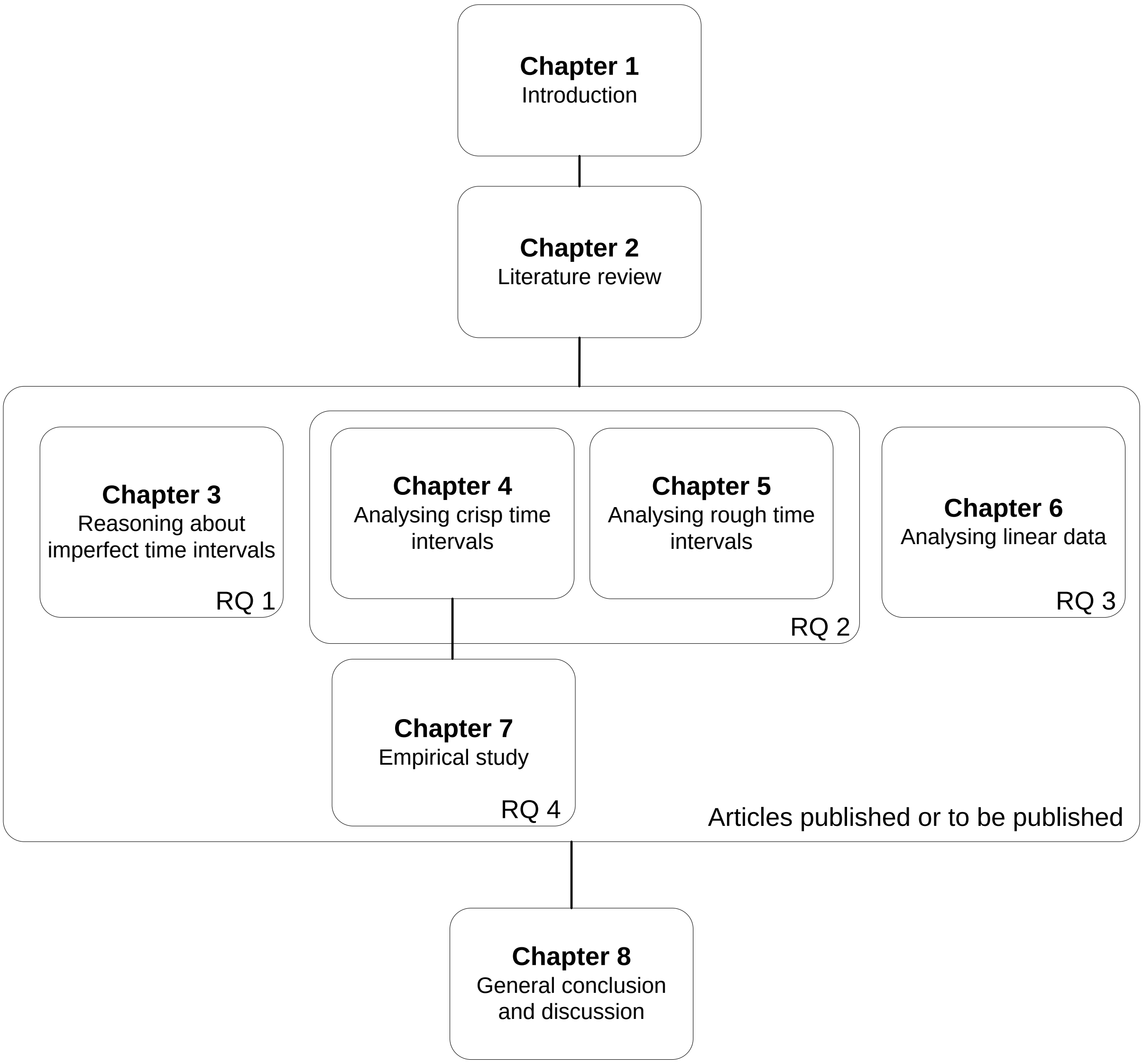


Figure 1-1: Thesis outline

References

Aigner, W., Miksch, S., Thurnher, B. and Biffl, S. (2005). PlanningLines: novel glyphs for representing temporal uncertainties and their evaluation. Proceedings of Ninth International Conference on Information Visualisation, London, UK, pp.457-463.

Allen, J. F. (1983). Maintaining Knowledge about Temporal Intervals. Communications of the ACM, 26(11), pp.832-843.

Andrienko, G., Andrienko, N., Demsar, U., Dransch, D., Dykes, J., Fabrikant, S. I., Jern, M., Kraak, M.-J., Schumann, H. and Tominski, C. (2010). Space, time and visual analytics. International Journal of Geographical Information Science, 24(10), pp.1577-1600.

Andrienko, N. and Andrienko, G. (2004). Interactive visual tools to explore spatio-temporal variation. Proceedings of the working conference on Advanced visual interfaces, Gallipoli, Italy, pp.417-420, ACM.

Andrienko, N. and Andrienko, G. (2006). Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach. Springer-Verlag Berlin, Heidelberg.

Andrienko, N., Andrienko, G. and Gatalsky, P. (2003). Exploratory spatio-temporal visualization: an analytical review. Journal of Visual Languages & Computing, 14(6), pp.503-541.

Chakraborti, A., Patriarca, M. and Santhanam, M. S. (2007). Financial time-series analysis: a brief overview. ed. by Chatterjee, A. and Chakrabarti, B. K., 3rd Workshop on Econophysics and Sociophysics of Markets and Networks, Kolkata, INDIA, pp.51-67, Springer-Verlag Italia.

Chasin, R., Woodward, D. and Kalita, J. (2011). Extracting and Displaying Temporal Entities from Historical Articles. in Machine Intelligence, pp.1-13, Narosa Publishing.

Cohn, A. G., Bennett, B., Gooday, J. and Gotts, N. M. (1997). Qualitative Spatial Representation and Reasoning with the Region Connection Calculus. GeoInformatica, 1(3), pp.275-316.

Dechter, R., Meiri, I. and Pearl, J. (1991). Temporal constraint networks. Artificial Intelligence, 49(1–3), pp.61-95.

Demsar, U. and Virrantaus, K. (2010). Space-time density of trajectories: exploring spatio-temporal patterns in movement data. International Journal of Geographical Information Science, 24(10), pp.1527-1542.

Dykes, J., MacEachren, A. M. and Kraak, M.-J. (2005). Exploring Geovisualization. Elsevier, Oxford, UK.

Freksa, C. (1990). Cognitive and linguistic aspects of geographic space. ed. by Mark, D. M. and Frank, A. U., in Cognitive and linguistic aspects of geographic space, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Freksa, C. (1992). Temporal Reasoning Based on Semi-Intervals. Artificial Intelligence, 54(1-2), pp.199-227.

Galton, A. (1990). A Critical Examination of Allen's Theory of Action and Time. Artificial Intelligence, 42(2-3), pp.159-188.

Galton, A. (1991). Reified temporal theories and how to unreify them. Proceedings of the 12th international joint conference on Artificial intelligence - volume 2, Sydney, New South Wales, Australia, pp.1177-1182, Morgan Kaufmann Publishers Inc.

Gatalsky, P., Andrienko, N. and Andrienko, G. (2004). Interactive analysis of event data using space-time cube. Proceedings of the 8th International Conference on Information Visualisation (IV 2004), London, UK, pp.145-152.

Goodchild, M. F., Yuan, M. and Cova, T. J. (2007). Towards a general theory of geographic representation in GIS. International Journal of Geographical Information Science, 21(3), pp.239-260.

Huisman, O., Santiago, I. F., Kraak, M.-J. and Retsios, B. (2009). Developing a Geovisual Analytics Environment for Investigating Archaeological Events: Extending the Space-time Cube. Cartography and Geographic Information Science, 36(3), pp.225-236.

Hurter, C., Tissoires, B. and Conversy, S. (2009). FromDaDy: Spreading Aircraft Trajectories Across Views to Support Iterative Queries. Visualization and Computer Graphics, IEEE Transactions on, 15(6), pp.1017-1024.

Iwanska, L. (1996). Natural (language) temporal logic: Reasoning about absolute and relative time. International Journal of Expert Systems, 9(1), pp.113-149.

Kapler, T. and Wright, W. (2005). GeoTime Information Visualization. Information Visualization, 4(2), pp.136-146.

Kraak, M.-J. and Ormeling, F. J. (2003). Cartography : visualization of geospatial data. Pearson Education Limited, Essex, UK.

Kraak, M.-J. and van de Vlag, D. E. (2007). Understanding Spatiotemporal Patterns: Visual Ordering of Space and Time. Cartographica: The International Journal for Geographic Information and Geovisualization, 42(2), pp.153-161.

Kulpa, Z. (1997). Diagrammatic Representation of Interval Space in Proving Theorems about Interval Relations. Reliable Computing, 3(3), pp.209-217.

Kulpa, Z. (2001). Diagrammatic representation for Interval Arithmetic. Linear Algebra and Its Applications, 324(1-3), pp.55-80.

Kulpa, Z. (2006). A Diagrammatic Approach to Investigate Interval Relations. Journal of Visual Languages and Computing, 17(5), pp.466-502.

Langran, G. (1993). Time in Geographic Information Systems. ed. by Peuquet, D. J. and Marble, D. F., in Technical Issues in Geographic Information System, Taylor & Francis Ltd, London.

Leith, M. and Cunningham, J. (1997). Representing and reasoning with events from natural language. Springer-Verlag Berlin, Berlin 33.

Ligozat, G. (1994). Towards a general characterization of conceptual neighborhoods in temporal and spatial reasoning. Paper presented in Workshop on Spatial and Temporal Reasoning (at AAAI-94) Seattle, WA, pp.745-750.

Ligozat, G. (1997). Figures for thought: temporal reasoning with pictures. Paper presented in Workshop on Spatial and Temporal Reasoning (at AAAI-97) Providence, RI, pp.31-36.

McCarthy, J. and Hayes, P. J. (1968). Some philosophical problems from the standpoint of artificial intelligence. in Readings in nonmonotonic reasoning, pp.26-45, Morgan Kaufmann Publishers Inc.

McDermott, D. (1982). A temporal logic for reasoning about processes and plans. Cognitive Science, 6(2), pp.101-155.

Monmonier, M. (1989). Geographic Brushing: Enhancing Exploratory Analysis of the Scatterplot Matrix. Geographical Analysis, 21(1), pp.81-84.

Moon, I., Powers, G. J., Burch, J. R. and Clarke, E. M. (1992). Automatic verification of sequential control systems using temporal logic. AIChE Journal, 38(1), pp.67-7575.

Nagypál, G. and Motik, B. (2003). A Fuzzy Model for Representing Uncertain, Subjective, and Vague Temporal Knowledge in Ontologies. Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics, (ODBASE), Catania, Italy, pp.906-923.

Ohlbach, H. J. (2004). Relations between fuzzy time intervals. Proceedings of 11th International Symposium on Temporal Representation and Reasoning (TIME 2004), Montréal, Québec, pp.44-51.

Peuquet, D. J. (2002). Representations of space and time. The Guilford Press, New York.

Plaisant, C., Shneiderman, B. and Mushlin, R. (1998). An information architecture to support the visualization of personal histories. Information Processing & Management, 34(5), pp.581-597.

Priesley, J. (1765). A Chart of Biography.

Schockaert, S. and De Cock, M. (2008). Temporal Reasoning about Fuzzy Intervals. Artificial Intelligence, 172(8-9), pp.1158-1193.

Schockaert, S., De Cock, M. and Kerre, E. E. (2008). Fuzzifying Allen's Temporal Interval Relations. IEEE Transactions on Fuzzy Systems, 16(2), pp.517-533.

Shi, M. Z. and Winter, S. (2010). Detecting Change in Snapshot Sequences. ed. by Fabrikant, S. I., Reichenbacher, T., VanKreveld, M. and Schlieder, C., in Geographic Information Science, pp.219-233, Springer-Verlag Berlin, Berlin.

Tominski, C., Schulze-Wollgast, P. and Schumann, H. (2005). 3D Information Visualization for Time Dependent Data on Maps. Proceedings of the Ninth International Conference on Information Visualisation, London, UK, IEEE Computer Society.

Van Beek, P. (1989). Approximation algorithms for temporal reasoning. Proceedings of the 11th international joint conference on Artificial intelligence - Volume 2, Detroit, Michigan, pp.1291-1296, Morgan Kaufmann Publishers Inc.

Van de Weghe, N., Docter, R., De Maeyer, P., Bechtold, B. and Ryckbosch, K. (2007). The Triangular Model as an Instrument for Visualising and Analysing Residuality. Journal of Archaeological Science, 34(4), pp.649-655.

Vilain, M., Kautz, H. and Van Beek, P. (1990). Constraint propagation algorithms for temporal reasoning: a revised report. in Readings in Qualitative Reasoning about Physical Systems, pp.377-382, Morgan Kaufmann Publishers Inc.

Worboys, M. and Duckham, M. (2006). Monitoring qualitative spatiotemporal change for geosensor networks. International Journal of Geographical Information Science, 20(10), pp.1087-1108.

2 A Brief History of Time: The Reasoning and Visualization Perspectives

2.1 Introduction

This chapter includes a review of the important literature and enlightening work in the areas of temporal reasoning and temporal visualization, which are closely related to the research of this thesis. Writing this review was a beneficial experience, which offered me a comprehensive overview of the development trajectories, current trends and future challenges in these areas. It also assisted me to conclude how the piece of work in this thesis can fit into the big picture of the background research. The remainder part of this chapter starts with a discussion of diverse notions of time. Next, it reviews the theoretical cornerstones of time representation and reasoning in computer science and artificial intelligence (AI). Following this, the emphasis will be placed on the visualization of temporal data and then the visualizations of temporal data in a geographical information system (GIS).

2.2 Notions of Time

The representations of time in human activities are derived from human perception about the changes in the world. Due to the different viewpoints on real-world phenomena, humans came up with diverse conceptualizations of time. In ancient times, the perception of time is mostly based on the observation of natural events, such as the alternation of day and night, the lunar phases and the cycles of changing seasons. Time is thus conceived as a cyclic notion in many ancient civilizations. Later, the inventions of hourglass, candle and mechanic clock enable people to measure and describe the flow of time. This gave rise to the linear view of time, which depicts time as a flow of events that moves from past towards the future in a linear manner. While the linear time becomes dominant in the modern sciences, the cyclic time still retains its importance in many aspects of our live. Many intelligent means have been invented to deal with the coexistence of the two competing views. On the one hand, the integrated concepts portrait time as a wave (Li and Kraak, 2008) or spiral (Hewagamage et al., 1999; Weber et al., 2001) that moves forward in cyclic patterns. On the other hand, the natural cycles in time (e.g. day, week and year) can be treated as units for the aggregations of events along the time line (Silvestrini and Veredas, 2008). Besides cyclic and linear time, the notion of branching time concerns about different possibilities of temporal processes, which is often used in concurrent system modelling.

Another important distinction can be made between absolute and relative time. The absolute time is considered as a container of events, which flows independently of human consciousness. In contrast, the relative view of time claims that the universe is made up of a relative sequence of events within a frame of reference. In the area of physics, Newton established the classic mechanics with the absolute view of time, which works well when dealing with things like falling apples that travel comparatively slowly. However, it has been challenged by things that move near the speed of light. This problem has been effectively tackled by the theory of special relativity of Einstein, which postulates that time is dependent on the relative speeds of the observers' reference frames (Einstein, 1961; Hawking, 1988). Besides the debates on the nature of time in the fundamental physics, in practical, whether time should be modelled absolutely or relatively depends on the tasks that need to be tackled. Generally speaking, absolute time is machine-oriented as it is measured by clocks and can be easily processed by computers. Therefore, computer engineering is in favour of absolute time because the communications between different functional elements within a distributed system relies on the synchronization according to a universal clock. Relative time is human-oriented because human beings do not have a precise internal clock and their perception of flowing time is based on the observation of external events. This is reflected in many common expressions in human languages, such as after, before and meanwhile. Thus, in information science and AI, which focus on higher-level abstractions of human knowledge, time is usually modelled in the relative manner and qualitative reasoning is a more important issue.

There exist other controversies about the notion of time, for instance, whether the primitive of time is instant or interval, or both of them. In the first view, time is modelled as an infinite and dense set of instants, which is similar to the set of real numbers. Every event is composed of infinite instantaneous actions and states. Other researchers (Allen, 1983) claim that intervals should be considered as the primitive of time since it is more close to common-senses temporal concepts such as duration of an event and existence period of an object. In this view, a time instant can be considered very short intervals. However, solely using interval as the time primitive loses the expressivity for instantaneous temporal concepts, such as the event or action that trigger transitions between different states. For instance, the action of turning off the light is intuitively instantaneous, which is just in-between the two continuous states that the room is illuminated and the room is dark. To express both the instantaneous and durative temporal concepts, some researchers take the neutral way and argue that both interval and instant can be considered as the primitives of time. However, the coexistence of instant and interval leads to the contradiction of the belongingness of the transitional instant, which is called the Division Instant Problem (DIP). If the action of turning of the light is an instant, whether it belongs to the previous interval in which the room is illuminated, or the later interval in which the room is dark, or is independent from both of them becomes a tricky issue (Van Benthem, 1983).

Moreover, controversial views of time exist in many other aspects such as the structure of time, finiteness of time and boundaries of time primitives (Schreiber, 1994; Vila, 1994). These different views of time arise in daily life and sciences because there are different levels of abstraction, different viewpoints, and different ways of thought. It is unnecessary and also impossible to establish a consensus notion of time that meets the requirements of all different applications. As a result, in most disciplines, the handling of time is fairly task-oriented. In spite of the lack of top-level consensus on what is time, most time-related issues can be effectively and efficiently tackled with a variety of domain-specific representations of time.

2.3 Temporal Reasoning

The problem of representing and reasoning about temporal knowledge arises in wide range of disciplines, including computer science, philosophy, psychology, and linguistics. This chapter takes special care of this problem in the context of computer science and AI. In general, the research of temporal reasoning aims at formalizing common-sense temporal knowledge in a way that can be reasoned by computer systems. Over the past decades, considerable work has been done in temporal reasoning in order to meet the increasing demands from a variety of applications such as planning with durative actions (Coddington, 2002), specification of clinical guidelines for temporal symptoms (Augusto, 2005), handling of time in temporal databases (Ozsoyoglu and Snodgrass, 1995), processing temporal information in human language (Leith and Cunningham, 1997). The emphasis of this area is on the qualitative reasoning of temporal knowledge, which is derived from the tradition how human conceptualise the physical world. The remainder of this section first reviews a number of well-known temporal logic formalizations of temporal knowledge. Then it reviews the algebra of temporal primitives. At the end of this section, these reviewed approaches are categorised in Table 2-1, according to a set of criteria.

2.3.1 Temporal Logics

The formalization of knowledge in an inferable logic framework is of vital importance for reasoning by computers. The term temporal logic is often narrowly referred to the modal logic system that was originally introduced by Arthur Prior in the 1950s (Prior, 1957). This system extended the proportional logic with temporal modal operators. The four classic operators include *P* (the proposition is true in some past time), *F* (the proposition is true in some future time), *H* (the proposition is true in all past time) and *G* (the proposition is true in all future time). Later, in 1968, he refined this work and summarized various temporal logics by postulating a glossary of time axioms based on modal operators in his book “Past, Present and Future” (Prior, 1967). The first application of MTL in reasoning about computer programs appears in the landmark paper of Pnueli (1977), which argued that Temporal Logic could be a useful formalism for specifying and verifying correctness of computer programs. Afterwards, many extensions and further developments have been done on the basis of modal temporal logic. In addition to the traditional qualitative operators, attempts have been made in incorporating metric temporal operators into modal temporal logic. Koymans (1990) proposed metric operators to express quantitative temporal reference of propositions, such as “Proposition *A* will be true in an interval of a certain length since now”. Reichgelt (1989) defined a modal logic that incorporated the operator *AT*(*t*), where *t* is a temporal constant, to express that a formula is true with reference to time *t*. While most modal temporal logics interpret formulas over time instants, some work has been down in modal logics based on time intervals, e.g. (Halpern and Shoham, 1991).

In general, modal temporal logic is regarded as a relative representation of temporal knowledge, since all propositions are referenced to a benchmark moment, which could be in the past, at present or in the future. An alternative approach that represents time as an additional argument in a first order predicate appeared to be a better treatment of absolute time (Haugh, 1987). In this logic, one can express “Mike met John at 10:00am” as a three-argument predicate *Meet* (Mike, John, 10:00am). In AI, the Situation Calculus (McCarthy and Hayes, 1968) and the recent work by Bacchus et al (Bacchus et al., 1991) are examples of this approach. While this approach can link an event with an absolute time instant in a very natural way, it cannot express specific statuses of assertions in time. For example, if the proposition "Mike drives" is true in the interval between 10:00am to 11:00am, it is also true at every subinterval between 10:00am and 11:00am. However, if "Mike drives from Ghent to Brussels" is true between 10:00am and 11:00am, it may not be true in a subinterval between 10:00am and 11:00am, because this action is only partially finished in the subinterval. To address this problem, reified temporal logic is proposed, in which one predicate can be generalised as an argument of the other predicate. In this kind of logic, a non-temporal predicate is linked to a temporal concept by a meta-predicate in the form of *TemporalOperator* (atemporal predicate, time). This way, the proposition "Mike drives from Ghent to Brussels between 10:00am and 11:00am" can be formalized as *Occur* (Drive (Mike, Ghent, Brussels), (10:00am, 11:00am)), in which *Occur* is the temporal operator that expresses the particular linkage between the event and the time interval. Reification has significantly increased the expressive power of the classic predicate logic, and thus, has been used in a wide range of temporal logic systems (Allen, 1984; Ma and Knight, 1996; McDermott, 1982). An issue of this approach is that the de-reification of a proposition can easily affect its validity. For instance, the proposition "John met Mike in Ghent at 10:00am" cannot be simplified as "John met Mike at 10:00am", which is a natural inference from the previous one. Galton (1991) argues that the temporal reification is "philosophically suspect and technically unnecessary", and have tackled this problem by specifying a temporal reference to every elementary predicate with a token. This idea is a part of the Event Calculus by Kowalski and Sergot (1986). Using this approach, the proposition "John met Mike in Ghent at 10:00am" can be represented as *Meet* (John, Mike, *e*) $\wedge$ Time (*e*, 10:00am) $\wedge$ Place (*e*,Ghent). In this combined proposition, *e* is the token of the *Meet* predicate, which is independently referenced to to the spatial and temporal elements by two other predicates. Removing either element does not affect the validity of the remaining part of the proposition.

Nowadays temporal logic has become a vast and active research area in AI and computer science. The trade-off between the expressivity and the computational efficiency is the key issue for the development of temporal logics and reasoners. Another typical challenge is to indicate the objects that are unaffected by irrelevant events or actions instead of explicitly formalising all the irrelevance. This issue is usually referred to as the Frame Problem (McCarthy and Hayes, 1968). Besides this, the non-determinism is an important issue in concurrent programs (Pnueli, 1981), which triggered the invention of systems based on the branching time model. Comprehensive reviews of the implementational issues in temporal logics can be found in (Galton, 1995) and (Kröger, 1987).

2.3.2 Temporal Algebra

This section reviews time algebra, which is another important topic in temporal reasoning. Time primitive is the fundamental of time algebra. In the early work, instant has been taken as the primitive of time and the relations between instants include *before*, *equal* and *after* (Bruce, 1972). However, only instant cannot sufficiently represent durative temporal concepts such as processes and events that last for a while. In 1983, Allen defined an interval algebra that includes thirteen mutually exclusive relations. These relations or their combinations can be used to express any relationships that can hold between two intervals. Allen also summarized the transitivity between every pair of the thirteen relations, and proposed an algorithm to add new temporal relations into an existing knowledge system according to the transitivity. Allen's work has triggered a variety of research on interval algebras (Allen and Hayes, 1985; Galton, 1990; Ladkin, 1987; Schockaert et al., 2008). Freksa (1992) has introduced the notion of conceptual neighbourhood based on Allen's relations. He also extended the thirteen relations into twenty-nine relations that can also applied to semi-intervals, namely, intervals with unknown start points or end points. In order to represent instantaneous events, Ma and Knight (1994) incorporated time instants into Allen's interval algebra, where interval-interval, interval-instant, instant-interval and instant-instant relations have been formalised. Most work of temporal algebra is based on linear time, except the cyclic interval algebra proposed by Osmani et al.(1999). The research on temporal algebra has even inspired the research on regional algebra for spatial reasoning (Cohn et al., 1997; Freksa, 1990; Güsgen, 1989). Regarding computational efficiency, Vilan et al. (1990) claimed that the interval algebra is NP-hard and suggested several strategies to increase the efficiency of the computation. One of the suggestions is translating interval algebra into point algebra due to the computational advantages of point algebra. This translation is carried out by representing an interval as a pair of instants, which has been applied in many reasoning systems (Dechter et al., 1991; Van Beek, 1989).

In addition to crisp time intervals, many disciplines are faced with intervals that do not have an exact start points and end points, which is generally called imperfect time intervals. The theories of fuzzy set and rough set are two common approaches to handle such time intervals. Bittner (2002) noticed that the uncertainty of time intervals is often caused by approximate descriptions in a course granularity. To tackle this problem, he proposed a rough set approach to derive the possible relations between intervals from the descriptions in a coarse granularity. Besides this, Bassiri et al. (2009) directly modelled an imperfect interval as a one-dimensional rough set, which is called rough time interval, and proposed an incomplete algebra of rough time intervals. In his approach, an arbitrary instant in the time line can be categorized to 'definitely in', 'possibly in' or 'definitely not in' a rough time interval. Later, Qiang et al. (2010) refined the algebra of rough time intervals using a diagrammatic representation and systematically studied the semantics behind these relations. On the other hand, the fuzzy set approach has been applied to represent imperfect intervals and relations. Dubois and Prade (1989) use fuzzy sets to represent the possibility distributions of ill-known events. For example, the event 'John leaves for vacation around mid-June' can be defined as a possibility distribution on the time line, and then the statement 'John leaves for vacation at 13 June' has a certain possibility between zero and one. In their approach, a fuzzy time interval derives from the possibility distributions of the start point and end point. The other approaches (Nagypál and Motik, 2003; Ohlbach, 2004; Schockaert and De Cock, 2008) focus on the events with inherent fuzziness. For example, one cannot define a specific interval for the Middle Ages, because it has a gradual starting and ending process. In these approaches, whether a time point belongs to the imperfect interval is quantified as a real number between zero and one according to a membership function. The relations between two fuzzy time intervals are also expressed by degree of truthiness between zero and one. The formula that derives temporal relations between fuzzy intervals is the major distinction among these fuzzy models. While each of them has strong points and weak points, either in the expressivity or computation efficiency, Schockaert's approach shows better performance in major reasoning aspects, including transitivity, reflexivity and symmetry (Schockaert and De Cock, 2008; Schockaert et al., 2006; Schockaert et al., 2008).

Table 2-1 categorises all approaches reviewed in Section 2.3. It becomes clear that the temporal logics have diverse focuses on the competing views of time interval and time instant, absolute time and relative time. While most logics treat time linearly, a few approaches can also deal with branching processes in concurrent systems. First order logic is the dominant logic base for most approaches. However, whether reification is optimal is still a controversial issue. The research about temporal algebra has an obvious emphasis on time interval against time instant. The relations and transitivity of relations between time intervals is the major issue of temporal algebra. It is noticeable that the algebra of indeterminate time intervals has received incredible attention since the 21st century.

Table 2-1: Categorization of the temporal logics and algebra reviewed in Section 2.3.

		Time ontology									Logic system			
		Time primitive		Time reference		Time arrangement			Determinacy		Logic framework		Reification	
		Interval	Instant	Relative	Absolute	Linear	Branching	Cyclic	Determinate	Uncertain	Modal logic	First-order logic	Yes	No
Temporal logics	McCarthy and Hayes (1968)		■		■	■			■			■		■
	Pnueli (1977)		■	■		■			■		■			■
	Allen (1984)	■		■	■	■			■			■	■	
	Kowalski and Sergot (1986)	■	■	■	■	■			■			■		■
	Haugh (1987)	■	■		■	■	■		■			■		■
	Reichgelt (1989)		■	■	■	■			■		■			■
	Koymans (1990)		■	■		■			■		■			■
	Halpern and Shoham (1991)	■			■	■	■		■		■		■	
	Galton (1991)	■			■	■			■			■		■
	Bacchus (1991)	■	■		■	■			■			■		■
	McDermott (1992)		■	■			■		■			■	■	
	Ma and Knight (1996)	■	■	■	■	■	■		■			■	■	
Temporal algebra	Bruce (1972)		■	■		■			■					
	Allen (1983)	■		■		■			■					
	Dubois and Prade (1989)		■	■		■				■				
	Freksa (1992)	■		■		■			■					
	Ma and Knight (1994)	■	■	■		■			■					
	Osmani et al.(1999)	■		■				■	■					
	Bittner (2002)	■		■		■				■				
	Nagypál and Motik (2003)	■		■		■				■				
	Ohlbach (2004)	■		■		■				■				
	Bassiri et al. (2009)	■		■		■				■				
	Schockaert et al. (2008)	■		■		■				■				
	Qiang et al. (2010)	■		■		■				■				

2.4 Temporal Visualization

Analysis of temporal data plays a very important role in many disciplines. This section focuses on visualizations of temporal data, which has been proven to be an effective analytical tool in many areas. Due to the sophisticated nature of time, visualization approaches for temporal data appear to be highly diverse and customized according to particular analytical tasks. A wide variety of techniques have been applied to visualise temporal data, including traditional static diagrams, innovative visual metaphors, interactive queries and manipulations. The remainder of this section first reviews the traditional approaches that are based on static diagrams. Second, it introduces the innovative approaches that rely on the advanced computational and graphic capabilities of computers. Third, it focuses on the visualization of spatio-temporal data, which has recently become a burgeoning realm of geographical information science (GIScience). Certainly, an exhaustive enumeration of all visualization approaches is impossible. Therefore, I select the landmarks and the representative ones to the best of my knowledge. In the end of this section, all the reviewed visualization approaches will be categorised in a table, according to a set of criteria.

2.4.1 Traditional Approaches

As time is something perceivable but invisible, the visual representation of time is a key issue in the analysis of temporal information. In the past, visualizations of temporal data were limited to diagrams on paper. Despite this limitation, there are still many remarkable pieces of work that have greatly influenced the subsequent development of visualization techniques. In 1786, the line chart was invented by William Playfair (1786) to represent economic time series (Figure 2-1a). The line chart follows the linear notion of time and represents time series as curves of varying height above the time axis. One of William Playfair's comments about the line chart is profound:

As the eye is the best judge of proportion, being able to estimate it with more quickness and accuracy than any other of our organs, it follows, that wherever relative quantities are in question … this mode of representing it is peculiarly applicable; it gives a simple, accurate, and permanent idea, by giving form and shape to a number of separate ideas, which are otherwise abstract and unconnected.

In addition to line charts with a single variable, Playfair made a further step to arrange multiple line charts in one layout to illustrate the commercial statistics of multiple countries (Figure 2-1d). This approach revealed another fundamental advantage of the line chart, i.e. easy and rapid comparison among a large number of variables. Another prominent line chart was created by a French civil engineering Charles Joseph Minard in 1869 (Figure 2-1c). This chart presents an excellent overview of the size and trajectory of the French army that invaded Russia and retreated back to Western Europe. The line chart and its variants have been used for centuries and are still prevalent in the state-of-the-art visualization systems. The applications of the line chart have spread to a wide range of disciplines such as economic and financial analysis, signal analysis, system monitoring and climate data visualization.

Another approach that has profound effect is the timeline. The use of timeline can be at least dated back to 1765, when Joseph Priestley created a chart in which lifetimes of empires were represented as linear segments (Figure 2-1b) (McLachlan, 1987). Priestley believed that this chart would allow students to “trace out distinctly the dependence of events to distribute them into such periods and divisions as shall lay the whole claim of past transactions in a just and orderly manner” (McLachlan, 1987). In the early 1910s, Henry Gantt designed a similar chart to illustrate the elementary processes and activities in a project schedule. This kind of chart is named after his name, i.e. Gantt chart. Both the timeline and Gantt chart aim to visualise time intervals of events and provide an image of the temporal distribution of events. Their applications can be found in many areas and the representation of time intervals as linear segments becomes a convention.

Besides visualizations for linear time, people have also attempted to visualize temporal data with a cyclic structure. The most classic example is the Coxcomb drawn by Florence Nightingale in 1858, in which the mortality of the British Army of different causes is represented in a cyclic fan chart (Figure 2-1e). Each fan has the same angle and represents a specific month of the year. The distance between the edge of the fan sector and the centre of the circles indicates the number of the mortalities. From this chart, people can easily observe the fact that the most deaths of the army were caused by the infections of the Preventable or Mitigable Zymotic diseases, which had an outbreak between December 1854 and June 1855.

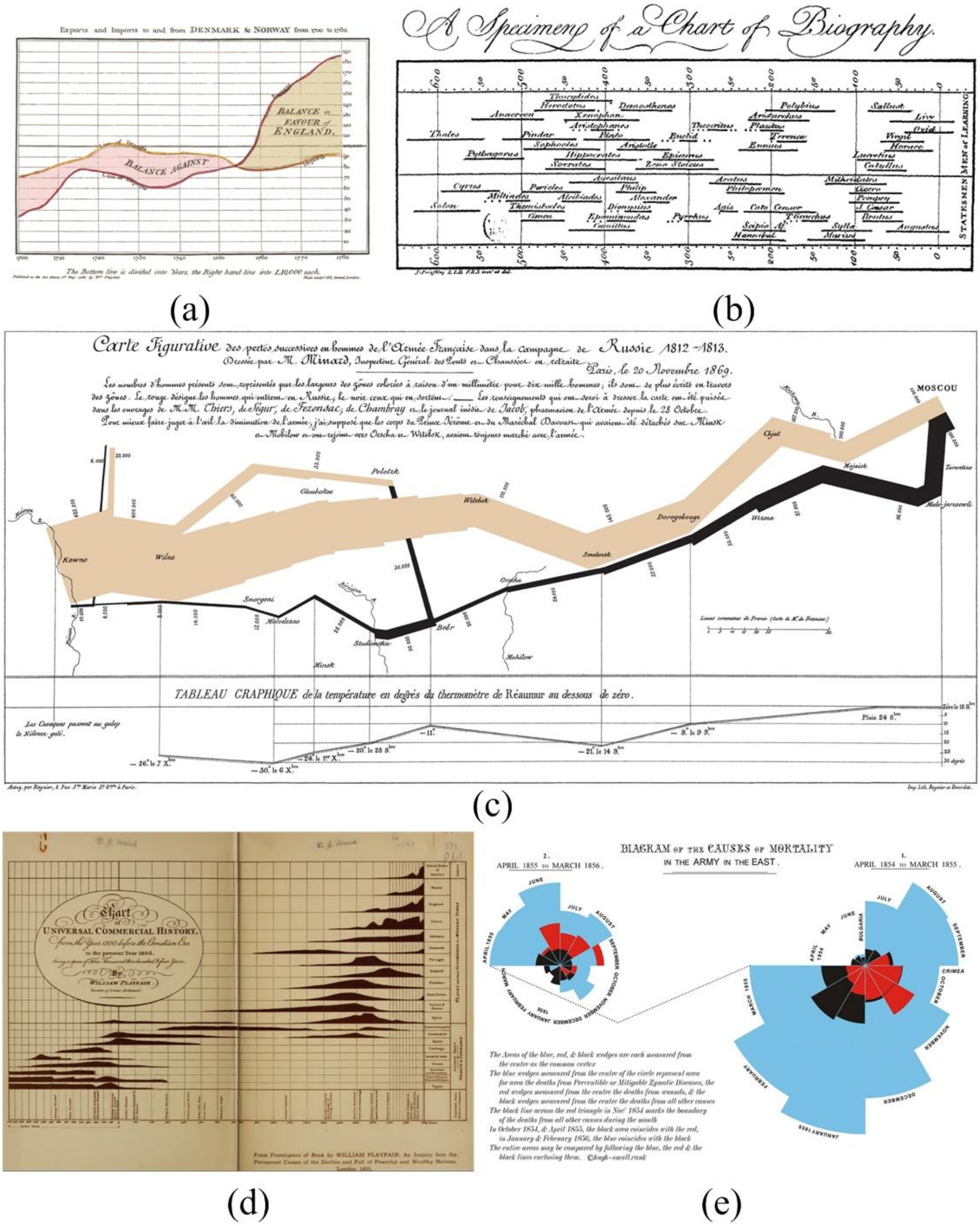


Figure 2-1: Traditional approaches of temporal visualization. (a): William Playfair’s line chart of the export and import to and from Denmark & Norway from 1700 to 1780 (Playfair, 1786). (b): A redacted version of Joseph Priestley's Chart of Biography, Cited in (McLachlan, 1987). (c): Charles Minard's flow map of Napoleon's March, cited in (Tufte, 2001). (d): William Playfair’s chart of Universal Commercial History (Playfair, 1805).(e): Florence Nightingale’s diagram of the causes and mortality in the army in the East, cited in (Funkhouser, 1937).

2.4.2 Innovative Approaches

The invention of computer has liberated people from manually chart drawing by automatic graphic techniques. The advanced graphic ability of computers has greatly stimulated people's inspiration of creating new visualization approaches for temporal data. Some approaches extend the aforementioned conventional diagrams with interactive functionalities, while the others are based on newly-invented visual metaphors, particularly those 3D metaphors that are hard to create and display without the help of computers. Time series is the data type that receives most attention. While the line chart is still frequently used in many visualization systems, the new display and printing techniques allow the use of large amounts of colours and raster images in visualizations. Compared with traditional line charts, colour lines (Kincaid and Lam, 2006; Matkovic et al., 2007) are considered as a more efficient visualization for a large number of time series (Figure 2-2i). It uses a straight line with a certain colour scheme to represent a time series so that numerous time series can be compactly arranged within a rectangular region. With the support of interactive sorting tools, users are able to visually identify clusters of time series with similar patterns. Other than using a single colour scale, the two-tone pseudo colouring (Saito et al., 2005) technique can visualise both the overview and details of a very large linear dataset in a single image (Figure 2-2j). The other well-known approach is ThemeRiver, which uses a continuous flow of colour bands to represent multiple time series (Havre et al., 2000). In Figure 2-2a, each band in the ThemeRiver represents a specific theme and the bandwidth indicates the number of its appearance in selected documents at a certain point of time. The ThemeRiver combined the advantages of the line chart and the bar chart, allowing users to observe the variation of individual time series and the difference among multiple series as well.

Integrating other visualization component with line charts can bring new insights into temporal data that are hard to observe merely within the line charts. For instance, Van Wijk & Selow (1999) developed a tool that combines a line chart and a calendar, in order to display pattern clusters of time series (Figure 2-2b). The line chart visualizes the averages of different time series clusters, while the colour of the dates in the calendar indicates which cluster the dates belong to. This system enables visualization of time series in two different time scales, i.e. daily and hourly, and is particularly interesting to temporal data that have strong links to the pattern of human activities. Besides, Lin et al. (2004) developed a visualization tool (VisTree) that combines a tree representation with line charts for visual detection of special patterns in time series (Figure 2-2e). In the tree representation, every branch represents a specific subsequence of time series and the thickness of the branch indicates the occurrence of identical subsequences. Using this approach, the frequently appeared patterns can be visually recognized from the thick branches.

Some tools are specially designed to visualize the time series with cyclic characteristics. Most of them derive from the idea of the spiral graph, which portraits time as a spiral line flowing from the inside out. For instance, Carlis & Konstan (1998) developed a visualization tool that represents monthly data as dots of different sizes or bars of different heights along a spiral line (Figure 2-2c left). Weber et al's tool (Weber et al., 2001) represents time series as spirals with gradual colours (Figure 2-2c right). In these spiral visualizations, the linear variation of time series can be observed along the spiral, while the periodical patterns can be detected in the straight line that extends from the centre of the spiral outward. On the other hand, Li and Kraak (2008) proposed an intermediate approach between the time line and the spiral, which depicts the time line as a wave (Figure 2-2d). A promising feature of this approach is that the time circles of different scales (e.g. yearly and monthly) can be displayed in one diagram. The other two intriguing approaches are the TimeWheel and the MutliCombo invented by Tominski et al. (2004) (Figure 2-2h). In TimeWheel, the time axis is arranged in the centre and surrounded by axes of time-dependent variables. The values in the attribute axes and the time point in the time axis are linked by lines. The patterns of the variables are thus expressed by the density of the lines. The idea of the MultiCombo is relatively simple, which arranges multiple line charts around a circle or star glyph.

Besides time series, many tools of timeline have been developed for the purpose of patient record analysis, schedule management and visualization of chronological events. Thanks to the advancement of computer graphics, these timelines evolve from static diagrams into interactive user interfaces, which support a variety of functionalities for data manipulation and visualization. For instance, the LifeLines (Plaisant et al., 1998) system has gained considerable success in assisting physicians with clinical tasks and medical decision-making (Figure 2-2f). In this system, the physicians are able to reorganize, reformulate data into hierarchical, problem-centric views, and render the display to a target presentation. Some approaches are developed to represent uncertain intervals of events and plans. Kosara and Miksch (Kosara and Miksch, 2001) uses the time annotation glyph to visualise medical treatment plans with flexible durations. Moreover, PlanningLines is a planning tool that has capability of visualizing and scheduling uncertain activities (Aigner et al., 2005) (Figure 2-2g).

2.4.3 Spatio-Temporal Visualization

Due to the increasing demands of analysing dynamics of spatial processes, considerable amount of work has been done in the representation of time in GIS and the ability of visualizing time becomes a common feature of up-to-date GIS. On the one hand, the visualization of time in GIS is greatly dependent on the conceptualization of the space and time. Time can either be considered as an attribute or a reference of an entity, resulting in various spatio-temporal data models and corresponding visualization strategies. On the other hand, the visualization of time in GIS is driven by the analytical questions that need to be answered. According to Peuquet (1994), analytical questions can be asked from the space, time and thematic points of view, which leads to a triad representational framework of spatiotemporal data. In order to effectively answer these analytical questions, GIS researchers have endeavoured to incorporate visualizations of information from these different perspectives. Andrienko et al. (2003) further categorised analytical questions into the general or elementary level, for identification or comparison tasks. Accordingly, GIS tend to support visualization in multiple temporal scales and comparisons between different time frames. With several decades' development, there have been a glossary of techniques to visualise time in GIS, aiming at different sorts of spatio-temporal processes.

Cartographical Approaches

As geovisualization originates from cartography, the conventional visualization of time in GIS is based on cartographic metaphors. While the static maps are mainly used to represent static view of the world, some strategies can be employed to represent temporal dynamics. One strategy is representing time with cartographical variables such as symbols, colours, texture and labels (Vasiliev, 1997). For instance, the movement of an object in map can be represented by a set of points, each of which has a time label that denotes the location of the object at a certain time stamp. Massive movements can be represented by flow maps (Figure 2-3b), which show abstractions of movement clusters instead of individual movements (Andrienko, G and Andrienko, N, 2008; Kraak and Ormeling, 2003). The approach of space-time composition (Langran, 1993) represents regions with different historical states as multiple layers and compose them into one map. The informative icons, which are usually called temporal glyphs, can represent more complex temporal information such as time intervals and positioned time series (Monmonier, 1990). More recently, the development of 3D geovisualization techniques enable the 3D glyphs to be applied to display multivariate (Figure 2-3c) and cyclic time series (Figure 2-3d) (Tominski et al., 2005). The major drawback of cartographic metaphors is that the messages conveyed by the representations are subjective to the intentions of cartographers, which, therefore, is rather limited in exploratory analysis. Another strategy is to use a sequence of discrete snapshots to express the changes in time. These maps are either arranged in a structured layout that readers can easily observe the changes (i.e. small multiples Figure 2-3a, see (Tufte, 1983)), or can be animated (Monmonier, 1989). This snapshot approach provides an objective presentation of the dynamics, allow readers to detect the phenomena of their own interests. However, its disadvantage is that the missing information between snapshots can lead to wrong interpretation of the spatio-temporal phenomena.

The Space-Time Cube

The most famous work in spatio-temporal representation is probably the space-time model introduced by Hagerstrand in late 1960s (Hägerstrand, 1970). The core concept is the space-time cube that adds the time axis as the third dimension on top of a 2D spatial plain. On the one hand, this concept has filled the gaps between discrete snapshots and represents the movement of objects in space and time in a continuous manner. Using this approach, trajectories of human activities or moving objects can be explicitly visualised in a 3D space (Figure 2-3e). The classic space-time cube is based on an absolute and linear time frame. To visualise trajectories with numerous cyclic patterns, the time frame of the space-time cube can also be transformed to relative and cyclic, so that temporal cycles in a trajectory can be better compared (Andrienko, G and Andrienko, N, 2011). On the other hand, the space-time cube provides a perfect integration of space and time (Figure 2-3f), which greatly facilitates the observation of the spatio-temporal distribution of discrete entities (Gatalsky et al., 2004; Huisman et al., 2009). Moreover, the advancement of graphics capability of computers has promoted the applications of 3D visual metaphors in GIS (Döllner, 2005). Many interactive functionalities, such as flexible browsing, selection, zooming in the spatial and temporal dimensions, movable plane, adaptable transparency and multiple views have release the potential of space-time cube in exploratory spatio-temporal analysis

(Kapler and Wright, 2005; Kraak, 2003). The invention of the space-time cube has overcome some inherent difficulties of 2D cartography. However, it has not led to a thorough revolution of geovisualization from 2D to 3D. The space-time cube has some limitations that derive from the typical problems of 3D visualization (Kjellin et al., 2009; Lind et al., 2003; Ware and Franck, 1996; Ware and Mitchell, 2008). For instance, the changes of view points and angles may cause variations of the structure of the 3D display, which may result in a biased perception of the data being displayed. Also, the objects in front obstruct the objects behind them. This problem is particularly serious when a large amount of objects are displayed in the space-time cube.

Linked Views

In some cases, multiple 2D visualisations lead to better performance than integrating space and time in the 3D environment. For example, the combination of a map view with line charts (Figure 2-3g) or parallel coordinates (Figure 2-3h) can be useful for the analysis of time series that are referenced to spatial locations (Andrienko, G and Andrienko, N, 2005; Jern and Franzen, 2006). Because the time series are plotted in a coordinate space, it is possible to recognise and compare the patterns within a large amount of time series. The combination of a map view with a bar chart or a calendar view can visualise the number of spatio-temporal events aggregated in days and weeks (Fredrikson et al., 1999). A common feature of such linked views is Linked-Brushing. Namely, when the user makes a selection in one view, the corresponding items in the other views are instantaneously highlighted. The response time has to be minimal so that one view can immediately react to the user's manipulation in the other view.

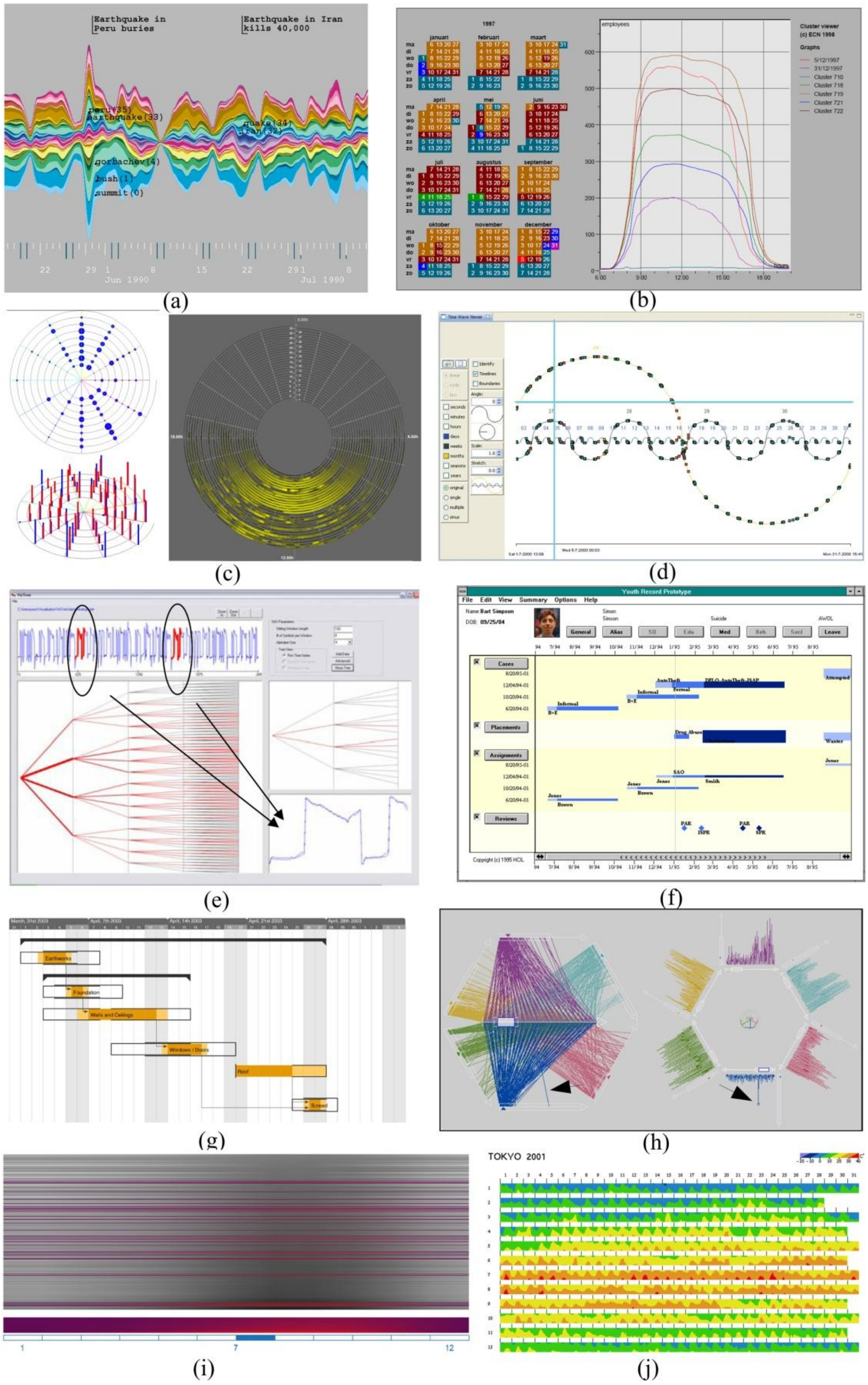


Figure 2-2: Innovative approaches of temporal visualization. (a): ThemeRiver (Havre et al., 2000). (b): Calendar-based visualization (Van Wijk and Selow, 1999). (c): Time spirals (Carlis and Konstan, 1998; Weber et al., 2001). (d): TimeWave (Li and Kraak, 2008). (e): VisTree (Lin et al., 2004). (f): LifeLines (Plaisant et al., 1998). (g): PlanningLines (Aigner et al., 2005). (h): TimeWheel and TimeCombo (Tominski et al., 2004). (i) Colour lines (Kincaid and Lam, 2006; Matkovic et al., 2007). (j) two-tone pseudo colouring (Saito et al., 2005).

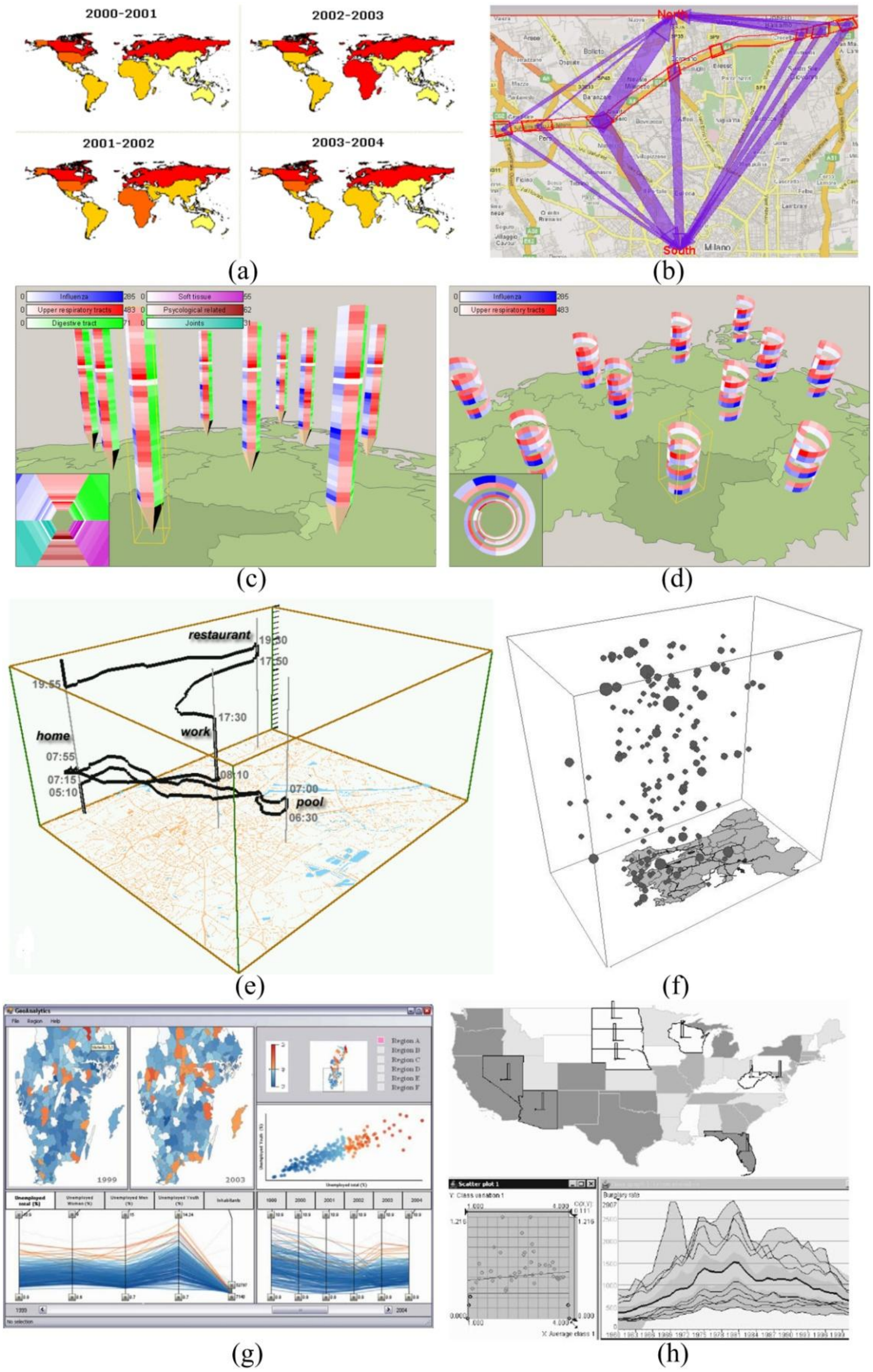


Figure 2-3: Spatio-temporal visualizations. (a): Small multiples (Tufte, 1983). (b): Flow map, cited in (Andrienko, G and Andrienko, N, 2008). (c): Pencil glyphs on 3D map (Tominski et al., 2005). (d): Helix glyphs on 3D map (Tominski et al., 2005). (e):Trajectories in space-time cube, cited in (Kraak, 2003). (f): Discrete events in space-time cube, cited in (Gatalsky et al., 2004). (g): Maps linked with parallel coordinates visualization (Jern and Franzen, 2006). (h): Maps linked with a line chart (Andrienko, G and Andrienko, N, 2005).

Table 2-2: Categorization of the visualization approaches reviewed in Section 2.4.1 - 2.4.3.

			Time				Data						Visualization			
			Time primitive		Time arrangement		Reference frame		Determinacy		Variable		Data display		Dimensionality	
			Interval	Instant	Linear	Cyclic	Spatial	Non-spatial	Determinate	Uncertain	Univariate	Multivariate	Standalone view	Linked views	2D	3D
Temporal visualization	Traditional approaches	Timelines (Priestley, 1765)	■		■			■	■		■		■		■	
		Line chart (Playfair, 1786)		■	■			■	■		■	■	■		■	
		The coxcombs (Nightingale, 1858)	■			■		■	■			■	■		■	
		The Chart of Napoleon's March (Minard, 1869)		■	■		■	■	■		■		■		■	
		Gantt chart (Gantt, the early 1910s)	■		■			■	■		■		■		■	
	Innovative approaches	Icons on spiral (Carlis & Konstan, 1998)	■			■		■	■		■	■	■		■	■
		LifeLines (Plaisant et al., 1998)		■	■			■	■		■	■	■		■	
		Calendar-based view (Van Wijk & Selow, 1999)	■	■	■			■	■		■			■	■	
		ThemeRiver (Havre et al., 2000)		■	■			■	■			■	■		■	
		Time Annotation glyph (Kosara & Miksch, 2001)	■		■			■		■	■	■	■		■	
		Spiral colour line (Weber et al., 2001)		■		■		■	■		■	■	■		■	
		TimeWheel (Tominski et al., 2004)		■		■		■	■			■	■		■	
		MultiCombo (Tominski et al., 2004)		■		■		■	■			■	■		■	
		VisTree (Lin et al., 2004)		■	■			■	■		■			■	■	
		PlanningLines (Aigner et al., 2005)	■		■			■		■	■		■		■	
		Two-tone presudo colouring (Saito et al., 2005)		■	■			■	■		■		■		■	
		Colour lines (Kincaid & Lam, 2006; Matkovic et al., 2007)		■	■			■	■			■		■	■	
		Time Wave (Li & Kraak, 2008)	■	■	■	■	■	■	■			■		■	■	
Spatio-temporal visualization		Temporal animation		■	■		■		■			■	■		■	
		Small multiples (Tufte, 1983)		■	■		■	■	■		■	■	■		■	
		Flow map	■		■		■		■		■		■		■	
		Space-time cube (Kraak, 2003; Gatalsky et al., 2004; Huisman et al., 2009)	■	■	■		■		■		■	■	■	■		■
		3D icons on map (Tominski et al., 2005)		■	■	■	■		■			■	■			■
		Multiple linked views (Andrienko, N. & Andrienko, G. 2005; Jern & Franzen, 2006)		■	■		■		■			■		■	■	

Table 2-2 categorized the temporal visualization approaches reviewed in Section 2.4.1 - 2.4.3, which considers three key criteria: time, data and visualization. For each key criterion, further sub-criteria are introduced. From this table one can see that different time models have been treated by a variety of visualization approaches. It is noticeable that not many approaches have the ability of handling the uncertainty issue in temporal information. Recent approaches tend to use several linked views rather a standalone view, particularly those approaches that deal with spatio-temporal data. Although the advancement of computer graphics makes 3D visualization increasingly convenient, 2D approaches still play an important role in temporal visualization.

2.4.4 Current Trends and Future Challenges

In recently years, an increasing number of people have recognized the value of dynamic and interactive techniques in assisting data visualization and analysis. Interaction may help people to form a comprehensive mental image of the data, detecting patterns and relationships in the data, formalizing mental hypotheses and testing them. Almost all the aforementioned visualization systems apply some sorts of interactions. One of the most common interaction techniques is the browsing technique, which enable users to pan, scroll, rotate and navigate in the visualization in order to observe the data in different ranges and different levels of details. On the other hand, many systems support dynamic temporal queries that are carried out by intuitive and convenient manipulations of graphic controls. The sliding bar is a typical example, in which the position of the bar in the linear track indicates a certain time point. With the manipulation of such controls, the answers of the queries can be expressed by instantaneous changes of visualization variables in the data display, such as appearance and disappearance, the changes of transparency, colours and symbol sizes. Moreover, some queries are directly implemented in the data visualization, for instance, when clicking or selecting objects in one view, the corresponding information will be displayed or highlighted in the other views. The merit of querying in visualization is that queries can be specified in a view of the characteristics of the dataset being displayed.

Incorporating multiple views is a trend in current visualization systems. One reason is that no visual representation of time can exhibit all aspects of the temporal data. Multiple views can supplement each other and display the temporal data in different visual representations or different levels of granularities. The other reason is that the interactions between the views can obtain extra insights that are not easy to observe from independent static images. Dynamic and visual query is mutually supported among these views, which is referred to as “Linked Brushing”. When queries are made in one view, the corresponding information or changes will be reflected in the other views. With Linked Brushing, users are able to explore the datasets by successively making queries in different views and observing the visual answers in the other views. Not only implemented as standalone systems, temporal visualization is often integrated in other information systems, for example, geographical information systems, and acts as an assisting component for spatio-temporal analysis. Recent research has focussed on providing principles for multiple views and developing flexible frameworks that allow unforeseen combinations of visualizations (Boukhelifa et al., 2003; Roberts, 2007). This research is opening the way for users to customise the coordination and linkage of the views that suit their specific requirements (North, 2000).

Visualising large amounts of temporal data will be a common issue in the future, which does not only require faster graphic rendering techniques, but also improved visualization techniques that facilitate analytical tasks such as classification, comparison, clustering and pattern detection. These techniques should take the best of computational capacity of computers and human intelligence to make the analysis more effective and efficient. Also, the visualization of temporal data should be flexibly transferable between different levels of granularity, different ways of aggregations and abstraction, which can better support the analysis and decision making in different temporal frames and scales. Currently, the issues of uncertainties and vagueness in temporal data have not yet received enough attention in visualization area. How to deal with the uncertainties in temporal data should be taken into account by the future development of visualization systems. Moreover, the visualization approaches in the future will be more targeted on specialized analytical tasks, which is often referred to as expert systems. Novel visual metaphors and interaction techniques need to be created to tackle special phenomena in datasets. It is preferable that an appropriate guidance is provided to users, which specifies targeted data types and analytical tasks that the visualization systems aim to solve. Also, evaluations are needed to quantify the efficiency, effectiveness and appropriateness of these systems in solving these tasks.

2.5 Conclusion

This chapter has briefly reviewed the most important theories and techniques in the areas of temporal reasoning and visualization. From this review, it has become clear that the handling of time is one of the most important concerns in human activities. Particularly in the last several decades, the invention of computers and the rapid development of information technologies have provided great opportunities for people to represent and analyse temporal information with sophisticated characteristics. On the one hand, people have endeavoured to make computers able to infer and reason against temporal knowledge in the way like they do, leading to a rich body of theories in temporal reasoning. On the other hand, thanks to the advancement of processing and graphic capability of computers, people are able to create a variety of visualization techniques that facilitate their observation and analysis of temporal phenomena in an information-rich environment. Moreover, the concepts of space and time can hardly be separated, not only in terms of the fundamental theories of physics, but also in mthe research that deals with dynamic spatial processes. Therefore, incorporating time with space has become one of the most important tasks in GIScience. Ample evidence shows that the integrated or tightly linked visualizations of space and time can effectively help the analysis of spatio-temporal data, especially exploratory analysis that greatly relies on visual observation and human intelligence.

References

Aigner, W., Miksch, S., Thurnher, B. and Biffl, S. (2005). PlanningLines: novel glyphs for representing temporal uncertainties and their evaluation. *Proceedings of Ninth International Conference on Information Visualisation*, London, UK, pp.457-463.

Allen, J. F. (1983). Maintaining Knowledge about Temporal Intervals. *Communications of the ACM*, 26(11), pp.832-843.

Allen, J. F. (1984). Towards a general theory of action and time. *Artificial Intelligence*, 23(2), pp.123-154.

Allen, J. F. and Hayes, P. J. (1985). A Common-sense Theory of Time. *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, Los Angeles, California, pp.528-531, Morgan Kaufmann Publishers Inc.

Andrienko, G. and Andrienko, N. (2005). Visual exploration of the spatial distribution of temporal behaviors. *Proceedings of Ninth International Conference on Information Visualisation*, London, UK, pp.799-806.

Andrienko, G. and Andrienko, N. (2008). Spatio-temporal aggregation for visual analysis of movements. *IEEE Symposium on Visual Analytics Science and Technology, 2008. VAST '08*, pp.51-58.

Andrienko, G. and Andrienko, N. (2011). Dynamic Time Transformations for Visualizing Multiple Trajectories in Interactive Space-Time Cube. *Proceedings of the International Cartographic Conference (ICC)*, Paris, France.

Andrienko, N., Andrienko, G. and Gatalsky, P. (2003). Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing*, 14(6), pp.503-541.

Augusto, J. C. (2005). Temporal reasoning for decision support in medicine. *Artificial Intelligence in Medicine*, 33(1), pp.1-24.

Bacchus, F., Tenenberg, J. and Koomen, J. A. (1991). A non-reified temporal logic. *Artificial Intelligence,* 52(1), pp.87-108.

Bassiri, A., Malek, M., Alesheikh, A. and Amirian, P. (2009). Temporal Relationships between Rough Time Intervals. ed. by Gervasi, O., Taniar, D., Murgante, B., Laganà, A., Mun, Y. and Gavrilova, M., in *Computational Science and Its Applications – ICCSA 2009,* pp.543-552, Springer Berlin / Heidelberg.

Bittner, T. (2002). Approximate Qualitative Temporal Reasoning. *Annals of Mathematics and Artificial Intelligence*, 36(1), pp.39-80.

Boukhelifa, N., Roberts, J. C. and Rodgers, P. J. (2003). A coordination model for exploratory multiview visualization. *Proceedings of International Conference on Coordinated and Multiple Views in Exploratory Visualization,* London, England, pp.76-85, IEEE Computer Society, Los Alamitos, California.

Bruce, B. C. (1972). A model for temporal references and its application in a question answering program. *Artificial Intelligence,* 3(0), pp.1-25.

Carlis, J. V. and Konstan, J. A. (1998). Interactive visualization of serial periodic data. *Proceedings of the 11th annual ACM symposium on User interface software and technology,* San Francisco, California, United States, ACM.

Coddington, A. (2002). A Continuous Planning Framework with Durative Actions. ed. by Fisher, M. and Artable, A., *Proceedings of the Ninth International Symposium on Temporal Representation and Reasoning (TIME'02),* Manchester, UK, pp.108-115, IEEE Computer Society.

Cohn, A. G., Bennett, B., Gooday, J. and Gotts, N. M. (1997). Qualitative Spatial Representation and Reasoning with the Region Connection Calculus. *GeoInformatica,* 1(3), pp.275-316.

Dechter, R., Meiri, I. and Pearl, J. (1991). Temporal constraint networks. *Artificial Intelligence,* 49(1–3), pp.61-95.

Döllner, J. (2005). Geovisualization and Real-Time 3D Computer Graphics. ed. by Dykes, J., MacEachren, A. M. and Kraak, M.-J., in *Exploring geovisualization,* pp.325-344, Elsevier Ltd, Oxford, UK.

Dubois, D. and Prade, H. (1989). Processing fuzzy temporal knowledge. *IEEE Transactions on Systems, Man and Cybernetics,* 19(4), pp.729-744.

Einstein, A. (1961). *Relativity: The Special and the General Theory*. Three Rivers Press, New York.

Fredrikson, A., North, C., Plaisant, C. and Shneiderman, B. (1999). Temporal, geographical and categorical aggregations viewed through coordinated displays: a case study with highway incident data. *Proceedings of the 1999 workshop on new paradigms in information visualization and manipulation in conjunction with the eighth ACM internation conference on Information and knowledge management*, Kansas City, Missouri, United States, pp.26-34, ACM, New York.

Freksa, C. (1990). Cognitive and linguistic aspects of geographic space. ed. by Mark, D. M. and Frank, A. U., in *Cognitive and linguistic aspects of geographic space*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Freksa, C. (1992). Temporal Reasoning Based on Semi-Intervals. *Artificial Intelligence,* 54(1-2), pp.199-227.

Funkhouser, H. G. (1937). Historical Development of the Graphical Representation of Statistical Data. *Osiris,* 3, pp.269-404.

Galton, A. (1990). A Critical Examination of Allen's Theory of Action and Time. *Artificial Intelligence,* 42(2-3), pp.159-188.

Galton, A. (1991). Reified temporal theories and how to unreify them. *Proceedings of the 12th international joint conference on Artificial intelligence - volume 2*, Sydney, New South Wales, Australia, pp.1177-1182, Morgan Kaufmann Publishers Inc.

Galton, A. (1995). *Temporal Logics and their Applications*. Academic Press, London.

Gatalsky, P., Andrienko, N. and Andrienko, G. (2004). Interactive analysis of event data using space-time cube. *Proceedings of the 8th International Conference on Information Visualisation (IV 2004)*, London, UK, pp.145-152.

Güsgen, H. W. (1989). Spatial reasoning based on Allen's temporal logic. International Computer Science Institute Berkeley.

Hägerstrand, T. (1970). What about people in Regional Science? *Papers in Regional Science,* 24(1), pp.6-21.

Halpern, J. Y. and Shoham, Y. (1991). A propositional modal logic of time intervals. *Journal of the ACM,* 38(4), pp.935-962.

Haugh, B. A. (1987). Non-standard semantics for the method of temporal arguments. *Proceedings of the 10th international joint conference on Artificial intelligence - Volume 1*, Milan, Italy, pp.449-455, Morgan Kaufmann Publishers Inc.

Havre, S., Hetzler, B. and Nowell, L. (2000). ThemeRiver: Visualizing theme changes over time. *Proceedings of the IEEE Symposium on Information Vizualization 2000*, Salt Lake City, Utah, pp.115-123, IEEE Computer Society.

Hawking, S. (1988). *A Brief History of Time*. Bantam Dell Publishing Group, New York.

Hewagamage, K. P., Hirakawa, M. and Ichikawa, T. (1999). Interactive visualization of spatiotemporal patterns using spirals on a geographical map. *Visual Languages, 1999. Proceedings. 1999 IEEE Symposium on*, pp.296-303.

Huisman, O., Santiago, I. F., Kraak, M.-J. and Retsios, B. (2009). Developing a Geovisual Analytics Environment for Investigating Archaeological Events: Extending the Space-time Cube. *Cartography and Geographic Information Science*, 36(3), pp.225-236.

Jern, M. and Franzen, J. (2006). "GeoAnalytics" - Exploring spatio-temporal and multivariate data. *Proceedings of Tenth International Conference on Information Visualization (IV 2006)*, London, UK, pp.25-31, IEEE Computer Society

Kapler, T. and Wright, W. (2005). GeoTime Information Visualization. *Information Visualization*, 4(2), pp.136-146.

Kincaid, R. and Lam, H. (2006). Line graph explorer: scalable display of line graphs using Focus+Context. *Proceedings of the working conference on Advanced visual interfaces*, Venezia, Italy, pp.401-411, ACM, New York.

Kjellin, A., Winkler Pettersson, L., Seipel, S. and Lind, M. (2009). Different levels of 3D: An evaluation of visualized discrete spatiotemporal data in space-time cubes. *Information Visualization*, 9(2), pp.152-164.

Kosara, R. and Miksch, S. (2001). Metaphors of movement: a visualization and user interface for time-oriented, skeletal plans. *Artificial Intelligence in Medicine*, 22(2), pp.111-131.

Kowalski, R. and Sergot, M. (1986). A logic-based calculus of events. *New Generation Computing*, 4(1), pp.67-95.

Koymans, R. (1990). Specifying real-time properties with metric temporal logic. *Real-Time Systems*, 2(4), pp.255-299.

Kraak, M.-J. (2003). The space-time cube revisited from a geovisualization perspective. *Proceedings of the 21st International Cartographic Conference*, Durban, South Africa, pp.1988-1996.

Kraak, M.-J. and Ormeling, F. (2003). *Cartography: Visualization of Geospatial Data.* Pearson Education, Harlow, England.

Kröger, F. (1987). *Temporal logic of programs*. Springer-Verlag, New York.

Ladkin, P. B. (1987). *The Logic of Time Representation.* Ph.D. thesis, University of California, Berkeley.

Langran, G. (1993). Time in Geographic Information Systems. ed. by Peuquet, D. J. and Marble, D. F., in *Technical Issues in Geographic Information System*, Taylor & Francis Ltd, London.

Leith, M. and Cunningham, J. (1997). *Representing and reasoning with events from natural language*. Springer-Verlag Berlin, Berlin 33.

Li, X. and Kraak, M.-J. (2008). The Time Wave. A New Method of Visual Exploration of Geo-data in Time–space. *The Cartographic Journal,* 45(3), pp.193-200.

Lin, J., Keogh, E., Lonardi, S., Lankford, J. P. and Nystrom, D. M. (2004). Visually mining and monitoring massive time series. *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining*, Seattle, WA, USA, pp.460-469, ACM.

Lind, M., Bingham, G. P. and Forsell, C. (2003). Metric 3D structure in visualizations. *Information Visualization,* 2(1), pp.51-57.

Ma, J. and Knight, B. (1994). A General Temporal Theory. *The Computer Journal,* 37(2), pp.114-123.

Ma, J. and Knight, B. (1996). A Reified Temporal Logic. *The Computer Journal,* 39(9), pp.800-807.

Matkovic, K., Gracanin, D., Konyha, Z. and Hauser, H. (2007). Color Lines View:An Approach to Visualization of Families of Function Graphs. *11th International Conference on Information Visualization, 2007. IV '07,* Vienna, pp.59-64.

McCarthy, J. and Hayes, P. J. (1968). Some philosophical problems from the standpoint of artificial intelligence. in *Readings in nonmonotonic reasoning,* pp.26-45, Morgan Kaufmann Publishers Inc.

McDermott, D. (1982). A temporal logic for reasoning about processes and plans. *Cognitive Science,* 6(2), pp.101-155.

McLachlan, J. (1987). Joseph Priestley and the Study of History. *Transactions of the Unitarian Historical Society,* 19(1987-90), pp.252-263.

Monmonier, M. (1989). Geographic Brushing: Enhancing Exploratory Analysis of the Scatterplot Matrix. *Geographical Analysis,* 21(1), pp.81-84.

Monmonier, M. (1990). Strategies for the Visualization of Geographic Time-Series Data. *Cartographica: The International Journal for Geographic Information and Geovisualization,* 27(1), pp.30-45.

Nagypál, G. and Motik, B. (2003). A Fuzzy Model for Representing Uncertain, Subjective, and Vague Temporal Knowledge in Ontologies. *Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics, (ODBASE)*, Catania, Italy, pp.906-923.

North, C. (2000). *A User Interface for Coordinating Visualizations based on Relational Schemata: Snap-Together Visualization.* Ph.D. thesis, University of Maryland.

Ohlbach, H. J. (2004). Relations between fuzzy time intervals. *Proceedings of 11th International Symposium on Temporal Representation and Reasoning (TIME 2004)*, Montréal, Québec, pp.44-51.

Osmani, A., Imam, I., Kodratoff, Y., El-Dessouki, A. and Ali, M. (1999). Introduction to Reasoning about Cyclic Intervals. *Proceedings of 12th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems IEA/AIE-99*, Cairo, Egypt, pp.698-706, Springer Berlin / Heidelberg.

Ozsoyoglu, G. and Snodgrass, R. T. (1995). Temporal and real-time databases: a survey. *IEEE Transactions on Knowledge and Data Engineering,* 7(4), pp.513-532.

Peuquet, D. J. (1994). It's About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. *Annals of the Association of American Geographers,* 84(3), pp.441 - 461.

Plaisant, C., Shneiderman, B. and Mushlin, R. (1998). An information architecture to support the visualization of personal histories. *Information Processing & Management,* 34(5), pp.581-597.

Playfair, W. (1786). *The Commercial and Political Atlas*.

Playfair, W. (1805). *An Inquiry Into the Permanent Causes of the Decline and Fall of Powerful and Wealthy Nations*. London.

Pnueli, A. (1977). The temporal logic of programs. *18th Annual Symposium on Foundations of Computer Science*, Providence, Rhode Island, USA, pp.46-57.

Pnueli, A. (1981). The temporal semantics of concurrent programs. *Theoretical Computer Science,* 13(1), pp.45-60.

Prior, A. N. (1957). *Time and Modality*. Oxford University Press Inc., New York.

Prior, A. N. (1967). *Past, present and future*. Oxford University Press Inc., New York.

Qiang, Y., Delafontaine, M., Asmussen, K., Stichelbaut, B., De Tré, G., De Maeyer, P. and Van de Weghe, N. (2010). Modelling Imperfect Time Intervals in a Two-Dimensional Space. *Control and Cybernetics,* 39(4), pp.983-1010.

Reichgelt, H. (1989). A comparison of first-order and modal logics of time. ed. by Jackson, P., Reichgelt, H. and Harmelen, F. v., in *Logic-based Knowledge Representation,* pp.143-176, The MIT press, Cambridge, MA, USA.

Roberts, J. C. (2007). State of the Art: Coordinated & Multiple Views in Exploratory Visualization. *Proceedings of Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV '07)* pp.61-71.

Saito, T., Miyamura, H. N., Yamamoto, M., Saito, H., Hoshiya, Y. and Kaseda, T. (2005). Two-Tone Pseudo Coloring: Compact Visualization for One-Dimensional Data. Paper presented in Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization, pp.173-180.

Schockaert, S. and De Cock, M. (2008). Temporal Reasoning about Fuzzy Intervals. *Artificial Intelligence,* 172(8-9), pp.1158-1193.

Schockaert, S., De Cock, M. and Kerre, E. (2006). Imprecise Temporal Interval Relations. in *Fuzzy Logic and Applications,* pp.108-113.

Schockaert, S., De Cock, M. and Kerre, E. E. (2008). Fuzzifying Allen's Temporal Interval Relations. *IEEE Transactions on Fuzzy Systems,* 16(2), pp.517-533.

Schreiber, F. A. (1994). Is time a real time? An overview of time ontology in informatics. ed. by Halang, W. A. and Stoyenko, A. D., in *Real Time Computing*, pp.283-307, Springer Verlag, Berlin.

Silvestrini, A. and Veredas, D. (2008). Temporal Aggregation of Univariate and Multivariate Time Series Models: a Survey. *Journal of Economic Surveys,* 22, pp.458-497.

Tominski, C., Abello, J. and Schumann, H. (2004). Axes-based visualizations with radial layouts. *Proceedings of the 2004 ACM symposium on Applied computing*, Nicosia, Cyprus, ACM.

Tominski, C., Schulze-Wollgast, P. and Schumann, H. (2005). 3D Information Visualization for Time Dependent Data on Maps. *Proceedings of the Ninth International Conference on Information Visualisation*, London, UK, IEEE Computer Society.

Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, U.S.

Tufte, E. R. (2001). *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, U.S.

Van Beek, P. (1989). Approximation algorithms for temporal reasoning. *Proceedings of the 11th international joint conference on Artificial intelligence - Volume 2*, Detroit, Michigan, pp.1291-1296, Morgan Kaufmann Publishers Inc.

Van Benthem, J. (1983). *The Logic of Time*. Kluwer, Dordrecht, The Netherlands.

Van Wijk, J. J. and Selow, E. V. (1999). Cluster and Calendar-based Visualization of Time Series Data. ed. by Wills, G. and Keim, D., *IEEE Symposium of Information Visualization (InfoVis '99)*, San Francisco, USA, pp.4-9, IEEE Computer Society.

Vasiliev, I. (1997). Mapping Time. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 34(2), pp.1-51.

Vila, L. (1994). A survey on temporal reasoning in artificial intelligence. *AI Communications*, 7(1), pp.4-28.

Vilain, M., Kautz, H. and Beek, P. v. (1990). Constraint propagation algorithms for temporal reasoning: a revised report. in *Readings in qualitative reasoning about physical systems*, pp.373-381, Morgan Kaufmann Publishers Inc.

Ware, C. and Franck, G. (1996). Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Transactions on Graphics*, 15(2), pp.121-140.

Ware, C. and Mitchell, P. (2008). Visualizing graphs in three dimensions. *ACM Transactions on Applied Perception*, 5(1), pp.1-15.

Weber, M., Alexa, M. and Muller, W. (2001). Visualizing time-series on spirals. *Proceedings of IEEE Symposium on Information Visualization 2001*, Washington, DC, U.S.A., pp.7-13.

3 Reasoning about Imperfect Time Intervals

Modified from: Yi Qiang, Matthias Delafontaine, Katrin Asmussen, Birger Stichelbaut, Guy De Tré, Philippe De Maeyer, Nico Van de Weghe, (2010). Modelling Imperfect Time Intervals in a Two-Dimensional Space. *Control and Cybernetics*, 39(4), pp. 983-1010.

Abstract: Every event has an extent in time, which is usually described by crisp time intervals. However, under some circumstances, temporal extents of events are imperfect, and thus cannot be adequately modelled by crisp time intervals. Rough sets and fuzzy sets are two frequently used tools for representing imperfect temporal information. In this chapter, we apply a two-dimensional representation of crisp time intervals, the Triangular Model (TM), to investigate rough time intervals (RTIs) and fuzzy time intervals (FTIs). With this model, RTIs and FTIs, as well as their temporal relations, can be represented as graphics (i.e. discrete geometries or continuous fields) in a two-dimensional time space. Compared to the conventional linear representation of time intervals, the TM offers a simpler representation of imperfect time intervals and thus allows more direct perception of temporal relations. Moreover, temporal queries involving imperfect intervals can be addressed by operations of graphics in the TM, which is potentially easier for humans to understand than formal expressions. We also demonstrate how the TM can be used for the visualization and analysis of RTIs. All in all, we contend that the TM can be applied as a valuable assistant tool for the research of reasoning and analysis of imperfect time intervals.

Keywords: Rough set, rough time interval, fuzzy set, fuzzy time interval, the Triangular Model, temporal relation

3.1 Introduction

Currently, a lot of research has been done on representing and reasoning about time intervals. Most of this work has focused on crisp time intervals (CTI), namely, time intervals bounded by a start point and an end point (Allen, 1983; Allen and Hayes, 1985; Freksa, 1992; Galton, 1990; Ladkin, 1987). However, under some circumstances, CTIs cannot always adequately describe temporal occupations of events and processes. For example, due to imprecise information, the start and end of an interval is known to be within certain ranges. However, the exact start and end cannot be defined. In these cases, time intervals of events can be represented by rough sets (Pawlak, 1982; Pawlak, 1992) which classify the upper and lower approximations of the interval. The yet obtained time intervals are called rough time intervals (RTI). On the other hand, some events may start or end gradually and therefore their start and end times cannot be pinned to exact time points. Intervals of this kind of event can be described by fuzzy sets (Zadeh, 1965, 1975) through the quantification of the graded truth of whether a time point is a member of the interval, leading to the concept of fuzzy time interval (FTI). Currently, a lot of disciplines (e.g. archaeology, geography, psychology, and philosophy) are faced with the problem of handling imperfect temporal information, as is reflected in many contributions about representing and reasoning about RTIs (Bassiri et al., 2009; Bittner, 2002) and FTI (De Caluwe et al., 1999; De Caluwe et al., 1997; Nagypál and Motik, 2003; Ohlbach, 2004; Schockaert and De Cock, 2008; Schockaert et al., 2008). However, most work is based on the linear representation in which time intervals are modelled as linear segments on the numerical line. Kulpa (1997, 2006) proposed an alternative representation of CTIs in which time intervals are mapped to points in a two-dimensional space. He showed this representation can be used as a diagrammatic tool to facilitate the research of reasoning of CTIs. Later, based on Kulpa's work, Van de Weghe et al. (2007) introduced the Triangular Model (TM) and applied it into an archaeological context. De Tré et al. (2006) made the preliminary trial of representing imperfect time intervals by the TM and applied this approach to visualise a historical database. In these researches, the TM has shown its potential in expressing abstract concepts and relations in temporal reasoning. Till present, the power of the TM in handling FTIs and RTIs has not yet been fully exploited. Therefore, in this chapter, we systematically investigate the use of the TM in representing and reasoning about RTIs and FTIs.

In the remainder of the chapter, we first discuss the two types of imperfect time intervals, namely fuzzy time intervals and rough time intervals (Section 2). Further, after a brief introduction of the Triangular Model (Section 3), we show how the TM can be used to represent rough time intervals (Section 4) and fuzzy time intervals (Section 5), as well as their temporal relations. In Section 6, this new model is then applied to an archaeological use case in order to illustrate its potential in visualization. Finally, conclusions are drawn and future work is proposed.

3.2 Imperfect Time Intervals

In philosophy, there are two opposite theories of time, namely, absolute time and relative time (Lin, 1991). An interesting overview of other theories is given in Knight and Ma's (Knight and Ma, 1993) work, which proposed a consensus glossary of temporal concepts. In the absolute concept, time is decided by numbers on the time line and is independent of anything else. The basic time units are time intervals which are defined by two numbers on the time line. Because the time line underlying most calendar systems can be modelled as a single numerical axis which is isomorphic to real numbers $\mathbb{R}$, a time interval is usually understood as an interval of $\mathbb{R}$, bounded by two real numbers I^- and I^+, with $I^- < I^+$. Absolute time is a machine-oriented concept, because it is usually measured by clocks, and can be easily recorded and processed by computers. The absolute concept of time is thus widely applied in computer science and engineering (Shoham and Goyal, 1988; Vila, 1994). In contrary to absolute time, the relative concept claims that time is determined by events and properties of time must be defined by investigating properties of events. This concept is based on human's perception of time, which is often expressed by *when*-clauses in human language. For example, in the sentence 'Tornado Katrina happened *when* George W. Bush was the president', the event 'Bush was the president' indicates a period during which Katrina tornado happened. Theoretically, relative time may be expressed in terms of absolute time. In other words, an event perceived by human beings always corresponds to a specific interval in the absolute time line. For example, the period when George W. Bush was the president corresponds to the interval between January 20, 2001 and January 20, 2009. Such matches between relative time and absolute time broadly exist in human knowledge and activities. All historical events are linked to specific past time intervals and all plans in people's schedules occupy specific intervals in the future. However, difficulties in matching relative time (intervals defined by events) to absolute time (intervals defined by numbers on the time line) are likely to happen. In these cases, the temporal location of events cannot be perfectly described by a CTI in the absolute time. In order to solve these problems, scientists applied rough sets and fuzzy sets to represent time intervals of these events, bringing the concepts of RTI and FTI. In the following two sub-sections, we introduce the concepts of RTI and FTI.

3.2.1 Rough Time Interval

Under some circumstances, the temporal location of an event cannot be decisively matched to a CTI in absolute time due to incomplete information. It may be known that the event started within a certain range and ended within a certain range. However, information of the exact starting time and ending time is not available. Rough set theory can be used to describe intervals of such events (Bassiri et al., 2009; Bittner, 2002), by defining upper approximations and lower approximations of intervals occupied by events. These intervals described by rough sets are called rough time intervals (RTI). Generally speaking, the incomplete information may have two causes. The first one is the granularity of descriptions. In daily life, time intervals are usually described with certain partitions of the time line (e.g. year, month, days, and hours). Under some circumstances, these partitions may be too coarse to express the exact intervals of events. For instance, the sentence ‘Bush started his presidency in 2001 and ended it in 2009’ is correct with respect to the yearly granularity. Nevertheless, this sentence lacks information on the exact date when Bush started and finished his presidency. Bittner (Bittner, 2002) applied rough set theory to represent the relations between time intervals and partitions of the time line. Cells of the partition that are definitely occupied by the time interval form the lower approximation. Cells that may be occupied by the time interval constitute the upper approximation. The limitation of this approach is that the approximations of intervals have to be referenced to equal time partitions and the syntactic notations have not been well linked to common-sense temporal semantics. The second cause of incomplete information stems from the data acquisition process. In many observation activities, data are acquired at discrete time stamps. Through these snapshots, the time interval of an event can only be approximately decided. Remote sensing, for instance, relies on images or photographs taken at discrete time stamps by which one can determine the state of an object. A specific state of an object, for instance the existence of an object, can be understood as an event. From discrete images, one can determine whether an object exists at specific time stamps. However, the object’s existence in between two time stamps is uncertain. With these discrete observations, the interval of the object’s existence can be described by an upper approximation and a lower approximation. Bassiri et al's approach deals with this situation (Bassiri et al., 2009). However, they have not exhausted all possible relations between rough time intervals, probably because of the limitation of the linear representation. Also, Bassiri et al. have not explained the semantics behind these enumerated relations.

The basic idea of rough set is classifying a set R into a lower approximation $\underline{R}$ and an upper approximation $\overline{R}$ according to a subset of its attributes. Within $\underline{R}$, elements are definitely members of a target set A; while outside $\overline{R}$, elements are not members of A, where $\underline{R} \subseteq \overline{R}$. The difference between $\overline{R}$ and $\underline{R}$ forms boundary regions. If the boundary region is nonempty, i.e. $\overline{R} \neq \underline{R}$, the set R is said to be rough; otherwise R is a crisp set. In the boundary region, elements cannot be decisively classified as members or non-members of A. Since time intervals are considered as convex subsets of real numbers $\mathbb{R}$, an imprecise interval I can also be modelled by an upper approximation $\overline{I}$ and a lower approximation $\underline{I}$, with $\underline{I} \subseteq \overline{I}$ (Figure 3-1). We call such a pair of $\overline{I}$ and $\underline{I}$ a Rough Time Interval (RTI), which is denoted as $R(I)$. Both $\overline{I}$ and $\underline{I}$ are closed sets because their left and right extremes are parts of them. In our notion, a rough time interval is not a special kind of time interval. It is essentially an approximate description of a crisp interval that cannot be precisely described due to lack of information. Since a CTI can be considered as a crisp subset of $\mathbb{R}$, an RTI can be considered as a rough subset of $\mathbb{R}$. Time points in $\underline{I}$ definitely belong to I, whereas time points out of $\overline{I}$ definitely do not belong to I. In between $\underline{I}$ and $\overline{I}$, there are two boundary intervals, i.e. $R(I^-)$ and $R(I^+)$, in which time points that cannot be decided whether they belong to I or not. $\overline{I}$ is the largest possibility of I, which is bounded by the earliest possibility of the start point $\overline{I}^-$ and the latest possibility of the end point $\overline{I}^+$. $\underline{I}$ is the shortest possibility of I, which is bounded by the latest possibility of the start point $\underline{I}^-$ and the earliest possibility of the end point $\underline{I}^+$.

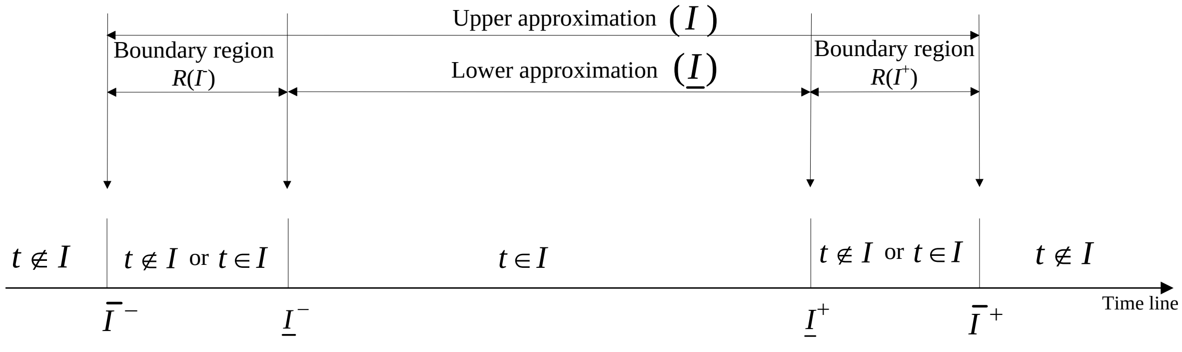


Figure 3-1: The linear representation of rough time intervals

3.2.2 Fuzzy Time Interval

Difficulties in matching events to intervals in absolute time may originate from the fuzzy nature of events. Some events may start or end gradually, and thus lack distinct starting or ending time. For example, it is difficult to decide when the Industrial Revolution started and finished. Though some historians like to use the invention of the steam engine to mark its start, it is unnatural to understand that the revolution suddenly started when the steam engine was invented. Other examples are clauses like 'when I was young'. It does not make sense to consider a specific day after which one is suddenly old. Intervals of these fuzzy events and processes cannot be adequately described by CTIs. Fuzzy set theory is a frequently used tool for modelling intervals of such fuzzy events (Nagypál and Motik, 2003; Ohlbach, 2004; Schockaert and De Cock, 2008; Schockaert et al., 2008). It extends conventional (crisp) set theory and handles the concept of partial truth, i.e. graded truth values between 0 (completely false) and 1 (completely true) (Zadeh, 1965, 1975). A fuzzy set A is modelled by a membership function $\mu_A(x)$ that maps every real number x in $\mathbb{R}$ to a real number between 0 and 1, representing the truth of whether x is a member of A, or to what extent x belongs to A. Besides fuzzy events with gradual starts or ends, fuzzy sets can also express uncertainties of non-fuzzy events in the framework of possibility theory (Dubois et al., 2003; Dubois et al., 1991; Garrido et al., 2009). But in this chapter, we will not investigate FTIs in the possibilistic context. This will be left for future work.

Following the principles of rough set theory, time intervals of fuzzy events are modelled by fuzzy sets, which quantify the truth of whether time points on the time line are occupied by events. Intervals described by fuzzy sets are called fuzzy time intervals (FTI) and denoted as $\tilde{I}$. For an arbitrary FTI $\tilde{I}$, every time point t on the time line is mapped to $\tilde{I}$ by a membership function $\tilde{I}(t)$, which returns the truth value of whether t is a member of $\tilde{I}$. All time points t that satisfy $\tilde{I}(t) = 1$ form the core of $\tilde{I}$, denoted as Core($\tilde{I}$), while all time points satisfy $\tilde{I}(t) > 0$ form the support of $\tilde{I}$, denoted as Support($\tilde{I}$) (Figure 3-2). If $\mathrm{Core}(\tilde{I}) = \mathrm{Support}(\tilde{I})$, $\tilde{I}$ reduces to a CTI. In order to start from the simplest situation, we assume that an FTI $\tilde{I}$ must have a non-empty core, and both Support($\tilde{I}$) and Core ($\tilde{I}$) are convex intervals. Intervals in between Support($\tilde{I}$) and Core ($\tilde{I}$) are called fuzzy start and fuzzy end (Figure 3-2). Unlike relations between CTIs, relations between FTI and another interval (CTI or FTI) cannot be decided by a yes or no answer, but quantified as a truth value between 0 and 1.

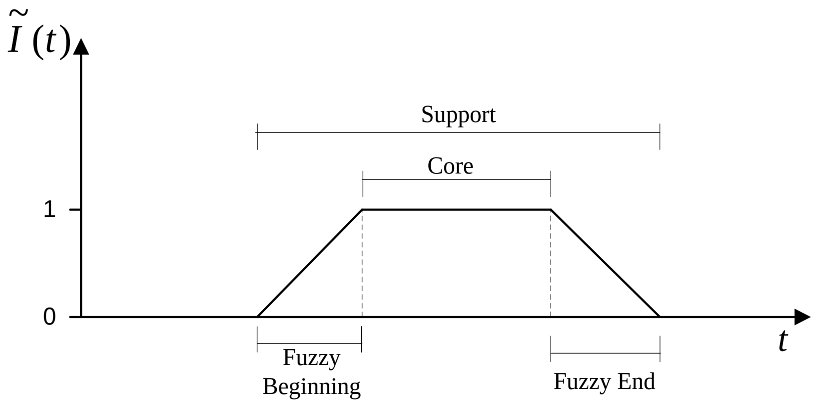


Figure 3-2: An FTI illustrated by a line chart

3.3 Triangular Model

3.3.1 Representing Crisp Time Intervals in the TM

In the classical representation, a time interval is represented as a finite linear segment on a horizontal time line (Figure 3-3 top-left). From the paralleled scales on the time line, one may read the numbers of I^- and I^+ of the interval. The vertical dimension is solely used to differentiate multiple overlapping intervals, if used at all. The linear representation of time intervals is widely used in our daily life, for example Gantt charts, geological and historical timelines. In different reasoning systems, whether an interval is open (at one side or both sides) is defined differently (Vila, 1994). the TM does not intend to solve reasoning issues that concerns this controversy. Whether an interval is open does not affect its representation in the TM. In this chapter, we define that an crisp interval is closed at both sides and denote it as $[I^-, I^+]$. Different from the classic representation, the basic concept of the Triangular Model (TM) is the construction of two lines through the extremes of an interval (Figure 3-3 top-right). For each time interval I, two straight lines (L_1 and L_2) are constructed, with L_1 passing through I^- and L_2 passing through I^+. α_1 is the angle between L_1 and the horizontal axis and α_2 is the angle between L_2 and the horizontal axis, with $\alpha_1 = -\alpha_2 = \alpha$. The intersection of L_1 and L_2 is called the interval point. The start point of the interval I^-, the end of the interval I^+ and interval point form an isosceles triangle. That is why we call this model the Triangular Model (TM). The angle α is a pre-defined constant which is identical for all interval points, in order to ensure that each time interval is mapped to a unique 2D point. In this work, we keep consistency with previous research of the TM (Kulpa, 1997, 2006; Van de Weghe et al., 2007) and set α = 45°. Of course, α can also alter to other angles (between 0 and 90) for specific purposes. In the TM, all time intervals are represented as such interval points in a two-dimensional space (Figure 3-3 bottom). In other words, the position of an interval point I completely determines both, the start and the end of the interval. The two-dimensional space of interval points is called the Interval Space. Because all time intervals can be considered as subsets of real numbers $\mathbb{R}$, the interval space is denoted as $I\mathbb{R}$ (Kulpa, 2006).

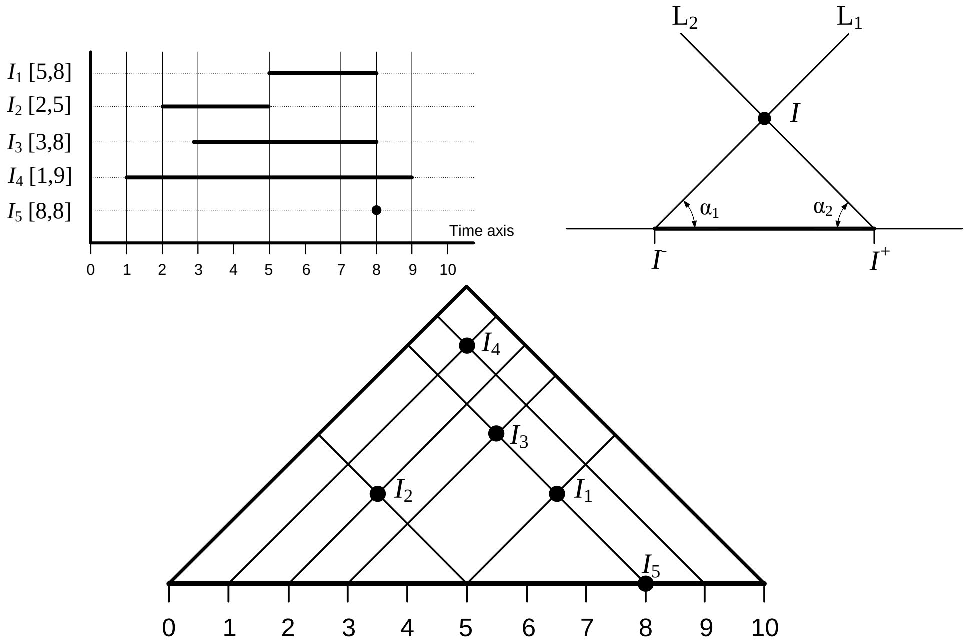


Figure 3-3: Transformation from the linear representation to the TM. Top-left: the classic linear model. Top-right: construction of an interval point. Bottom: time intervals in the TM.

3.3.2 Representing Crisp Temporal Relations in the TM

Let the start I^- and end I^+ of two intervals have the following three possible relations: smaller than (<), equal to (=) and larger than (>). Then, according to Allen (Allen, 1983), thirteen possible relationships between two CTIs can be defined (see Table 3-1). In the TM, every Allen relation corresponds to a specific zone (Kulpa, 1997). Given a study period from 0 to 100, all examined intervals are located within the isosceles triangle of I [0, 100]. Let us consider a reference interval I_2 [33,66] and several intervals (I_{1a}, I_{1b}, I_{1c}) existing before interval I_2 (Figure 3-4a). In the TM, I_{1a}, I_{1b}, I_{1c} are located in the triangular zone in the left corner of the study area (Figure 3-4b). Therefore, it is easy to deduce that all intervals before I_2 must be located in the black zone (Figure 3-4c). Namely, this zone encloses all intervals that are before I_2. All Allen relations with respect to a CTI can be represented by such zones in $I\mathbb{R}$ (Figure 3-5). For each individual figure in Figure 3-5, the reference CTI I has been chosen in the centre of the study period in order to avoid visual bias. Because all zones in Figure 3-5 occupy a unique area in $I\mathbb{R}$, $I\mathbb{R}$ can be divided into thirteen zones with respect to thirteen Allen relations (Figure 3-6) (Kulpa, 1997, 2006). We call these zones Crisp Relational Zones (CRZ). Each CRZ represents the set of CTIs that are in a specific relation with respect to the reference interval I. Such set of CTIs is denoted as $Rel(I)$. For example, the set of CTIs that are before I is denoted as *before*(I). According to the number of common extremes (i.e. start and end), Allen relations can be categorized into two types. Type 1 intervals have no common extremes (*before*, *overlaps*, *during*, *contains*, *overlapped-by*, *after*). Type 2 intervals have common extremes (*meets*, *met-by*, *starts*, *started-by*, *finishes*, *finished-by* and *equal*). In the TM, Type 1 CRZs have a triangular or rectangular shape (two-dimensional); while Type 2 CRZs have a point or linear geometry (zero-dimensional or one-dimensional).

Table 3-1: Thirteen Allen Relations (Allen 1983)

I_1 *equal* I_2	if $I_1^- = I_2^- \wedge I_1^+ = I_2^+$
I_1 *starts* I_2	if $I_1^- = I_2^- \wedge I_1^+ < I_2^+$
I_1 *started−by* I_2	if $I_1^- = I_2^- \wedge I_2^+ < I_1^+$
I_1 *finishes* I_2	if $I_1^+ = I_2^+ \wedge I_1^- > I_2^-$
I_1 *finished−by* I_2	if $I_1^+ = I_2^+ \wedge I_2^- > I_1^-$
I_1 *meets* I_2	if $I_1^+ = I_2^-$
I_1 *met−by* I_2	if $I_2^+ = I_1^-$
I_1 *overlaps* I_2	if $I_2^- > I_1^- \wedge I_1^+ < I_2^+ \wedge I_1^+ > I_2^-$
I_1 *overlapped−by* I_2	if $I_1^- > I_2^- \wedge I_1^- < I_2^+ \wedge I_2^+ < I_1^+$
I_1 *during* I_2	if $I_1^- > I_2^- \wedge I_1^+ < I_2^+$
I_1 *contains* I_2	if $I_2^- > I_1^- \wedge I_2^+ < I_1^+$
I_1 *before* I_2	if $I_1^+ < I_2^-$
I_1 *after* I_2	if $I_2^+ < I_1^-$

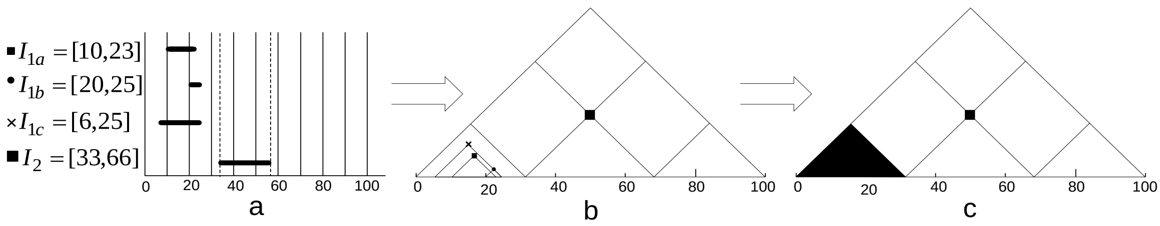


Figure 3-4: Temporal relations in the linear model and the TM, taking *before* as an example

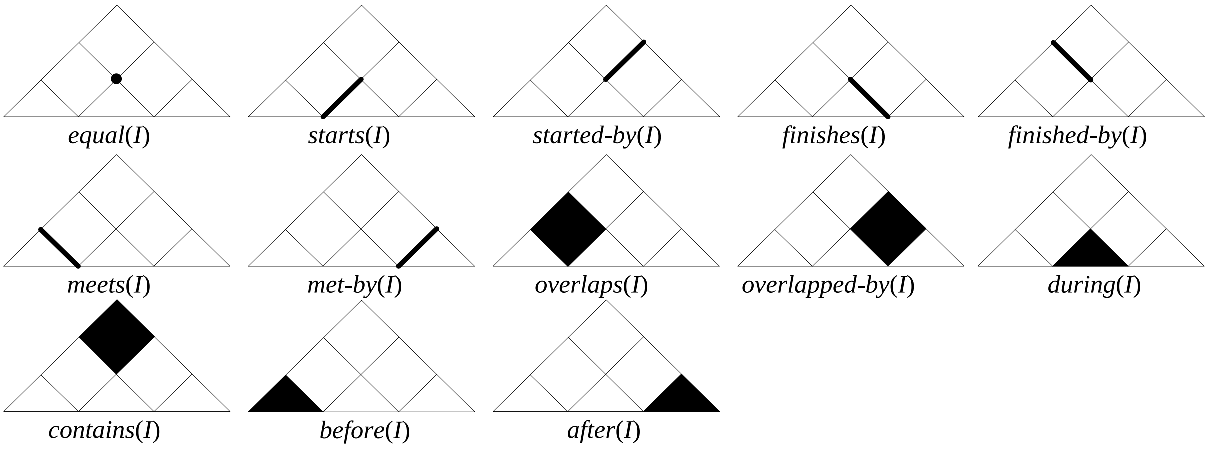


Figure 3-5: CRZs in individual interval spaces, representing sets of intervals in Allen relations to the reference interval *I*

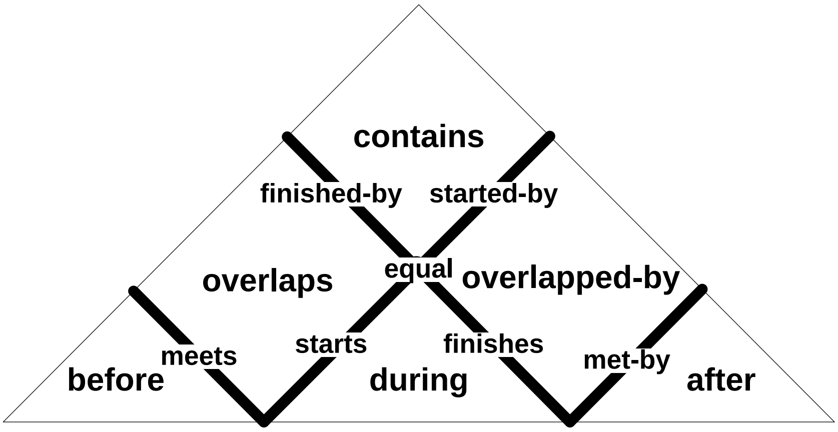


Figure 3-6: CRZs of an interval in the TM

3.4 Modelling Rough Time Intervals in the TM

3.4.1 Representing Rough Time Intervals in the TM

Different from the linear model in which an RTI is represented as a tripartite linear segment with an uncertain start and an uncertain end (Figure 3-7), the TM represents RTIs as geometries in a two-dimensional space. In order to construct an RTI $R(I)$ in the TM, four lines are projected respectively from $\overline{I}^-$, $\underline{I}^-$, $\underline{I}^+$and $\overline{I}^+$, forming a rectangle (Figure 3-8). This rectangle indicates a zone where the exact CTI *I* can be found. In other words, this zone represents the set of CTIs that are possibly equal to *I*. Therefore, the zone is called the *maybe equal* (*ME*) zone of *I*. Other relational zones of RTIs will be discussed in detail in the next section. We note that the *ME* zone is the only relational zone that express all the information of $R(I)$. The shape, position and area of this zone fully reflect the characteristics of $R(I)$. Thus $R(I)$ can be represented by its

ME zone which covers all possible locations of I. In this way, $R(I)$ is represented by a 2D geometry (Figure 3-9). Apart from a rectangle, the 2D geometry can be in other shapes. For instance, if $\underline{I} = \emptyset$, the *ME* zone becomes a triangle on the horizontal axis (e.g. I_5 in Figure 3-9) . If either $R(I^-) = \emptyset$ or $R(I^+) = \emptyset$, the *ME* zone becomes a line (e.g. I_2 in Figure 3-9). If both $R(I^-) = \emptyset$ and $R(I^+) = \emptyset$, $R(I)$ reduces to a CTI and its *ME* zone becomes an point. In this work, we emphasize on RTIs whose *ME* zone is a rectangle.

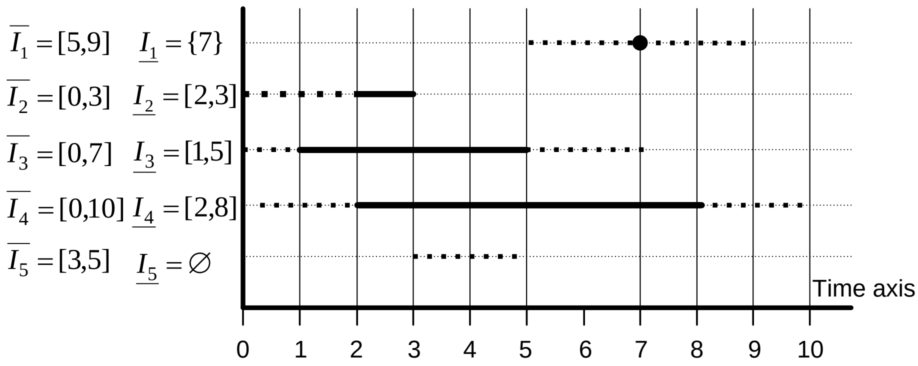


Figure 3-7: Rough time intervals in the linear concept. Solid lines denote $\underline{I}$, and dotted lines denote $R(I^-)$ and $R(I^+)$. The combination of the solid line and dotted lines forms $\underline{I}$.

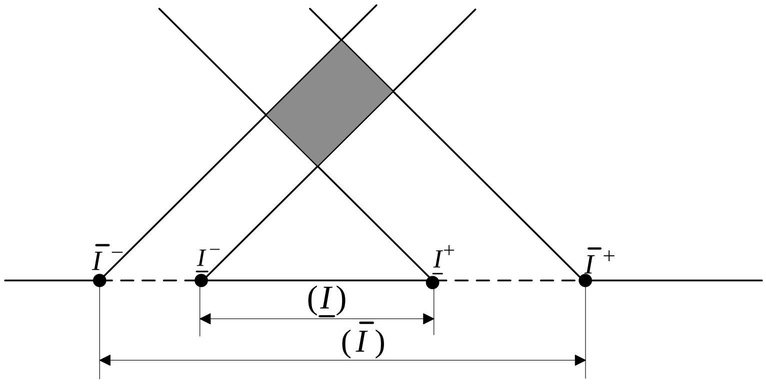


Figure 3-8: The construction of an RTI in the TM

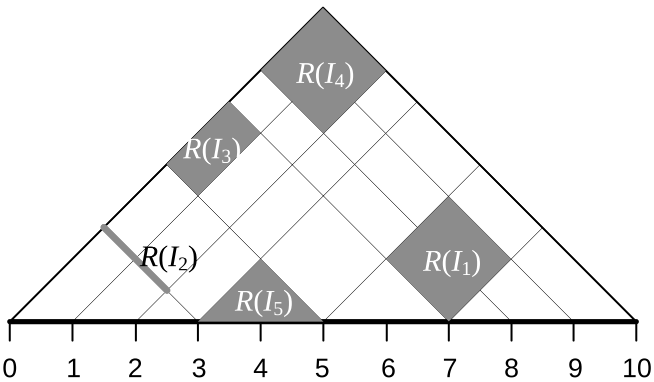


Figure 3-9: Using the TM to represent RTIs in Figure 3-7.

3.4.2 Rough Relational Zones

According to the upper approximation $\overline{I}$ and the lower approximation $\underline{I}$ of an RTI, the interval space ($I\mathbb{R}$) can be divided into zones. The number of zones depends on whether the lower approximation or boundary regions are empty. First, we focus on RTIs with $R(I^{-}) \neq \emptyset$, $R(I^{+}) = \emptyset$ and $\underline{I} \neq \emptyset$. Regarding this kind of RTI, $I\mathbb{R}$ are divided into 15 zones (Figure 3-10). We call these zones Rough Relational Zones (RRZ). Table 3-2 lists details of the 15 RRZs in Figure 3-10. When comparing CRZs and RRZs, we can see polygons in CRZs (i.e. zones of Type 1 relations) remain polygons in RRZs, whereas the point and lines in CRZs (i.e. zones of Type 2 relations) have expanded to polygons in RRZs. Two new RRZs are expanded from the start and end point of *I*. In contrary to CRZs which express determinate relations, these expanded RRZs express more than one possible relation. For example, intervals in the *Maybe Meets* (*MM*) zone have three possible relations to *I*, i.e. *meets*, *before* and *overlaps* (Table 3-2). In the RRZs that are not expanded from points and lines, only one relation is possible. Boundaries between RRZs do not occupy any space, but belong to the neighbouring RRZs. Arrows in Figure 3-10 indicate the belongingness of boundaries. Moreover, RTIs with an empty lower approximation or an empty boundary region result in different numbers and structures of RRZs. For example, RTIs with an empty boundary region (i.e. $R(I^{-}) = \emptyset$) have 14 RRZs (Figure 3-11). For this type of RTIs, some RRZs shrink to lines or points, and the RB zone does not exist anymore. RTIs with an empty lower approximation ($\underline{I} = \emptyset$) have only 6 RRZs (Figure 3-12).

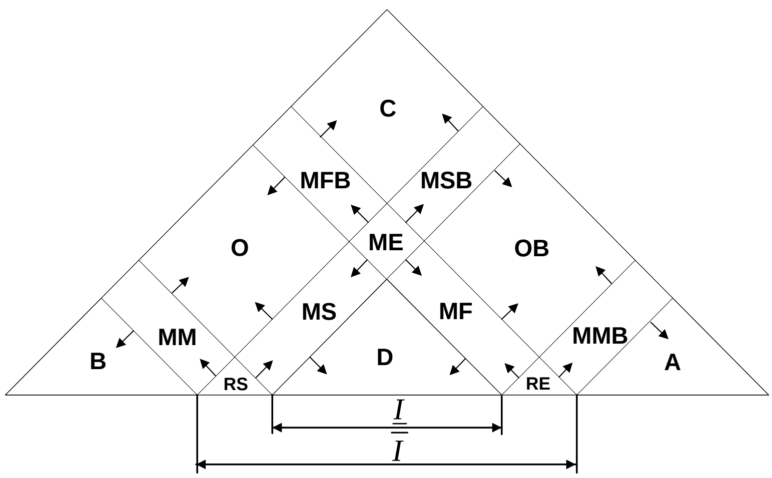


Figure 3-10: Rough Relational Zones of an RTI, with $R(I^{-}) \neq \emptyset$, $R(I^{+}) \neq \emptyset$ and $\underline{I} \neq \emptyset$

Table 3-2: Details of RRZs with respect to Figure 3-10

Name of RRZ	Abbr.	Possible Relations to I	Original Name in CRZ
before	B	before	before
overlaps	O	overlaps	overlaps
contains	C	contains	contains
during	D	during	during
overlapped-by	OB	overlapped-by	overlapped-by
after	A	after	after
maybe Meets	MM	meets, before, overlaps	meets
maybe Starts	MS	overlap, starts, during	starts
maybe finished-by	MFB	overlaps, finished-by, contains	finished-by
maybe equal	ME	contains, started-by, overlapped-by, finishes, during, starts, overlaps, finished-by, equal	equal
maybe started-by	MSB	contains, overlapped-by, started-by	started-by
maybe finishes	MF	overlapped-by, finishes, during	finishes
maybe met-by	MMB	overlapped-by, met-by, after	met-by
rough start	RS	before, meets, overlaps, starts,	Not available
rough end	RE	during, finishes, overlapped-by, met-by, after	Not available

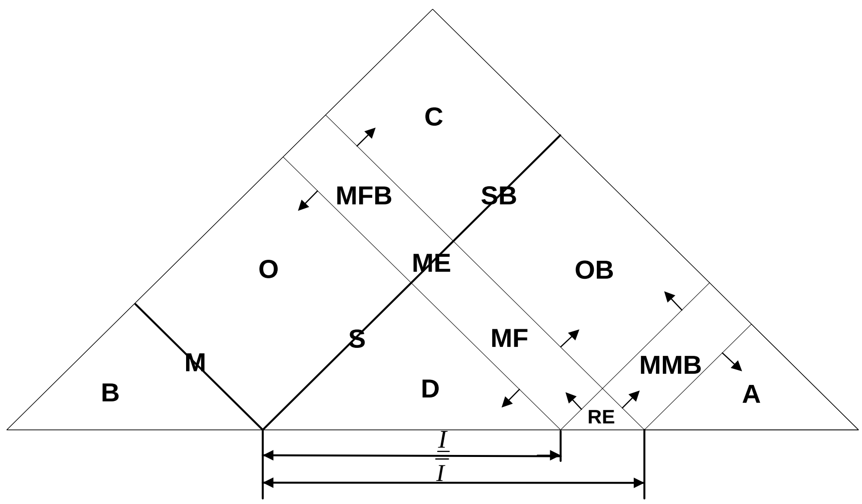


Figure 3-11: Rough Relational Zones of an RTI, with $R(I^{-}) = \emptyset, R(I^{+}) \neq \emptyset$ and $\underline{I} \neq \emptyset$

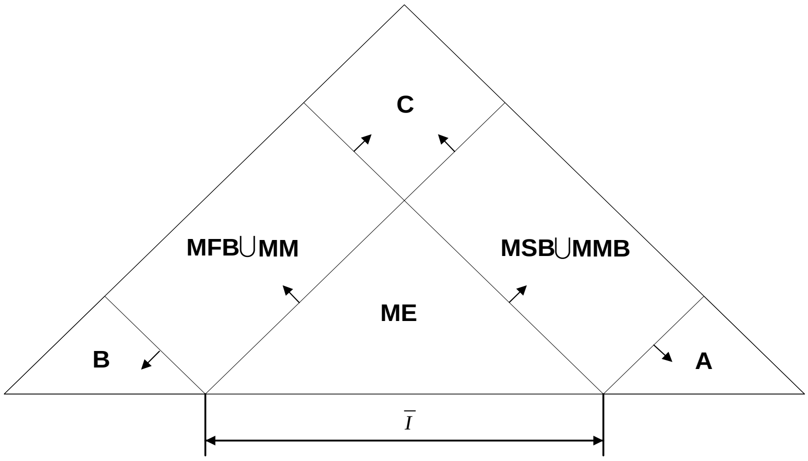


Figure 3-12: Rough Relational Zones of an RTI, with $\underline{I} = \emptyset$

3.4.3 Temporal Relations between Rough Time Intervals

Bassiri et al. (2009) enumerate 68 topological relations between two RTIs, e.g. $R(I_1)$, $R(I_2)$, by 2×2 matrixes of relations between $\overline{I_1}$ and $\overline{I_2}$, $\overline{I_1}$ and $\underline{I_2}$, $\underline{I_1}$ and $\overline{I_2}$, $\underline{I_1}$ and $\underline{I_2}$. However, these topological relations between RTIs do not directly deliver practical meanings. Because RTIs are approximated descriptions of CTIs, the useful information comes from temporal relations between the exact CTIs that are represented by RTIs. Therefore, in this section, we investigate how to use topological relations between RTIs to deduce possible relations between the exact CTIs underlying. If $R(I_1)$ and $R(I_2)$ are represented in the linear model (Figure 3-13), the possible temporal relations between I_1 and I_2 cannot be directly captured by human beings. But by the TM, relations between I_1 and I_2 can be easily recognised by intersecting $R(I_1)$ and RRZs of $R(I_2)$. For example, in Figure 3-14, when intersecting the rectangle of $R(I_1)$ and RRZs of $R(I_2)$, $R(I_1)$ 'touches' four RRZs of $R(I_2)$, namely, the *maybe meets*, *rough start*, *overlaps* and *maybe starts* zones. Note that, within this context, 'touch' means that two zones have common parts. By checking Table 3-2, we can obtain the possible relations between I_1 and I_2 as the union of possible relations of these four RRZs, i.e. {*meets*, *before*, *overlaps*} ∪ {*meets*, *before*, *overlaps*, *starts* ,*during*} ∪ {*overlaps*, *starts*, *during*}∪{*overlaps*} = {*meets*, *before*, *overlaps*, *starts*, *during*}. With this approach, one can easily deduce possible temporal relations between two RTIs. If $R(I_2)$ has a non-empty lower approximation and non-empty boundary regions (i.e. $R(I_2^-) \neq \emptyset$, $R(I_2^+) \neq \emptyset$ and $\underline{I_2} \neq \emptyset$), there are in total 80 topological relations between I_1 and RRZs of I_2 (Figure 3-15). 21 different sets of possible relations between I_1 and I_2 can be inferred from these 80 relations (Table 3-3). There are 12 more relations than Bassiri et al's (2009), because they missed situations when $R(I_1)$ has an empty lower approximation and $R(I_1)$ is totally in the boundary region of $R(I_2)$. We marked these situations with stars in Figure 3-15.

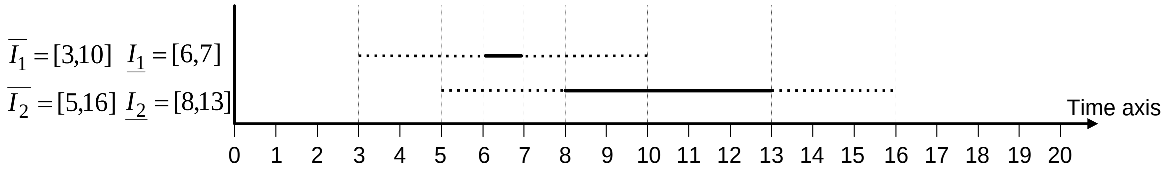


Figure 3-13: The linear representation of $R(I_1)$ and $R(I_2)$.

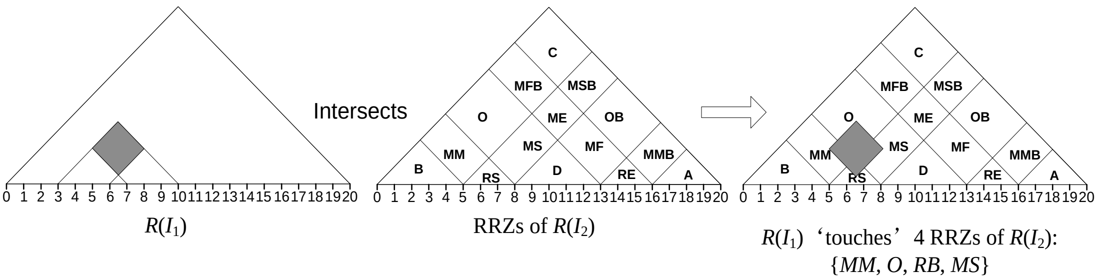


Figure 3-14: Using the TM to deduce possible relations between I_1 and I_2

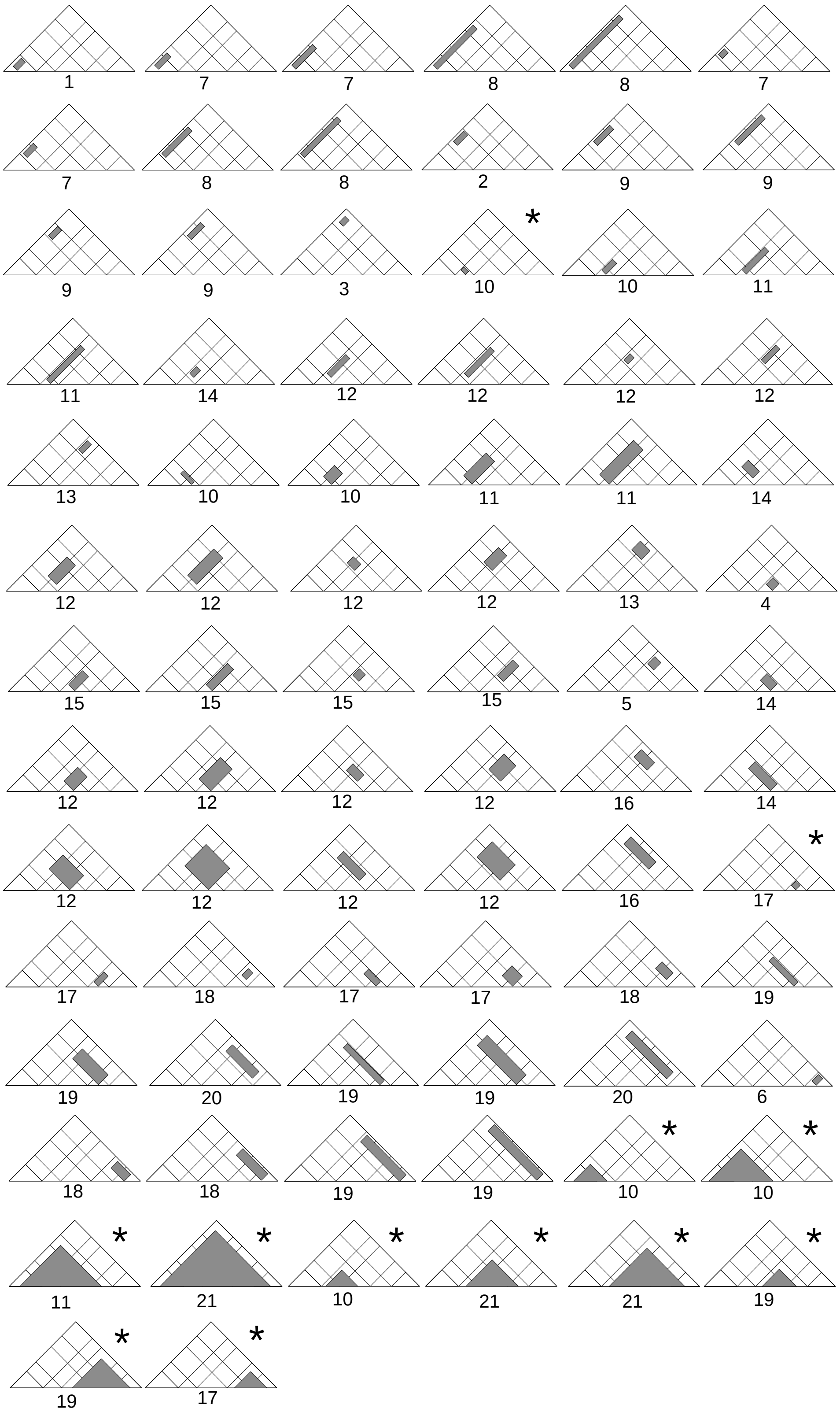


Figure 3-15: All possible topological relations between $R(I_1)$ and RRZs of $R(I_2)$, where $R(I_2^-) \neq \emptyset$, $R(I_2^+) \neq \emptyset$ and $\underline{I_2} \neq \emptyset$. Numbers denotes the set of possible relations between I_1 and I_2 in Table 3-3. (* denotes situations that are not included in the relations defined by Bassiri et al. (2009))

Table 3-3: Twenty one sets of possible relations between I_1 and I_2 with respect to Figure 3-15

No.	Possible	No.	Possible Relations	No.	Possible Relations
1	before	8	before, meets, finished-by, contains, overlaps	15	overlapped, finishes, during
2	overlapped	9	overlaps, finished-by, contains	16	contains, started-by, overlapped-by
3	contains	10	before, meets, overlaps, starts, during	17	finishes, during, overlapped-by, met-by, after
4	during	11	before, meets, overlaps, overlapped-by, starts, started-by, during, equal, finishes, finished-by,	18	overlapped-by, met-by, after,
5	overlapped-by	12	overlaps, starts, during ,finishes, equal, overlapped-by, started-by,	19	during, finishes, overlapped-by, met-by, after, starts, equal, overlaps, started-by,
6	after	13	contains, started-by, overlapped-by	20	contains, started-by, overlapped-by, met-by, after
7	before, meets, overlaps	14	overlaps, starts, during	21	All 13 Allen relations

3.4.4 Temporal Relations between Crisp and Rough Time Intervals

If an interval I is described by an upper approximation $\overline{I}$ and a lower approximation $\underline{I}$, the set of CTIs that are in a specific Allen relation to I also will also have an upper approximation $\overline{Rel(I)}$ and a lower approximation $\underline{Rel(I)}$. Figure 3-16 displays rough sets counterparts of Allen relations in the TM. Black zones represent lower approximations $\underline{Rel(I)}$, while combinations of the black zone and the grey zone are upper approximations $\overline{Rel(I)}$. More exactly, CTIs in the black zone ($\underline{Rel(I)}$) are definitely in the relation Rel to I, while CTIs in the blank zones ($\neg\overline{Rel(I)}$) are definitely not in the relation Rel to I. In the grey region ($\overline{Rel(I)} - \underline{Rel(I)}$), intervals that may or may not be in the relation Rel to I. Note that for Type 1 relations, both $\overline{Rel(I)}$ and $\underline{Rel(I)}$ are non-empty. Whereas for Type 2 relations, $\underline{Rel(I)}$ is empty.

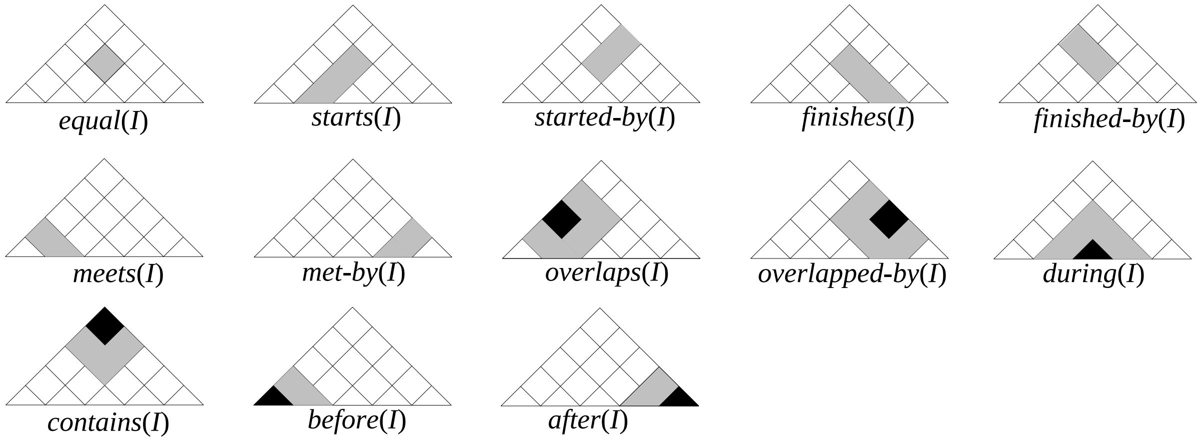


Figure 3-16: Rough sets of intervals in Allen relations with respect to *I*

3.4.5 Queries with Rough Constraints

As discussed in Sections 3.3.2, the set of CTIs in a certain relation to another CTI *I* are modelled as zones in the TM, i.e. $Rel(I)$. These zones can be used to model constraints of temporal relations between CTIs and *I*. By operations on these zones, CTIs that satisfy these constraints can be obtained. Conjunctive queries can be answered by the intersection operation, e.g. $Rel(I_1) \cap Rel(I_2) \cap Rel(I_3)$. For example, CTIs during I_v and overlaping I_u, i.e. $\{I|overlaps(I, I_u) \wedge during(I, I_v)\}$, can be obtained by intersecting zones of $overlaps(I_u)$ and $during(I_v)$ in the TM. As discussed in Section 0, if I_u and I_v are described by RTIs, $overlaps(I_u)$, $during(I_v)$ also become rough sets. Figure 3-17 illustrates the process of obtaining the set of CTIs that *overlap* I_u and *during* I_v, when I_u and I_v are described by RTIs. $\overline{overlaps(I_u)}$ and $\overline{during(I_v)}$, $\underline{overlaps(I_u)}$ and $\underline{during(I_v)}$ are intersected respectively and then both intersections are combined together to form another rough set $< \overline{overlaps(I_u) \cap during(I_v)}$, $\underline{overlaps(I_u) \cap during(I_v)}$. Similarly, disjunctive queries can be expressed by the union operation, e.g. $Rel(I_1) \cup Rel(I_2) \cup Rel(I_3)$. Intervals containing I_w and after I_x are obtained by the union of $contains(I_w)$ and $after\ (I_x)$. If I_w and I_x have upper and lower approximations, the resulting set $\{I|overlaps(I, I_w) \vee during(I, I_x)\}$ will also be a rough set (Figure 3-18).

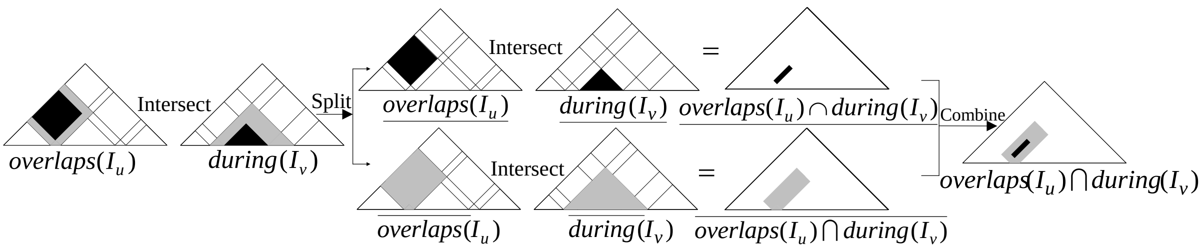


Figure 3-17: Querying for the set $\{I|overlap(I, I_u) \wedge during(I, I_v)\}$ in the TM

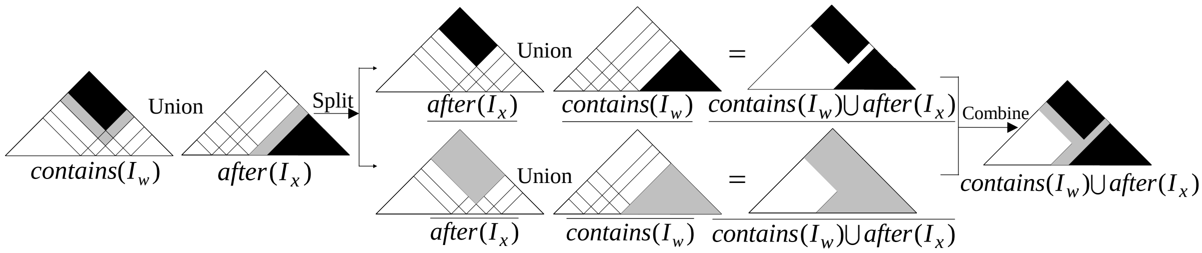


Figure 3-18: Querying for the set $\{I|contain(I, I_w) \vee after(I, I_x)\}$ in the TM

3.5 Modelling Fuzzy Time Intervals in the TM

According to the assumptions of Section 3.2.2, an FTI consist of a non-empty core and support, both of which occupy a convex CTI. If the fuzzy start and the fuzzy end are linear and monotonic, the intervals of the core and the support totally determine the structure of the FTI. De Tré *et al.* (De Tré et al., 2006) modelled such trapezoidal FTIs in the TM as linear segments between the interval point of the core and the interval point of the support (Figure 3-19). The position, inclination and length of the linear segment totally determine the characteristics of an FTI. This approach shows potential of the TM in displaying simple FTIs in historical databases (Figure 3-20). However, this approach only represents the configuration of the core and support of an FTI, rather than the functions of the fuzzy start and end. If FTIs are defined by more complex membership functions, then this approach cannot fully describe their differences. At this point, FTIs with arbitrary functions cannot be adequately modelled within the TM. Further extensions and modifications to the TM need to be considered to represent complex FTIs and the relations between them. This is left for future work. In this section, we will solely investigate the relations between FTIs and CTIs. In contrary to relations between RTIs and CTIs which are represented as discrete geometries, relations between FTIs and CTIs are modelled as continuous fields in the 2D space. In this work, 'fields' refer to 2D areas that are described in a continuous manner. In a field, every possible position is described by a variable or a set of variables (Longley et al., 2001). With operations on these continuous fields, queries about CTI-FTI relations can be answered in a more intuitive way.

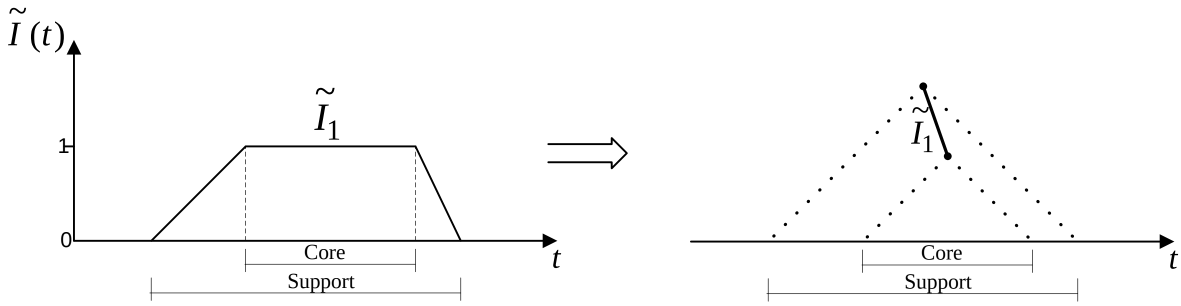


Figure 3-19: Representing a simple FTI as a linear segment in the TM (De Tré *et al.* 2006).

Number	Country	Name	*support* ([a, b])	*core* ([a, b])
1	Belgium	Leopold II	[1865, 1908]	[1865, 1908]
2	Belgium	Albert I	[1908, 1934]	[1908, 1934]
3	France	Loubet	[1898, 1907]	[1899, 1906]
4	France	Fallières	[1906, 1913]	[1907, 1913]
5	Italy	Vittorio Emanuele III	[1899, 1945]	[1900, 1944]
6	Spain	Maria Cristina	[1885, 1902]	[1885, 1902]
7	Spain	Alfonso XIII	[1902, 1931]	[1903, 1930]
8	United Kingdom	Victoria	[1837, 1901]	[1837, 1901]
9	United Kingdom	Edward VII	[1901, 1910]	[1902, 1909]
10	United Kingdom	George V	[1909, 1936]	[1910, 1935]

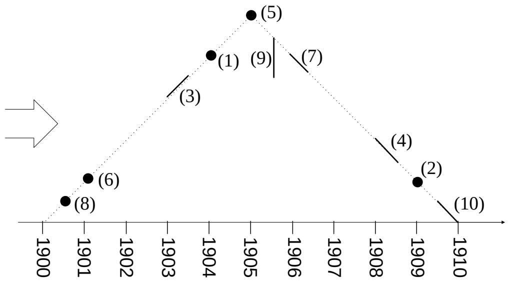


Figure 3-20: Representing trapezoidal FTIs in a historical database by the use of the TM (De Tré et al., 2006), restricted to the study period between 1900 and 1910.

3.5.1 Relations between Crisp and Fuzzy Time Intervals

Relations between two CTIs are determined by binary operators (<, = and >) between their starts and ends. However, these operators do not exist for fuzzy time intervals. Relations between an FTI and another interval (either CTI or FTI) are expressed by a real number between 0 and 1, which quantifies the truth of whether the two intervals are in this relation. Different functions have been proposed to obtain truth values of FTI relations (Nagypál and Motik, 2003; Ohlbach, 2004; Schockaert and De Cock, 2008). Every function takes special care of certain aspects of expressivity and reasoning ability. All these functions can be represented in the TM, yielding continuous fields with different patterns. In this work, we adopt Schockaert's function (Schockaert and De Cock, 2008) of FTI relations for its advantages in the reasoning aspect (e.g. reflectivity, symmetry and transitivity) which offers a better basis for further reasoning research. In the TM, every point representing a CTI, is allotted a value between 0 and 1, expressing the graded truth of whether this CTI is in a certain relation to an FTI. Let $Rel(I_1, \widetilde{I_2})$ denote the truth value of whether a CTI $I_1$is in the relation Rel to a FTI $\widetilde{I_2}$. $Rel(\widetilde{I_2})$ denotes the fuzzy set of CTIs in relation Rel to $\widetilde{I_2}$. Figure 3-21 (left) illustrates three typical situations of the *during* relation between a CTI I_1 and an FTI $\widetilde{I_2}$. We refer to Schockaert's work (Schockaert and De Cock, 2008; Schockaert et al., 2008) for a detailed explanation of the function applied. If the whole part of I_1 is within $\mathrm{core}(\widetilde{I_2})$, then $during(I_1, \widetilde{I_2}) = 1$. If there is a part of I_1 that is totally out of the $\mathrm{support}(\widetilde{I_2})$, then $during(I_1, \widetilde{I_2}) = 0$. In between these two situations, I_1 is partially during I_2, then $0 < during(I_1, \widetilde{I_2}) < 1$. In this way, every I in $I\mathbb{R}$ may have a value of $during(I, \widetilde{I_2})$. The set of CTIs $during(\widetilde{I_2})$ forms a fuzzy set, denoted as $during(\widetilde{I_2})$, and is modelled as a continuous field in the TM (Figure 3-21 right). Analogously, fuzzy sets of other Allen relations with respect to $\widetilde{I_2}$ can also be represented by continuous fields (Figure 3-22). This approach is not restricted to trapezoidal FTIs as we illustrated. It can also apply to FTIs described by other functions.

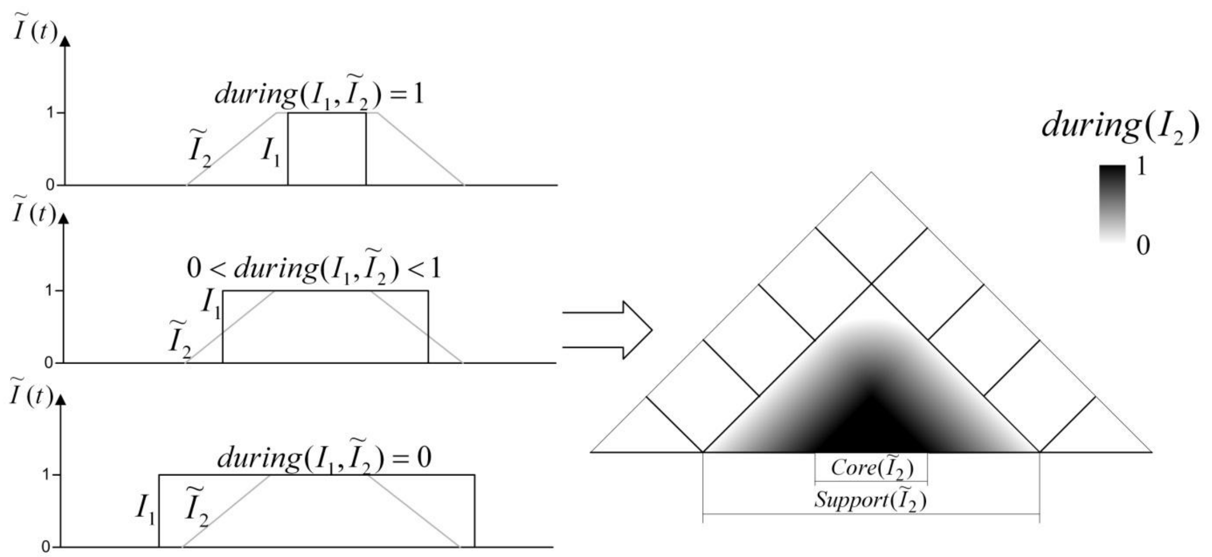


Figure 3-21: Representing the fuzzy set $during(\widetilde{I_2})$ in the TM

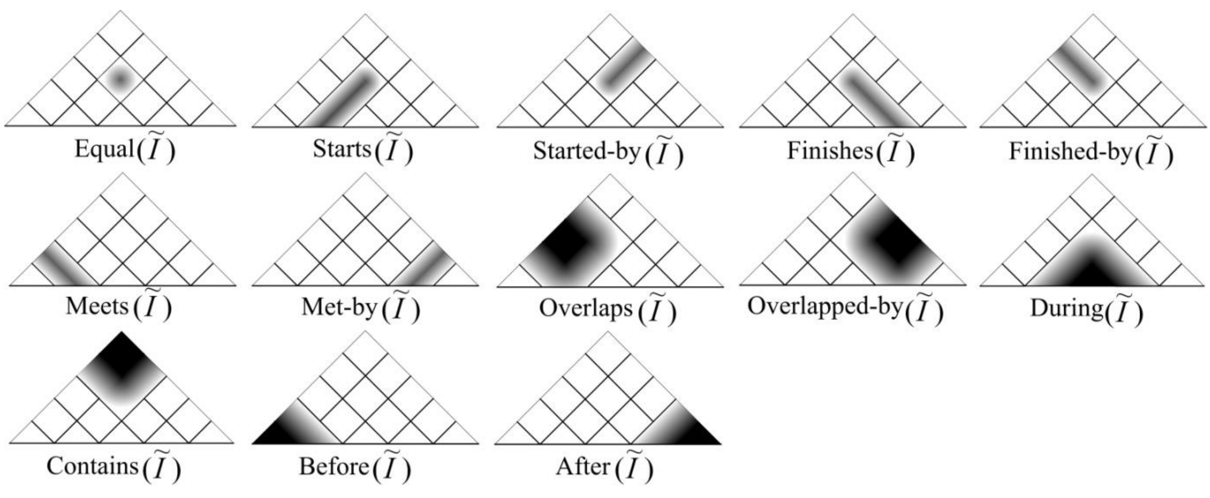


Figure 3-22: Fuzzy sets of Allen relations to $\widetilde{I_2}$ of Figure 3-21

3.5.2 Queries with Fuzzy Constraints

Temporal queries with fuzzy constraints can be modelled in the TM as well. As discussed in the previous subsection, constraints of relations between CTIs and an FTI can be modelled by continuous fields in the TM. By spatial operations on these continuous fields, one may obtain the set of CTIs that satisfy these constraints. The obtained set of CTIs is also a fuzzy set and modelled as a continuous field in the TM. For example, CTIs that *contain* $\tilde{I}_1$ constitute a fuzzy set, i.e. $contains\ (\tilde{I}_1)$, while CTIs that *contain* $\tilde{I}_2$ form another fuzzy set, i.e. $contains(\tilde{I}_2)$. Both $contains\ (\tilde{I}_1)$ and $contains\ (\tilde{I}_2)$ are modelled as continuous fields in the TM. The intersection of $contains\ (\tilde{I}_1)$ and $contains\ (\tilde{I}_2)$ yields the fuzzy set of CTIs that both contain $\tilde{I}_1$ and $\tilde{I}_2$, i.e. $contains\ (\tilde{I}_1) \cap contains\ (\tilde{I}_2)$ (Figure 3-23). In the TM, the intersection of fuzzy sets can be obtained by applying fuzzy intersection to every point in the two-dimensional $I\mathbb{R}$. In Figure 3-23, the fuzzy intersection uses the minimum t-norm (Dubios and Prade 2000). Of course, this approach is also compatible to other t-norms. Similarly, unions of fuzzy sets can also be obtained by operations on the continuous fields. For example, in Figure 3-24, the union of $contains\ (\tilde{I}_1)$ and $contains\ (\tilde{I}_2)$ is obtained by the *maximum t-conorm* operation (Dubios and Prade, 2000) applied to values in the two continuous fields of $contains\ (\tilde{I}_1)$ and $contains\ (\tilde{I}_2)$.

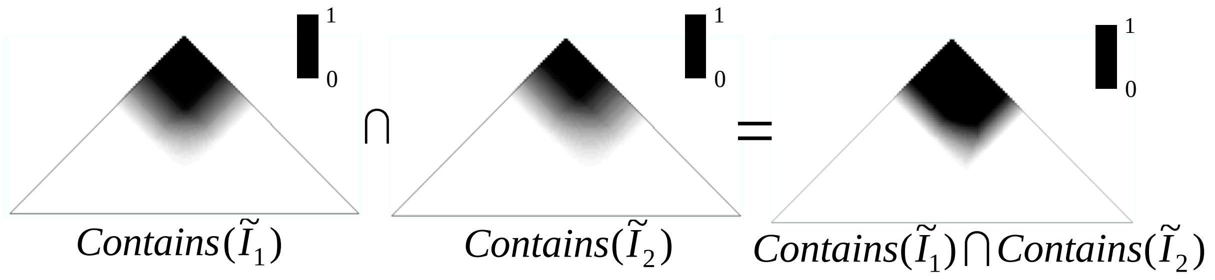


Figure 3-23: Temporal query for CTIs that *contain* both $\tilde{I}_1$ and $\tilde{I}_2$, using the *minimum t-norm*

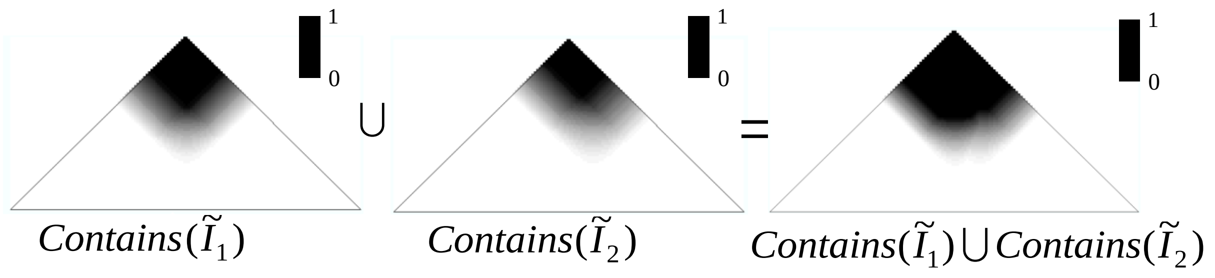


Figure 3-24: Temporal query for CTIs that either *contain* $\tilde{I}_1$ or $\tilde{I}_2$, using the *maximum t-conorm*

3.6 Solving Practical Problems

Because of the remaining problems and difficulties with FTI modeling in the TM, as described in the previous section, in this section we only deal with RTIs to illustrate the usefulness and applicability of the TM for the representation of RTIs in a two-dimensional space. During World War One, aerial photos covering the Belgian-German front line in West-Flanders (Belgium) were taken at discrete time stamps. From these aerial photos, we can observe whether a military feature (e.g. a fire trench, gun position or barrack) was not yet present, present, or destroyed. Although the state of a feature is uncertain in between two time stamps, we assume that it does not change in between two snapshots which show similar states, such that the uncertainty only remains in between two snapshots showing different states. Certainly, this assumption relies on our knowledge that snapshots are dense enough to capture most of features' changes. When more volatile entities are considered, an appropriate temporal resolution will be required. In this context, RTIs are excellently suited to handle time modelling. Indeed, as no information is available about the state of a feature in between two time stamps, the construction of adequate membership functions for FTIs would induce an extra overhead and difficulties in the time modelling and handling. Together with the unsolved representation difficulties for FTIs in the TM this justifies the choice to focus the case study under consideration on the use of RTIs.

We can consider a period of snapshots showing similar states as a lower approximation for this state, its neighbouring uncertain intervals as boundary region, and all of them form the upper approximation (Figure 3-25). Thus, a feature's lifetime can be meaningfully represented by an RTI. Basically, at least four photos are required to determine the lifetime of a feature (Figure 3-25): (1) the last photo in which the feature is not yet present, (2) the first photo in which the feature is present, (3) the last photo in which the feature is present, and (4) the first photo in which the feature is destroyed or abandoned. The interval between the dates of photo (2) and (3) is the lower approximation of the feature's lifetime, while interval between the dates of photo (1) and photo (4) is the upper approximation. Intervals between the dates of photo (1) and (2), and intervals between the dates of photo (3) and (4) are boundary regions, which indicate respectively the range of the feature's construction and destruction dates. There are a few exceptions, where a feature was not yet present in one photo and already destroyed in the following photo. Since photo (2) and (3) are missing, the RTI for these features has an empty lower approximation. As described in Section 4.2, such RTIs are represented as triangles on the horizontal axis. The dates of the photos have been obtained from the database of the military features in which they were originally stored. From the database, people cannot easily capture distributions of features' RTIs. However, the TM can display these RTIs in a more visible way.

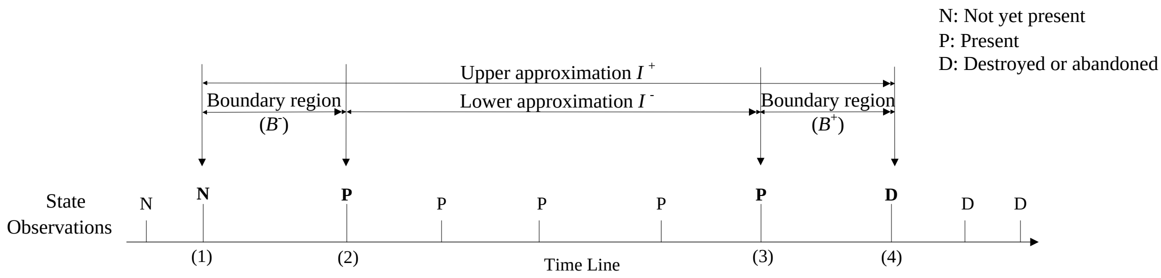


Figure 3-25: The RTI of a military feature

Figure 3-26 displays RTIs of a type of fire trenches (i.e. FT1) in the TM. We assign darker colours to areas that have more overlapped polygons. From this figure we can see that most of FT1s are present in the beginning of the war, because there is a dark zone distributed at the left corner of the study area. Some other FT1s were randomly distributed through the war. However, in the very right corner of the study area, we can see a small group of FT1s that were specifically present at the end of the war. Compared with the data stored in database tables, the TM provides a more direct visualisation of the distribution of RTIs.

Moreover, the TM may be combined with traditional geographical maps in order to support spatiotemporal analysis. In Figure 3-27, RTIs of barracks' lifetimes are displayed in the TM where we can see barracks are temporally distributed in two clusters. The construction dates of barracks in these two clusters overlap in most of the cases, whereas their destruction dates are clearly distributed in two distinctive periods. Barracks in the first cluster are mostly destroyed in 1916, while barracks in the second cluster are mostly destroyed in the second half of 1917. When checking the geographical distribution of these barracks (Figure 3-28), one may observe that most barracks of the second cluster are further away from the front line than barracks of the first cluster. From this observation, people may infer that barracks near the front line were destroyed earlier than barracks further away from the front line. According to records of the war, fighting along the front line was getting increasingly intensive during the period of the two clusters. This fact can probably explain our finding in the TM: barracks near the front line were destroyed or abandoned due to intensive fighting;

while barracks further away from the front line survived for a longer time. In general, the TM offers people a preliminary perception of the distribution of RTIs, which cannot be easily done by traditional representations. This perception can guide people to set up specific hypothesis to be tested by further studies.

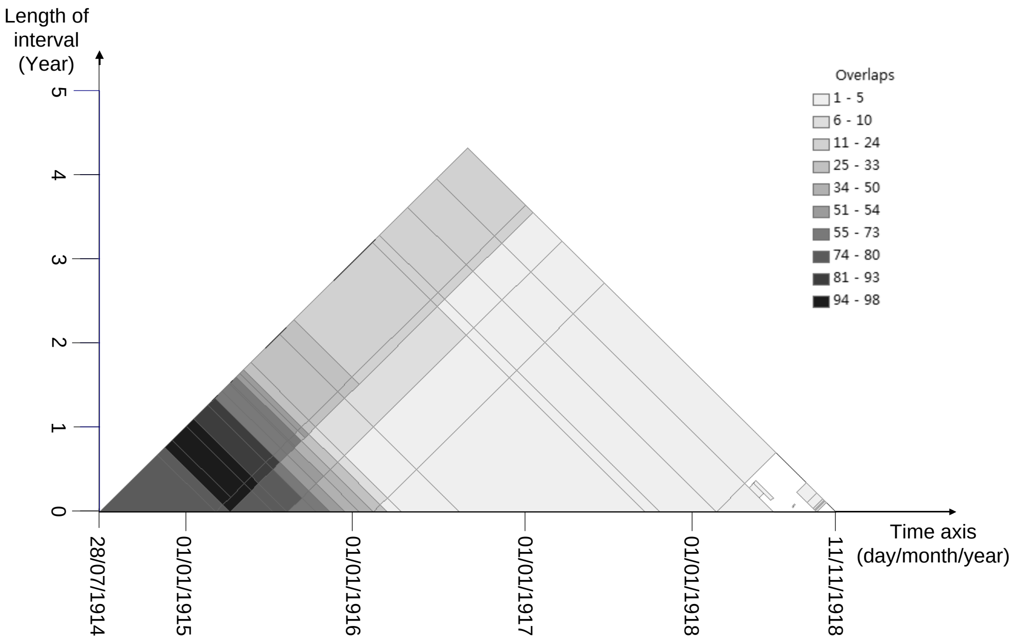


Figure 3-26: RTIs of FT1 features in the TM

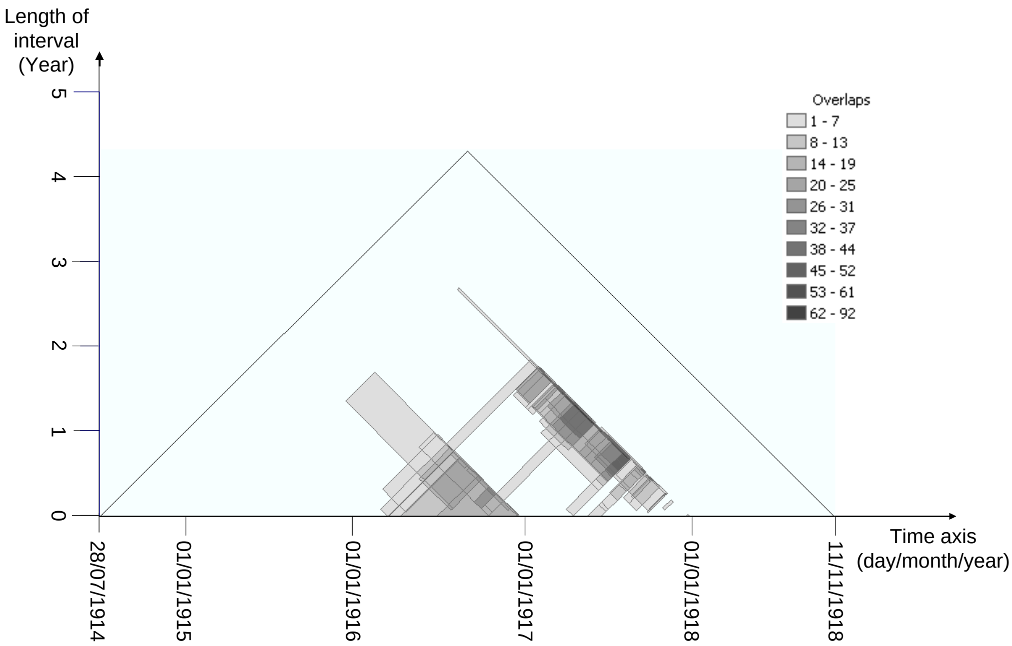


Figure 3-27: RTIs of barracks in the TM

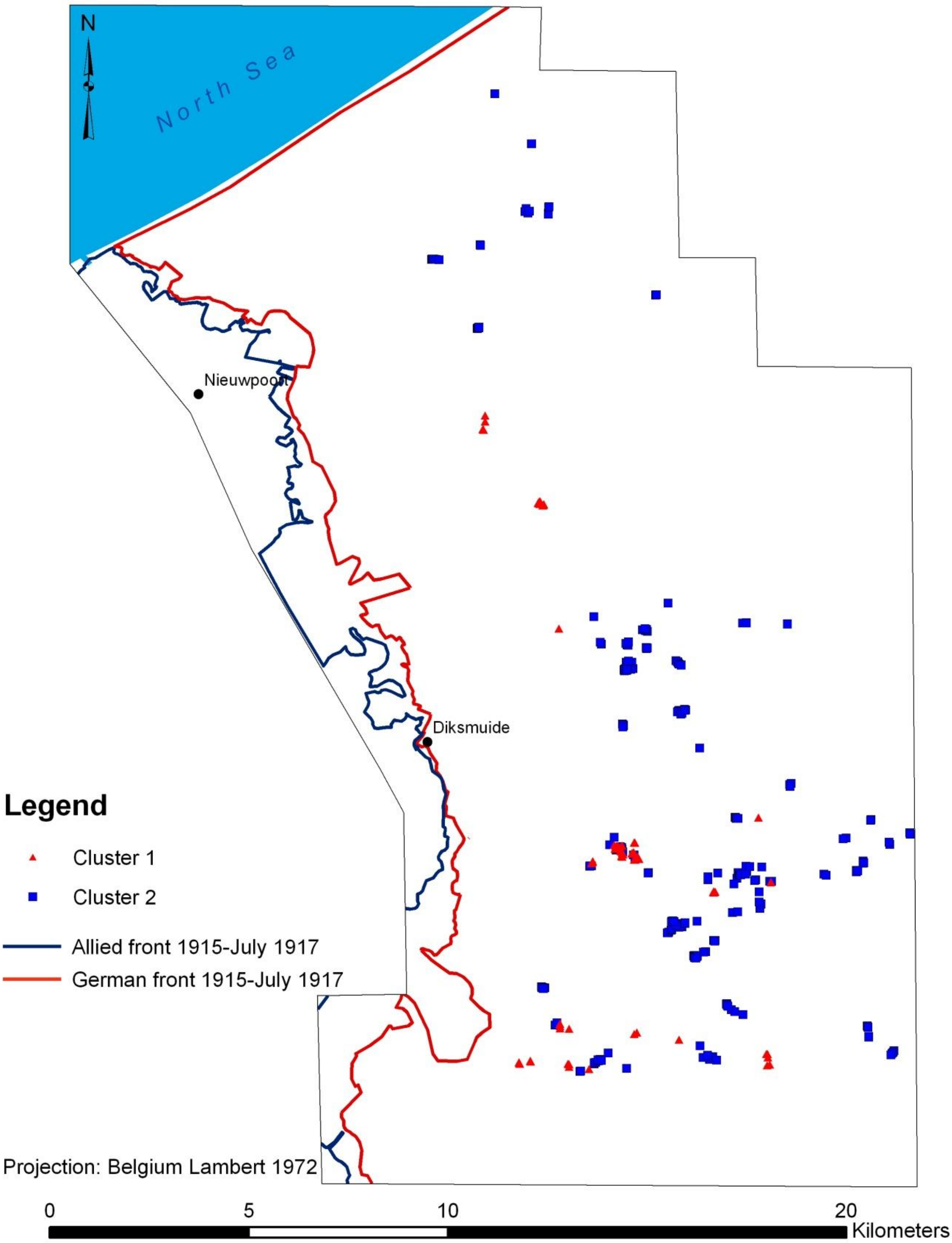


Figure 3-28: Geographical locations of the two barrack clusters

3.7 Conclusion and Future Work

In this chapter, we have discussed two types of imperfect time intervals (i.e. rough time intervals and fuzzy time intervals), and investigated their representations in the TM. In the TM, RTIs are represented by discrete geometries in a two-dimensional space. According to the shapes and locations of the geometries, one can easily read the properties of an RTI. Compared with the linear representation, the temporal relations between RTIs are explicitly expressed by the topological relations between the geometries. This representation allows more direct perception of temporal relations between RTIs and facilitates the finding of new relations and structures that are easy to perceive in the conventional linear representation. Also, according to the spatial configuration and topology of RRZs, one can easily deduce the possible relations between the CTIs that are approximated by the RTIs. Moreover, we have discussed how temporal queries involving RTIs can be modelled as 2D rough sets (i.e. rough regions) and the operations of multiple queries. Besides the reasoning issues, a concrete scenario is introduced to demonstrate the potential of the TM in visualising large amounts of RTIs in an archaeological data for exploratory spatio-temporal data analysis (Andrienko, G and Andrienko, N, 2006).

Due to the continuous and complex structure of FTIs, FTI reasoning has not been investigated as thoroughly as RTI reasoning. In the future, more work is needed to effectively represent complex FTIs and their relations in the TM. Possibly, more dimensions and visual metaphors have to be added to the TM. In this chapter, the discussion is limited to the representation of relations between CTIs and FTIs and temporal queries involving FTI constraints. In the TM, temporal query with FTI constraints are represented as 2D fuzzy sets (i.e. continuous field), in which people can observe the range of the answers with the degree of truthiness. If the FTIs are modelled in a possibilistic or probabilistic context, the fuzzy areas can be interpreted as possibility or probability densities. This feature is potentially useful for the development of a visual calendar for people to make plans with respect to fuzzy temporal constraints.

The TM provides an alternative representation of temporal information, where people can investigate imperfect time intervals or relations of imperfect time intervals from a different perspective. Compared to the conventional linear model, the TM offers a more efficient and descriptive representation of imperfect time intervals. The shapes and locations of every time interval is fixed and unique, offering the possibility of displaying the distribution of large numbers of intervals. Moreover, sets (crisp, rough or fuzzy) of intervals can be represented as graphics in the TM, and the operations of the set can be modelled as spatial operations of the graphics, which is more perceptible compared to mathematical formulas. Objectively speaking, the TM does not create a new nor extend an existing temporal calculus or logic, however, it takes special care of intuitive and visualisation aspects. Since many evidences have shown that diagrams can be effective means for humans to understand, communicate and studying abstract knowledge (Giere, 1999; Tversky, 2005), we contend that the TM can further benefit the research of temporal reasoning and analysis.

Future research will take account of both theoretical and application aspects. The former issues include the continued work on the representation of complex FITs in the TM (or extensions thereof). The representation of FTI-FTI and FTI-RTI relations in the TM will also be studied. Furthermore, we will investigate whether the TM can express temporal semantics based on possibility theory and probability theory. Regarding the application issues, we plan to elaborate more use cases in order to evaluate the TM within different contexts. Also, we will implement the TM as a software application with an interactive interface for people to handle temporal information in general and imperfect time intervals in particular. This interactive graphical tool may be a useful add-on for temporal database systems or information systems. The tool may assist users in analysing distributions of time intervals, or in defining and answering flexible queries by manipulating the graphic objects. For example, the TM can be incorporated with a geographical information system in order to display temporal information of geographical objects. By interactively linking the TM view and the map view, people can flexibly analyze geographical objects from both spatial and temporal aspects.

References

Allen, J. F. (1983). Maintaining Knowledge about Temporal Intervals. Communications of the ACM, 26(11), pp.832-843.

Allen, J. F. and Hayes, P. J. (1985). A Common-sense Theory of Time. Proceedings of the 9th International Joint Conference on Artificial Intelligence, Los Angeles, California, pp.528-531, Morgan Kaufmann Publishers Inc.

Andrienko, N. and Andrienko, G. (2006). Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach. Springer-Verlag Berlin, Heidelberg.

Bassiri, A., Malek, M., Alesheikh, A. and Amirian, P. (2009). Temporal Relationships between Rough Time Intervals. ed. by Gervasi, O., Taniar, D., Murgante, B., Laganà, A., Mun, Y. and Gavrilova, M., in Computational Science and Its Applications – ICCSA 2009, pp.543-552, Springer Berlin / Heidelberg.

Bittner, T. (2002). Approximate Qualitative Temporal Reasoning. Annals of Mathematics and Artificial Intelligence, 36(1), pp.39-80.

De Caluwe, R., Van Der Cruyssen, B., De Tré, G., Devos, F. and Maesfranckx, P. (1999). The Semantics and Modelling of Flexible Time Indications. in Computing with Words in Information/Intelligent Systems 1 (Foundations), pp.229-256, Springer/Physica-Verlag, Heidelberg, Germany.

De Caluwe, R., Van Der Cruyssen, B., De Tré, G., Devos , Filip and Maesfranckx, P. (1997). Fuzzy Time Indications in Natural Language Interfaces. in Flexible Query Answering Systems, pp.163-185, Kluwer Academic Publishers.

De Tré, G., Van de Weghe, N., De Caluwe, R. and De Maeyer, P. (2006). Towards a Flexible Visualization Tool for Dealing with Temporal Data. ed. by Larsen, H., Pasi, G., Ortiz-Arroyo, D., Andreasen, T. and Christiansen, H., in Flexible Query Answering Systems, pp.109-120, Springer Berlin, Heidelberg.

Dubios, D. and Prade, H. (2000). Fundamentals of Fuzzy Sets. Kluwer Academic Publisher Group, AH Dordrecht, The Netherlands.

Dubois, D., HadjAli, A. and Prade, H. (2003). Fuzziness and uncertainty in temporal reasoning. Journal of Universal Computer Science, 9(9), pp.1168-1194.

Dubois, D., Lang, J. and Prade, H. (1991). Timed possibilistic logic. Fundamenta Informaticae, 15(3-4), pp.211-234.

Freksa, C. (1992). Temporal Reasoning Based on Semi-Intervals. Artificial Intelligence, 54(1-2), pp.199-227.

Galton, A. (1990). A Critical Examination of Allen's Theory of Action and Time. Artificial Intelligence, 42(2-3), pp.159-188.

Garrido, C., Marín, N. and Pons, O. (2009). Fuzzy Intervals of Represent Fuzzy Valid Time In a Temporal Relational Database. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 17(1), pp.173-192.

Giere, R. N. (1999). Science Without Laws. The University of Chicago Press, Chicago.

Knight, B. and Ma, J. (1993). Time representation: A taxonomy of temporal models. Artificial Intelligence Review, 7(6), pp.401-419.

Kulpa, Z. (1997). Diagrammatic Representation for a Space of Intervals. Machine Graphics & Vision, 6, pp.5-24.

Kulpa, Z. (2006). A Diagrammatic Approach to Investigate Interval Relations. Journal of Visual Languages and Computing, 17(5), pp.466-502.

Ladkin, P. B. (1987). The Logic of Time Representation. Ph.D. thesis, University of California, Berkeley.

Lin, Y. (1991). Two Theories of Time. Journal of Applied Non-Classical Logics, 1(1), pp.37-63.

Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W. (2001). Geographic Information Systems and Science. John Wiley & Sons, Ltd., Chichester, England.

Nagypál, G. and Motik, B. (2003). A Fuzzy Model for Representing Uncertain, Subjective, and Vague Temporal Knowledge in Ontologies. Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics, (ODBASE), Catania, Italy, pp.906-923.

Ohlbach, H. J. (2004). Relations between fuzzy time intervals. Proceedings of 11th International Symposium on Temporal Representation and Reasoning (TIME 2004), Montréal, Québec, pp.44-51.

Pawlak, Z. (1982). Rough Sets. International Journal of Information and Computer Science, 11, pp.341-356.

Pawlak, Z. (1992). Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers.

Schockaert, S. and De Cock, M. (2008). Temporal Reasoning about Fuzzy Intervals. Artificial Intelligence, 172(8-9), pp.1158-1193.

Schockaert, S., De Cock, M. and Kerre, E. E. (2008). Fuzzifying Allen's Temporal Interval Relations. IEEE Transactions on Fuzzy Systems, 16(2), pp.517-533.

Shoham, Y. and Goyal, N. (1988). Temporal Reasoning in Artificial Intelligence. in Exploring Artificial Intelligence, pp.419-438, Morgan Kaufmann Publishers Inc.

Tversky, B. (2005). Some ways images express and promote thought. ed. by Grialou, P., Longo, G. and Okada, M., in Images and Reasoning, Keio University, Tokyo.

Van de Weghe, N., Docter, R., De Maeyer, P., Bechtold, B. and Ryckbosch, K. (2007). The Triangular Model as an Instrument for Visualising and Analysing Residuality. Journal of Archaeological Science, 34(4), pp.649-655.

Vila, L. (1994). A survey on temporal reasoning in artificial intelligence. AI Communications, 7(1), pp.4-28.

Zadeh, L. A. (1965). Fuzzy Sets. Information and Control, 8(3), pp.338-353.

Zadeh, L. A. (1975). Fuzzy logic and approximate reasoning. Synthese, 30(3), pp.407-428.

4 Analysing Crisp Time Intervals

Modified from: Yi Qiang, Matthias Delafontaine, Mathias Versichele, Philippe De Maeyer, Nico Van de Weghe, (2012). Interactive Analysis of Time Intervals in a Two-Dimensional Space. *Information Visualization* (In press)

Abstract: Time intervals are conventionally represented as linear segments in a one-dimensional space. An alternative representation of time intervals is the Triangular Model (TM), which represents time intervals as points in a two-dimensional space. In this chapter, the use of the TM in visualising and analysing time intervals is investigated. Not only does this model offer a compact visualisation of the distribution of intervals, it also supports an innovative temporal query mechanism that relies on geometries in the two-dimensional space. This query mechanism has the potential to simplify queries that are hard to specify using traditional linear temporal query devices. Moreover, a software prototype that implements the TM in a geographical information system (GIS) is introduced. This prototype has been applied in a real scenario to analyse time intervals that were detected by a Bluetooth tracking system. This application shows that the TM has potential of supporting a traditional GIS to analyse interval-based geographical data.

Keywords: Time interval, Temporal query, Triangular Model, GIS, Spatio-temporal data analysis

4.1 Introduction

Considerable work has been done in modelling and reasoning about time intervals, ranging from qualitative relational algebras (Allen, 1983; Freksa, 1992), to rough and fuzzy extensions (Bittner, 2002; Ohlbach, 2004; Schockaert et al., 2008), to applications in temporal queries (Bassiouni and Llewellyn, 1999; Toman, 1996) and temporal databases (Garrido et al., 2009; Ozsoyoglu and Snodgrass, 1995). Compared with the many contributions made to the reasoning aspect, methods and tools that visualise and analyse time intervals have received far less attention. This situation may be explained by the limitations of the conventional representation of time intervals, i.e. the linear representation. In this representation, time intervals are modelled as linear segments in a one dimensional space. The second dimension is often exploited as well (e.g. gantt chart and time table). However, this dimension is solely to differentiate the intervals of different entities and thus has no temporal meaning. Thus, in the linear representation, the visual distribution of intervals may vary, according to different ordering rules applied to the second dimension. This polymorphism prohibits the existence of a universal visual recognition approach for detecting patterns among time intervals. As a result, this hampers the development of visual techniques and tools to handle interval-based data, for instance, techniques of exploratory data analysis (EDA) (Tukey, 1977) that greatly rely on information visualisation. Note that there exist other representations of time intervals, e.g. cyclic representations (Li and Kraak, 2008; Weber et al., 2001) and calendars (Van Wijk and Selow, 1999). However, they focus on special aspects of time intervals, and are not applicable in a broader range of contexts.

This problem also exists in many information systems that deal with interval-based data, for instance, geographical information systems (GIS), which considers time as one of the most important components. In the recent development of GIS, great efforts have been made in spatio-temporal analysis (Longley et al., 2001; Smith et al., 2007), especially exploratory spatio-temporal data analysis (ESTDA) (Andrienko, G and Andrienko, N, 2006; Andrienko et al., 2003; Hardisty and Klippel, 2011). Nevertheless, the existing approaches and tools show limited ability in analysing geographical data with reference to time intervals. In traditional GIS, time intervals are usually represented by cartographical variables (e.g. symbols, labels and colours) (Stojanovic et al., 1999), animations (Hibbard et al., 1994) or spatio-temporal composites (Langran, 1993). These representations cannot explicitly display the distribution of time intervals, and thus limits their suitability for pattern detection. Also the lack of interactivity obstructs their application in ESTDA. Recently, thanks to the developments of computer science, a variety of innovative techniques have been introduced to analyse spatio-temporal data (Beard et al., 2007; Butkiewicz et al., 2008; Guo et al., 2006; Kapler and Wright, 2005; Tominski et al., 2005). In spite of that, since most techniques are still based on the linear time representation, their ability of analysing interval-based geographical data is not satisfactory. On the one hand, time is often directly integrated with the two-dimensional (2D) space into a three-dimensional (3D) space, i.e. the space-time cube (Delafontaine et al., 2011; Neutens et al., 2008; Neutens et al., 2007a; Neutens et al., 2007b; Versichele et al., 2012). Gatalsky et al. (2004) implemented the space-time cube into an interactive environment to display geographical events that occur in very short intervals (i.e. instantaneous events). This approach can also apply to events with a temporal extent, for example, representing such events as linear segments parallel with the time axis in the cube (Huisman et al., 2009). However, implementations of the space-time cube suffer from typical problems of 3D visualization (Forlines and Wittenburg, 2010; Kjellin et al., 2009) which will be specifically explained in the discussion section of this chapter. On the other hand, attempts have been made to display time and space through interactively coordinated visualizations (Monmonier, 1989). Fredrikson (1999), for instance, utilised a visualisation coordination system, i.e. Snap-Together Visualisation (North and Shneiderman, 2000), to analyse aggregates of spatio-temporal data. In this system, temporal, geographical and categorical attributes are displayed in several coordinated views. As the time view of this tool is still based on the linear representation, it is limited to instantaneous events and does not show any potential in analysing interval-based data.

To overcome these difficulties, an alternative representation of time intervals can be considered. Kulpa (1997) introduced a presentation of time intervals as points in a metric 2D space. He elaborated the use of this diagrammatic representation in interval reasoning and arithmetic (Kulpa, 2001). Later, Van de Weghe et al. (2007) named this representation the Triangular Model (TM) and applied it into an archaeological use case. Recently, Qiang et al. (2010) have extended the TM to represent and reason about rough and fuzzy time intervals. Beyond their theoretical extensions, these contributions have presumed considerable potential of the TM for visualising and analysing time intervals. However, until present, the analytical power of the TM has not been thoroughly investigated. To fill this gap, this chapter investigates the use of the TM for querying and analysing time intervals. In order to demonstrate the uses of TM, a prototype tool GeoTM is developed, which implements the TM in a geographical information system (GIS).

The remainder of this chapter first introduces the basic concept of the TM (Section 4.2). In Section 4.3, we describe how queries about time intervals can be defined in TM. Section 4.4 introduces the implementation of TM, i.e. GeoTM, including its graphical user interface and functionalities. In Section 4.5, this implementation is applied in a real scenario to analyse time intervals that were detected by a Bluetooth tracking system. In Section 4.6, the strengths and weaknesses of GeoTM are discussed and compared with alternative approaches. Finally, Section 4.7 draws conclusions and proposes the directions of future work.

4.2 Triangular Model

A crisp time interval I is usually modelled as a pair of real numbers $[I^-, I^+]$ with $I^- < I^+$, denoting the start and end of the interval respectively. In the linear representation, such a time interval is represented as a finite line segment (Figure 4-1 (a)). The two extremes of the segment respectively represent the start (I^-) and the end (I^+) of the interval, while the length of the segment expresses the duration of the interval ($dur(I)$). This linear representation of time intervals is widely used in our daily life, for example in Gantt charts and historical time lines. In this chapter, we assume a crisp time interval I to be closed at both sides, such that $I = [I^-, I^+]$. In different reasoning systems, whether an interval is open (at one or both sides) is defined differently (Vila, 1994). Since the TM does not intend to solve reasoning issues that concern this controversy, whether an interval is open does not affect its representation in TM.

The transformation from the linear representation to the Triangular Model (TM) starts from the construction of two lines through the extremes of an interval (Figure 4-1 (b)). For each time interval I, two straight lines (L_1 and L_2) are constructed, with L_1 passing through I^- and L_2 passing through I^+. α_1 is the angle between L_1 and the horizontal axis and α_2 is the angle between L_2 and the horizontal axis, with $\alpha_1 = -\alpha_2 = \alpha$. The intersection of L_1 and L_2 is called the interval point. The start of the interval (I^-), the end of the interval (I^+) and the interval point form an isosceles triangle. The angle α is a predefined constant that is identical for all time intervals, in order to ensure that each time interval is mapped to a unique point in the 2D space. The position of an interval point completely determines both, the start and the end of the interval. α can be different values for specific purposes. In this chapter, we set $\alpha = 45°$, as to be consistent with earlier work (Kulpa, 1997; Kulpa, 2001; Qiang et al., 2010; Van de Weghe et al., 2007). In this way, all time intervals can be represented as 2D points in the TM (Figure 4-1 (c)). With $\alpha_1 = -\alpha_2$, it is straightforward to deduce that, in the horizontal dimension, the position of the interval point indicates the middle point of the interval, i.e. $mid(I)$. In the vertical dimension, the height of the interval point (h) is proportional to the length of the linear interval (l), i.e. $h = \frac{\tan\alpha}{2} \cdot l$. As a result, the height of an interval point in the TM indicates the duration of the interval, i.e. $dur(I)$. Note that the TM is compatible with time instants, which can be characterised as time intervals with a zero length. Time instants are represented as points that coincide with the time axis, e.g. I_5 in Figure 4-1(c). Given that time instants are also modelled through horizontally aligned points in the linear representation, both the TM and the linear representation can be considered identical for this specific class of intervals.

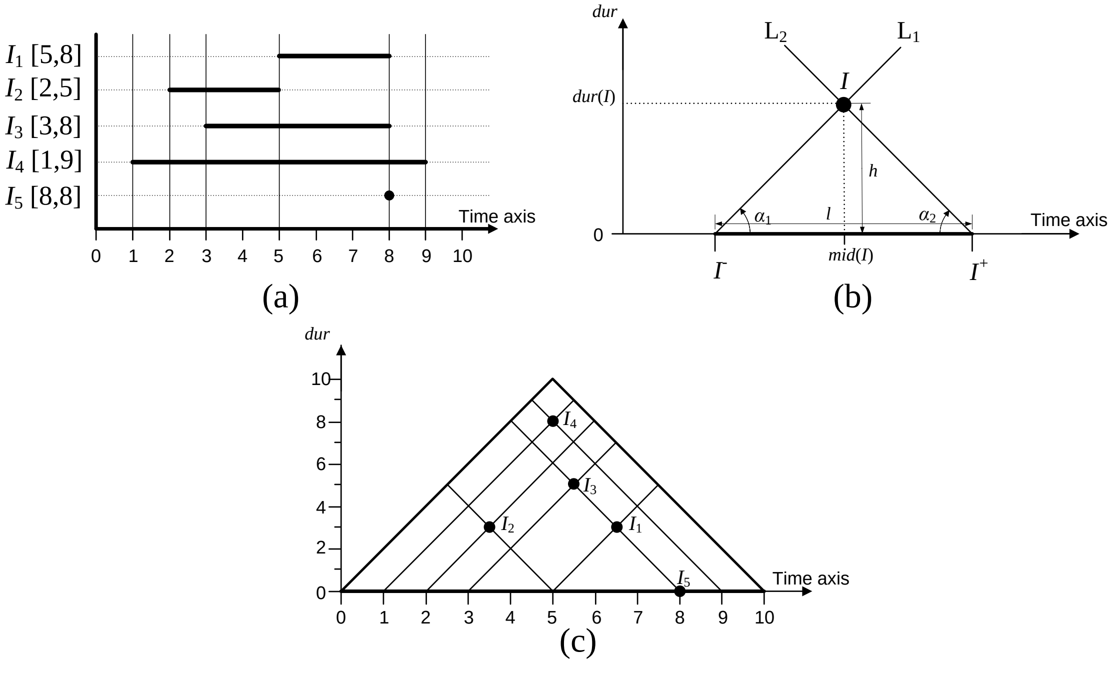


Figure 4-1: The transformation from the linear representation to the Triangular Model (TM). (a): The linear representation of time intervals. (b): The construction of an interval point in TM. (c): The TM representation of time intervals.

4.3 Temporal Queries in the Triangular Model

Since the TM represents time intervals as points in a 2D space, the TM coordinates have temporal semantics. In this section, we will describe how temporal queries can be modelled in TM. The implementation of these temporal queries will be illustrated in Section 4.4.

4.3.1 Queries of Allen Relations

According to the outcome of the logical comparison (i.e. smaller than, equal to and larger than) between the starts and ends of two time intervals, James F. Allen (Allen, 1983) defined thirteen relations between two time intervals (Table 3-1), which are referred to as Allen relations. In TM, Allen relations can be transferred to spatial relations (Kulpa, 1997). Given a study interval I [0, 100], all examined intervals are located within the isosceles triangle formed by I^-, I^+ and the interval point of I. Let us consider a reference interval I_1 [33,66] and several intervals (I_2, I_3, I_4) that are *before* interval I_1 (Figure 3-4 (a)). In TM, I_2, I_3, I_4 are located in the zone in the left corner of the study interval (Figure 3-4 (b)). Therefore, it is easy to deduce that this zone (i.e. the black zone in Figure 3-4 (c)) encloses all intervals that are before I_1. In like manner, all Allen relations with respect to an interval can be represented by such zones in the TM (see Figure 3-5) (Kulpa, 1997, 2006). In each diagram in Figure 3-5, the reference interval I_1 has been chosen in the centre of the study period to avoid visual bias. Each black zone represents the set of intervals that are in a specific relation to I_1. Such a set of intervals is denoted as $R(I_1)$, where R stands for a certain temporal relation (e.g. *before*, *after*, etc.). Thus, temporal queries based on Allen relations can be modelled as zones in TM. The query constraint is expressed by the extent of the zone which encloses all intervals satisfying the query. The zones expressing temporal queries are called query zones.

Table 4-1: Thirteen Allen Relations (Allen 1983)

I_1 *equal* I_2	if $I_1^- = I_2^- \wedge I_1^+ = I_2^+$
I_1 *starts* I_2	if $I_1^- = I_2^- \wedge I_1^+ < I_2^+$
I_1 *started−by* I_2	if $I_1^- = I_2^- \wedge I_2^+ < I_1^+$
I_1 *finishes* I_2	if $I_1^+ = I_2^+ \wedge I_1^- > I_2^-$
I_1 *finished−by* I_2	if $I_1^+ = I_2^+ \wedge I_2^- > I_1^-$
I_1 *meets* I_2	if $I_1^+ = I_2^-$
I_1 *met−by* I_2	if $I_2^+ = I_1^-$
I_1 *overlaps* I_2	if $I_2^- > I_1^- \wedge I_1^+ < I_2^+ \wedge I_1^+ > I_2^-$
I_1 *overlapped−by* I_2	if $I_1^- > I_2^- \wedge I_1^- < I_2^+ \wedge I_2^+ < I_1^+$
I_1 *during* I_2	if $I_1^- > I_2^- \wedge I_1^+ < I_2^+$
I_1 *contains* I_2	if $I_2^- > I_1^- \wedge I_2^+ < I_1^+$
I_1 *before* I_2	if $I_1^+ < I_2^-$
I_1 *after* I_2	if $I_2^+ < I_1^-$

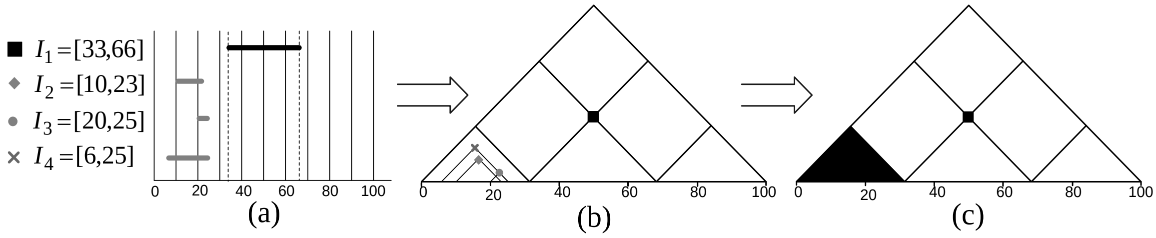


Figure 4-2: Temporal relations in the linear model and in TM, taking *before* as an example.

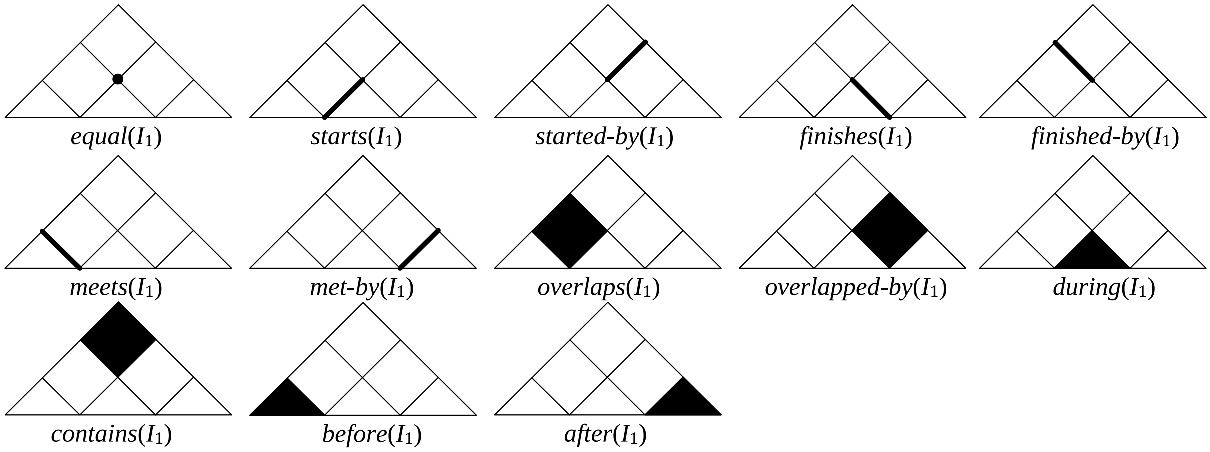


Figure 4-3: The representation of the thirteen Allen relations in TM. Each black zone represents a set of intervals in a specific Allen relation to the reference interval.

4.3.2 Other Temporal Queries

Besides queries of Allen relations, a variety of queries based on other temporal relations can be modelled in TM. In this section, we introduce four of them that have been implemented in the current version of GeoTM. First, intervals *in-between* two different time intervals are within a rectangle with sides in α or $-\alpha$ angle to the horizontal axis (Figure 4-4) (Kulpa, 2006). $\ll I_1, I_2 \gg$ is used to denote the set of intervals *in-between* I_1 and I_2. $\ll I_1, I_2 \gg$ consists of all intervals whose starts are in-between the start of I_1 and the start of I_2, and ends are in-between the end of I_1 and the end of I_2. The formal definition of $\ll I_1, I_2 \gg$ can be found in Eq. 1. A set of intervals *in-between* two time intervals is a convex set of intervals, because it can be directly interpreted by Allen relations or combinations of Allen relations. If the rectangle intersects the horizontal axis, the query zone is only the part above the horizontal axis, because there is no valid time interval below the horizontal axis according to our definition. Additionally, *starts-within* (I_1) is defined as a set of intervals whose starts are in I_1 (Eq. 4-2). Analogously, *ends-within* (I_1) is defined as a set of intervals whose ends are in I_1 (Eq. 4-3). The query zones of *starts-within* (I_1) and *ends-within* (I_1) are illustrated in Figure 4-5 (a) and Figure 4-5 (b). Moreover, a set of intervals whose durations are in a specific range are located within a range along the vertical axis. For example, in Figure 4-5 (c), the set of intervals $\{I|a < dur(I) < b\}$ is within the zone between a and b along the vertical axis.

$$\ll I_1, I_2 \gg \equiv \{I|\, min(I_1^-, I_2^-) < I^- < max(I_1^-, I_2^-) \wedge min(I_1^+, I_2^+) < I^+ < max(I_1^+, I_2^+)\} \qquad \text{Eq. 4-1}$$

$$\textit{starts-within}\ (I_1) \equiv \{I|I_1^- < I^- < I_1^+\} \qquad \text{Eq. 4-2}$$

$$\textit{ends-within}\ (I_1) \equiv \{I|I_1^- < I^+ < I_1^+\} \qquad \text{Eq. 4-3}$$

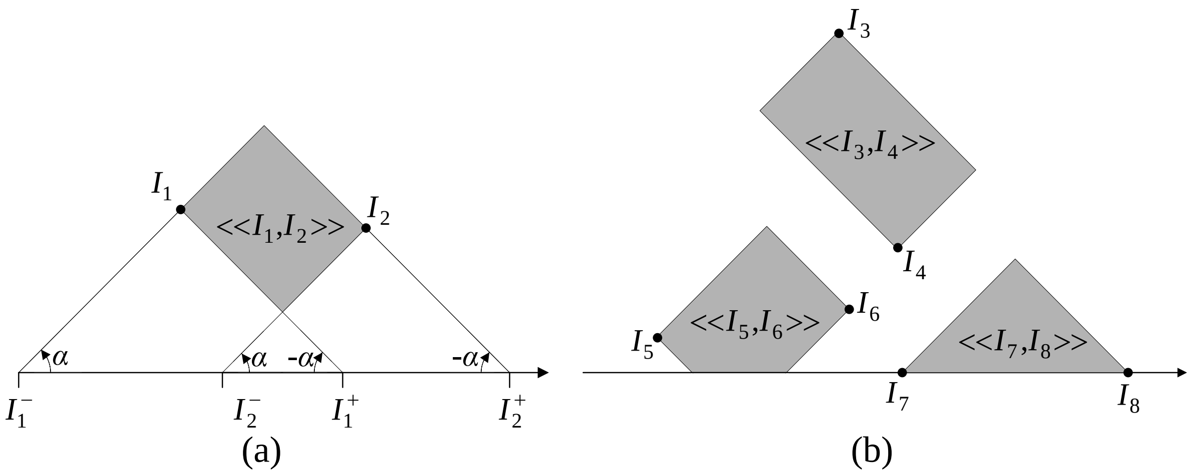


Figure 4-4: The query for intervals *in-between* two intervals. (a): The formation of the query zone of *in-between* I_1 and I_2. (b): Three examples of *in-between* query zones.

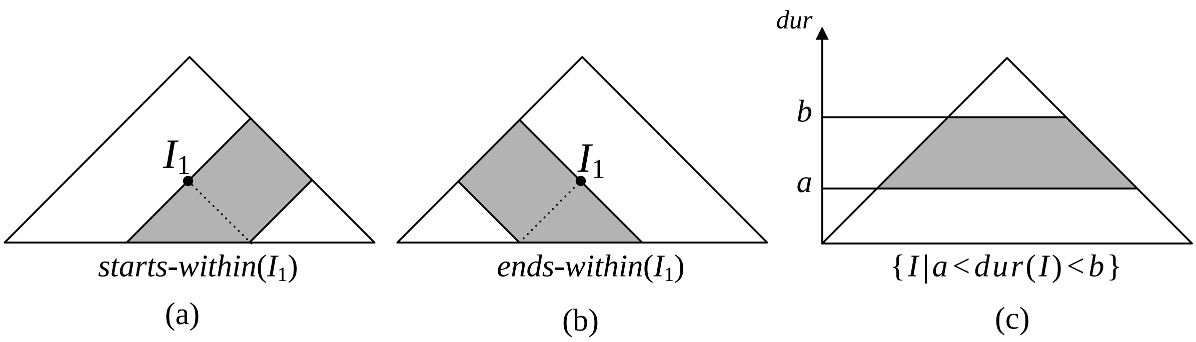


Figure 4-5: Queries of three non-Allen relations in TM. (a): The query zone of *starts-within* (I_1), containing intervals that *start-within* I_1. (b): The query zone of *ends-within* (I_1), containing intervals that *ends-within* I_1. (c): The query zone of $\{I|a < dur(I) < b\}$, containing intervals whose durations are greater than a and smaller than b.

4.3.3 Composite Queries

Since temporal queries are expressed as 2D geometries in TM, composition of temporal queries can be modelled by spatial operations. In TM, the query composition follows the same principles of Venn Diagrams (Venn, 1881). The intersection of queries is represented by the intersection of the query zones (Figure 4-6(a)). The union of queries is represented by the union of the query zones (Figure 4-6(b)) Furthermore, the subtraction of two queries is modelled as subtracting one query zone from the other, which selects intervals that satisfy only one out of two queries (Figure 4-6(c)).

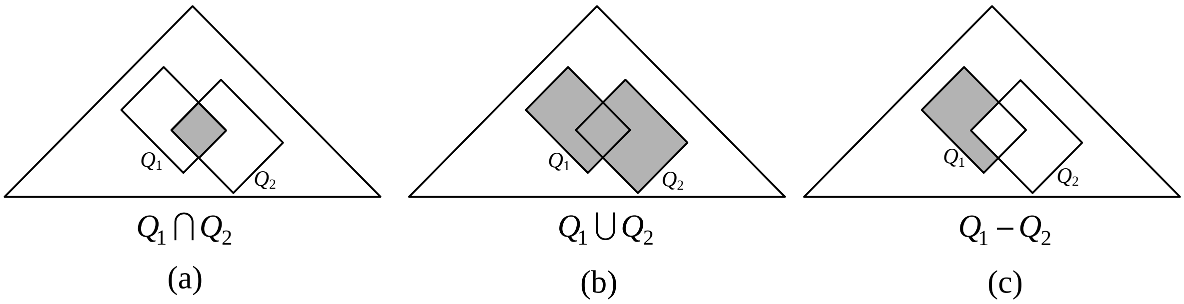


Figure 4-6: The composition of temporal queries in TM. (a): The intersection of the two queries. (b): The union of the two queries. (c): The subtraction of one query from the other query.

4.4 Implementation

To support and exploit practical uses of TM, we implemented the TM into a prototype software tool called GeoTM. In this section, we introduce its user interface and functionalities.

4.4.1 GeoTM

GeoTM supports visualising, querying and analysing spatio-temporal data. More specifically, it has the ability to analyse entities that are referenced to spatial locations and temporal extents. In GeoTM, the spatial locations are modelled as vector geometries (i.e. point, line and polygon), while the temporal extents are modelled as crisp time intervals. Besides their spatial and temporal components, entities may have other attributes that are stored in the attribute table. The spatial locations, temporal extents and other attributes are linked to entities by unique identifiers (ID). GeoTM is built on top of ArcGIS™, which is a desktop GIS produced by ESRI®. In GeoTM, the spatial locations, time intervals and other attributes are stored in ArcGIS™ compatible formats, i.e. ESRI shapefile and dBase. Figure 4-7 gives an overview of how the spatio-temporal information is modelled and represented in GeoTM.

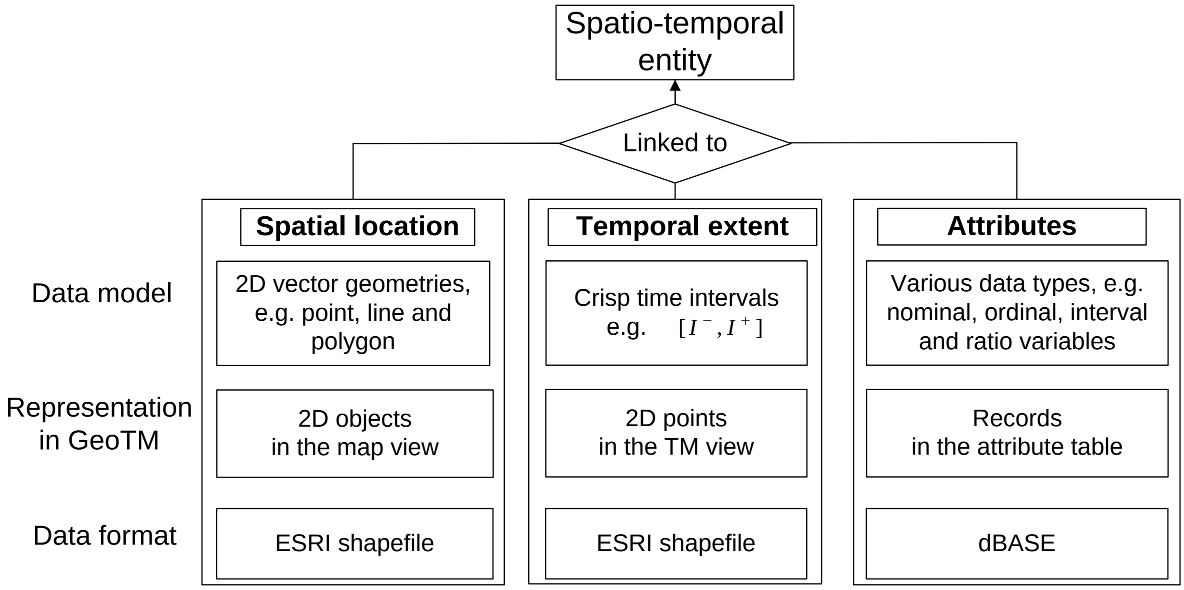


Figure 4-7: The models and representations of spatio-temporal information in GeoTM.

The graphical user interface of GeoTM consists of a map view, the TM view and other controls (Figure 4-8). The map view is used to display spatial locations of entities, which are modelled as points, lines or polygons. The TM view represents time intervals of entities, using the TM representation. In GeoTM, users can zoom in and out in both views by scrolling or dragging a rectangle. In the map view, objects can be selected by dragging an orthogonal rectangle. In the TM view, special selection tools have been designed to select intervals according to specific temporal queries. The other attributes of entities are stored in the attribute table that can pop up when the user clicks the corresponding button. The map view, the TM view and the attribute table are connected through linked brushing. This means that when selecting objects from any of these three views, the other two views dynamically update according to the selection. Additionally, many common GIS functions are supported in GeoTM, such as the use of multiple data layers. Furthermore, objects in the map view and the TM view can be displayed in different sizes, colours and symbols, in order to express attribute information. In Figure 4-8, the map view displays the locations of 26 Bluetooth scanners that have been installed in the centre of Ghent, Belgium, the TM view displays the time intervals of Bluetooth devices that were detected by these scanners between 1:03:00 and 1:06:00 on 19th, July, 2010 (note: 24-hour clock is used in this chapter). In the remainder of this section, we use these time intervals to demonstrate the query functions of GeoTM. The complete scenario of this dataset will be explained in the next section.

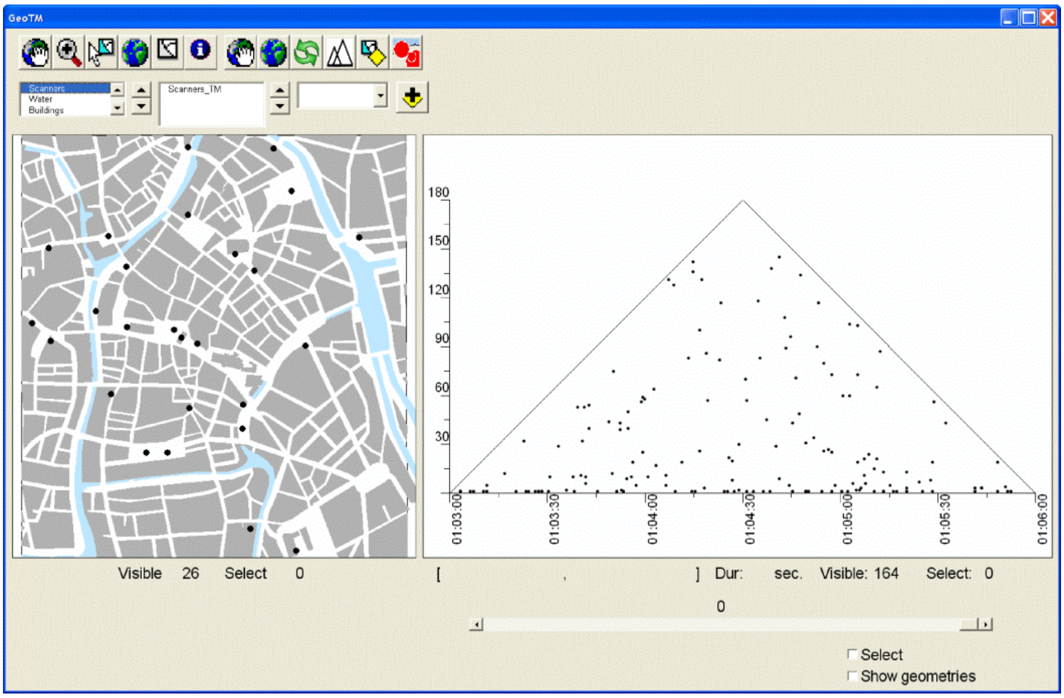


Figure 4-8: The graphical user interface of GeoTM.

4.4.2 Temporal Queries in GeoTM

Traditionally, queries on time intervals are determined by formal expressions (e.g. SQL statements) or by manipulating controls (e.g. buttons, sliding bars, and check boxes) in predefined graphical user interfaces. At present, visual and interactive query tools are being increasingly supported in software (Catarci et al., 1997; Derthick et al., 1997; Egenhofer, 1997; Hochheiser and Shneiderman, 2004), including those that define queries by sketching or manipulating graphics on the screen (Ahlberg et al., 1992). These kinds of query tools have significantly enhanced human-computer interactivity and therefore play an important role in EDA (Andrienko et al., 2010; Andrienko et al., 2003). Despite the large variety of tools that have been developed to present and analyse temporal information (Aigner et al., 2007; Muller and Schumann, 2003; Silva and Catarci, 2000), most of them are designed for visualisation and do not support abundant temporal query functions. As a graphical representation of time intervals, the TM provides a promising platform for visual temporal queries. In GeoTM, temporal queries can be defined by clicking and dragging geometries in the TM view. Users can indicate an interval by moving the mouse cursor to a specific position, and then query for intervals that satisfy an Allen relation to the indicated interval (Figure 4-9 (a) and (b)). In addition, the query zones of the thirteen Allen relations of indicated interval can be drawn (Figure 4-9 (c)), which offers an overview of the distribution of intervals in different relations to the indicated interval.

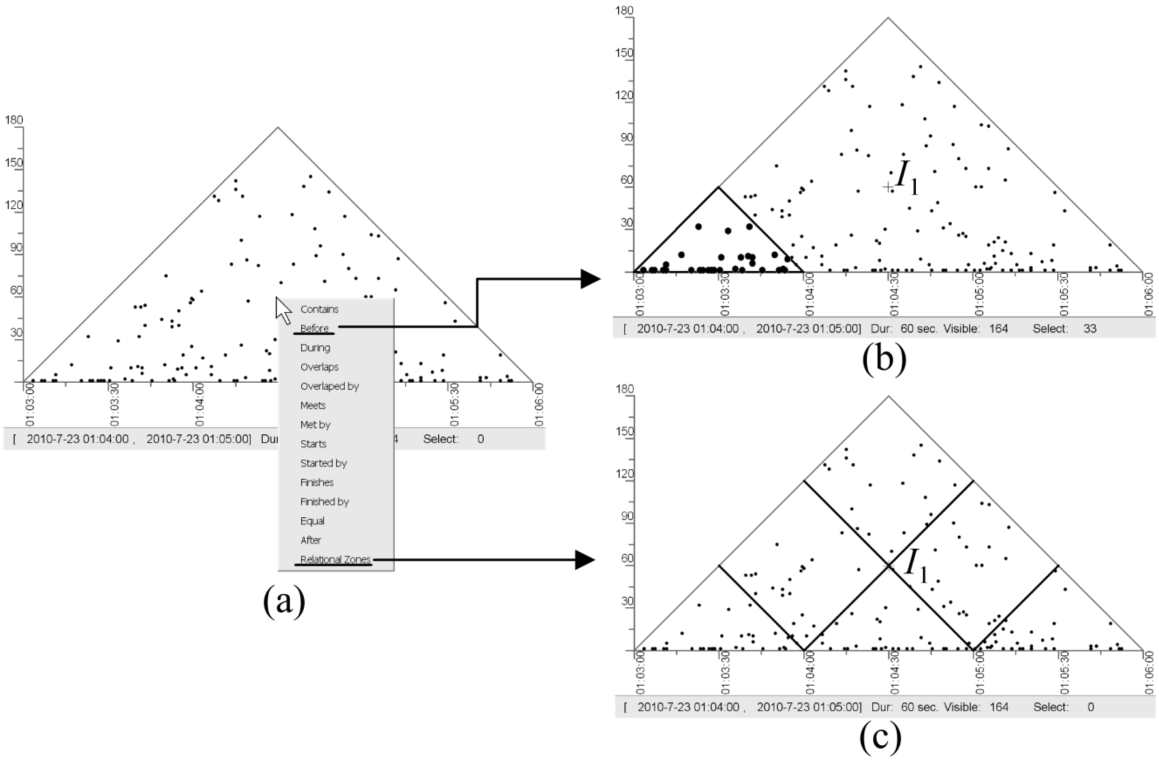


Figure 4-9: Querying intervals that are *before* $I_1[01{:}04{:}00, 01{:}05{:}00]$ on 19 July 2010 and drawing the query zones of I_1. (a): Move the mouse cursor to I_1 and right click to trigger the drop-down menu of Allen relations. (b): With the 'Before' option clicked, the zone of $before(I_1)$ is drawn and intervals in this zone are selected. The larger dots represent the selected intervals and the cross represents I_1. (c): With the 'Relational Zones' option clicked, the query zones of the thirteen Allen relations of I_1 are drawn.

Moreover, temporal queries can be defined by dragging a zone in the TM view. In line with the characteristics of the TM space, several special dragging tools have been designed in order to select intervals that satisfy certain temporal queries. The default dragging box is a non-orthogonal rectangle that selects intervals *in-between* two time intervals. With this dragging box, every selected set of intervals is a convex interval set that can be interpreted by Allen relations or combinations thereof. Dragging a rectangle selects a convex interval set *in-between* two intervals (e.g. Figure 4-10 (a)), and dragging a triangle on the horizontal axis selects a convex interval set *in-between* two time instants (Figure 4-10 (b)). Besides convex interval sets, GeoTM supports range queries according to interval duration by dragging a range along the vertical axis (e.g.

Figure 4-10 (c)). Furthermore, a range of two parallel lines can be dragged at α or $-\alpha$ angle with respect to the horizontal axis to select the intervals that *start-within* or *end-within* an interval (Figure 4-10 (d)).

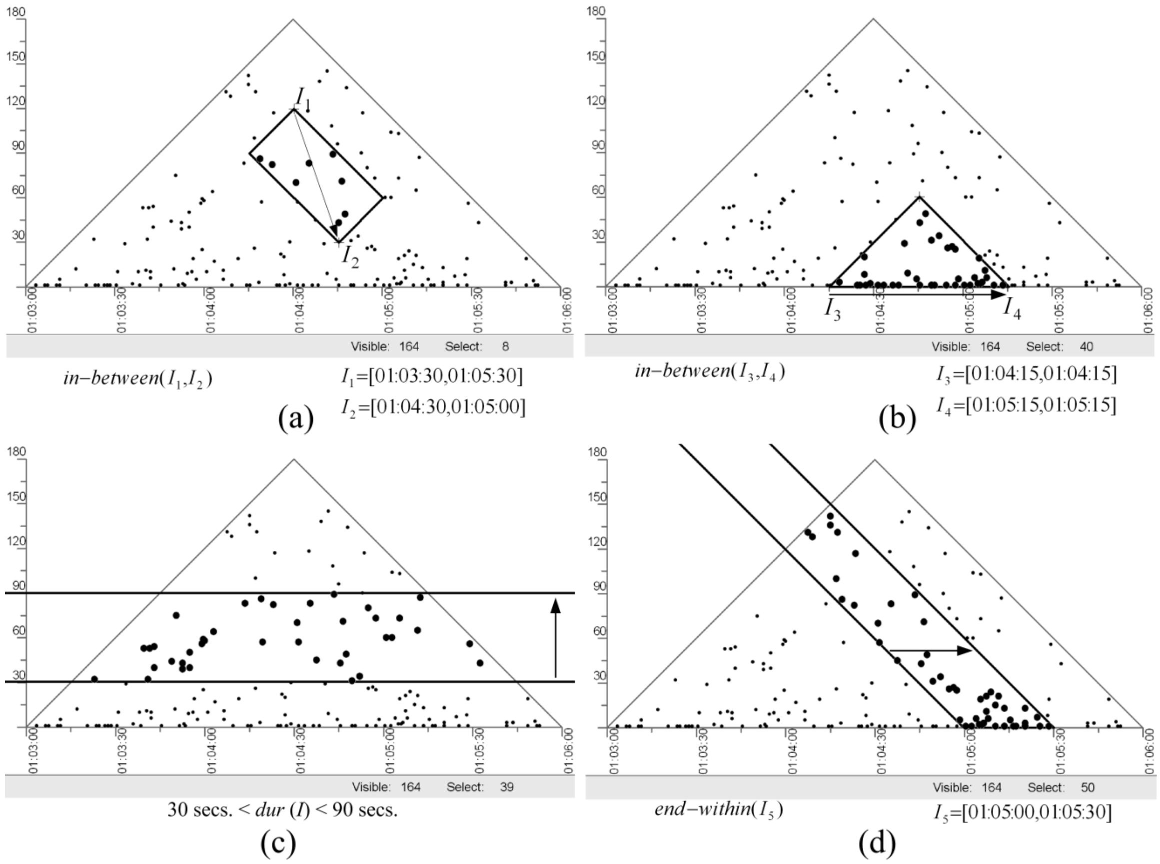


Figure 4-10: Temporal queries by dragging geometries in TM. (a): selecting intervals in a convex set in-between two time intervals. (b): selecting intervals in a convex set in-between two time instants. (c): selecting intervals that are longer than 30 seconds and shorter than 90 seconds. (d): selecting intervals end-within [01:05:00, 01:05:30].

4.4.3 Composite Queries

All the above mentioned queries can be composed in GeoTM. Users can make queries successively and compose them by logical operators (e.g. union, intersection or subtraction). Figure 4-11 illustrates how the composite query $contains(\ I_1) \cap \ll I_2, I_3 \gg \cap \{I|dur(I) > 90\ seconds\} \cap$ *end-within*(I_4) is constructed in GeoTM, given the intervals I_1, I_2, I_3, and I_4. The querying process consists of four steps. In each step, a component query is made, with a selected logical operator. As in this case component queries are composed with intersection, the result is the intersected area of the query zones. The intervals located in the intersected area are selected, which satisfy the entire composite query.

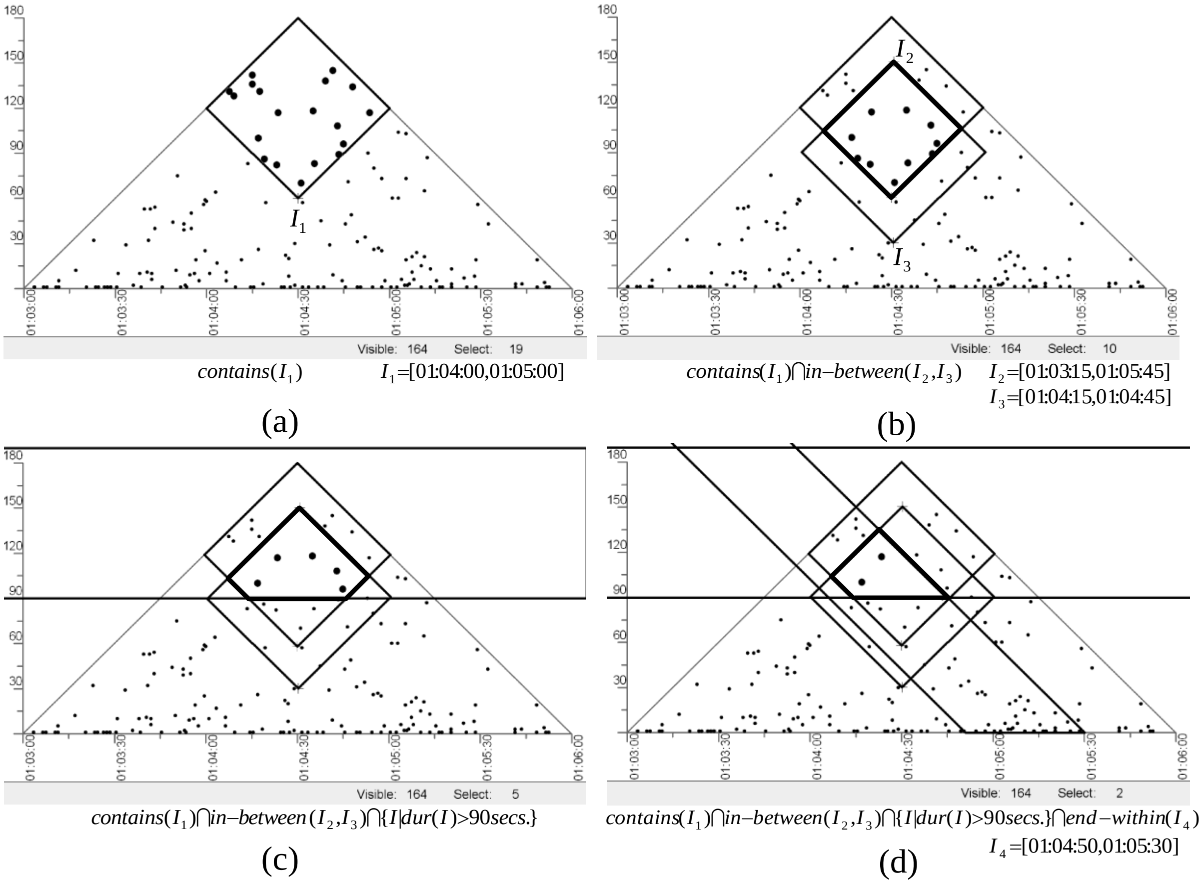


Figure 4-11: The creation of the composite query: $contains(I_1) \cap \ll I_2, I_3 \gg \cap \{I|dur(I) > 90\ seconds\} \cap end\text{-}within(I_4)$. In each step, the intersected areas of the queries are marked by bold boundaries and selected intervals are marked by larger dots.

4.5 Case Study

Although existing as a communication technology since the mid-nineties, Bluetooth has only recently been employed for positioning and tracking individuals (Delafontaine et al., 2011; Hay et al., 2009; Nicolai and Kenn, 2007; O'Neill et al., 2006; Stange et al., 2011; Versichele et al., 2012). Despite its limited positional accuracy, Bluetooth tracking is a low cost alternative for true location-aware technologies. Nowadays, due to the increasingly widespread use of Bluetooth-integrated devices, e.g. mobile phones, laptops and handsets, Bluetooth tracking has been more and more used for scientific experiments and observations. In this use case, the data were collected from the most basic form of a Bluetooth tracking system, in which, a number of Bluetooth scanners were installed at certain strategic points. These scanners continuously make inquiries, listen for responses of other devices in their detection ranges and log the results with a time resolution of one second. Therefore, the period when a Bluetooth device is in the detection range of the scanner is a time interval. Technically, this interval starts at the first second that the Bluetooth device is logged by the scanner and ends at the last second that it is logged. Moreover, if a Bluetooth device has left the detection range of a scanner but it re-enters the range again within 30 minutes, the two time intervals as well as the gap between them are merged to one time interval. The purpose of doing this it to eliminate the fractional intervals that are caused by unstable signals or Bluetooth devices moving on the border of the detection range. In this way, the scanners generate a dataset of time intervals, in which every single record consists of three major components: the MAC address of the Bluetooth device, the ID of the scanner logging this device and the time interval.

The dataset was collected during the Ghent Festivities 2010, from 10:00 on 17 July to 10:00 on 26 July, 2010. 26 Bluetooth scanners were installed in selected locations (e.g. concert stages, parking lots, the tourist information office and major accesses) in the centre of Ghent. When the dataset is loaded to GeoTM, the time intervals are displayed as dots in the TM view and the Bluetooth scanners are displayed in the map view (Figure 4-12). In the TM view, some regular daily patterns can be observed: the number and durations of logged intervals increases from the afternoon, reach a peak around midnight and start to decrease after midnight. In GeoTM, if specific Bluetooth scanners are selected in the map view, the time intervals that are logged by these scanners instantaneously turn to red in the TM view. With this function, users can observe and compare the distribution of time intervals detected at different locations. During the Ghent Festivities 2010, two concert stages were located at Sint-Baafs Cathedral and Sint-Jacob Church respectively. By successively selecting the scanners installed near Sint-Baafs Cathedral (i.e. Scanner 1, 2 and 3 in Figure 4-12) and the scanners installed near Sint-Jacob Church (i.e. Scanner 8 and 9 in Figure 4-12) in the map view, from the TM view one can observe that the distributions at these two sites are similar, which gradually increase from the afternoon or evening, and decrease more rapidly in the deep night or even the early morning in the next day. However, the time intervals detected near Sint-Baafs Cathedral are distributed earlier than the intervals detected near Sint-Jacob Church (Figure 4-13). This finding reveals the fact that the crowd near Sint-Baafs Cathedral gathered and dispersed earlier than the crowd near Sint-Jacob church.

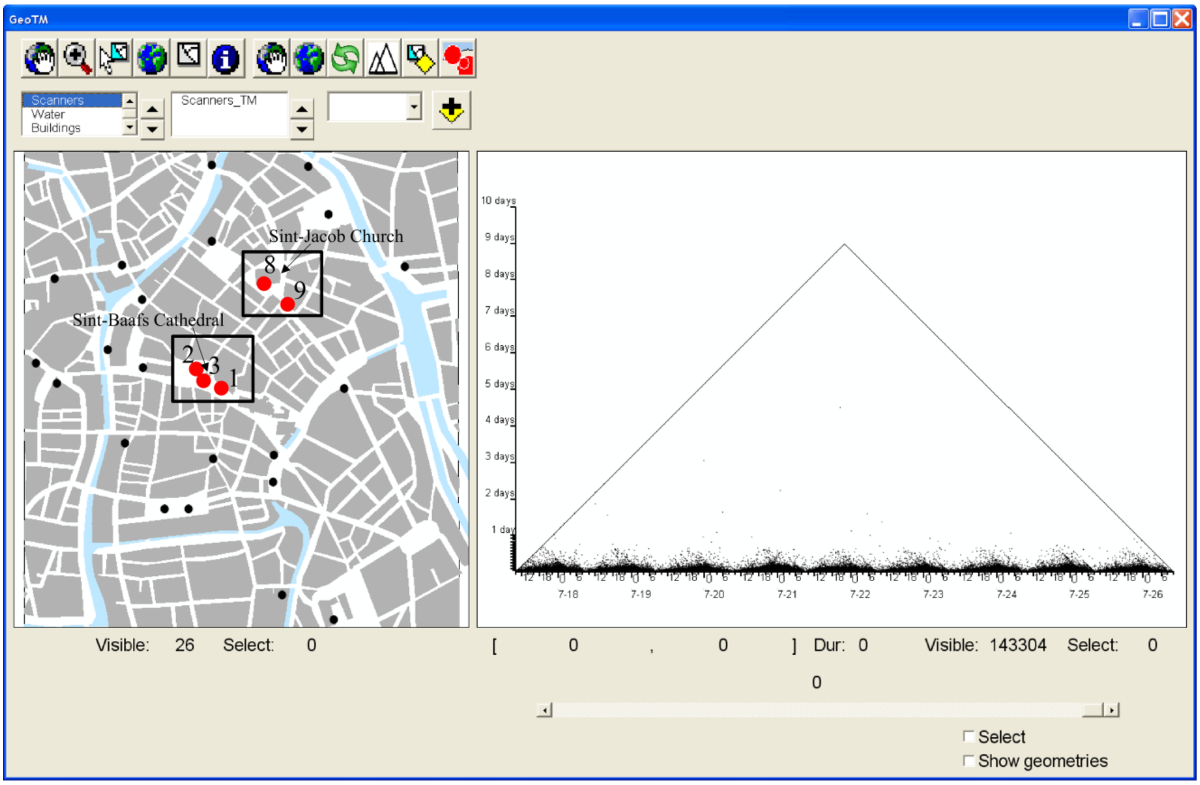


Figure 4-12: The GeoTM interface with the complete Bluetooth dataset loaded. The squares and red dots in the map view indicate the two selections of scanners that generate the results in Figure 4-13.

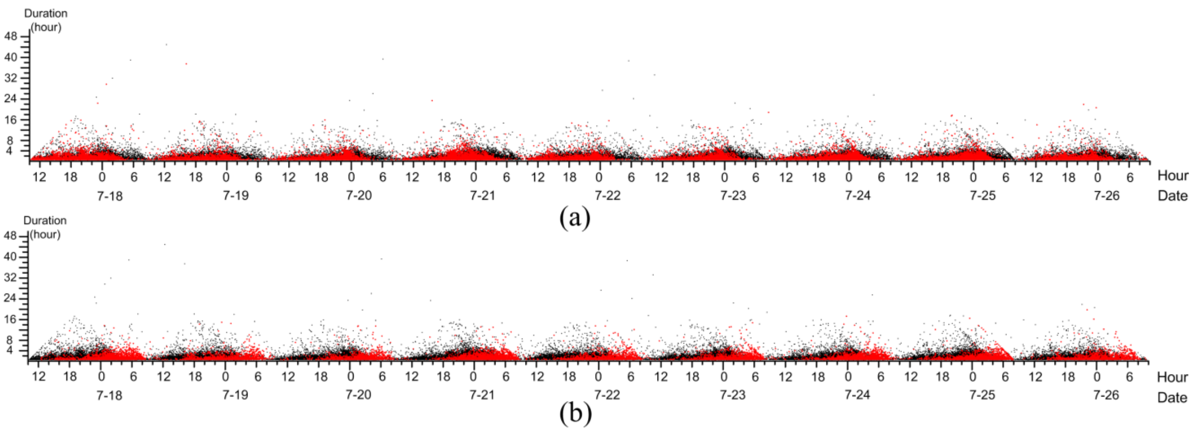


Figure 4-13: The time intervals logged by the two groups of scanners in Figure 4-12 become red in the TM view. (a): The time intervals logged by Scanner 1, 2 and 3 are red, and the rest remain black. (b): The time intervals logged by Scanner 8 and 9 are red, and the rest remain black.

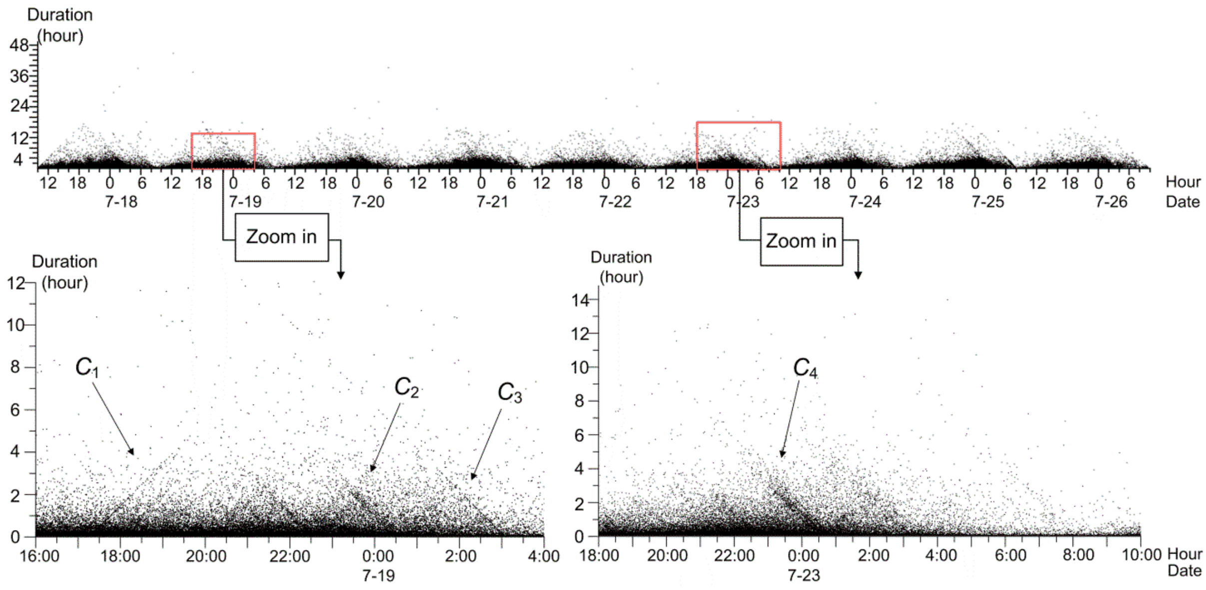
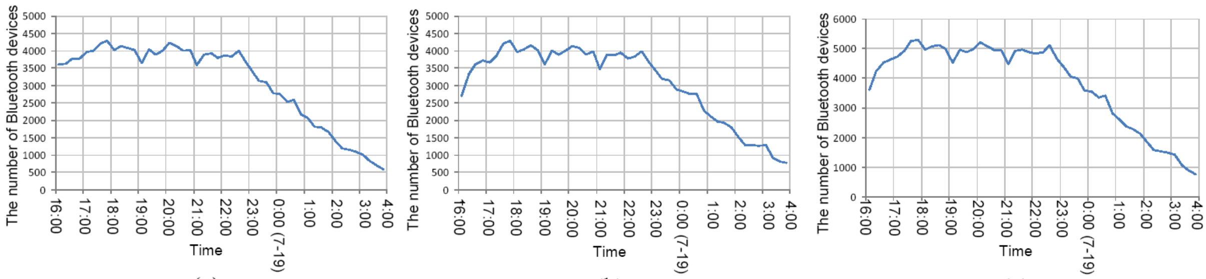


Figure 4-14: Zooming into the period from 16:00 on 18 July to 4:00 on 19 July and the period from 18:00 on 22 July to 10:00 on 23 July in the TM view.

In the TM view, users can also zoom into smaller areas to observe the interval distribution during specific hours. Zooming into the period from 16:00 on 18 July to 4:00 on19 July and the period from 18:00 on 22 July to 10:00 on 23 July (Figure 4-14), one can see that most interval points are distributed on the bottom, meaning that most detected time intervals in these two periods are relatively short. This reflects the high dynamics of people during the events. Moreover, some linear clusters that extend to the higher area can be observed (Figure 4-14). Some of these clusters are in $-\alpha$ ($\alpha = 45°$) angle to the horizontal axis (such as $C1$ in Figure 4-14). This kind of clusters are formed by intervals that *start-within* a very short period and *end-within* a much longer period, which can be interpreted as that a large number of people having Bluetooth devices intensively entered the detection range of the scanners within a very short period, and these people gradually left the ranges of the scanners during a much longer period. On the other hand, some other linear clusters are in α angle to the horizontal axis (such as $C2$, $C3$ and $C4$ in Figure 4-14), which are formed by intervals that *start-within* a long period and *end-within* a very short period. This kind of clusters can be interpreted as a number of people with Bluetooth devices gradually entering the range of the scanners within a long period and intensively leaving within a short period. These linear patterns cannot be easily detected from traditional visualisation approaches, for example, line diagrams. In Figure 4-15, the detected patterns in the period from 16:00 on 18 July to 4:00 on 19 July (i.e. $C1$, $C2$ and $C3$) do not clearly appear in the corresponding line diagrams, because the clusters of long intervals can be easily diluted by the large number of short intervals.



(a) (b) (c)

Figure 4-15: Representing the dataset from 16:00 on 18 July to 4:00 on 19 July with line diagrams. (a): The numbers of Bluetooth devices that entered the detection range of scanners in every 15 minutes. (b): The number of Bluetooth devices that left the range of scanners in every 15 minutes. (c): The total number of detected Bluetooth devices in every 15 minutes.

Using the query tools in GeoTM, users can select intervals in these observed clusters, and analyse the spatial distribution of the selected intervals. In GeoTM, when a selection is made in the TM view, the dot size of scanners in the map view is instantaneously updated, indicating the proportion of intervals logged by every scanner. Since we observe that the intervals in Cluster $C1$ more or less *start-within* [17:00, 17:30] on 18 July, we can drag a query zone of *start-within* [17:00, 17:30] to select intervals in $C1$ (Figure 4-16). And then, the updated map view shows dots corresponding to the different scanners in similar sizes, meaning that the selected intervals are more or less evenly logged by the 26 scanners. In order to focus on the people that come for the events of the Ghent Festivities, the user may want to eliminate the short intervals of people passing by and the very long intervals of people living in the neighbourhood. To this end, another query zone is drawn to select intervals that have a duration between 15 minutes and 10 hours upon the previous selection (Figure 4-17). In the map view, the dot sizes of Scanner 4 and 10 become much bigger than the other scanners, meaning that large proportions of retrieved intervals are logged by these two scanners (21.5% and 21.6%). Considering Scanner 4 and 10 are installed respectively near the stage at Graslei and the stage at Baudelopark, it is straightforward to infer that a large number of people intensively went to the events at these two stages between 17:00 and 17:30 and gradually left afterwards. Consulting the schedule of Ghent Festivities 2010 (2010), one can find out that the concert of the band Maximum Basie at Graslei and an oriental dance activity at Baudelopark started at 17:00, which probably attracted people and kept them staying around for a long time.

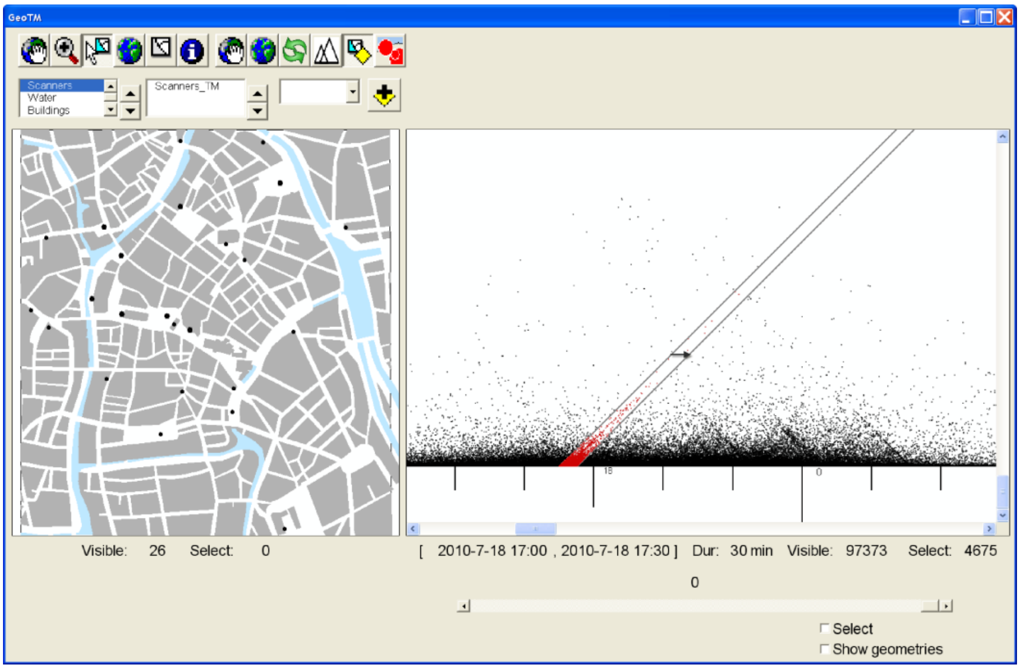


Figure 4-16: Select intervals that *start-within* [17:00, 17:30] on 18 July.

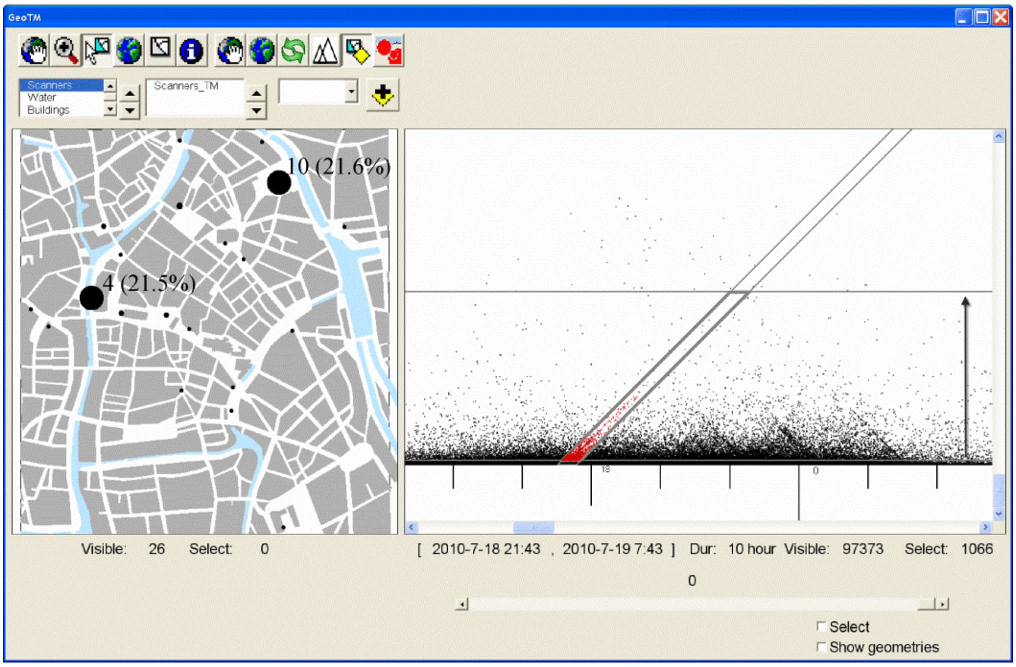


Figure 4-17: Select intervals that *start-within* [17:00, 17:30] on 18 July and have a duration between 15 minutes and 10 hours. The numbers in the map view indicate the scanner IDs, and the numbers between brackets indicate the proportions of logged intervals.

In the same way, the user can successively drag query zones of '*end-within* [0:00, 1:00] on 23 July' and 'have a duration between 15 minutes and 10 hours', in order to select the intervals in the cluster *C*4, excluding the short intervals of passing people and long intervals of people living around (Figure 4-18). Here, the map view shows that a large proportion of intervals are logged by Scanner 1 (23.5%) and Scanner 4 (23.7%). Considering that Scanner 1 and 4 are respectively installed near two stages at Sint-Baafs Cathedral and Graslei, one can infer that a large number of people gradually gathered at these two stages and then intensively left these two stages between 0:00 and 1:00 on 23 July. Consulting the schedule of the Ghent Festivities 2010 (2010), we can see that the performances at these two stages ended during this period, which probably caused a wave of leaving people. These revealed patterns may provide valuable information for police dispatch, transportation and activity management.

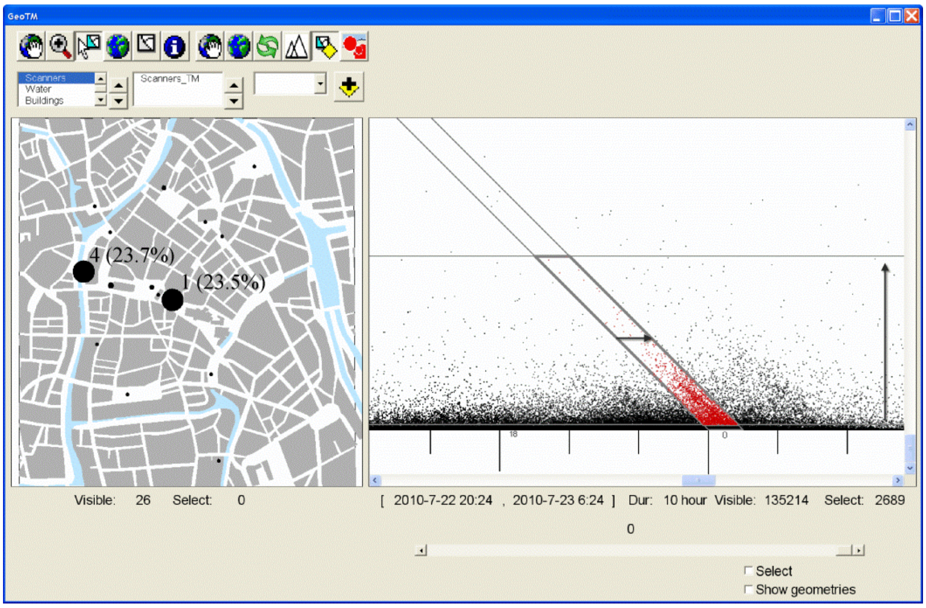


Figure 4-18: Select intervals that *end-within* [0:00, 1:00] on 23 July and have a duration between 15 minutes and 10 hours.

4.6 Discussion

As an alternative representation of time intervals, the TM shows a promising performance in some aspects that are inherent difficulties of the traditional linear representation. In the linear representation, a given set of time intervals may show different patterns, when different ordering rules are applied (e.g. Figure 4-19). The characteristics of the distribution are not easy to observe from an individual display. However, the TM offers a compact and fixed visualisation of time intervals. Because every interval has a unique spatial position in TM, the structure of a set of intervals is fixed, which potentially benefits pattern detection. When a dataset of intervals is represented in TM, one can have a first image of the interval distribution. For example, in the dataset of the case study in Section 4.5, the distribution of time intervals can be directly observed from the TM visualisation, which not only shows where the intervals are located in time, but also their durations. In this visualization, some interesting patterns (e.g. linear clusters) can be detected, which gives clues to analyse the phenomena behind these patterns. Moreover, points are apparently more space-saving than linear segments. Within the same display extent, the TM is capable of displaying a larger amount of intervals than the linear model. For instance, it would be much more difficult to use linear segments represent the time intervals in the datasets similar to that of the case study that consists of millions of time intervals.

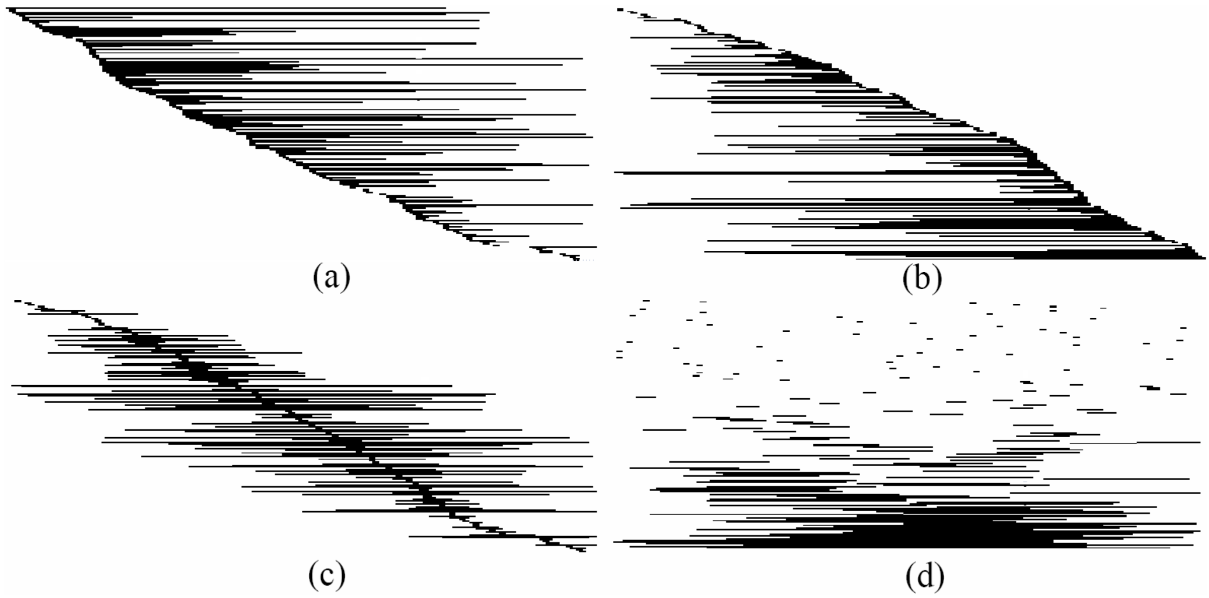


Figure 4-19: Time intervals in the linear representation, with different ordering rules. (a) Ascending order by the start. (b) Ascending order by the end. (c) Ascending order by the midpoint. (d) Ascending order by the duration.

Furthermore, the TM provides a platform for visual temporal queries. In TM, temporal queries can be specified by creating 2D zones. These zones give a visual impression of query range. This unique feature stems from the fact that every interval has an absolute position in the TM and intervals in the same temporal relation are grouped in convex zones. Moreover, in TM, the composition of temporal queries can be defined by spatial operations, following the concept of Venn diagrams. This graphic representation of temporal queries is likely to be more intuitive than alphanumerical expressions, because previous research shows that humans have excellent ability in processing spatial and graphical information (Tversky, 2005), and the transformation from alphanumeric expressions into graphic representations has gained considerable success in various disciplines (Giere, 1999). Moreover, the TM supports an integration of the temporal query and the visualisation. In other words, temporal queries in the TM can be directly carried out by creating zones on top of the visualisation of time intervals. In this way, one can directly observe the range of the query over the dataset and instantaneously adapt the query. This feature is valuable when the query constraints are vaguely specified, or tailor-made queries need to be designed to select intervals in special clusters, e.g. the observed clusters in the case study. One can instantaneously redefine the query range with respect to the distribution of intervals. On the contrary, the traditional query tools (e.g. dynamic sliders, drop-list and text boxes) are standalone and do not visualise data. They normally cooperate with a data visualisation to do analytical jobs. For example, when using dynamic sliders, one needs to query with the slider and keep eyes on the graph next to it.

With these promising features, the TM shows potential in assisting a GIS to analyse interval-based geographical data. In this work, the TM is implemented in a GIS in order to better exploit this potential. For the moment, not many tools have the capability of analysing this kind of geographical data. As one of the few possibilities, the tools based on the Space-Time Cube concept represent spatio-temporal entities in a 3D space in which the time axis is added to a flattened topography. For example, the tool developed by Huisman et al. (2009) represents geographical entities referenced to time intervals as linear segments in the Space-Time Cube. The extents of these linear segments along the time axis indicate the time intervals. However, this tool suffers from some typical problems of 3D visualization (Forlines and Wittenburg, 2010; Kjellin et al., 2009). For example, the structure of objects varies when the view point varies. Moreover, foreground objects block one’s view of background objects. Also, compared with 2D environments, it is cumbersome to manipulate objects (e.g. selecting objects) in a 3D environment given that the prevalent computer screens, touch pads and mouse pads are 2D. Since GeoTM is based on a 2D visualisation, it does not suffer from these difficulties. We contend that GeoTM can bring the advantage of the TM in visualisation and temporal queries to the GIS, enhancing its capabilities regarding, for instance, ESTDA.

4.7 Conclusion and Future Work

In this chapter, we have investigated the use of the TM in visualizing and analysing time intervals. In order to better demonstrate the use of TM, we have implemented it into a software prototype (i.e. GeoTM) that integrates the TM with a GIS. In addition, GeoTM has been applied in a real scenario to analyse time intervals that were detected by Bluetooth scanners in the Ghent Festivities of 2010. Since abstracting Bluetooth tracking data as time intervals is a new methodology and there are no mature approaches to analyse these kinds of datasets, GeoTM can be considered as an assistant tool that provides ESTDA functionalities. Although this analytical approach needs further investigation and validation, it provides unique insights into interval distributions, and can reveal interesting patterns that cannot be clearly presented by traditional approaches. Moreover, the TM offers a mechanism for temporal queries that we contend is innovative and likely to support analysis effectively due to its graphic, flexible and direct representation on top of the visualisation. GeoTM shows that the querying and analytical power of the TM can be better exploited when it is implemented into an interactive tool. Furthermore, the integration with a GIS brings the merits of the TM in visualisation and temporal query to the GIS, enhancing its capabilities of analysing interval-based geographical data.

In the future, the analytical methodology based on the TM needs to be further investigated, including the interpretations of atomic patterns and metric measurement of patterns. Also, this methodology will be tested in more use cases, preferably covering different research contexts. On the other hand, the usability of GeoTM needs to be further evaluated. To this end, empirical experiments will be systematically designed and conducted to assess whether GeoTM can help people, particularly non-expert people, to solve practical questions. In addition, more user-friendly and interactive functions will be added to GeoTM to improve its usability, helping users understand the TM concepts and make use of the corresponding functionalities. The query tool can be implemented in a more dynamic manner. When the user is dragging a query zone, the other views can instantaneously update to display the spatial locations and other information of the selection set. Instead of using linear scales in the interval space, it would be interesting for users to manipulate the scales. For instance, users can apply a logarithmic scale to magnify the intervals in the bottom area in Figure 4-13. Other possible extensions exist in supporting the representation of rough and fuzzy time intervals, or the representation of continuous temporal information (e.g. time series) other than discrete time intervals.

References

Ahlberg, C., Williamson, C. and Shneiderman, B. (1992). Dynamic queries for information exploration: an implementation and evaluation. *Proceedings of the SIGCHI conference on Human factors in computing systems*, Monterey, California, United States, pp.619-626, ACM.

Aigner, W., Miksch, S., Muller, W., Schumann, H. and Tominski, C. (2007). Visualizing time-oriented data - A systematic view. *Computers & Graphics-UK,* 31(3), pp.401-409.

Allen, J. F. (1983). Maintaining Knowledge about Temporal Intervals. *Communications of the ACM*, 26(11), pp.832-843.

Andrienko, G., Andrienko, N., Demsar, U., Dransch, D., Dykes, J., Fabrikant, S. I., Jern, M., Kraak, M.-J., Schumann, H. and Tominski, C. (2010). Space, time and visual analytics. *International Journal of Geographical Information Science,* 24(10), pp.1577-1600.

Andrienko, N. and Andrienko, G. (2006). *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer-Verlag Berlin, Heidelberg.

Andrienko, N., Andrienko, G. and Gatalsky, P. (2003). Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing,* 14(6), pp.503-541.

Bassiouni, M. A. and Llewellyn, M. (1999). Extending temporal query languages to handle imprecise time intervals. *Computer Languages,* 25(1), pp.39-54.

Beard, K., Deese, H. and Pettigrew, N. R. (2007). A framework for visualization and exploration of events. *Information Visualization,* 7(2), pp.133-151.

Bittner, T. (2002). Approximate Qualitative Temporal Reasoning. *Annals of Mathematics and Artificial Intelligence,* 36(1), pp.39-80.

Butkiewicz, T., Dou, W., Wartell, Z., Ribarsky, W. and Chang, R. (2008). Multi-Focused Geospatial Analysis Using Probes. *IEEE Transactions on Visualization and Computer Graphics*, 14(6), pp.1165-1172.

Catarci, T., Costabile, M. F., Levialdi, S. and Batini, C. (1997). Visual Query Systems for Databases: A Survey. *Journal of Visual Languages & Computing*, 8(2), pp.215-260.

Delafontaine, M., Versichele, M., Neutens, T. and Van de Weghe, N. (2011). Analysing spatial and temporal sequences in Bluetooth tracking data. Paper presented in Mobile positioning, Spatial Mobility, Geography, LBS Tartu, Estonia.

Derthick, M., Kolojejchick, J. and Roth, S. F. (1997). An interactive visual query environment for exploring data. *Proceedings of the 10th annual ACM symposium on User interface software and technology*, Banff, Alberta, Canada, pp.189-198.

Egenhofer, M. (1997). Query Processing in Spatial-Query-by-Sketch. *Journal of Visual Languages & Computing*, 8(4), pp.403-424.

Forlines, C. and Wittenburg, K. (2010). Wakame: sense making of multi-dimensional spatial-temporal data. *Proceedings of the International Conference on Advanced Visual Interfaces*, Roma, Italy, pp.33-40, ACM.

Fredrikson, A., North, C., Plaisant, C. and Shneiderman, B. (1999). Temporal, geographical and categorical aggregations viewed through coordinated displays: a case study with highway incident data. *Proceedings of the 1999 workshop on new paradigms in information visualization and manipulation in conjunction with the eighth ACM internation conference on Information and knowledge management*, Kansas City, Missouri, United States, pp.26-34, ACM, New York.

Freksa, C. (1992). Temporal Reasoning Based on Semi-Intervals. *Artificial Intelligence*, 54(1-2), pp.199-227.

Garrido, C., Marín, N. and Pons, O. (2009). Fuzzy Intervals of Represent Fuzzy Valid Time In a Temporal Relational Database. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 17(1), pp.173-192.

Gatalsky, P., Andrienko, N. and Andrienko, G. (2004). Interactive analysis of event data using space-time cube. *Proceedings of the 8th International Conference on Information Visualisation (IV 2004)*, London, UK, pp.145-152.

Giere, R. N. (1999). *Science Without Laws*. The University of Chicago Press, Chicago.

Guo, D., Chen, J., MacEachren, A. M. and Liao, K. (2006). A Visualization System for Space-Time and Multivariate Patterns (VIS-STAMP). *IEEE Transactions on Visualization and Computer Graphics*, 12(6), pp.1461-1474.

Hardisty, F. and Klippel, A. (2011). Analysing spatio-temporal autocorrelation with LISTA-Viz. *International Journal of Geographical Information Science*, 24(10), pp.1515-1526.

Hay, S., Harle, R., Choudhury, T., Quigley, A., Strang, T. and Suginuma, K. (2009). Bluetooth Tracking without Discoverability. *Proceedings of the 4th International Symposium on Location and Context Awareness (LoCA 2009)*, Tokyo, Japan, pp.120-137, Springer-Verlag Berlin, Heidelberg.

Hibbard, W. L., Paul, B. E., Santek, D. A., Dyer, C. R., Battaiola, A. L. and Voidrot-Martinez, M. F. (1994). Interactive visualization of Earth and space science computations. *Computer*, 27(7), pp.65-72.

Hochheiser, H. and Shneiderman, B. (2004). Dynamic query tools for time series data sets: timebox widgets for interactive exploration. *Information Visualization,* 3(1), pp.1-18.

Huisman, O., Santiago, I. F., Kraak, M.-J. and Retsios, B. (2009). Developing a Geovisual Analytics Environment for Investigating Archaeological Events: Extending the Space-time Cube. *Cartography and Geographic Information Science,* 36(3), pp.225-236.

Kapler, T. and Wright, W. (2005). GeoTime information visualization. *Information Visualization,* 4(2), pp.136-146.

Kjellin, A., Winkler Pettersson, L., Seipel, S. and Lind, M. (2009). Different levels of 3D: An evaluation of visualized discrete spatiotemporal data in space-time cubes. *Information Visualization,* 9(2), pp.152-164.

Kulpa, Z. (1997). Diagrammatic Representation for a Space of Intervals. *Machine Graphics & Vision,* 6, pp.5-24.

Kulpa, Z. (2001). Diagrammatic representation for Interval Arithmetic. *Linear Algebra and Its Applications,* 324(1-3), pp.55-80.

Kulpa, Z. (2006). A Diagrammatic Approach to Investigate Interval Relations. *Journal of Visual Languages and Computing,* 17(5), pp.466-502.

Langran, G. (1993). Time in Geographic Information Systems. ed. by Peuquet, D. J. and Marble, D. F., in *Technical Issues in Geographic Information System*, Taylor & Francis Ltd, London.

Li, X. and Kraak, M.-J. (2008). The Time Wave. A New Method of Visual Exploration of Geo-data in Time–space. *The Cartographic Journal,* 45(3), pp.193-200.

Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W. (2001). *Geographic Information Systems and Science*. John Wiley & Sons, Ltd., Chichester, England.

Monmonier, M. (1989). Geographic Brushing: Enhancing Exploratory Analysis of the Scatterplot Matrix. *Geographical Analysis,* 21(1), pp.81-84.

Muller, W. and Schumann, H. (2003). Visualization methods for time-dependent data - An overview. *Proceedings of the 2003 Winter Simulation Conference*, New Orleans, U.S.A., pp.737-745.

Neutens, T., Van de Weghe, N., Witlox, F. and De Maeyer, P. (2008). A three-dimensional network-based space-time prism. *Journal of Geographical Systems,* 10(1), pp.89-107.

Neutens, T., Witlox, F., Van de Weghe, N. and De Maeyer, P. (2007a). Human Interaction Spaces Under Uncertainty. *Transportation Research Record: Journal of the Transportation Research Board,* 2021(1), pp.28-35.

Neutens, T., Witlox, F., Van de Weghe, N. and De Maeyer, P. (2007b). Space-time opportunities for multiple agents: a constraint-based approach. *International Journal of Geographical Information Science,* 21(10), pp.1061-1076.

Nicolai, T. and Kenn, H. (2007). About the relationship between people and discoverable Bluetooth devices in urban environments. *Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology,* Singapore, pp.72-78, ACM.

North, C. and Shneiderman, B. (2000). Snap-Together Visualization: A User Interface for Coordinating Visualizations via Relational Schemata. Paper presented in Andvanced Visual Interfaces Palermo, Italy.

O'Neill, E., Kostakos, V., Kindberg, T., Schiek, A., Penn, A., Fraser, D., Jones, T., Dourish, P. and Friday, A. (2006). Instrumenting the City: Developing Methods for Observing and Understanding the Digital Cityscape. ed. by Dourish, P. and Friday, A., in *UbiComp 2006: Ubiquitous Computing*, pp.315-332, Springer Berlin / Heidelberg.

Ohlbach, H. J. (2004). Relations between fuzzy time intervals. *Proceedings of 11th International Symposium on Temporal Representation and Reasoning (TIME 2004)*, Montréal, Québec, pp.44-51.

Ozsoyoglu, G. and Snodgrass, R. T. (1995). Temporal and real-time databases: a survey. *IEEE Transactions on Knowledge and Data Engineering,* 7(4), pp.513-532.

Qiang, Y., Delafontaine, M., Asmussen, K., Stichelbaut, B., De Tré, G., De Maeyer, P. and Van de Weghe, N. (2010). Modelling Imperfect Time Intervals in a Two-Dimensional Space. *Control and Cybernetics,* 39(4), pp.983-1010.

Schiettekatte, J. and Rogiest, H. (2010). *Gentse Feesten 17-26.07.2010 167ste Cultureel Volksfeest.* Roularta Media Group. Ghent, Belgium.

Schockaert, S., De Cock, M. and Kerre, E. E. (2008). Fuzzifying Allen's Temporal Interval Relations. *IEEE Transactions on Fuzzy Systems,* 16(2), pp.517-533.

Silva, S. F. and Catarci, T. (2000). Visualization of Linear Time-Oriented Data: A Survey. *Proceedings of the First International Conference on Web Information Systems Engineering (WISE'00),* Hong Kong, China, pp.310 - 319, IEEE Computer Society.

Smith, M. J. d., Goodchild, M. F. and Longley, P. A. (2007). *Geospatial Analysis: a Comprehensive Guide to Principles, Techniques and Software Tools*. Troubador Publishing Limited, Leicester, UK.

Stange, H., Liebig, T., Hecker, D., Andrienko, G. and Andrienko, N. (2011). Analytical Workflow of Monitoring Human Mobility in Big Event Settings using Bluetooth. Paper presented in The 3rd International Workshop on Indoor Spatial Awareness (ISA2011) Chicago, IL, U.S.A.

Stojanovic, D., Djordjevic-Kajan, S., Mitrovic, A. and Stojanovic, Z. (1999). Cartographic Visualization and Animation of the Dynamic Geographical Processes and Phenomena. *Proceedings of 19th International Cartographic Conference*, Ottawa, Canada, pp.739-746.

Toman, D. (1996). Point vs. interval-based query languages for temporal databases. *Proceedings of the Fifteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems. PODS 1996*, Montreal, Canada, pp.58-67.

Tominski, C., Schulze-Wollgast, P. and Schumann, H. (2005). 3D Information Visualization for Time Dependent Data on Maps. *Proceedings of the Ninth International Conference on Information Visualisation*, London, UK, IEEE Computer Society.

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley, Massachusetts, U.S.A.

Tversky, B. (2005). Some ways images express and promote thought. ed. by Grialou, P., Longo, G. and Okada, M., in *Images and Reasoning,* Keio University, Tokyo.

Van de Weghe, N., Docter, R., De Maeyer, P., Bechtold, B. and Ryckbosch, K. (2007). The Triangular Model as an Instrument for Visualising and Analysing Residuality. *Journal of Archaeological Science,* 34(4), pp.649-655.

Van Wijk, J. J. and Selow, E. V. (1999). Cluster and Calendar-based Visualization of Time Series Data. ed. by Wills, G. and Keim, D., *IEEE Symposium of Information Visualization (InfoVis '99)*, San Francisco, USA, pp.4-9, IEEE Computer Society.

Venn, J. (1881). *Symbolic Logic*. Macmillan and Company, London.

Versichele, M., Neutens, T., Delafontaine, M. and Van de Weghe, N. (2012). The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities. *Applied Geography*, 32(2), pp.208-220.

Vila, L. (1994). A survey on temporal reasoning in artificial intelligence. *AI Communications*, 7(1), pp.4-28.

Weber, M., Alexa, M. and Muller, W. (2001). Visualizing time-series on spirals. *Proceedings of IEEE Symposium on Information Visualization 2001*, Washington, DC, U.S.A., pp.7-13.

5 Analysing Rough Time Intervals

Modified from: Yi Qiang, Matthias Delafontaine, Tijs Neutens, Birger Stichelbaut, Guy De Tré, Philippe De Maeyer, Nico Van de Weghe. (2012). Analysing Imperfect Temporal Information in GIS using the Triangular Model. *The Cartographic Journal*. (in press)

Abstract: Rough set and fuzzy set are two frequently-used approaches for modelling and reasoning about imperfect time intervals. In this chapter, we focus on imperfect time intervals that can be modelled by rough sets and use an innovative graphic model (i.e. the Triangular Model) to represent this kind of imperfect time intervals. This work shows that the TM is potentially advantageous in visualising and analysing such rough time intervals, and its analytical power can be better exploited when it is implemented in a computer application with graphical user interfaces and interactive functions. Moreover, a probabilistic framework is proposed to handle the uncertainty issues in temporal queries. We use a case study to illustrate how the unique insights gained by the TM can assist a geographical information system for exploratory spatio-temporal analysis.

Keyword: Rough sets, imperfect time interval, the Triangular Model, Geographical Information System, spatio-temporal analysis

5.1 Introduction

The temporal extents of entities, such as events and processes, are usually described by crisp time intervals bounded by a well-defined start point and end point. However, under some circumstances, the temporal extent of an entity is imperfect, and cannot be adequately modelled by a crisp time interval. On the one hand, some events may start or end gradually and therefore their start and end cannot be pinned to exact time points. For example, it is difficult to decide when the Cold War started and finished. Intervals of this kind of events are usually modelled by fuzzy sets (Zadeh, 1965) through the quantification of the graded truth of whether a time point is in the interval, bringing the concept of fuzzy time interval. On the other hand, in some other cases, it is only known that the start and end of a crisp interval are within certain ranges, but no extra information or assumptions about the distribution of the start and the end is available. Modelling this kind of imperfect time intervals with fuzzy sets would induce extra overhead and unnecessary complexity. In these cases, the alternative approach, i.e. rough sets (Pawlak, 1982), can excellently suit the modelling and handling of time intervals. Currently, a lot of disciplines are faced with the issue of imperfect time intervals, which is reflected in many contributions in modelling intervals by fuzzy sets (De Caluwe et al., 1999; De Caluwe et al., 1997; Garrido et al., 2009; Nagypál and Motik, 2003; Ohlbach, 2004; Schockaert et al., 2008) and rough sets (Bassiri et al., 2009; Bittner, 2002). However, while most of this work focuses on modelling and reasoning about imperfect time intervals, techniques and tools for visualising and analysing imperfect time intervals are still lacking. This limitation probably stems from the conventional representation of time intervals (e.g. Gantt chart, time table and historical timeline) which represents time intervals as linear segments along a one-dimensional time axis. The second dimension is often exploited merely to differentiate intervals of entities and has no temporal meaning. In this representation, the visual distribution of intervals is variable, according to the application of different ordering rules in the second dimension, for example, ordering intervals from the shortest to the longest, or from the earliest started to the latest started. The characteristics of the distribution of intervals cannot be observed in one single display. This is not convenient for visual pattern detection of time intervals, not to mention imperfect time intervals with more complex structure.

This issue also exists in geographical information science (GIScience), which considers time as one of the most important components of geographical information (Li and

Kraak, 2008; MacEachren et al., 1999; Neutens et al., 2008; Peuquet, 2002). In the recent development of GIScience, considerable effort has been made in handling the temporal aspect of geographical data (Andrienko et al., 2003; Neutens et al., 2007; Smith et al., 2007). Due to the limitations of data acquisition techniques, spatial data are often linked to imperfect temporal information. Dealing with imperfect temporal information becomes an increasingly significant issue in spatio-temporal data analysis, particularly in exploratory spatio-temporal data analysis (ESTDA) which greatly relies on graphical representations and visualisations (Andrienko, G and Andrienko, N, 2006). Since most prevalent techniques and tools of ESTDA represent time in the linear form, their ability in dealing with temporal imperfectness in spatio-temporal data is still not satisfactory.

To address these issues, attempts have been made to represent time in a two-dimensional (2D) space. For example, Keim et al. (2006) arranged time series of financial investments in a 2D representation called the Growth Matrix. This matrix is able to display the growth rates of investments in all possible sub-intervals of the time series. The TT-plot introduced by Imfeld (2000) applied a similar idea to analyse movement patterns. Besides these approaches, Kulpa proposed the MR Diagram that represents time intervals as 2D points, and investigated the use of this diagram for interval reasoning and arithmetic (Kulpa, 1997). Later, Van de Weghe *et al*. (2007) named this representation the Triangular Model (TM) and applied it in an archaeological context. Recently, Qiang *et al*. (Qiang et al., 2009; Qiang et al., 2010) have extended the TM to represent time intervals modelled by rough sets and fuzzy sets. This work majorly focused on the use of the TM in temporal reasoning. The practical value of the TM in visualising and analysing imperfect time intervals is still yet to be exploited.

To fill this gap, this chapter investigates the use of the TM in visualising and analysing imperfect time intervals. The focus is on imperfect time intervals that can be modelled by rough set theory. A probabilistic framework has been proposed to model the uncertainties in the temporal relations of such roughly-described intervals. In order to better demonstrate the use of TM, a prototype tool is introduced, which implements the TM in a geographical information system (GIS). This tool supports analysis of geospatial features that are referenced to imperfect intervals. We show that the advantages of the TM in visualising and querying roughly-described intervals can be better exploited through a computer application with a graphical user interface (GUI) and interactive functionalities. In addition, a case study is used to illustrate how the unique insights gained by the TM can assist a GIS for ESTDA involving imperfect time intervals.

In the remainder of the chapter, we first introduce the basic concept of the TM and how rough approximations of intervals are represented in TM. In Section 5.3, we apply probability theory to model uncertainties in temporal queries of imperfect intervals. Section 5.4 presents the implementation GeoTM, including its GUI, functionalities and supported data model. In section 5.5, the use of GeoTM is demonstrated in a concrete use case. The chapter ends with a brief conclusion and an outline of avenues for future work. All notations and symbols used in this chapter are summarised in the table in the appendix.

5.2 Triangular Model

5.2.1 Representation of Time Intervals

A time interval I is usually modelled as a convex set of real numbers, i.e. $[I^-, I^+]$ with $I^- < I^+$. I^- and I^+ respectively denote the start and end of I. In the traditional linear representation, a time interval is represented by a finite linear segment bounded by I^- and I^+ (see Figure 5-1a). This linear representation of time intervals is widely used in our daily life, for example time tables and historical time lines. The transformation from the linear representation of a time interval to the TM can be achieved by constructing two lines through the extremes of an interval (Figure 5-1b). For each interval I, two straight lines (L_1 and L_2) are constructed, with L_1 passing through I^- and L_2 passing through I^+. α_1 is the angle between L_1 and the horizontal axis and α_2 is the angle between L_2 and the horizontal axis, with $\alpha_1 = -\alpha_2 = \alpha$. The intersection of L_1 and L_2 is called the interval point. The angle α is a predefined constant which is identical for all intervals to ensure that every interval is mapped to a unique point in the 2D space. Although α can be any value between 0 and 90°, we consider α = 45° for consistency with earlier work (Kulpa, 1997, 2006; Qiang et al., 2010; Van de Weghe et al., 2007). In this way, the TM represents all time intervals as points in a 2D space, which is called the interval space (Figure 5-1c). The interval space is denoted as $I\mathbb{R}$ (Kulpa, 2006). In $I\mathbb{R}$, given an interval point I, the horizontal position indicates its midpoint (e.g. $mid(I)$) and the vertical position indicates its duration (e.g. $dur(I)$).

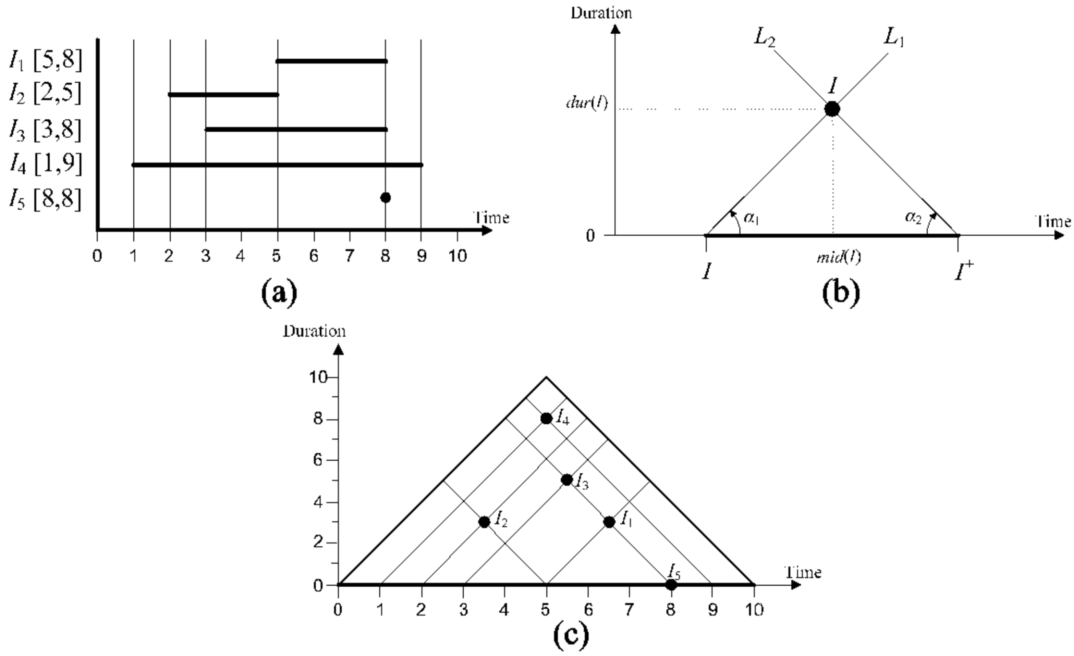


Figure 5-1: The transformation from the linear representation to TM. (a): The linear representation of time intervals. (b): The construction of an interval point in TM. (c): The TM representation of time intervals.

5.2.2 Representation of Temporal Relations

James F. Allen (1983) specified thirteen possible relations between two time intervals (see Table 3-1), which are referred to as Allen relations. In TM, every Allen relation can be represented as a specific zone (Kulpa, 1997). Given a study interval I_S = [0, 100], in the TM all examined intervals are located within the isosceles triangle formed by I_S^-, I_S^- and I. Let us consider, for example, a reference interval I_2 [33,66]. Any intervals (e.g. I_{1a}, I_{1b}, I_{1c}) *before* I_2 (Figure 3-4a) are located in the triangular area in the left corner of the study area (Figure 3-4b). Therefore, it is easy to deduce that all intervals *before* I_2 must be located in the black zone in Figure 3-4c, which is called the *before* zone of I_2. Likewise, all Allen relations with respect to an interval can be represented by zones in $I\mathbb{R}$ (Figure 5-3), which are called relational zones. For an arbitrary relation in Figure 5-3, the reference interval I has been chosen in the centre of the study period in order to avoid visual bias. Each relational zone represents the set of intervals that are in a specific relation to the reference interval I, which are denoted as $Rel(I)$. For example, the *during* zone of I represents the set of intervals that are *during* I and is denoted as $during(I)$. On the other hand, $Rel(I_1, I_2)$ expresses the statement that I_1 is in the relation Rel to I_2. For example, $overlaps(I_1, I_2)$ expresses that I_1 overlaps I_2.

Table 5-1: Thirteen Allen Relations (Allen 1983)

I_1 *equal* I_2	if $I_1^- = I_2^-$	$\wedge\ I_1^+ = I_2^+$	
I_1 *starts* I_2	if $I_1^- = I_2^-$	$\wedge\ I_1^+ < I_2^+$	
I_1 *started–by* I_2	if $I_1^- = I_2^-$	$\wedge\ I_2^+ < I_1^+$	
I_1 *finishes* I_2	if $I_1^+ = I_2^+$	$\wedge\ I_1^- > I_2^-$	
I_1 *finished–by* I_2	if $I_1^+ = I_2^+$	$\wedge\ I_2^- > I_1^-$	
I_1 *meets* I_2	if $I_1^+ = I_2^-$		
I_1 *met–by* I_2	if $I_2^+ = I_1^-$		
I_1 *overlaps* I_2	if $I_2^- > I_1^-$	$\wedge\ I_1^+ < I_2^+$	$\wedge\ I_1^+ > I_2^-$
I_1 *overlapped–by* I_2	if $I_1^- > I_2^-$	$\wedge\ I_1^- < I_2^+$	$\wedge\ I_2^+ < I_1^+$
I_1 *during* I_2	if $I_1^- > I_2^-$	$\wedge\ I_1^+ < I_2^+$	
I_1 *contains* I_2	if $I_2^- > I_1^-$	$\wedge\ I_2^+ < I_1^+$	
I_1 *before* I_2	if $I_1^+ < I_2^-$		
I_1 *after* I_2	if $I_2^+ < I_1^-$		

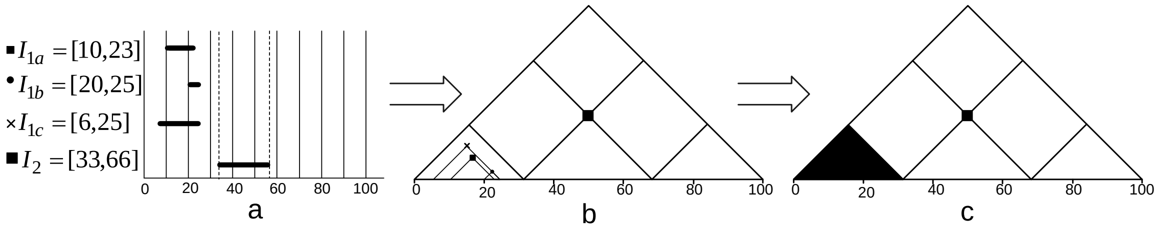


Figure 5-2: Temporal relations in the linear model and TM, taking *before* as an example. (a): I_{1a}, I_{1b}, I_{1c} and I_2 in the linear representation. (b): I_{1a}, I_{1b}, I_{1c} and I_2 in TM. (c): The *before* zone of I_2.

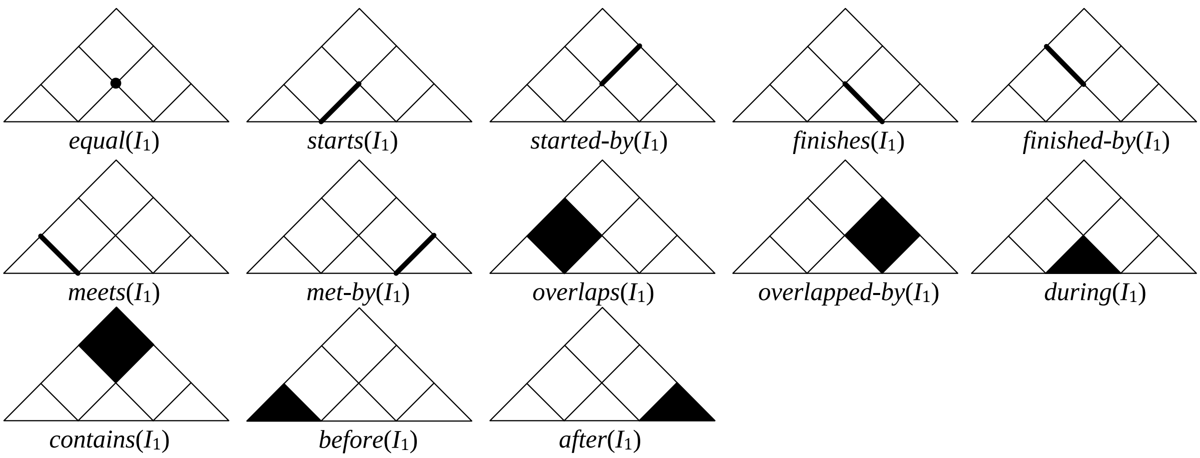


Figure 5-3: Relational zones representing sets of intervals in certain Allen relations to the reference interval *I*.

5.2.3 Representation of Rough Approximation

Incomplete information may result in uncertainties about the exact start and end of a time interval. This may happen in many observation activities in which data are acquired at discrete time stamps, such as images or photographs in remote sensing. From a sequence of images, for example, one can determine whether a feature exists at specific time stamps. However, the feature’s status between two time stamps is unknown. With these discrete snapshots, the interval of the feature’s existence is thus imperfect. If there is no prior knowledge about the distribution of the start and end, modelling this kind of imperfect time intervals with membership functions of fuzzy sets would induce extra overhead and unnecessary complexity. In these cases, rough sets can be considered as an appropriate and adequate solution (Bassiri et al., 2009; Bittner, 2002).

In the rough sets approach, an imperfect time interval I is described by an upper approximation $\overline{I}$ and a lower approximation $\underline{I}$, where $\underline{I} \subseteq \overline{I}$ (Figure 3-1). Such a pair of $\underline{I}$ and $\overline{I}$ is the rough approximation of I, which is denoted as $R(I)$. The rough approximation can also be called a rough time interval. Time points in $\underline{I}$ are *definitely* in I, whereas all time points not in $\overline{I}$ are *definitely not* in I. $\overline{I}$ is bounded by the earliest possible start $\overline{I}^{-}$ and the latest possible end $\overline{I}^{+}$, while $\underline{I}$ is bounded by the latest possible start $\underline{I}^{-}$ and the earliest possible end $\underline{I}^{+}$. In between $\underline{I}$ and $\overline{I}$, there are two rough boundaries $R(I^{-})$ and $R(I^{+})$ gathering the time points that are *possibly* in I. Unlike fuzzy time intervals which define the extent to which the time points are possibly in I, the rough approximation of I classifies time points into *definitely in*, *definitely not in* and *possibly in* I.

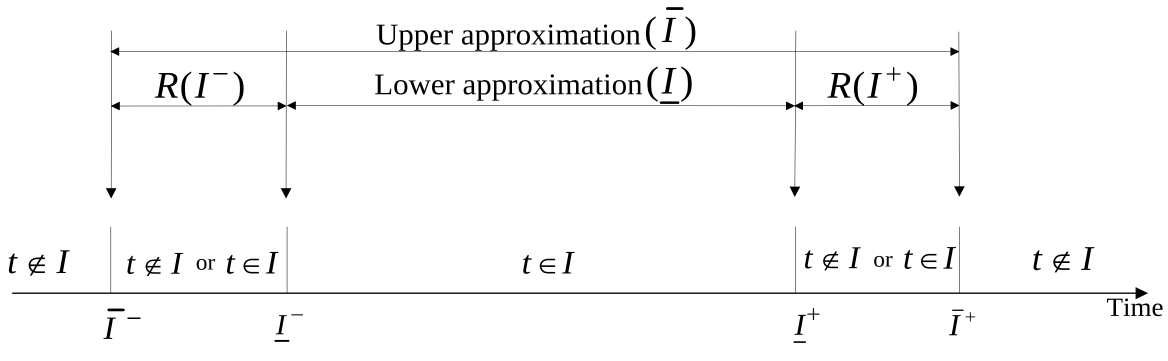


Figure 5-4: The linear representation of the rough approximation of I.

In the linear representation, the rough approximation of an interval is usually represented as a tripartite linear segment (Figure 5-5a). However, in TM, a rough approximation of an interval is represented by a convex 2D geometry in $I\mathbb{R}$. To construct a rough approximation $R(I)$ in TM, two parallel lines are projected from $\overline{I}^-$ and $\underline{I}^-$ with angle α to the horizontal axis; and the other two lines are projected from $\underline{I}^+$ and $\overline{I}^+$ with angle $-\alpha$ to the horizontal axis. These four lines form a rectangle (Figure 5-5b) which indicates a zone where the exact interval I can be found. In other words, this zone represents the set of intervals that are possibly equal to I. The shape and location of the rectangle completely express the characteristics of $R(I)$. In this way, the rough approximation of a time interval can be represented by such a rectangle (Figure 5-5c). If the lower approximation is empty, the rough approximation becomes a triangle on the horizontal axis, e.g. I_2 in Figure 5-5.

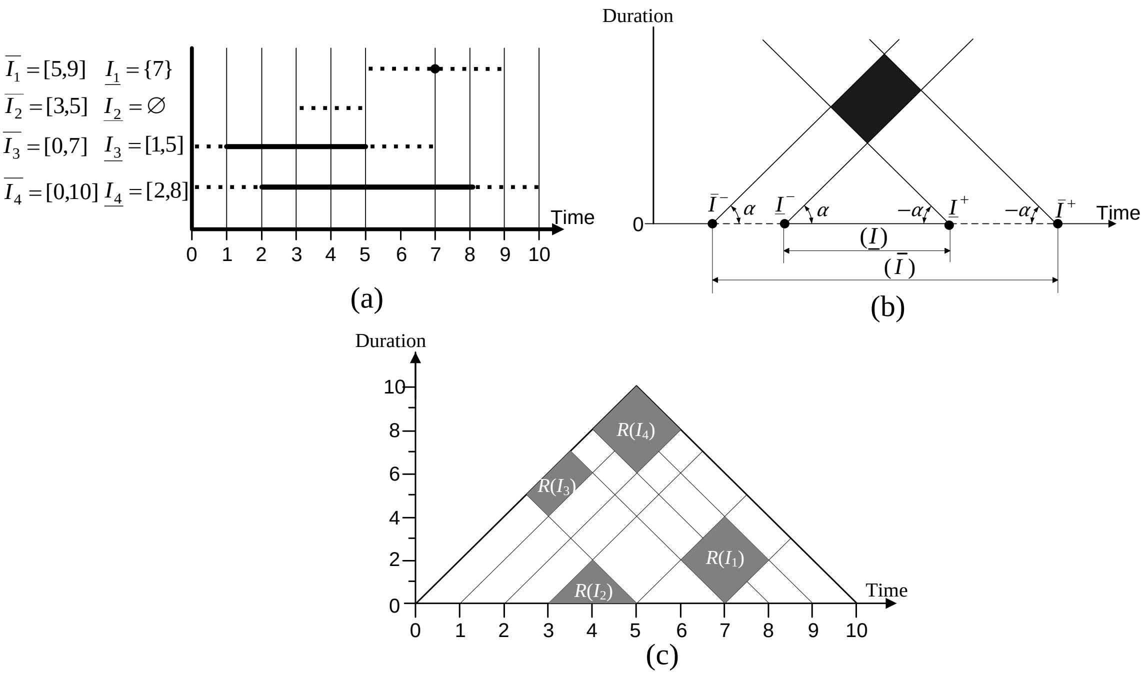


Figure 5-5: The transformation from the linear representation to the TM representation of rough approximations of intervals. (a): Rough approximations of intervals in the linear representation. Full lines denote $\underline{I}$, and dashed lines denote $R(I^-)$ and $R(I^+)$. The combination of full lines and dotted line forms $\overline{I}$. (b): The construction of the rough approximation of an interval in TM. (c): Rough approximations of intervals in TM.

5.3 Probabilities in Rough Approximations

Because time intervals described by rough approximations have an uncertain start and an uncertain end, temporal queries about such rough time intervals may result in uncertain answers. To handle the uncertainties in temporal queries, a probabilistic framework is needed. Given a rough approximation $R(I)$, $R(I^-)$ and

$R(I^+)$ define the ranges where the exact start I^- and the exact end I^+ can be found (Eq. 5-1 and Eq. 5-2). As there is no further knowledge about the distribution of I^- and I^+ in $R(I^-)$ and $R(I^+)$, we assume that the probability distribution in $R(I^-)$ and $R(I^+)$ are uniform. That is to say, every time point in $R(I^-)$ has the same probability of being equal to I^- (i.e. Eq. 5-3). Therefore, the probability density of $x = I^-$ (where $x \in R(I^-)$) is the quotient of 1 and the cardinality of $R(I^-)$ (Eq. 5-4). In the same way, every time point in $R(I^+)$ has the same probability of being equal to I^+ (Eq. 5-5), and therefore the probability density of $x = I^+$ (where $x \in R(I^+)$) is the quotient of 1 and the cardinality of $R(I^+)$ (Eq. 5-6). Based on these prerequisites, it can be deduced that every interval in $R(I)$ has the same probability of being equal to I (Eq. 5-7). Because $R(I)$ defines the interval set that includes I (Eq. 5-8), the probability density of $x = I$ (where $x \in R(I)$) is the quotient of 1 and the cardinality of $R(I)$ (Eq. 5-9). Therefore, $R(I)$ can be considered as an interval set in which every interval has the same probability of being equal to I. Because the TM maps time intervals into points in a 2D space, the cardinality of an interval set is proportional to its area, and therefore the probability density of $x = I$ (where $x \in R(I)$) is inversely proportional to the area of $R(I)$ in the TM (see Eq. 5-9).

According to Eq. 5-9, we can deduce that, given a set of intervals A and a rough approximation $R(I)$ (such as Figure 5-6), the probability that I is in A is the ratio between the cardinality of $A \cap R(I)$ and the cardinality of $R(I)$ (Eq. 5-10). In the TM representation, this probability is the ratio between the area of $A \cap R(I)$ and the area of $R(I)$ (Eq. 5-10). If $R(I)$ is totally contained in A (i.e. $R(I) \subseteq A$) the probability that I is in A is 1. If $R(I)$ and A do not intersect (i.e. $R(I) \cap A = \emptyset$), the probability that I is in A is 0. The interval set A can also be a relational zone of a certain interval, e.g. $before(I')$. In this case, the probability that I is in A can be interpreted as the probability that I is *before* I'. With respect to the principles of probability theory (Jaynes and Bretthorst, 2003), the probabilities of multiple rough approximations can be deduced. For example, given an interval set A and two independent rough approximations $R(I_1)$ and $R(I_2)$ (Figure 5-7), the probability that only one of I_1 and I_2 is in A can be obtained from Eq. 5-11, and the probability that both I_1 and I_2 are in A can be obtained by Eq. 5-12. Analogously, given an interval set A and n independent rough approximations i.e. $R(I_1)$, $R(I_2) \dots R(I_n)$, the probability that m intervals $(0 \le m \le n)$ are in A can be obtained by Eq. 5-13. Since interval sets and rough approximations of intervals are represented as 2D geometries in TM, the probability that an interval is in an interval set is expressed by the overlap ratio of their geometries. This feature of the TM are similar to Venn diagrams (Venn, 1881) in which the areas represents the occurrences of a certain event and the intersection of areas expresses the coincidence of events. Compared with mathematic formulas, we contend that such graphic representation is more perceivable and intuitive for human beings. Hence, the TM can be considered as a promising platform for visual queries of intervals described by rough approximations.

$$\sum_{x \in R(I^-)} P(I^- = x) = 1$$ Eq. 5-1

$$\sum_{x \in R(I^+)} P(I^+ = x) = 1$$ Eq. 5-2

$$P(I^- = x_1) = P(I^- = x_2), \qquad x_1, x_1 \in R(I^-)$$ Eq. 5-3

$$f_{I^-}(x) = P(I^- = x) = \frac{1}{|R(I^-)|} \quad , \qquad x \in R(I^-)$$ Eq. 5-4

$$P(I^+ = x_1) = P(I^+ = x_2), \qquad x_1, x_2 \in R(I^+)$$ Eq. 5-5

$$f_{I^+}(x) = P(I^+ = x) = \frac{1}{|R(I^+)|}, \qquad x \in R(I^+)$$ Eq. 5-6

$$P(I = I_1) = P(I = I_2), \qquad I_1, I_2 \in R(I)$$ Eq. 5-7

$$\sum_{x \in R(I)} P(I = x) = 1$$ Eq. 5-8

$$f_I(x) = P(I = x) \; = \frac{1}{|R(I)|} \propto \frac{1}{Area(R(I))} \quad , \qquad x \in R(I)$$ Eq. 5-9

$$P(I \in A) = \sum_{x \in A} \frac{1}{|R(I)|} = \frac{|A \cap R(I)|}{|R(I)|} = \frac{Area(A \cap R(I))}{Area(R(I))}$$ Eq. 5-10

$$P(either\ I_1\ or\ I_2\ is\ in\ A) = P(I_1 \in A \wedge I_2 \notin A) + P(I_1 \notin A \wedge I_2 \in A)$$
$$= P(I_1 \in A) \times P(I_2 \notin A) + \mathrm{P}(I_1 \notin A) \times \mathrm{P}(I_2 \in A)$$
$$= \frac{|A \cap R(I_1)|}{|R(I_1)|} \times \frac{\left|A - \left(A \cap R(I_2)\right)\right|}{|R(I_2)|} + \frac{|A \cap R(I_2)|}{|R(I_2)|} \times \frac{\left|A - \left(A \cap R(I_1)\right)\right|}{|R(I_1)|}$$
$$= \frac{Area\left(A \cap R(I_1)\right) \times Area\left(A - A \cap R(I_2)\right) + Area\left(A \cap R(I_2)\right) \times Area\left(A - A \cap R(I_1)\right)}{Area\left(R(I_1)\right) \times Area\left(R(I_2)\right)}$$

Eq. 5-11

$$P(both\ I_1\ and\ I_2\ are\ in\ A) = P(I_1 \in A) \times P(I_2 \in A)$$
$$= \frac{|A \cap R(I_1)|}{|R(I_1)|} \times \frac{|A \cap R(I_2)|}{|R(I_2)|}$$
$$= \frac{Area(A \cap R(I_1))}{Area(R(I_1))} \times \frac{Area(A \cap R(I_2))}{Area(R(I_2))}$$

Eq. 5-12

$$\mathrm{P}(m\ intervals\ of\ I_1,\ I_2, \ldots I_n\ are\ in\ A) = \sum_{Q \subseteq S} [\prod_{i \in Q} P(I_i \in A) \times \prod_{j \in (S-Q)} P(I_j \notin A)],$$
$$S = (1,2,3 \ldots n) \wedge Q \subseteq S \wedge |Q| = m$$

Eq. 5-13

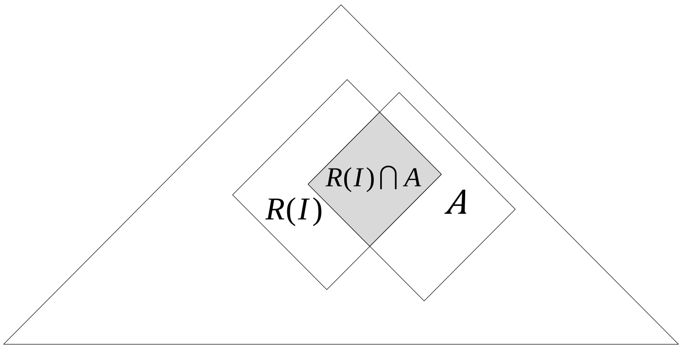


Figure 5-6: An arbitrary interval set A and an arbitrary rough approximation $R(I)$.

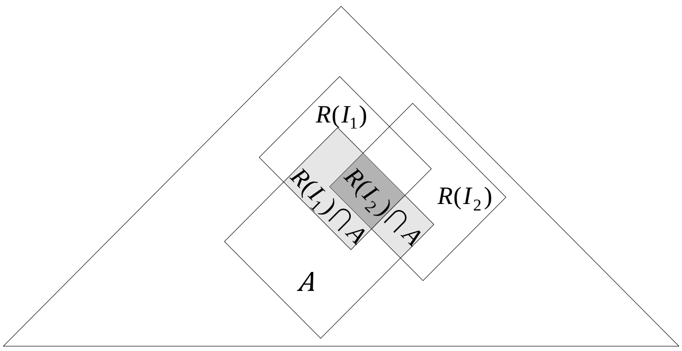


Figure 5-7: An arbitrary interval set A and two arbitrary rough approximations $R(I_1)$ and $R(I_2)$.

5.4 Implementation

To demonstrate and exploit practical use of TM, we have implemented the TM into a prototype tool (i.e. GeoTM) which incorporates the TM into a GIS. GeoTM allows visualising, querying and analysing spatio-temporal data. More specifically, it is able to handle discrete entities with spatial locations and temporal extents. The spatial locations are modelled as vector geometries (i.e. point, line and polygon), while the temporal extents are modelled as time intervals. In this chapter, we will take special care of time intervals described by rough approximations. Besides spatial and temporal extents, these geographical entities may have other attributes. The spatial locations, temporal extents and other attributes are linked to geographical entities by unique identifiers (ID) of entities. GeoTM is built on top of ArcGIS™ which is a desktop GIS produced by ESRI®. Within its object model (ArcObjects), developers can call on existing functions and components of ArcGIS™ to develop customised applications. Consequently, GeoTM is compatible with the supported data formats in ArcGIS™, such as shapefile and dBASE files. In the representation of geographical entities, spatial locations and geometries of time interval are stored in two shapefiles, while the attributes are stored in a database table (i.e. a dBASE file). Figure 5-8 gives an overview of how geographical entities are modelled and represented in GeoTM.

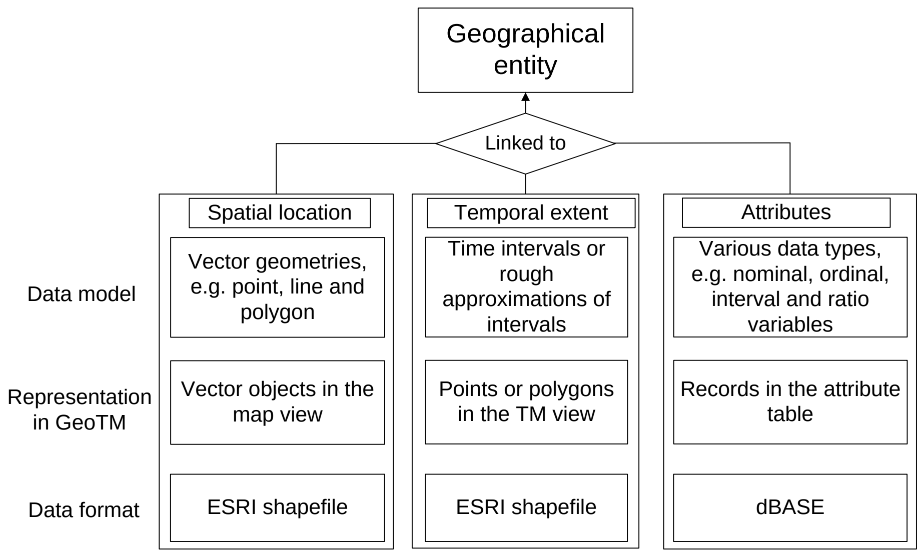


Figure 5-8: Representation of spatio-temporal data in GeoTM.

5.4.1 Graphical User Interface of GeoTM

The user interface of GeoTM consists of a map view, the TM view and controls that can trigger specific functions or user forms (Figure 5-9). The map view is used to display spatial locations of geographical entities, which are modelled as points, lines or polygons. The TM view displays time intervals of entities, using the TM representation. Rough approximations of intervals are represented as polygons. In the TM view, there may exist overlaps of polygons. In order to display the pattern of overlapped polygons, gradual colours are assigned to areas according to the numbers of overlaps: a darker colour is assigned to the areas with more overlaps. In Figure 5-9, a dataset of military features in the First World War (WWI) is displayed in GeoTM. The spatial locations of these features are displayed in the map view, while their lifetimes (represented by rough approximations of intervals) are displayed in the TM view. In this section, we will use this dataset to illustrate the functionalities of GeoTM. The entire scenario of this dataset will be described in Section 5.5. The temporal resolution of the dataset is one day. Thus, in this context, a time interval is the period between two specific dates formatted as Year/Month/Day.

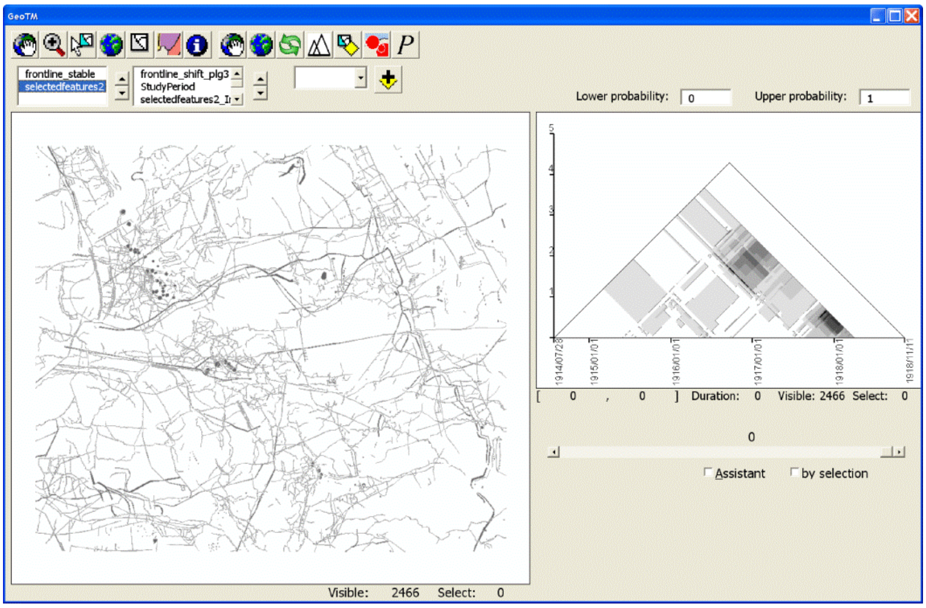
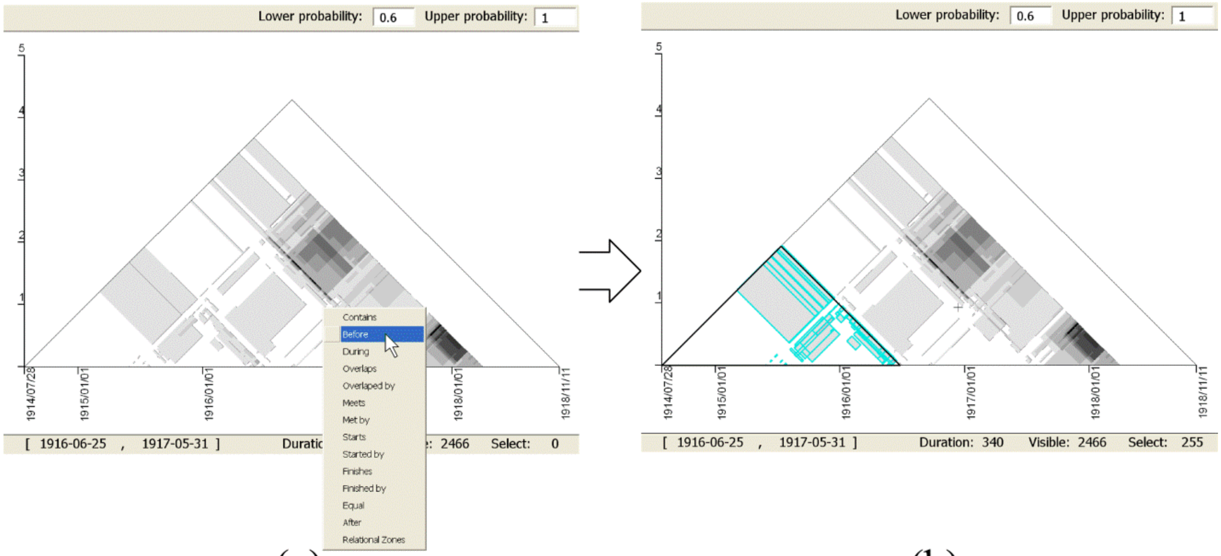


Figure 5-9: The user interface of GeoTM, consisting of a map view and a TM view.

Users can interactively browse the map view and TM view by zooming and panning. In the map view, objects can be selected by dragging a rectangle. In the TM view, special selection tools have been designed to select intervals according to specific temporal queries, which will be described in detail in the next sub-section. Other attributes of entities are stored in the attribute table that can pop up when the user clicks the corresponding button. 'Linked brushing' is supported among the map view, the TM view and the attribute table. This function allows selecting objects from any of these three views and dynamically updating the other two views to highlight the corresponding objects. With this function, geographical entities in GeoTM can be queried with spatial, temporal and attribute constraints. Additionally, many common functions of a conventional GIS are supported in GeoTM. For example, several datasets can be loaded into GeoTM as multiple layers.

5.4.2 Temporal Queries on Rough Approximations

As introduced in Section 5.2, intervals that satisfy a certain temporal relation are located within a relational zone. Thus, queries based on temporal relations can be modelled as specific zones in TM, i.e. query zones. A query zone can be a relational zone or combinations (e.g. intersection and union) of multiple relational zones. Whether an interval satisfies the temporal query depends on whether the interval point lies within the query zone. However, when intervals are described by rough approximations, in the TM the rough approximation can be partially in the query zone. If this is the case, temporal queries can be answered with probability thresholds. For example, within a dataset of rough approximations of intervals, one can select all intervals that have more than 90% probability of being *during* an indicated interval. In TM, this query is expressed by selecting rough approximations that have more than 90% of its area in the *before* zone of the indicated interval. Therefore, in the TM view of GeoTM, querying tools are available to define temporal queries by creating query zones. On the one hand, users can indicate an interval by moving the cursor to a specific position in the TM view. When right-clicking on this position, a drop-down menu of Allen relations appears. Next, by clicking an Allen relation in the menu, all intervals in this relation are selected, according to a pre-defined probability threshold. Figure 5-10 shows an example of selecting intervals that are probably *before* [1916/06/25,1917/05/31]. In this example, the lower probability threshold is 0.6 and the upper probability threshold is 1. As a result, intervals with more than 60% probability *before* [1916/6/25,1917/5/31] are selected.



(a) (b)

Figure 5-10: The selection of intervals probably *before* [1916/06/25, 1917/05/31], with more than 60% probability. (a): Moving the mouse cursor to the interval [1916/06/25, 1917/05/31], and right-clicking to trigger the menu of Allen relations, and then right-clicking the ‘before’ option. (b): Intervals that are probably before [1916/6/25,1917/5/31] are selected.

Besides temporal queries of Allen relations, some other queries can be made by dragging specific geometries in the TM view. For example, a convex set of intervals can be selected by dragging a rectangle whose sides are in α or $-\alpha$ angle to the horizontal axis (Figure 5-11a). A convex interval set is defined as the set of time intervals *in-between* two different time intervals (Kulpa, 2006). We developed this query tool because convex interval sets can be easily interpreted by Allen relations or combinations of Allen relations. $\ll I_1, I_2 \gg$ is used to denote a convex interval set *in-between* I_1 and I_2. The formal definition of $\ll I_1, I_2 \gg$ can be found in Eq. 5-14. Moreover, two parallel lines can be dragged in α or $-\alpha$ angle with the horizontal axis, in order to select intervals that *start-within* or *end-within* a certain interval (e.g. Figure 5-11b and c). *starts-within* (I_1) is defined as a set of intervals whose start points are in I_1 (Eq. 5-15). Analogously, *ends-within* (I_1) is defined as a set of intervals whose end points are in I_1 (Eq. 5-16). Though the *starts-within* and *ends-within* can be expressed by unions of Allen relations, we design such query tools for people that are more accustomed to the expressions of ‘starts within’ and ‘ends within’ than unions of Allen relations. Furthermore, by dragging a range along the vertical axis, users can select intervals whose durations are in a specific range (Figure 5-11d). In GeoTM, all queries are carried out with respect to pre-defined probability thresholds. In the examples in Figure 5-11, the lower threshold is set to 0.6 and the upper threshold is set to 1, meaning that the selected intervals have more than 60% probability of satisfying these queries.

$$\ll I_1, I_2 \gg \equiv \{I \mid min(I_1^-, I_2^-) < I^- < max(I_1^-, I_2^-) \wedge min(I_1^+, I_2^+) < I^+ < max(I_1^+, I_2^+)\} \qquad \text{Eq. 5-14}$$

$$starts\text{-}within\ (I_1) \equiv \{I \mid I_1^- < I^- < I_1^+\} \qquad \text{Eq. 5-15}$$

$$ends\text{-}within\ (I_1) \equiv \{I \mid I_1^- < I^+ < I_1^+\} \qquad \text{Eq. 5-16}$$

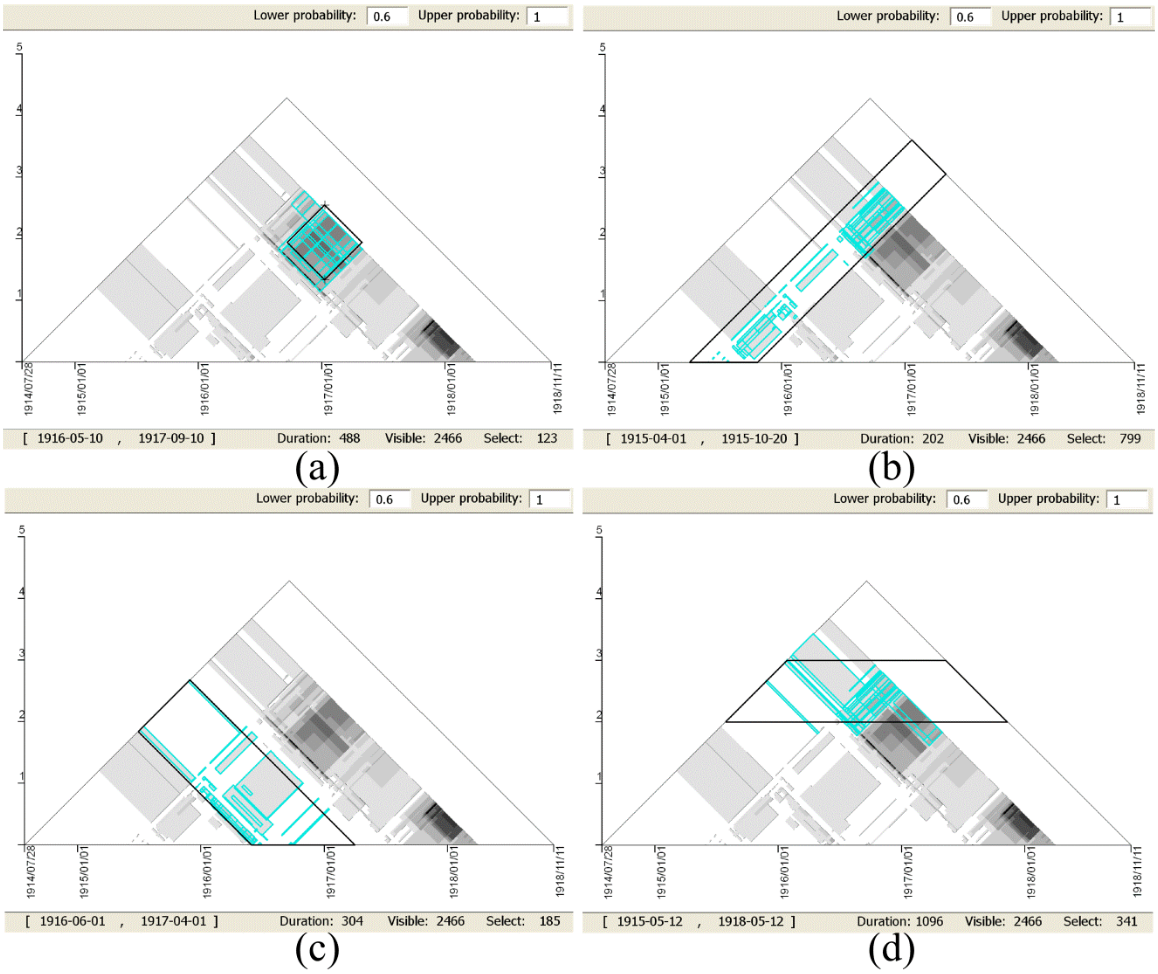


Figure 5-11: Making temporal queries by dragging geometries in the TM view. (a): Selecting intervals *in-between* [1915/10/01, 1918/04/20] and [1916/05/10, 1917/09/10]. (b): Selecting intervals that *start-within* [1915/04/01, 1915/10/20]. (c): Selecting intervals that *end-within* [1916/06/01, 1917/04/01]. (d) Selecting intervals that are longer than 2 years and shorter than 3 years.

Based on Eq. 5-13, when a temporal query is defined, users may obtain the probability that a given number of intervals satisfy this query. For example, GeoTM can return the probabilities that a certain number of intervals are *before* [1915/11/21, 1916/08/30]. In this case, 67 rough approximations have common parts with (intersect or within) the query zone, while 43 out of these 67 rough approximations are totally within the query zone. This means that there are minimum 43 intervals *before* [1915/11/21,1916/08/30] and maximum 67 intervals *before* [1915/11/21, 1916/08/30]. The probability that n intervals ($43 < n < 67$) are *before* [1915/11/21, 1916/08/30] is between 0 and 1. GeoTM can automatically generate the probability that n intervals are *before* [1915/11/21, 1916/08/30], from n= 43 to n= 67 (Figure 5-12). In this way, one can know how many intervals satisfy the query, with respect to a confidence level. According to Figure 5-12, there are 59 intervals *before* [1915/11/21, 1916/08/30], at 0.9 confidence level. This function is also useful in distinguishing 'real clusters' and 'fake clusters' of intervals. Because the colours in the TM represent the number of overlaps, which is actually the maximum number of intervals within the area, an area with a darker colour indicates a potential cluster of intervals. However, the probability that intervals are clustered in this area may vary. By dragging a query zone over the dark area, one can be aware of the numbers of intervals clustered in this area with respect to specific probabilities.

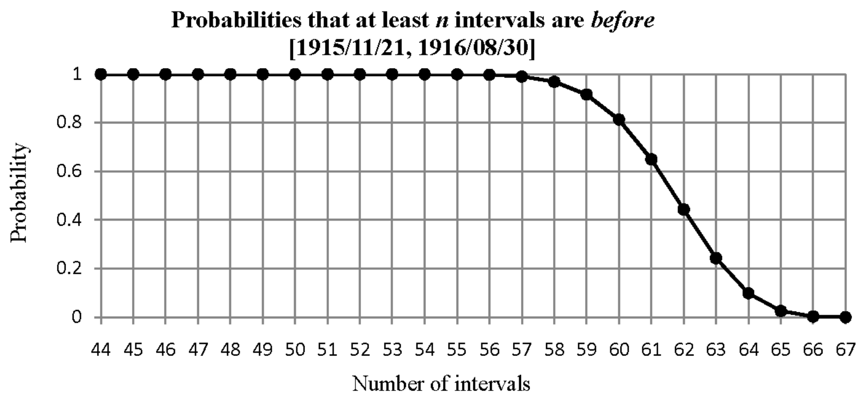


Figure 5-12: The line diagram of probabilities that at least n intervals are *before* [1915/11/21, 1916/08/30].

5.5 Case Study

Having introduced basic functionalities of GeoTM in the previous section, this section will illustrate the use of the TM in supporting GIS for ESTDA by means of a case study.

5.5.1 Dataset

During World War One (WWI), a large number of aerial photos in West-Flanders (Belgium) were taken at discrete time stamps by all combating nations as an intelligent tool to collect information on the enemy's intentions. These aerial photographs are preserved in archives all over the world. The largest collections are held at the Belgian Royal Army Museum, the Imperial War Museum, the Australian War Memorial and the Bavarian Military archive. From these aerial photos, one can observe whether a military feature (e.g. a fire trench, gun position or barrack) was not yet present, present, or destroyed (Stichelbaut and Bourgeois, 2009). Although the state of a feature is uncertain between two snapshots, we assume that it does not change between two snapshots which show similar states. Thus, the uncertainty only exists between two snapshots showing different states. Certainly, this assumption relies on our knowledge that snapshots are dense enough to capture most of features' changes. When more volatile entities are considered, an appropriate temporal resolution will be required. In this context, rough sets are excellently suited for temporal modelling, because no knowledge is available about the state of a feature between two snapshots. Thus, we can consider a period of snapshots showing similar states as the lower approximation, its neighbouring uncertain intervals as boundary regions, and all of them form the upper approximation (Figure 3-25). Thus, a feature's lifetime can be meaningfully represented by a rough approximation.

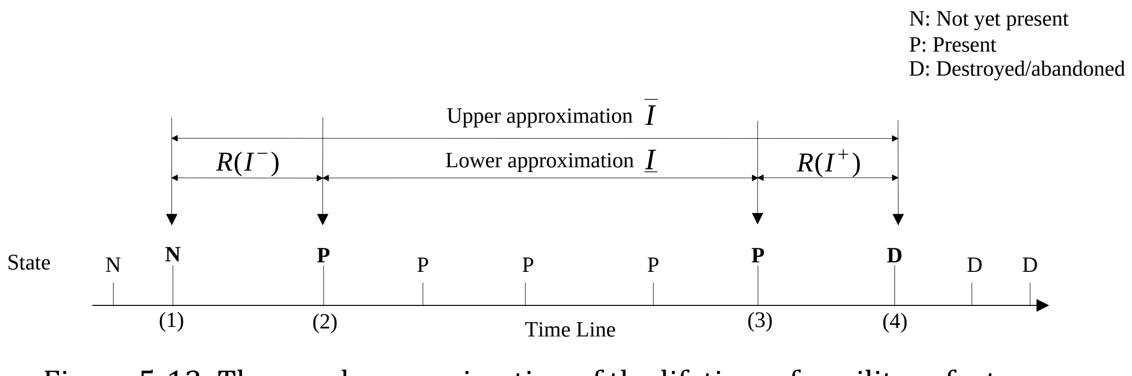


Figure 5-13: The rough approximation of the lifetime of a military feature.

More specifically, four photos determine the lifetime of a feature (Figure 3-25): (1) the last photo in which the feature is not yet present, (2) the first photo in which the feature is present, (3) the last photo in which the feature is present, and (4) the first photo in which the feature is destroyed or abandoned. The interval between the dates of photo (2) and (3) is the lower approximation of the feature's lifetime, while the interval between the dates of photo (1) and photo (4) is the upper approximation. Intervals between the dates of photo (1) and (2), and intervals between the dates of photo (3) and (4) are the boundary regions, which indicate respectively the range of the feature's construction and destruction/abandonment dates. There are a few exceptions, where a feature was not yet present in one photo and already destroyed in the following photo. In these cases, photo (2) and (3) are missing, and therefore the rough approximation has an empty lower approximation. As described in Section 5.2.3, such rough approximations are represented as triangles on the horizontal axis. In this case study, a rectangular area (around 3 kilometre × 3.5 kilometre) is selected near Ypres (Belgium) as the study area, containing 2466 military features (Figure 5-14). This study area near Ypres is one of the most important battlefields of WWI. From 1914 to 1918 there was constant artillery fighting going on. And more importantly, Ypres was the scene of several large offensive actions by both Germans and Allies. Therefore, it is an ideal area to test the potential of the TM in exploratory analysis.

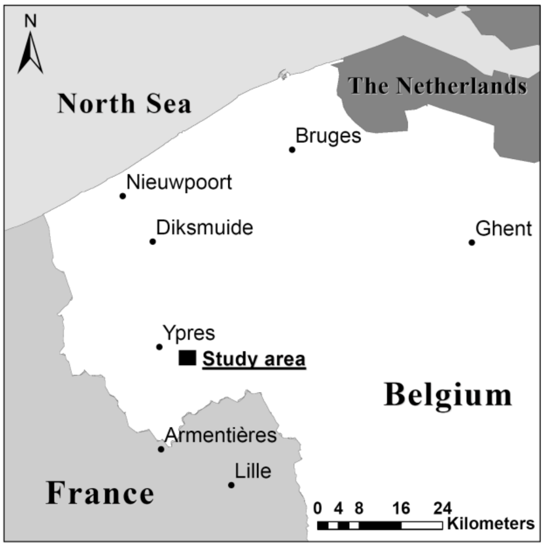


Figure 5-14: The location of the study area.

For analysis, we take account of spatio-temporal information of the frontlines during the war. There are 11 snapshots of the states of the German and Allied frontlines during WWI. From these snapshots we have observed that the frontlines were stable in I_2 [1915/5/25, 1916/6/14] (we format a date as year/month/date), but had significant shifts in three time intervals, i.e. I_1[1915/5/23, 1915/5/25], I_3[1916/6/14, 1917/10/1] and I_4[1917/10/1, 1918/4/15] (Figure 5-15). Note that the frontline shifts in I_3 and I_4 might only take a few days. But from snapshots of frontlines, it is only known that the shifts happened in these three intervals and the exact dates of the shifts are unavailable. Figure 5-15 displays the locations and shifts of frontlines in these four intervals. Note that the German army was always on the east side of the Allied army. If the positions of frontlines were out of the study area, such as in I_3 and I_4, we use arrows to indicate the direction of the frontline shifts.

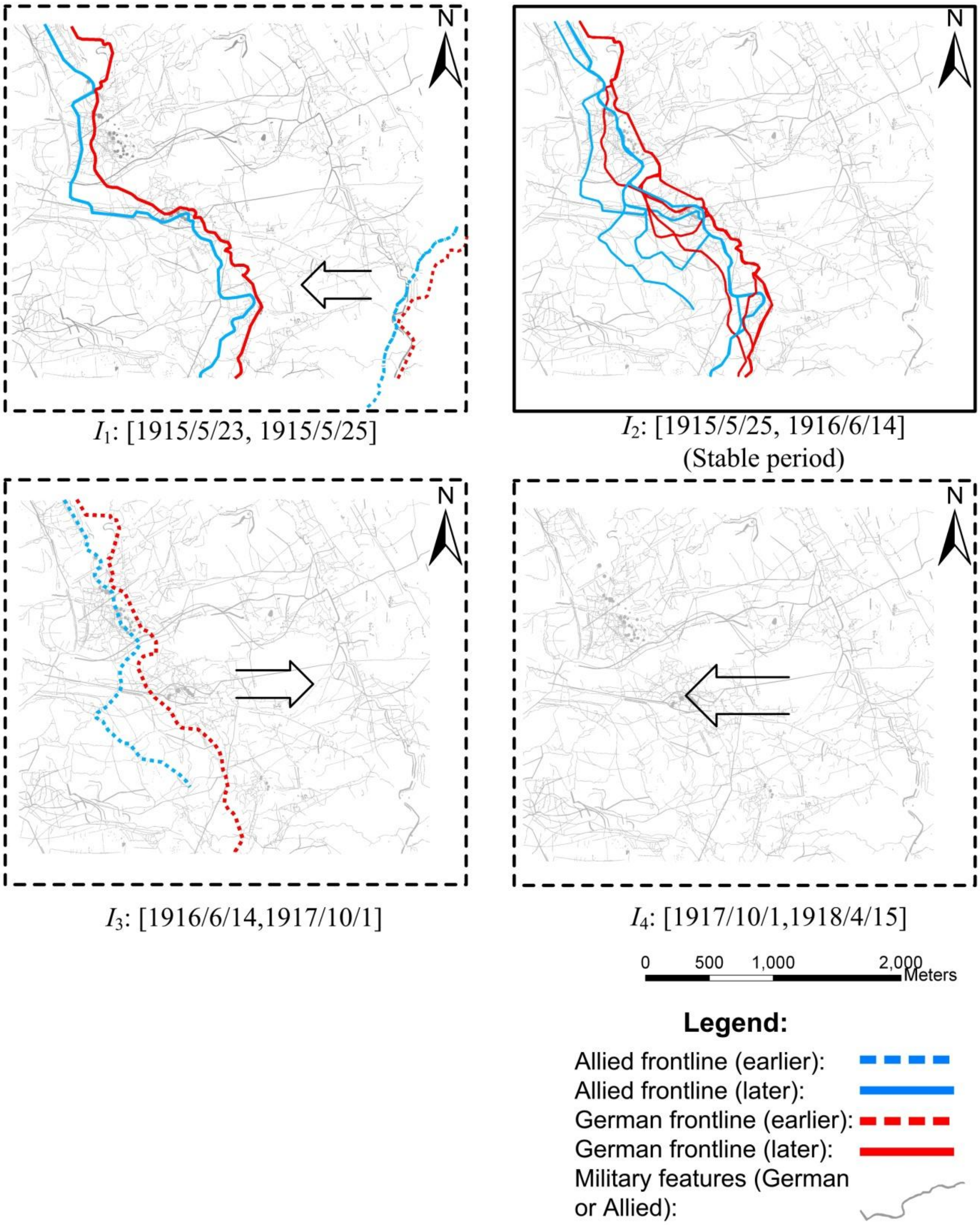


Figure 5-15: The states of German and Allied frontlines in WWI. The map in the solid box shows the period during which the frontlines were relatively stable, while the maps in the dashed boxes show periods during which the frontlines had significantly shifted.

5.5.2 Analysing Rough Time Intervals in GeoTM

First, the dataset of the military features is imported to GeoTM. The spatial locations of these features are displayed in the map view, while the rough approximations of their intervals are displayed in the TM view (Figure 5-16). To identify the intervals of frontline shifts, the interval points and *during* zones of I_1, I_2, I_3 and I_4 are added to the TM view (Figure 5-16). The zone with solid boundary is the *during* zone of the interval in which the frontlines were stable, i.e. I_2, while the zones with dashed boundaries are the *during* zones of the intervals in which the frontlines had significant shifts, i.e. I_1, I_3, and I_4. As I_1 is very short, its *during* zone is invisible in this scale. Using the selection tool we find that 99% of features (2433 out of 2466) in the study area were built after I_1. Considering the frontline shift in I_1, we can infer that the major military activities in the study area started after I_1 when the frontline moved from the south-eastern edge of the study area to the centre. Three dark areas that indicate potential clusters of intervals can be observed in the TM (Figure 5-16). In the next three sub-sections these potential clusters will be analysed with respect to frontline states in I_2, I_3 and I_4. Note that in the following three sub-sections, we set the lower probability threshold of all temporal queries as 0.9 and the upper probability threshold as 1, in order to select intervals that satisfy the query with more than 90% probability.

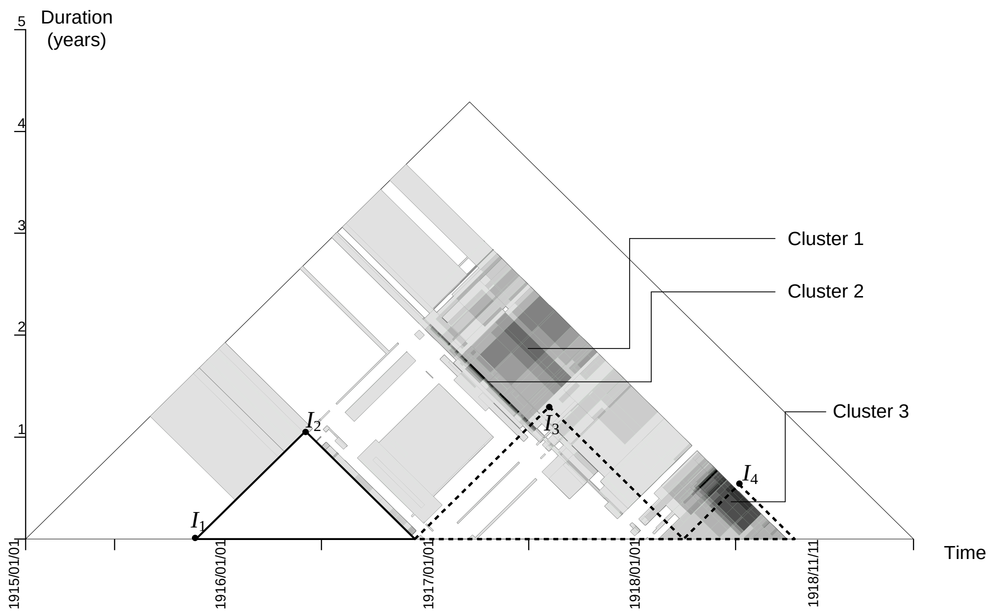


Figure 5-16: The rough approximations of the feature lifetimes in the TM view, with the *during* zones of I_1, I_2, I_3 and I_4 displayed.

Stable period of frontlines

Most intervals in Cluster 1 and Cluster 2 *start-within* I_2, when the frontlines were relatively stable. When selecting intervals that probably *start-within* I_2 (more than 90% probability), we obtain 1531 intervals (62% of all features in the study area), which reflects that a majority of military features were constructed during this stable period. There are also some intervals that *end-within* I_2, which are lifetimes of features that were destroyed within I_2. By selecting intervals that *end-within* I_2, we obtained 200 intervals (8% of all features) and the map view shows all these features were close to the frontlines (Figure 5-17). This is reasonable because military features close to the frontline were easier to be destroyed by artillery attacks or minor offensives. Therefore, it is natural to infer that during I_2 there were no significant military actions, and the armies at both sides were building military features in order to keep their positions. Only a small portion of features close to the frontlines were destroyed.

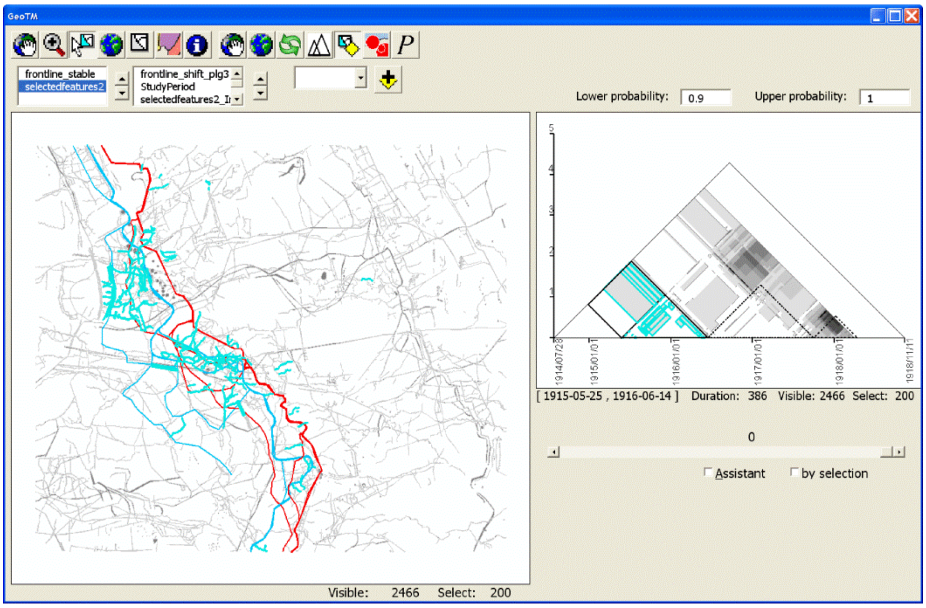


Figure 5-17: The selection of intervals that *end-within* I_2.

Shift of frontlines in I_3

During I_3, the frontlines shifted from their positions to the east of the study area, towards the German side. In the TM view, it is observable that the intervals in Cluster 2 *end-within* I_3. More specifically, the intervals in Cluster 2 *end-within* a quite short interval [1917/07/10, 1917/08/01], which results in a slim shape (see Cluster 2 in Figure 5-16). When selecting intervals that *end-within* [1917/07/10, 1917/08/01], we obtain 1050 features, which is 43% of all features (2466 features) in the study area. This finding reflects that a large number of features were intensively destroyed during [1917/07/10, 1917/08/01]. The map view shows that all these features were distributed in the eastern side of the old frontlines, which were earlier occupied by the German army (Figure 5-18). By checking the attribute table, one can see that 99% of these 1050 features were German features, which was 83% of all German features (1260 features) in the study area. With these findings, it becomes possible to infer, without consulting any historical documents, that an overwhelming and intensive destruction happened to German features in July 1917 in the study area. Probably due to this destruction, the German army lost its area and was pushed to the east. Referring to historical literature (Barton, 2005; Verbeke, 2006), these findings reflect the fact that during this period the Allies intensively destroyed German features using artillery fire, in a battle which became known as the Third Battle of Ypres. During the battle, Allied army took Passchendaele (Belgium), pushing the front line towards the east.

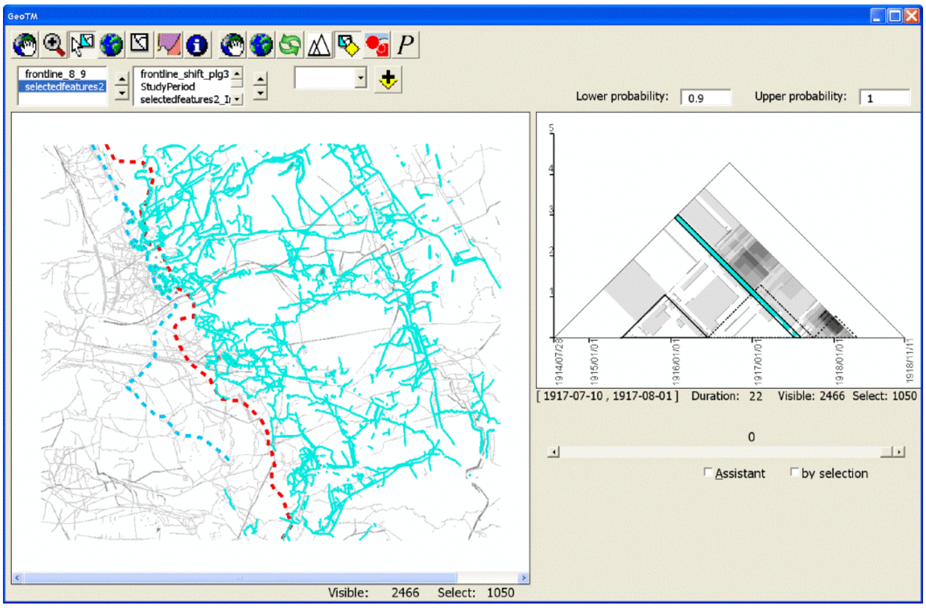


Figure 5-18: The selection of intervals that *end-within* [1917/7/10, 1917/8/1].

Shift of frontlines in I_4

Through observation in the TM view, most intervals in Cluster 1 and Cluster 3 *end-within* I_4. When selecting intervals probably *end-within* I_4 (with more than 90% probability), 817 features are obtained and the attribute table shows that all these features are Allied. This means that 68% Allied features (817 out of 1206) in the study area were probably destroyed within I_4. Considering that within I_4 the frontlines have moved from east to west towards the Allied side, it is straightforward to infer that these features were destroyed due to the offensive by the German army. Features in Cluster 1 were built earlier (i.e. during I_2), when the frontlines were relatively stable. By selecting features in Cluster 1, one can see these features are mostly located in the eastern part of the study area, which was occupied by the Allied army during I_2 (Figure 5-19). Features in Cluster 3 were built later (i.e. during I_4), and were evenly distributed over the entire study area (Figure 5-20). These findings reveal that after the Third Battle of Ypres, the Allied army controlled the whole study area and built military features over the area (i.e. features in Cluster 3). Later on, the German army attacked back and destroyed all Allied features (features in both Cluster 1 and Cluster 3) in the study area. After this military action, the frontlines shifted from the east of the study area to the west of the study area. Known from historical literatures (Howard, 2002), these findings probably reflect the Battle of Lys (a part of Spring Offensives of Germany) in April 1918, during which the German army attacked the Allied army and pushed the frontline back to the west.

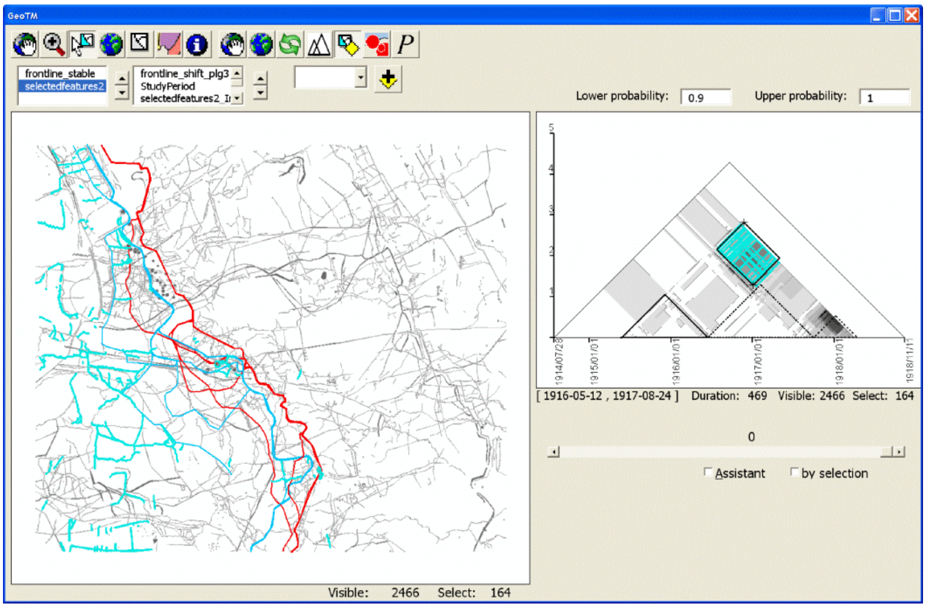


Figure 5-19: The selection of intervals in Cluster 1.

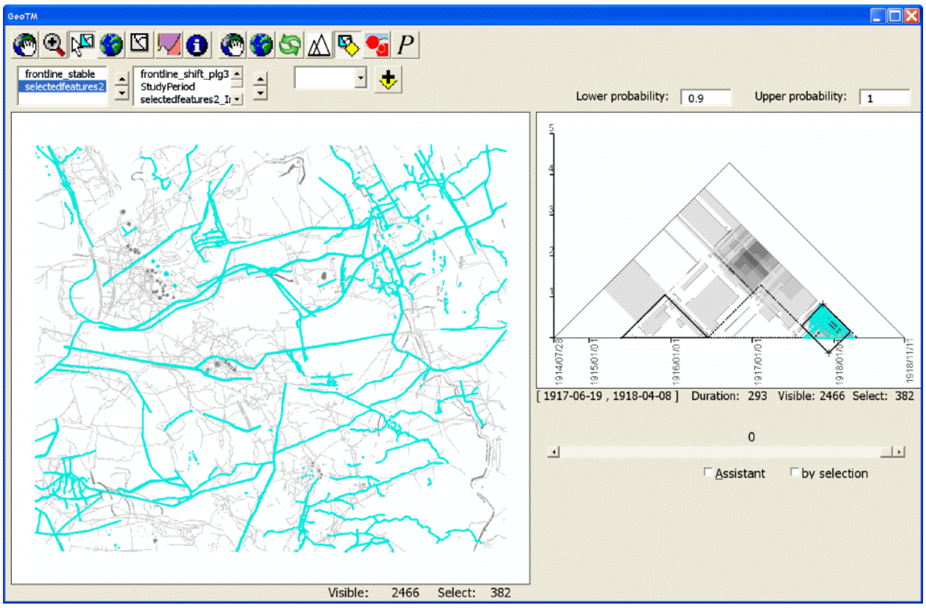


Figure 5-20: The selection of intervals in Cluster 3.

5.6 Conclusion and Future Work

This chapter has investigated the use of the TM in visualising and analysing time intervals modelled by rough sets. Compared with the fuzzy set approach, the rough set approach is excellently suited to imperfect time intervals with no prior knowledge and assumptions about the distribution of their starts and ends, which broadly exist in discrete data acquisitions. In TM, rough approximations of intervals are represented as polygons in a 2D space. We contend that this representation is advantageous in visualising the distribution of imperfect time intervals, because the patterns (e.g. clusters) in the distribution can be explicitly displayed. If an interval is described by rough approximation, its relation with other intervals may have uncertainties. Due to these uncertainties, temporal queries may generate uncertain answers. Therefore, we have proposed a probabilistic framework to model uncertainties in temporal relations of roughly-described intervals. Since the TM represents temporal relations as 2D geometries, the probability of temporal relations can be expressed by the overlap rate of the corresponding geometries, which is analogous to Venn diagrams. Compared to mathematical expressions, such graphical representation is potentially more intuitive to human beings and can offer a promising basis for visual query and analysis.

In order to evaluate the capabilities of TM, we have implemented it into a prototype tool (i.e. GeoTM) which incorporates the TM within a GIS. GeoTM shows that the advantages of the TM in visualisation and querying can be better exploited when it is implemented in a computer program. Instead of providing a static TM visualisation, GeoTM supports interactive functionalities that enable flexible manipulations of the TM display. These functionalities have the potential of assisting visual observation and pattern detections. Besides visualisation, GeoTM also offers possibilities to make temporal queries by creating 2D geometries. Whether intervals satisfy the query depends on the extent to which the intervals are part of the query zone. Moreover, query zones can be directly created on top of the visualisation of intervals so that users can select an observed cluster of intervals by designing a proper temporal query. In order to handle the uncertainties in temporal relations of roughly-described intervals, special functionalities based on the proposed probabilistic framework have been applied in GeoTM, allowing users to make temporal queries with consideration of probabilities. On the one hand, temporal queries can be defined with probability thresholds, which ensure that every retrieved interval satisfies the query in a certain probability. On the other hand, once a query is made, GeoTM can return the probability that a certain number of intervals satisfy the query. With this feature, users can be aware of how many intervals satisfy the query, with respect to a certain confidence level. With these functionalities, GeoTM is able to support ESTDA of geographical entities with reference to roughly-described intervals. As shown in the case study of WWI features, GeoTM can be used to explore the military features from spatial, temporal and attribute aspects. From the detected patterns and clusters, users can discover interesting phenomena in the war without consulting historical literatures.

Not only spatio-temporal data, we believe that the TM can be applied in other contexts that involve roughly-described intervals for purposes of information visualisation, exploratory analysis and data mining. In future work, the applicability of the TM needs to be further assessed by more use cases, preferably covering different research contexts. Future extensions will improve the usability and interactivity of the implementation. For instance, a line diagram can be added to dynamically display the probabilities that different numbers of intervals satisfy the query being sketched. Other than indicating the maximum number of overlapped rough approximations, the colour coding in the interval view can be manipulated with dynamic controls to display possibilities of different meanings. Before the implementation is released to a broader community, its scalability needs to be systematically evaluated. Furthermore, we plan to investigate the possibility of representing and analysing fuzzy time interval by TM. The fuzzy set approach is more suitable for modelling imperfect time intervals when assumptions or knowledge of the distributions of starting points and end points are available. This extension may rely on the application of more advanced visualisation techniques such as 3D visualisation.

Appendix: Summary of notations

Notation	Name of symbols	Examples
I	A crisp time interval	
I^-	The start of I	
I^+	The end of I	
$dur(I)$	The duration of I	
$mid(I)$	The midpoint of I	
$I\mathbb{R}$	The interval space, i.e. the universal set that contains all time intervals	
$\overline{I}$	The upper approximation of I	
$\underline{I}$	The lower approximation of I	
$\overline{I}^-$	The start of $\overline{I}$	
$\overline{I}^+$	The end of $\overline{I}$	
$\underline{I}^-$	The start of $\underline{I}$	
$\underline{I}^+$	The end of $\underline{I}$	
$R(I)$	The rough approximation of I	
$R(I^-)$	The earlier boundary region or the rough start of I	
$R(I^+)$	The later boundary region or the rough end of I	
$f_X(x)$	The probability density function of X	$f_{I^-}(x)$ is the probability density function of I^-.
$P(X)$	The probability that the statement X is true	$P(I_1 = I)$ denotes the probability that $I_1 = I$ is true.
$\|A\|$	The cardinality of the set A, i.e. the total number of elements in A. A can be a set of real numbers or a set of intervals.	$\|R(I)\|$ denotes the cardinality of $R(I)$.
$Rel(I)$	A relational zone of I or the set of intervals that are in a certain relation to I.	$before(I)$ is the *before* zone of I and denotes the set of intervals that are *before* I.
$Rel(I_1, I_2)$	The statement that I_1 and I_2 are in a certain relation.	$overlaps(I_1, I_2)$ expresses the statement that I_1 *overlaps* I_2.
$\ll I_1, I_2 \gg$	The set of intervals that are *in-between* I_1 and I_2.	The formal definition of *in-between* can be found in Eq. 5-14

References

Allen, J. F. (1983). Maintaining Knowledge about Temporal Intervals. Communications of the ACM, 26(11), pp.832-843.

Andrienko, N. and Andrienko, G. (2006). Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach. Springer-Verlag Berlin, Heidelberg.

Andrienko, N., Andrienko, G. and Gatalsky, P. (2003). Exploratory spatio-temporal visualization: an analytical review. Journal of Visual Languages & Computing, 14(6), pp.503-541.

Barton, P. (2005). De Slagvelden van Wereldoorlog 1. Van de Eerste Slag om Ieper to Passendale. Lannoo, Tielt.

Bassiri, A., Malek, M., Alesheikh, A. and Amirian, P. (2009). Temporal Relationships between Rough Time Intervals. ed. by Gervasi, O., Taniar, D., Murgante, B., Laganà, A., Mun, Y. and Gavrilova, M., in Computational Science and Its Applications – ICCSA 2009, pp.543-552, Springer Berlin / Heidelberg.

Bittner, T. (2002). Approximate Qualitative Temporal Reasoning. Annals of Mathematics and Artificial Intelligence, 36(1), pp.39-80.

De Caluwe, R., Van Der Cruyssen, B., De Tré, G., Devos, F. and Maesfranckx, P. (1999). The Semantics and Modelling of Flexible Time Indications. in Computing with Words in Information/Intelligent Systems 1 (Foundations), pp.229-256, Springer/Physica-Verlag, Heidelberg, Germany.

De Caluwe, R., Van Der Cruyssen, B., De Tré, G., Devos , Filip and Maesfranckx, P. (1997). Fuzzy Time Indications in Natural Language Interfaces. in Flexible Query Answering Systems, pp.163-185, Kluwer Academic Publishers.

Garrido, C., Marín, N. and Pons, O. (2009). Fuzzy Intervals of Represent Fuzzy Valid Time In a Temporal Relational Database. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 17(1), pp.173-192.

Howard, M. (2002). The First World War. Oxford University Press, Oxford.

Imfeld, S. (2000). Time, Points and Space: Towards a better analysis of wildlife data in GIS. Ph.D. thesis, University of Zurich.

Jaynes, E. T. and Bretthorst, G. L. (2003). Probability Theory: The Logic of Science. Cambridge University Press, Cambridge, United Kingdom.

Keim, D. A., Nietzschmann, T., Schelwies, N., Schneidewind, J., Schreck, T. and Ziegler, H. (2006). A Spectral Visualization System for Analyzing Financial Time Series Data. ed. by Ertl, T., Joy, K. and Santos, B., Proceedings of the

Eurographics/IEEE-VGTC Symposium on Visualization (EuroVis 2006), Lisbon, Portugal, pp.195-200.

Kulpa, Z. (1997). Diagrammatic Representation for a Space of Intervals. Machine Graphics & Vision, 6, pp.5-24.

Kulpa, Z. (2006). A Diagrammatic Approach to Investigate Interval Relations. Journal of Visual Languages and Computing, 17(5), pp.466-502.

Li, X. and Kraak, M.-J. (2008). The Time Wave. A New Method of Visual Exploration of Geo-data in Time–space. The Cartographic Journal, 45(3), pp.193-200.

MacEachren, A. M., Wachowicz, M., Edsall, R., Haug, D. and Masters, R. (1999). Constructing knowledge from multivariate spatiotemporal data: integrating geographical visualization with knowledge discovery in database methods. International Journal of Geographical Information Science, 13(4), pp.311-334.

Nagypál, G. and Motik, B. (2003). A Fuzzy Model for Representing Uncertain, Subjective, and Vague Temporal Knowledge in Ontologies. Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics, (ODBASE), Catania, Italy, pp.906-923.

Neutens, T., Van de Weghe, N., Witlox, F. and De Maeyer, P. (2008). A three-dimensional network-based space-time prism. Journal of Geographical Systems, 10(1), pp.89-107.

Neutens, T., Witlox, F., Van de Weghe, N. and De Maeyer, P. (2007). Space-time opportunities for multiple agents: a constraint-based approach. International Journal of Geographical Information Science, 21(10), pp.1061-1076.

Ohlbach, H. J. (2004). Relations between fuzzy time intervals. Proceedings of 11th International Symposium on Temporal Representation and Reasoning (TIME 2004), Montréal, Québec, pp.44-51.

Pawlak, Z. (1982). ROUGH SETS. International Journal of Computer & Information Sciences, 11(5), pp.341-356.

Peuquet, D. J. (2002). Representations of space and time. The Guilford Press, New York.

Qiang, Y., Asmussen, K., Delafontaine, M., De Tré, G., Stichelbaut, B., De Maeyer, P. and Van de Weghe, N. (2009). Visualising rough time intervals in a two-dimensional space. Proceedings of 2009 IFSA World Congress / EUSFLAT Conference, Lisbon, Portugal.

Qiang, Y., Delafontaine, M., Asmussen, K., Stichelbaut, B., De Tré, G., De Maeyer, P. and Van de Weghe, N. (2010). Modelling Imperfect Time Intervals in a Two-Dimensional Space. Control and Cybernetics, 39(4), pp.983-1010.

Schockaert, S., De Cock, M. and Kerre, E. E. (2008). Fuzzifying Allen's Temporal Interval Relations. IEEE Transactions on Fuzzy Systems, 16(2), pp.517-533.

Smith, M. J. d., Goodchild, M. F. and Longley, P. A. (2007). Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools Troubador Publishing Ltd, Leicester, UK.

Stichelbaut, B. and Bourgeois, J. (2009). The Aerial Imagery of World War One: A Unique Source for Conflict and Landscape Archaeology. Photogrammetrie - Fernerkundung - Geoinformation, (3), pp.235-244.

Van de Weghe, N., Docter, R., De Maeyer, P., Bechtold, B. and Ryckbosch, K. (2007). The Triangular Model as an Instrument for Visualising and Analysing Residuality. Journal of Archaeological Science, 34(4), pp.649-655.

Venn, J. (1881). Symbolic Logic. Macmillan and Company, London.

Verbeke, R. (2006). De Derde Slag bij Ieper-Passendale: de Honderddaagse slag in 1917. ed. by Chielens, P., Dendooven, D. and Decoodt, H., in De Laatste Getuige. Het Landschap van Wereldoorlog 1 in Vlaanderen, pp.63 - 66, Lannoo, Tielt.

Zadeh, L. A. (1965). Fuzzy Sets. Information and Control, 8(3), pp.338-353.

6 ANALYSING LINEAR DATA

Modified from: Yi Qiang, Steven Logghe, Philippe De Maeyer, Nico Van de Weghe. Multi-scale Analysis of Linear Data in a Two-Dimensional Space. In preparation for the submission to *Informaiton Visualization*

Abstract: Many disciplines are faced with the problem of handling linear data, such as time series and traffic speed along a road. This chapter introduces an innovative visual representation for linear data, the Continuous Triangular Model (CTM). In CTM, values of all subintervals within the linear data are displayed in a two-dimensional continuous field, where every position corresponds to a specific subinterval of the linear data. The value at a position is derived through a certain calculation (e.g., average or summation) of the linear data within the subinterval it represents. The CTM thus provides an explicit overview of the linear data in all different scales. This chapter shows how the CTM can bring added value to the visual analysis of different types of linear data. We also show how the coordinate interval space in the CTM can support the analysis of multiple linear datasets through the methods of spatial analysis, including map algebra and cartographic modelling. Real-world datasets and scenarios will be employed to demonstrate the usefulness of this representation in analytical tasks.

Keywords: Continuous Triangular Model, Linear data, Time Series, Multi-scale Analysis, Information Visualisation

6.1 Introduction

Many disciplines are faced with the problem of handling linear data, which is referred to data sequences arranged in a one-dimensional (1D) space. The most well-known example is time series, which is a data sequence collected at successive time points or aggregated in successive time intervals. Considerable efforts have been dedicated to researching time series data, which leads to a rich body of methodologies and theories for time series analysis (Enders, 2008; Hamilton, 1994; Liao, 2005). In addition to time series, linear data can be derived from a linear geographical space, such as traffic speed along a road and runoff along a river, or objects with a linear structure, such as voltage along power lines and DNA sequences. The same as when analysing other data types, visualisation has been proven to be an effective analytical approach for linear data (Aigner et al., 2007; Muller and Schumann, 2003). Although many approaches have been developed for the visualization of linear data (Havre et al., 2000; Hochheiser and Shneiderman, 2004; Lin et al., 2004; Weber et al., 2001), the line chart remains the most frequently used. In a line chart, the horizontal dimension indicates positions in the linear space, and the vertical dimension indicates the values at the positions. The linear data are represented as a curve, offering a direct view of the variation of the linear data along the linear space. Line diagrams usually only display the linear data in a certain scale. Displaying data in different scales would require drawing more curves, which makes the data display matted. Manipulating a sliding bar to shift the scales of the linear data to be displayed is an alternative approach. However, using such controlled animation, one cannot obtain an overall picture of the linear data in all different scales.

The Continuous Triangular Model (CTM) provides an alternative approach to represent linear data and overcomes the difficulty of traditional approaches in visualising linear data in multiple scales. The CTM is based on a diagrammatic representation of time intervals that is initially proposed by Kulpa (Kulpa, 1997, 2006). Later, Van de Weghe named it the Triangular Model (TM) and applied it to archaeological use cases (Van de Weghe et al., 2007). More recently, Qiang investigated its use in reasoning imperfect intervals (Qiang et al., 2010) and visual analytics (Qiang et al., 2012a; Qiang et al., 2012b). The basic idea of the TM is representing time intervals as points in a coordinated two-dimensional (2D) space. Evolved from TM, the CTM adds the third dimension to the interval space in the TM and forms a continuous field to display time series in all different intervals. In the continuous field, every point represents a specific interval and is referenced to a certain value of the interval it represents, such as the summation, average or standard deviation etc. In a similar way, other types of linear data can also be represented in CTM. On the one hand, the CTM can provide an overview of linear data in all different scales. On the other hand, as the CTM is based on a 2D coordinate space, the glossary of spatial analysis methods in geographical information science (GIScience) are now open to be employed to manipulate and analyse the CTM data (Goodchild et al., 2007; Smith et al., 2007). An idea similar to the CTM is the Growth Matrix, introduced by Keim (Keim et al., 2006), which visualises time series of stock prices in a 2D space. In the Growth Matrix, the horizontal axis indicates the time when the stock is purchased, and the vertical axis indicates when the stock is sold. Every point in the matrix is referenced to the price difference between the purchasing and selling times. Beyond Keim's research, this chapter demonstrates how other formulas (i.e., average and summations) can be applied to calculate the values of intervals, and how this representation can be useful for analysing other types of linear data (e.g., traffic data). Moreover, we show how the methods of map algebra and cartographic modelling (Tomlin, 1990) can be applied to the CTM to solve multi-criteria decision-making problems based on time series.

In the remainder of this chapter, we first review the basic concept of TM. Next, the concept of the CTM is introduced. We then demonstrate how the CTM can be applied to the analysis of traffic speed along a motorway. Afterwards, we show how map algebra and cartographic modelling can be applied to analyse linear data represented in CTM. Finally, conclusions are drawn and directions of future work are proposed.

6.2 Triangular Model

In the classical linear representation, a time interval I is represented as a linear segment bounded by a start point I^- and end point I^+. The properties of an interval are expressed by the location and extent of the linear segment in a 1D space. The basic idea of the TM is mapping the linear segment in the 1D space into points in a 2D space. Given an arbitrary time interval I, two straight lines (L_1 and L_2) are projected from the two extremes (I^- and I^+), with L_1 passing through I^-; L_2 through I^+ (Figure 6-1). α_1 is the angle between L_1 and the horizontal axis, while α_2 is the angle between L_2 and the horizontal axis, where $\alpha_1 = -\alpha_2 = \alpha$. The intersection point of L_1 and L_2 is called the interval point, which completely expresses the properties of the time interval I. The horizontal position indicates the midpoint of I, i.e., $mid(I) = (I^- + I^+)/2$, while the vertical position indicates the duration of I, i.e., $dur(I) = \tan\alpha \cdot (I^+ - I^-)/2$. The start of the interval I^-, the end of the interval I^+ and interval point I form an isosceles triangle. Therefore, this representation of time intervals is called the Triangular Model (TM). The angle α is a pre-defined constant that is identical to the construction of all interval points. Here, we set $\alpha = 45°$ to be consistent with previous work (Kulpa, 1997, 2006; Qiang et al., 2010), though α can be set to any value between 0° and 90° for specific purposes. In TM, every time interval can be represented as a unique point in the 2D space. The 2D space where interval points are located in is called the Interval Space ($I\mathbb{R}$).

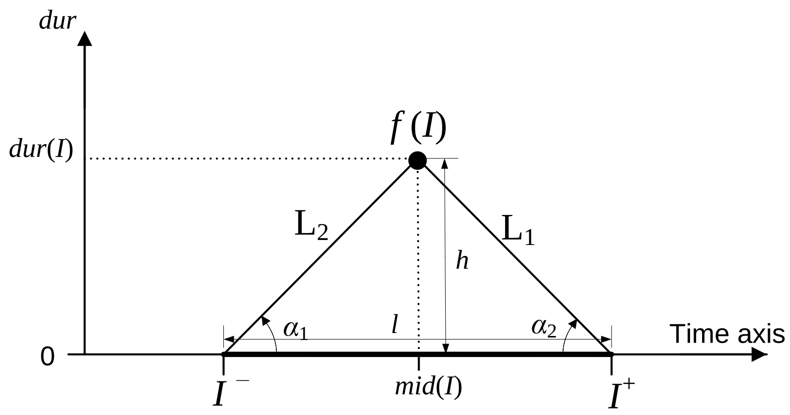


Figure 6-1. The configuration of the Triangular Model (TM).

According to the interval algebra introduced by Allen, there exist thirteen atomic relations between two time intervals. In TM, the intervals in a certain temporal relation to a referenced interval I are located in a specific zone in $I\mathbb{R}$. Figure 6-2 illustrates the zones of the thirteen Allen relations. In Figure 6-2, it is assumed that there exists a triangular study area that contains all intervals, and the referenced interval I is in the centre of the study area. The black zones contain the sets of intervals in certain temporal relations to the referenced interval. For instance, the black zone in the left corner of the study area represents the set of intervals that are $before\ I$, which is denoted as $before(I)$. Temporal constraints based on Allen relations can thus be modelled as such zones. The composition of temporal constraints is based on the same principle as that of the Venn Diagram. For instance, the set of intervals that satisfy several constraints are located in the intersection of the corresponding zones. The set of intervals that satisfy one of the several constraints is located in the union of the zones. For a more detailed description of the relational zones and their compositions please refer to (Qiang et al., 2010).

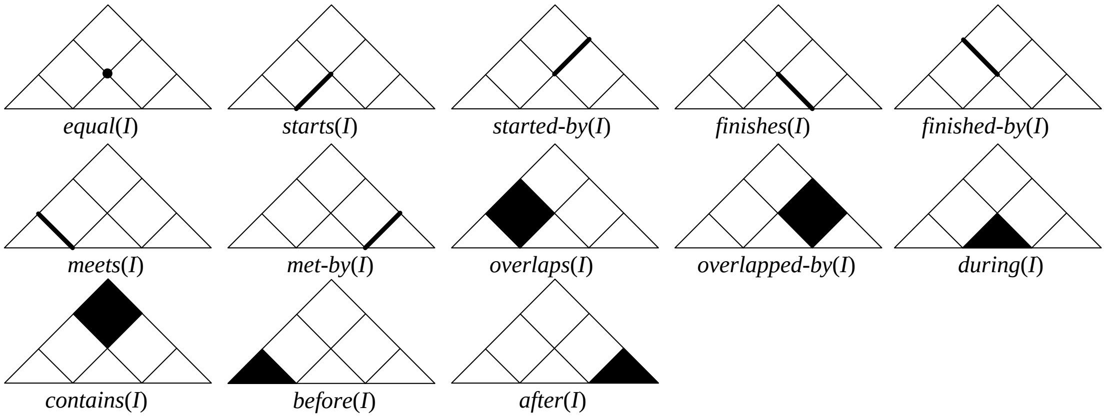


Figure 6-2: The representation of thirteen Allen relations in TM.

6.3 Continuous Triangular Model

In addition to discrete time intervals, the TM can be extended to represent continuous temporal data. Given a time interval I, all intervals *during* I are enclosed in a triangular zone below it (Figure 6-2). In other words, every interval I_n *during* I corresponds to a specific point in this triangular zone. Let us consider a linear dataset arranged within I. Every point in the triangular zone represents an interval I_n of the linear data. If every point is assigned a certain value of the interval it represents, i.e. $f(I_n)$, then the triangular area can be filled and become a continuous field. $f(I_n)$ is a certain formula dependent on I_n, such as the average, summation or standard deviation of the linear data in I_n. Through colour-coding, this continuous field can be displayed as an image, in which every point represents a specific interval I_n, and the colour at the point indicates the value of the linear data within this interval, i.e., $f(I_n)$.

Figure 6-3 demonstrates how the time series of the daily traffic jam lengths in Belgium is represented in CTM. The traffic jam lengths are observed in the Belgian motorway network at a frequency of every five minutes. The daily traffic jam lengths are averaged from all five-minute lengths during a day. Figure 6-3(a) is the line diagram of this time series, which exhibits apparent weekly fluctuations. However, variations in larger scales (e.g., monthly and quarterly) are hard to observe. Figure 6-3(b) shows the CTM representation of this time series, in which $f(I_n)$ is the average of the time series in I_n. In CTM, short-term fluctuations can be observed in the lower levels, while the long-term trend can be observed in the higher levels. Moreover, it explicitly displays a hierarchy of the time series in all different scales, in which one can observe the relationship between the short-term variations and long-term variations. In this CTM diagram, there is a major period of high traffic jams (busy period) between early April and middle July and a major period of low traffic jams (free period) between middle July and late September. Inside the busy period, the traffic jam length is not stable, and some weekly fluctuations can be observed at lower levels. Furthermore, there is a less obvious busy period between September and November, formed by two shorter busy periods within it (September to October, the end of November to the beginning of December). Compared to the Growth Matrix of Keim (2006), in which I^- and I^+ are, respectively, coordinated along the vertical and horizontal axes, the coordinate space of the CTM preserves the linear nature of time that flows from left to right. Mapping longer intervals in higher positions is also somehow more intuitive than the Growth Matrix.

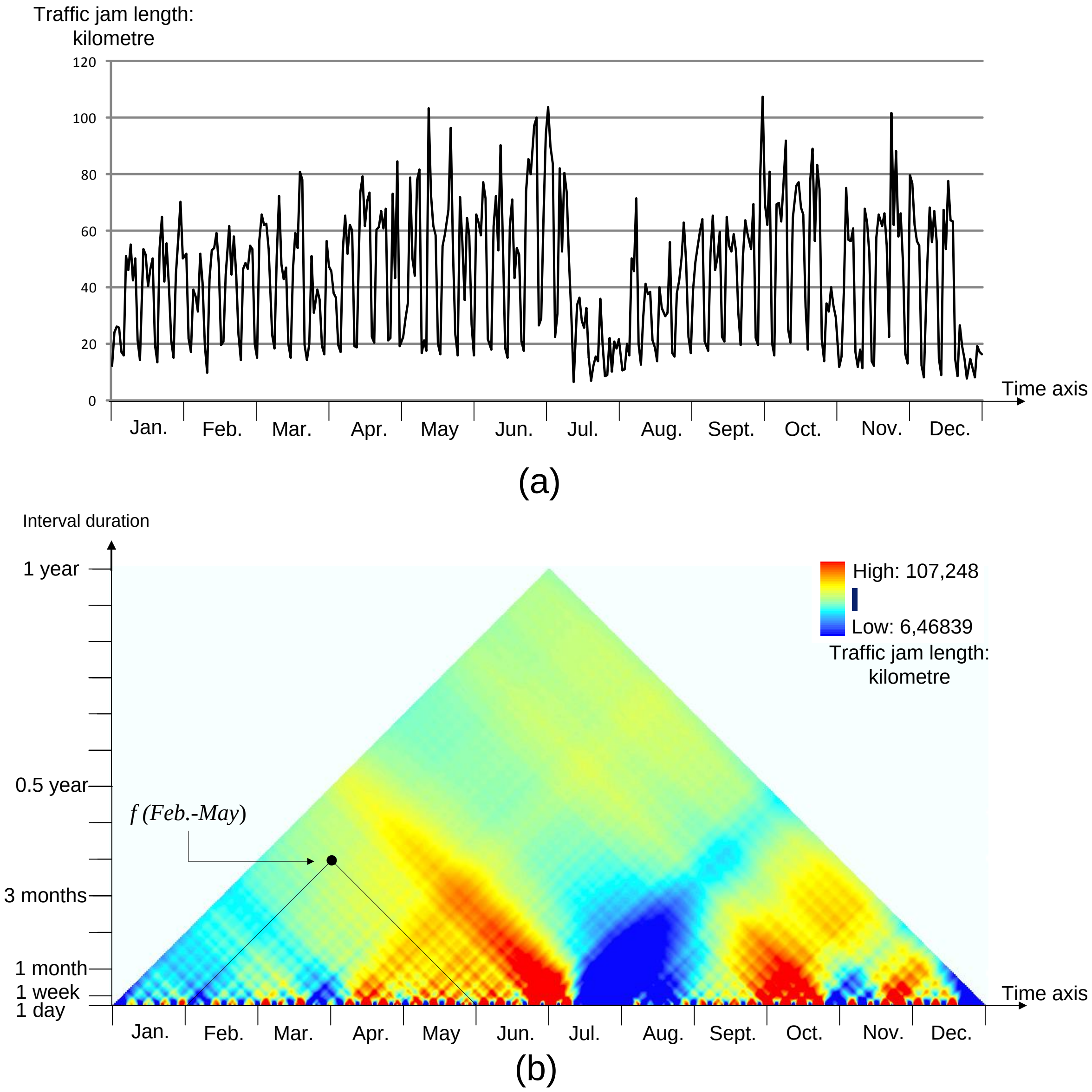


Figure 6-3: The representation of time series in a line diagram (a) and the CTM (b).

6.4 Traffic Analysis with CTM

CTM can also be applied to other types of linear data, including the traffic speed along a road, which is detected at a sequence of minimum road segments. In this case, every point in the CTM represents a specific road segment, and the colour at the point indicates the average speed of traffic in this segment. One CTM diagram can only represent the traffic in one direction, which is from left to right. In Figure 6-4, the four diagrams represent the traffic speeds along the E40 motorway in Belgium from Merelbeke (near Ghent) to Boerderijstraat (in Brussels) at four different timestamps. The ticks on the horizontal axis indicate the exits and entrances along the road. As the exit and entrance for one place is always less than 1 kilometre, a single tick is used to mark both the exit and entrance to the same place. Lines that are in α and $-\alpha$ to the horizontal axis are drawn from these ticks. The average speed from one place to the other can be read from the colour at the intersection point of two lines projected from the two places. In traditional representations, such as line charts and colour-coded polylines on maps, one can only read the traffic speed in road segments that are partitioned in a certain scale. The average speed across several partitions is obtained through mental estimation, which is not precise. In contrast, the CTM diagram provides an explicit overview of the average speed in all different segments of the road. One can observe the location of traffic jams along the road (i.e. low speed road segments) in the bottom of the triangular field, and also how much the traffic jams influence the average speed of longer distance from the higher levels. In Figure 6-4, one can observe that the average speed from Erpe-Mere to Ternat at 7:20 has fallen below 20 km/hour (the 24 hour clock is applied in this chapter), as the intersection point of the lines from these two places is in a very dark area. Thus, it is strongly not recommended that drivers take the motorway during this segment. In the higher level, one can identify that the average speed from Ghent to Brussels is approximately 60 km/hour. Thus, taking this motorway to travel from Ghent to Brussels is still feasible at this moment, as the secondary road nearby is limited to 50 km/hour. At 7:40 and 8:00, although some short segments with low traffic speed can be observed (e.g., at 8:00 from Erpe-Mere to Aalst and from Affligem to Groot-Bijgaarden), the average speed of most medium-distance and long-distance segments can reach 50 km/hour. At 8:20, the traffic speed from Aalst to Ternat is below 30 km.

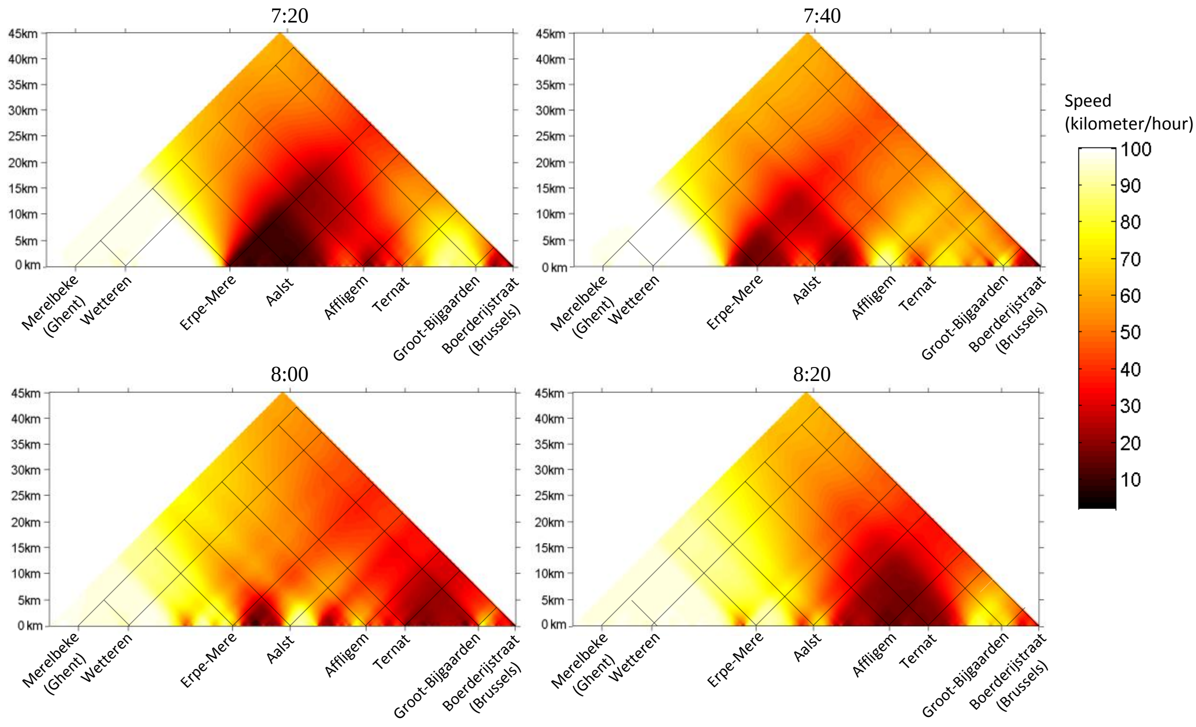


Figure 6-4: CTM diagrams of average speed on the E40 motorway from Ghent to Brussels.

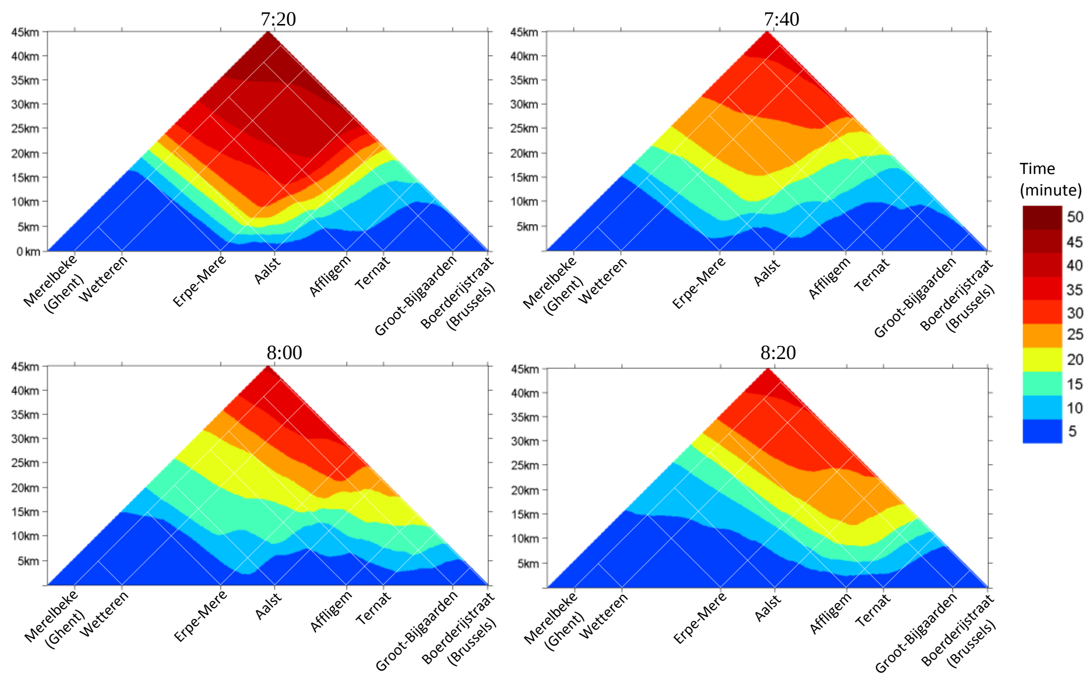


Figure 6-5: CTM diagrams of travel time on the E40 motorway from Ghent to Brussels.

In addition, the time that drivers need to spend on the road can also be represented in CTM. The linear data of travel time on the road can be calculated by dividing the length of the minimum road segments by the speed detected in these segments. The travel time of any road segment is the summation of the travel time of all road segments within it. The CTM diagrams in Figure 6-4 display the travel time of the same motorway at the same time stamps as that in Figure 6-5. The colour at every point indicates the time that travellers need to spend to travel from one place to another, according to the speed along the road detected at the time stamp. A discrete colour scale with 5-minute increments is applied for the ease of visual identification. From these diagrams, one can observe the time that one needs to travel between any two places. At 7:20, it takes 10 minutes to travel from Merelbeke (Ghent) to Erpe-Mere. However, for a similar distance, it takes 30 minutes to travel from Erpe-Mere to Affligem. Such a difference is less obvious at 7:40, which implies that the traffic jam reduced to some degree and is distributed more evenly. At 8:00 and 8:20, it takes more time to travel through the second half than the first half of the motorway, as the red area sinks in the right part of the triangle.

6.5 Analysing Multiple Time Series

CTM is based on a 2D coordinate interval spaces. Many spatial analysis techniques in GIScience can be employed to analyse CTM diagrams. This section demonstrates how the methods of map algebra and cartographic modelling are used to analyse multiple time series represented in CTM.

6.5.1 Map Algebra

Figure 6-3 displays the time series of daily traffic jam lengths in Belgium. The daily traffic jam lengths are averaged from observations taken every five minutes during a day, which loses the variations of traffic jams within a day. Instead of averaging all observations into the daily length, we combine the observations at the same time of different days into a time series. For instance, the observations at 6:00 over the year form a time series with 365 values. Applying to the observations at every five minutes, we obtain 288 time series, each of which contains the traffic jam lengths at a specific time of all days in the year. These time series can be transferred to 288 CTM diagrams, with the average formula applied to calculate the values of subintervals. An identical point in these diagrams represents an identical interval of the year. These CTM diagrams are then composed into one using map algebra: at every point, the time with the maximum traffic jam length is selected. Figure 6-6 illustrates an example of this algebra applied to three diagrams. In the composed CTM diagram, the value of every point is the time at which the average traffic jam length is the maximum during the interval represented by the point. Figure 6-7 shows the CTM diagram composed through map algebra. For ease of visual observation, a discrete colour scale is applied with the increment of one hour. The range of the colour scale is limited to display times between 6:00 to 22:00, as the maximum traffic jam length never appeared at the other times. This diagram shows that, although there are differences from day to day at the bottom of the diagram, the hour between 7:00 and 8:00 in the morning is peak hour for traffic jams in Belgium over the whole year. In the summer (i.e. from early June to mid-September), the peak hour changes to the afternoon, between 17:00 and 18:00. These findings can provide advice for police dispatch strategies and traffic management policies.

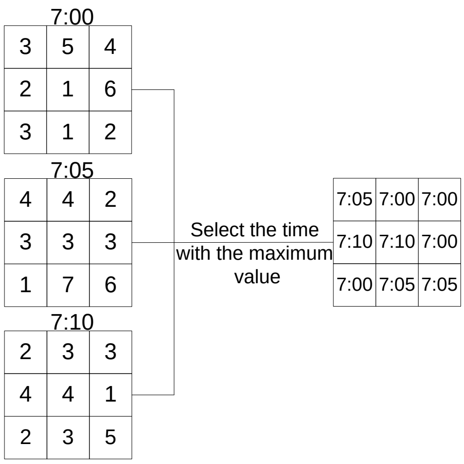


Figure 6-6: Map algebra of selecting the times with the maximum traffic jams.

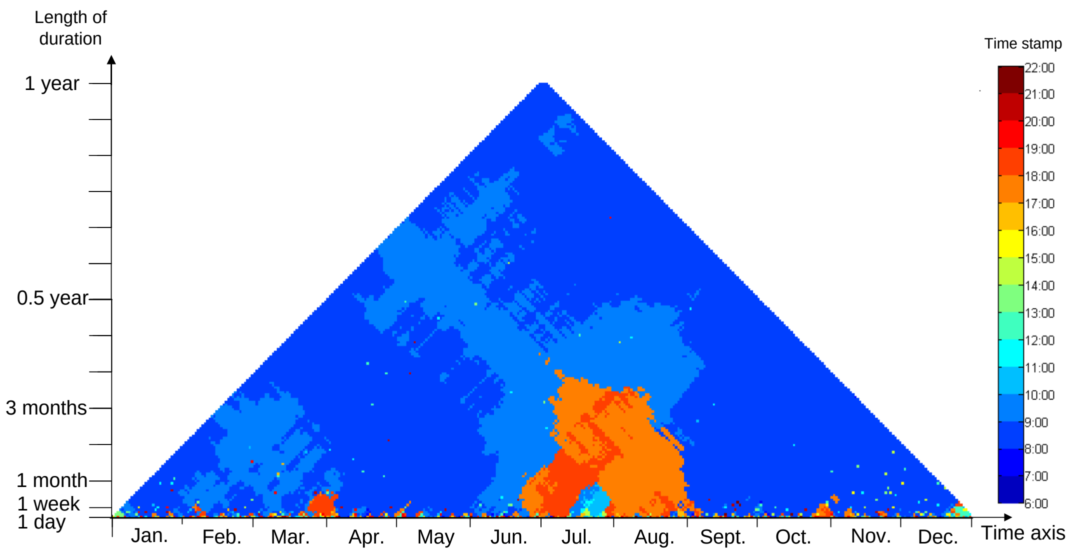


Figure 6-7: CTM diagram of the times with the maximum traffic jams, composed from CTM diagrams of 288 different time stamps.

6.5.2 Cartographic Modelling

Cartographic modelling addresses complex geographical problems by decomposing the problem into component criteria or constraints, which are usually modelled in different forms of geospatial datasets (e.g., raster, vector or TIN). Through a logical sequence of operations on these geospatial datasets, the final result (normally a map) is generated, indicating the solution for the problem. One typical application of cartographic modelling is site selection (e.g., selecting the site for a vineyard or windfarm), which takes account of many geographical criteria and constraints (e.g., climate, soil, topography and history) (Collins et al., 2001; Smith, 2002; Walsh et al., 1990). Through a series of operations, a suitability map is produced that indicates the suitable areas. This approach can also be referred to as multi-criteria decision making analysis (Jankowski, 1995; Jankowski et al., 2001). Using the same methods, cartographic modelling can also apply to CTM diagrams to solve multi-criteria decision making problems based on time series. Analogous to geographical site selection, cartographic modelling on the the CTM can help one to select time intervals that satisfy different criteria and constraints. In CTM, these suitable intervals are represented as areas in the interval space. Next, we use a concrete scenario to explain how cartographic modelling can be applied to CTM diagrams.

Suppose several professional surfers want to select a training site for the next year. There are four candidate surfing sites, including South Africa, Hawaii, Fiji and Australia. These four sites have different weather conditions throughout the year. Every site may be the best option during some specific periods of the year. Considering the surfers' requirements and preferences in Table 6-1, they need to select one of these four sites and also decide when and for how long they will arrange their training there. This question involves two sub-questions: which site and for which period? Given the annual weather statistics, including seasonal wave situations and sea temperature of all sites, this question is not easy to answer with traditional representations. However, by means of cartographic modelling, the CTM can give an explicit answer.

Table 6-1: Preferences and requirements of surfers and available statistics about the surfing sites

	Statistics of the site	Requirements of surfers	Preference of surfers
Ridable wave	Percentage of days that have ridable waves of every month.	At least 60 days having ridable wave.	The more, the better.
Ground wave	Percentage of days that have ground waves of every month.	N/A	The more, the better.
Sea temperature	Average sea temperature of every month.	N/A	The higher, the better.
Avoidance	N/A	Avoiding the international tournament in August.	

Generally speaking, there are two steps in the problem-solving procedure. In the first step, the suitability map of each surfing site is created, taking account of constraints and criteria, in this case, surfers' requirements and preferences (Table 6-1). In the second step, suitability maps of the four sites are combined into one summary map to answer which site is the best for which period.

Figure 6-8 illustrates the specific procedure of the first step. The monthly weather conditions are rated according to the three criteria, i.e., percentage of days with ridable waves[1], percentage of days with ground waves[2] and sea temperature. The ridable wave is the minimum condition on which the surfing training can be performed. The ground wave is more attractive for surfers because high-level skills can be trained. Furthermore, warmer sea temperatures increase surfing comfort. Different weights are given to these criterias according to their importance (i.e., ridable wave: 2, ground wave: 5, sea temperature: 2). Combining these weighted criteria produces a time series of suitability rates according to the general weather condition. The time series of suitability rates is represented by a CTM diagram with the average formula applied (CTM1). The number of days with ridable waves is represented by a CTM diagram with the summation formula applied (CTM2). In this case, all statistic data are recorded in a monthly scale. Therefore, the values of intervals across monthly partitions are obtained by interpolation. Considering the requirement of at least 60 days of ridable

[1] Ridable wave: waves last for 7 seconds period or more. (http://magicseaweed.com/)
[2] Ground wave: waves last for 10 seconds period or more and over 3ft. (http://magicseaweed.com/)

waves (Table 6-1), CTM2 is reclassified into a binary diagram (CTM3), where values above 60 are set to one and the remaining are set to zero. Multiplying CTM1 by CTM3, we obtain the suitability diagram (CTM4) of a surfing site, excluding intervals that do not have 60 days with ridable waves. Following the same procedure, the suitability diagrams of the other sites can be generated.

Figure 6-9 illustrates the specific procedure of the second step. The suitability diagrams (CTM4s) of all sites are combined together using a series of operations. At every position, the site with the highest value is selected, resulting in a nominal diagram with four zones. Each zone contains the intervals during which a certain site is the best in the four candidate sites. In other words, each candidate site has a set of intervals during which it is the best, which are in the corresponding zone in the CTM diagram. Due to the international tournament in August, only intervals before and after August can be used for training, while all other intervals have to be excluded. Referring to the relational zones of the TM in Section 6.2, relational zones that ‘touch’ August have been erased. Only intervals in the *before* and *after* zones are suitable. After all operations, the final output is produced, i.e., the rightmost CTM diagram in Figure 6-9.

The final output is enlarged in Figure 6-10. From this diagram, surfers may have an idea of which site is the best candidate during which intervals. Instead of providing a fixed choice, this diagram presents all possible intervals through the entire year. This diagram is flexible enough for surfers to make rough plans and remains open for modifications caused by other constraints. In this case, the output is a nominal map with four distinctive zones. Moreover, these four zones can become continuous fields to quantify to what extent a site is better than the others.

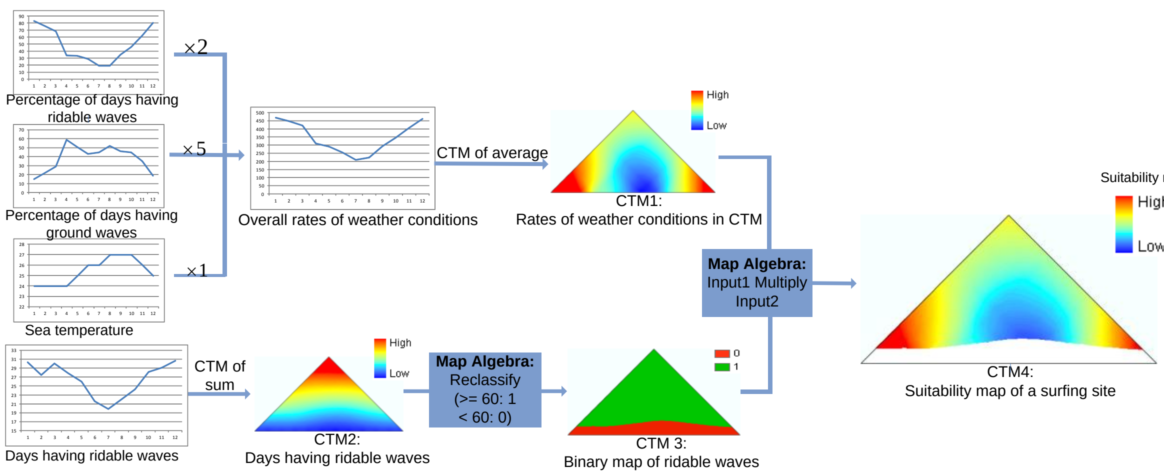


Figure 6-8: The procedure of generating the suitability CTM diagram for a site.

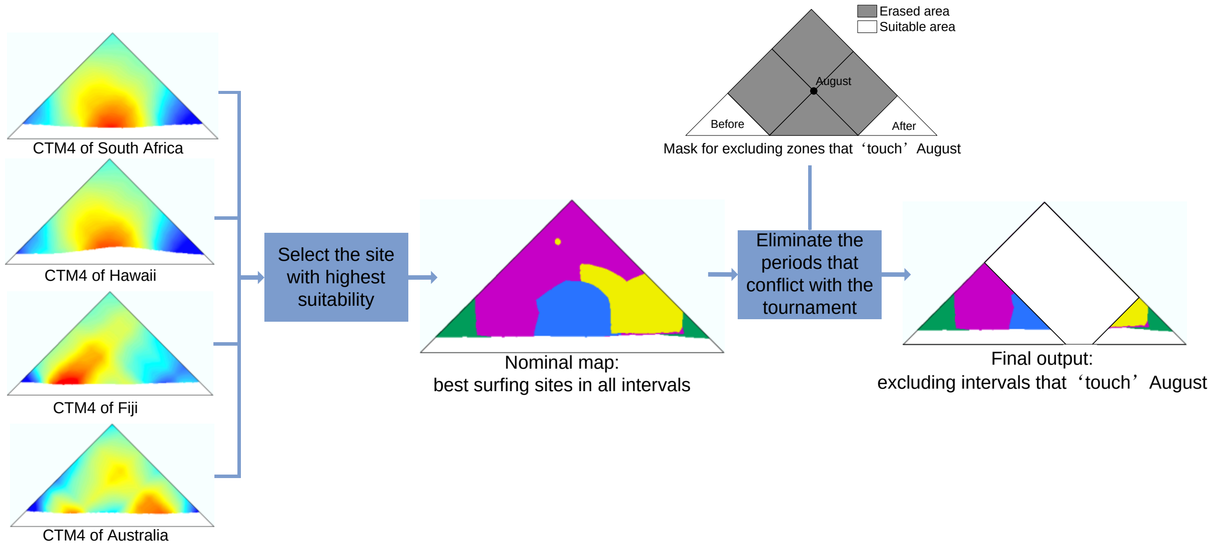


Figure 6-9: The composition of the suitability CTM diagrams of all sites.

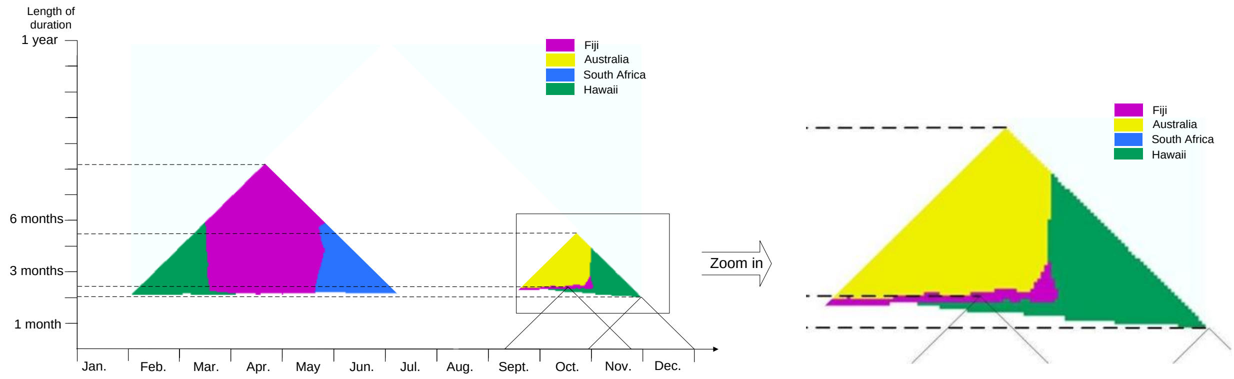


Figure 6-10: The enlarged final CTM diagram that visually answers surfers' questions (Left: the overall diagram. Right: zoom into the right part of the overall diagram).

According to the coordinate system of the CTM (described in Section 6.3), these zones can be easily interpreted. At the beginning of the year, Hawaii is the best site for surfing, while Fiji and South Africa are the best in the spring and summer, respectively, before the tournament in August. In autumn and early winter (the right part of Figure 16), the situation is more complex. There is a thin purple slice extending from September to November, which means that between September and November, Fiji is the best surfing site for approximately 2.5 months. If the period is longer or shorter than 2.5 months, either Australia (Yellow) or Hawaii (Green) is the best choice. Furthermore, if surfers would like to stay for the shortest period that guarantees 60 days with ridable waves, Hawaii is the very place, because the lowest position is in the green zone, at the right corner of the suitable area. In November and December, surfers only need to stay in Hawaii for just over 2 months to get 60 days with ridable waves. Moreover, if surfers would like to stay at one site as long as possible, Fiji is the best choice, because the highest point is in yellow. This means that Fiji has the best average weather condition during the whole period from the beginning of the year to the end of July. After August, Australia is the best site for a long stay, because it has the highest rate in the whole period between August and the end of the year. The traditional representation such as line charts or tables, can only give suggestions in a certain scale. In contrast, the CTM diagram offers multi-scale suggestions to surfers for choosing surfing sites and suitable periods. In this case, only several criteria and constraints are modelled by CTM. When it is applied to a more comprehensive analysis of surfing site, other time-related criteria can be added. Also, the weights of criteria can be adjusted for specific requirements.

6.6 Conclusion and Future Work

This chapter introduced an innovative representation of linear data, namely CTM. In CTM, the linear data in different intervals are uniformly displayed in a two-dimensional space, constituting a basis for a multi-scale analysis of these data. General speaking, the CTM has three major advantages. First, it provides an explicit and compact visualisation of linear data in multiple scales. Moving statistics (e.g., the average and summation) in intervals of different lengths can be explicitly displayed in one diagram. Both short-interval and long-interval values can be read from one diagram, and the relationships between these short-interval and long-interval variations can be directly observed. Second, the CTM is based on a coordinate space of linear intervals. Therefore, multiple linear data can be combined using the similar means of map algebra, which allows multi-scale analysis and comparison among linear data. Temporal and attribute constraints can also be modelled as areas. These areas are understood as sets of intervals that are in a certain temporal relation or attribute range, which can be used as masks to screen the intervals that satisfy the constraints. Using map algebra on CTM diagrams, one can obtain all intervals of different lengths that meet their requirements, which is difficult in the traditional representations, where analysis is performed on intervals of the same length. Third, due to similarities between CTM diagrams and geospatial maps, there is a vast toolbox of existing, high-level techniques in GIScience for manipulating and analysing CTM diagrams, such as map algebra and cartographic modelling, at its disposal. These techniques are abundantly integrated in current geographical information systems, which therefore provide an efficient platform for storing, managing, visualising and analysing CTM diagrams.

This chapter gave several examples to show possible applications of CTM. In future work, there are a lot of possible scenarios in which CTM can be applied. For instance, the CTM can represent the time series of the performance and physical conditions of a soccer player in all different time scales. Then time series of multiple players in the same soccer match can be compared and analysed using map algebra. The visualization in the CTM may support the coach to make strategies or adjustment accordingly in all different time intervals. The idea of cartographic modelling in the CTM can be also applied to agriculture and farming, which rely heavily on the analysis of time series of weather statistics. The analysis based on the CTM can potentially provide advice for the irrigation strategies or fish egg incubation control. Moreover, implementing the CTM into an interactive system can improve the analytical usability of CTM. In such a system, the operation and manipulation of the CTM diagrams can be performed more conveniently and automatically. Dynamic controls for colour ramps, the formulas in the CTM and area selection by attribute can facilitate visual exploration and analysis in CTM. The sensitivity of the cartographic modelling approach to different parameters and utility functions can also be better analysed within this system. Integrating the CTM to other information systems can also be interesting. For example, linking the CTM with a visualization based on the space-time cube (Gatalsky et al., 2004; Hägerstrand, 1970; Kraak, 2003) can support the analysis of space-time trajectories in all different time intervals. From a general perspective, the CTM can be considered as a conceptual model of generalisation. Many types of data can be described at different granularities. In a digital map, the border of a country can be displayed in different granularities, or a system can be decomposed into different levels. These types of data can also be plotted into a CTM diagram in the same manner as linear data, which could potentially benefit the analysis. These ideas will be further investigated in future work.

References

Aigner, W., Miksch, S., Muller, W., Schumann, H. and Tominski, C. (2007). Visualizing time-oriented data - A systematic view. Computers & Graphics-UK, 31(3), pp.401-409.

Collins, M. G., Steiner, F. R. and Rushman, M. J. (2001). Land-Use Suitability Analysis in the United States: Historical Development and Promising Technological Achievements. Environmental Management, 28(5), pp.611-621.

Enders, W. (2008). Applied Econometric Time Series (Second Edition). John Wiley & Sons, Inc, Hoboken, New Jersey.

Gatalsky, P., Andrienko, N. and Andrienko, G. (2004). Interactive analysis of event data using space-time cube. Proceedings of the 8th International Conference on Information Visualisation (IV 2004), London, UK, pp.145-152.

Goodchild, M. F., Yuan, M. and Cova, T. J. (2007). Towards a general theory of geographic representation in GIS. International Journal of Geographical Information Science, 21(3), pp.239-260.

Hägerstrand, T. (1970). What about people in Regional Science? Papers in Regional Science, 24(1), pp.6-21.

Hamilton, J. D. (1994). Time Series Analysis. Princeton University Press.

Havre, S., Hetzler, B. and Nowell, L. (2000). ThemeRiver: Visualizing theme changes over time. Proceedings of the IEEE Symposium on Information Vizualization 2000, Salt Lake City, Utah, pp.115-123, IEEE Computer Society.

Hochheiser, H. and Shneiderman, B. (2004). Dynamic query tools for time series data sets: timebox widgets for interactive exploration. Information Visualization, 3(1), pp.1-18.

Jankowski, P. (1995). Integrating geographical information systems and multiple criteria decision-making methods. International Journal of Geographical Information Systems, 9(3), pp.251-273.

Jankowski, P., Andrienko, N. and Andrienko, G. (2001). Map-centred exploratory approach to multiple criteria spatial decision making. International Journal of Geographical Information Science, 15(2), pp.101-127.

Keim, D. A., Nietzschmann, T., Schelwies, N., Schneidewind, J., Schreck, T. and Ziegler, H. (2006). A Spectral Visualization System for Analyzing Financial Time Series Data. ed. by Ertl, T., Joy, K. and Santos, B., Proceedings of the Eurographics/IEEE-VGTC Symposium on Visualization (EuroVis 2006), Lisbon, Portugal, pp.195-200.

Kraak, M.-J. (2003). The space-time cube revisited from a geovisualization perspective. Proceedings of the 21st International Cartographic Conference, Durban, South Africa, pp.1988-1996.

Kulpa, Z. (1997). Diagrammatic Representation for a Space of Intervals. Machine Graphics & Vision, 6, pp.5-24.

Kulpa, Z. (2006). A Diagrammatic Approach to Investigate Interval Relations. Journal of Visual Languages and Computing, 17(5), pp.466-502.

Liao, T. W. (2005). Clustering of time series data: a survey. Pattern Recognition, 38(11), pp.1857-1874.

Lin, J., Keogh, E., Lonardi, S., Lankford, J. P. and Nystrom, D. M. (2004). Visually mining and monitoring massive time series. Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining, Seattle, WA, USA, pp.460-469, ACM.

Muller, W. and Schumann, H. (2003). Visualization methods for time-dependent data - An overview. Proceedings of the 2003 Winter Simulation Conference, New Orleans, U.S.A., pp.737-745.

Qiang, Y., Delafontaine, M., Asmussen, K., Stichelbaut, B., De Tré, G., De Maeyer, P. and Van de Weghe, N. (2010). Modelling Imperfect Time Intervals in a Two-Dimensional Space. Control and Cybernetics, 39(4), pp.983-1010.

Qiang, Y., Delafontaine, M., Neutens, T., Stichelbaut, B., De Tré, G., De Maeyer, P. and Van de Weghe, N. (2012a). Analysing Imperfect Temporal Information in GIS using the Triangular Model. The Cartographic Journal, (in press).

Qiang, Y., Delafontaine, M., Versichele, M., De Maeyer, P. and Van de Weghe, N. (2012b). Interactive Analysis of Time Interval in a Two-Dimensional Space. Information Visualization, (in press).

Smith, L. (2002). Site Selection for Establishment & Management of Vineyards. Paper presented in The 14th Annual Colloquium of the Spatial Information Research Centre University of Otago Dunedin, New Zealand.

Smith, M. J. d., Goodchild, M. F. and Longley, P. A. (2007). Geospatial Analysis: a Comprehensive Guide to Principles, Techniques and Software Tools. Troubador Publishing Limited, Leicester, UK.

Tomlin, C. D. (1990). Geographic information systems and cartographic modeling. Prentice Hall, New Jersey.

Van de Weghe, N., Docter, R., De Maeyer, P., Bechtold, B. and Ryckbosch, K. (2007). The Triangular Model as an Instrument for Visualising and Analysing Residuality. Journal of Archaeological Science, 34(4), pp.649-655.

Walsh, S. J., Butler, D. R., Brown, D. G. and Bian, L. (1990). Cartographic Modelling of Snow Avalanche Path Location within Glacier-National-Park, Montana. Photogrammetric Engineering and Remote Sensing, 56(5), pp.615-621.

Weber, M., Alexa, M. and Muller, W. (2001). Visualizing time-series on spirals. Proceedings of IEEE Symposium on Information Visualization 2001, Washington, DC, U.S.A., pp.7-13.

7 EMPIRICAL STUDY

Modified from: Yi Qiang, Martin Valcke, Nico Van de Weghe, Representing Time Intervals in a Two-Dimensional Space: an Empirical Exploratory Study. Submitted to *Spatial Cognition and Computation*.

Abstract: Time intervals are conventionally represented as linear segments along the one-dimensional time line, which is labelled as the Linear Model in this chapter. However, this representation has some inherent difficulties in the visualization and analysis of time intervals. To overcome these difficulties, an alternative representation, i.e. the Triangular Model, can be considered, which represents time intervals as points in a two-dimensional space. Previous research has exhibited the potential of the Triangular Model in reasoning and visualization of time intervals. Till present, there remains a lack of empirical evaluation of user comprehension and usability of this representation. This chapter describes an empirical study that aims to evaluate whether people can easily learn the Triangular Model and whether they are able to use it for visual observation and judgments of time intervals. The result shows that people without prior knowledge about this representation quickly learned this model after a short training process and answered questions about time intervals more correctly and efficiently questions with the Triangular Model as compared to the Linear Model. The results obtained from this empirical study underpin the potential of the Triangular Model in visualizing time intervals and give support to its usability by non-expert users.

Keywords: time intervals, information visualization, the Triangular Model, empirical study

7.1 Introduction

A wide range of disciplines are confronted with the problem of handling information related to time, including information science (Hochheiser and Shneiderman, 2004), archaeology (Stichelbaut and Bourgeois, 2009) and geography (Neutens et al., 2007). While time can be conceptualized and represented in diverse ways, the linear concept is predominant, which is reflected in many graphical representations such as time tables, chronological time lines and even the recent application in Facebook® (i.e. the Facebook timeline). A segment of the time line is called a time interval, which is usually considered as the primitive of time. Up to now, considerable work has been done in handling time intervals in the areas of computer science and artificial intelligence (Bhatt et al., 2011; De Tré et al., 2006; Gottfried, 2008; Knauff, 1999). The most well-known work is the qualitative interval algebra introduced by James F. Allen (1983) and the extension by Freksa (1992). Much seminal work about temporal reasoning is based on their theories. On the other hand, the research on visualization and analysis of time intervals receives far less attention. The visual representation of time intervals remains limited to linear segments along a one-dimensional (1D) time line, which is labelled as the Linear Model (LM) in this chapter. Alternative representations of time intervals are available, e.g. the cyclic representations (Li and Kraak, 2008; Weber et al., 2001) and calendars (Weaver et al., 2006). They rather focus on the representation of specific aspects of time-dependent data, and are therefore not applicable in a broader range of contexts. In the LM, the second dimension is exploited solely to differentiate between the consecutive intervals of different events and thus has no metric temporal meaning. Therefore, the arrangement of linear segments can vary, depending on the sorting rules applied in the second dimension. This polymorphism prohibits the existence of a universal approach for visual analysis of time intervals. As a result, the linear time representation is most used for illustration, but rarely applied in analytical tasks of time intervals, especially exploratory data analysis, which greatly relies on data visualization.

To overcome these difficulties, a two-dimensional (2D) representation of time intervals has been considered. This representation maps a time interval to a unique point in a 2D space. This 2D representation of time interval was initially proposed by Ligozat (1994, 1997). Later, Kulpa (1997a; 1997b; 2006) have comprehensively elaborated its use in qualitative interval reasoning and interval arithmetic. Van de Weghe et al. (2007) labelled this representation the Triangular Model (TM) and applied it to an archaeological use case. More recently, Qiang et al. have investigated the use of the TM in reasoning about imperfect intervals (Qiang et al., 2010) and interval analysis (Qiang et al., 2012a; Qiang et al., 2012b). Since the TM displays a set of intervals within a fixed point structure, it offers special insights into interval distributions, particularly when a large amount of intervals are represented in it. Moreover, the TM supports a graphical query mechanism that relies on the manipulation of geometrical relationships in the 2D space of intervals. Available research stressed the potential of the TM in visualising and analysing time intervals. However, there is a lack of empirical evidence to ground its understandability and usability. To fill this gap, we have initiated an empirical study to evaluate the TM, as compared to the conventional LM as a reference model. This chapter summarizes the design and results of this empirical study.

The remainder of this chapter is structured as follows. Section 7.2 introduces the basic concepts of time intervals and the two models. Section 7.3 first introduces the theoretical base of the learning process, and then describes the design of the test and research process. In Section 7.4, the results obtained from the experiment are presented and analysed. In Section 7.5, we discuss the interesting findings derived in the study. Section 7.6 summarises the contribution of this study and proposes avenues for future research.

7.2 The Two Time Models

7.2.1 Linear Model

A time interval is an extent of time, which can be the duration of an event or the lifetime of a person. In physics and computer science, a time interval is usually abstracted as a pair of real numbers $[I^-, I^+]$ with $I^- < I^+$. I^- is the start point of I, I^+ is the end point, and the difference between I^- and I^+ (i.e. $I^+ - I^-$) is the duration of the interval, which is denoted as $dur(I)$. The LM is derived from the experience and interpretation of time linearity, and intuitively represents a time interval as a linear segment along the time line. The two boundary points of the segment respectively indicate the start point and the end point of the interval (Figure 7-1(a)). The length of the segment expresses the duration of the interval. Since time intervals may overlap and representing multiple overlapping segments in the time line may cause difficulties for visual observation, the linear segments are often arranged at different positions in the second dimension (Figure 7-1 (b)). The arrangement can be decided by the properties of the intervals (e.g. the start point, end point, and the duration) or the properties of entities referenced to the intervals (e.g. the types of the events). Therefore, the structure of linear segments is variable according to the arrangement in the second dimension. Figure 7-2 illustrates four different arrangements of the linear segments along the second dimension. The characteristics of time intervals are expressed by the location and extent of the linear segments in the time line.

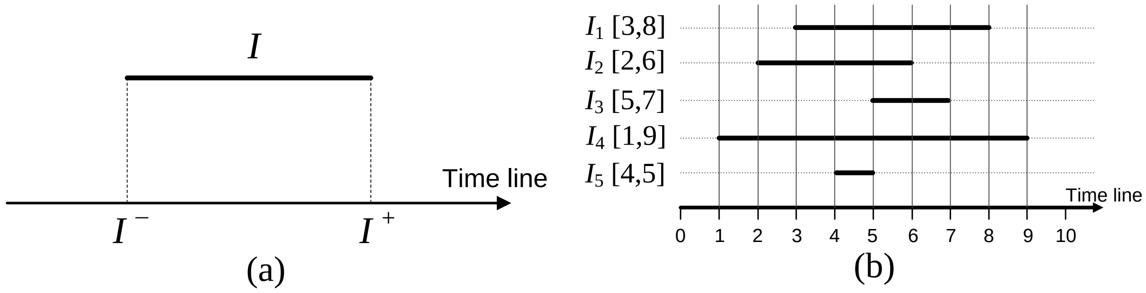


Figure 7-1: Representing time interval in the Linear Model

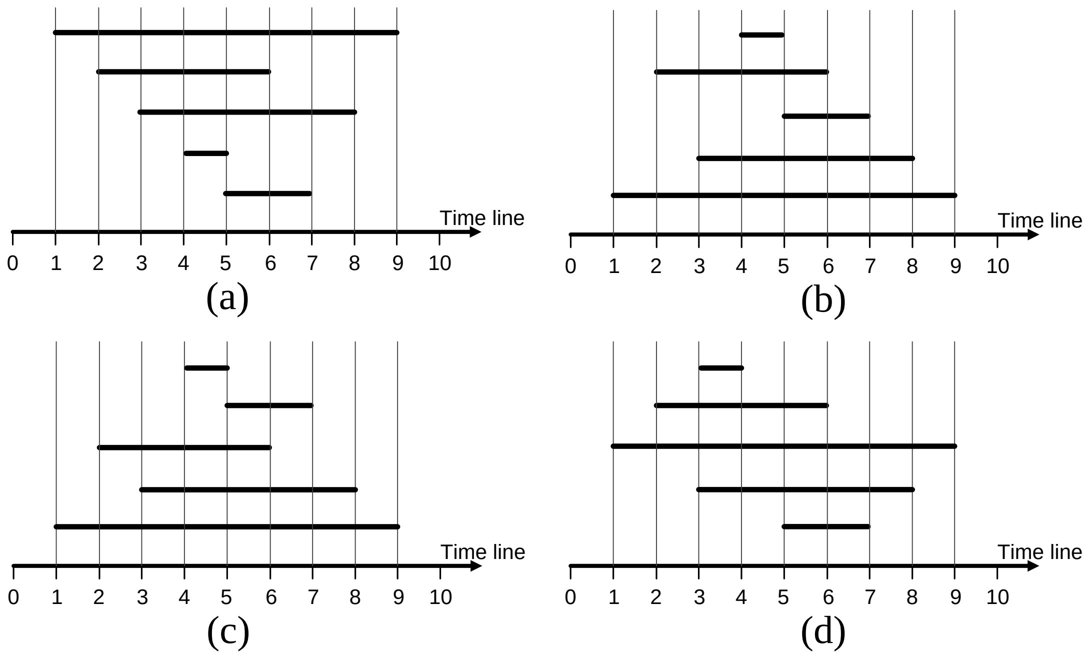
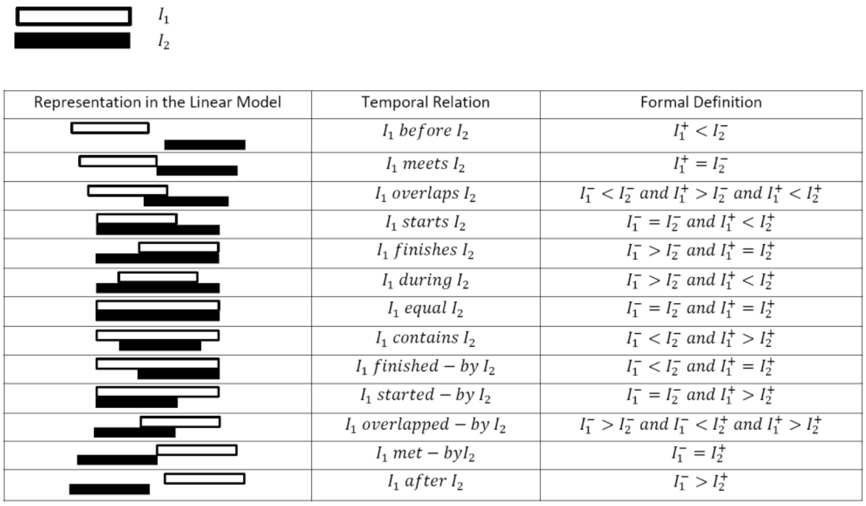


Figure 7-2: Four different arrangements of time intervals in the vertical dimension in the Linear Model. (a): ascending start point; (b): ascending end point; (c): ascending duration; (d): ascending midpoint.

Two time intervals may have different relations. In 1983, Allen specified thirteen relations between two time intervals (Allen, 1983), which have been considered as the cornerstone of many theories about temporal reasoning (Bittner, 2002; Freksa, 1992; Galton, 1990; Schockaert et al., 2008). These thirteen temporal relations are defined by the relations of the start points and end points between the two intervals. In the LM, these relations are expressed by the topological relations between two linear segments in a 1D space. Differing from the spatial topologies, e.g. the RCC calculus by Cohn et al. (1997), the topology of time intervals also take account of the direction of time. Therefore, there are six pairs of self-reflexive inverse relations, except the *equal* relation (Table 7-1). Figure 7-3 illustrates the formal definitions of the temporal relations and the corresponding representations in the LM.



Representation in the Linear Model	Temporal Relation	Formal Definition
	I_1 *before* I_2	$I_1^+ < I_2^-$
	I_1 *meets* I_2	$I_1^+ = I_2^-$
	I_1 *overlaps* I_2	$I_1^- < I_2^-$ and $I_1^+ > I_2^-$ and $I_1^+ < I_2^+$
	I_1 *starts* I_2	$I_1^- = I_2^-$ and $I_1^+ < I_2^+$
	I_1 *finishes* I_2	$I_1^- > I_2^-$ and $I_1^+ = I_2^+$
	I_1 *during* I_2	$I_1^- > I_2^-$ and $I_1^+ < I_2^+$
	I_1 *equal* I_2	$I_1^- = I_2^-$ and $I_1^+ = I_2^+$
	I_1 *contains* I_2	$I_1^- < I_2^-$ and $I_1^+ > I_2^+$
	I_1 *finished* − *by* I_2	$I_1^- < I_2^-$ and $I_1^+ = I_2^+$
	I_1 *started* − *by* I_2	$I_1^- = I_2^-$ and $I_1^+ > I_2^+$
	I_1 *overlapped* − *by* I_2	$I_1^- > I_2^-$ and $I_1^- < I_2^+$ and $I_1^+ > I_2^+$
	I_1 *met* − *by*I_2	$I_1^- = I_2^+$
	I_1 *after* I_2	$I_1^- > I_2^+$

Figure 7-3: The representations of temporal relations in the Linear Model.

Table 7-1: The pairs of inverse temporal relations. The abbreviations are put between brackets.

before (*b*)	↔	*after* (*a*)
meets (*m*)	↔	*met-by* (*mb*)
overlaps (*o*)	↔	*overlapped-by* (*ob*)
starts (*s*)	↔	*started-by* (*sb*)
finishes (*f*)	↔	*finished-by* (*fb*)
during (*d*)	↔	*contains* (*c*)
equal (*e*)		

Two relations between pairs of intervals are conceptual neighbours if they can directly transfer into one another by continuous deformation (Freksa, 1992). For example, *before* and *meets* are conceptual neighbours because extending the earlier interval towards the later interval may cause a direct transition from the *before* relation to *meets* relation. Such direct transition is not possible from the *before* relation to the *overlap* relation, because it must pass the *meets* relation. Based on this definition of conceptual neighbourhood, Freksa (1992) has mapped the thirteen temporal relations into a nested structure, called the conceptual neighbourhood structure (Figure 7-4). Furthermore, the thirteen temporal relations can be divided into three categories, i.e. open relations, semi-open relations and the closed relation. *After*, *before*, *overlaps*, *overlapped-by*, *contains* and *during* are open relations because the two intervals in these relations do not have common boundary points (i.e. the start point and the end point). *Meets*, *met-by*, *starts*, *started-by*, *finishes* and *finished-by* are semi-open relations because the two intervals in these relations have one common boundary point. The *equal* relation is the only closed relations as the two intervals have two common points. It is notable that all transitions from an open relation to another open relation must pass through a semi-open relation or the closed relation. Also, every pair of inverse relations is centrosymmetric about the *equal* relation.

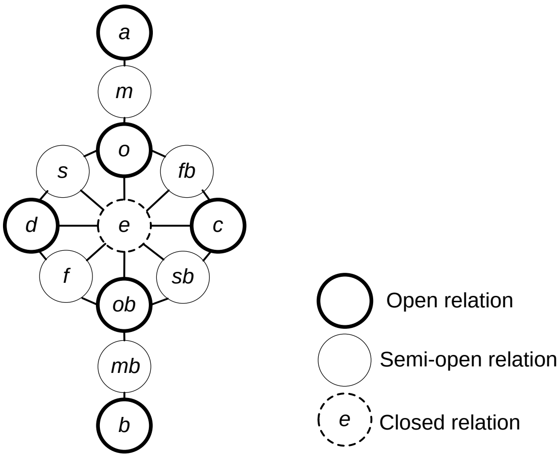


Figure 7-4: The conceptual neighbourhood structure of thirteen temporal relations.

7.2.2 Triangular Model

A time interval is defined by a pair of parameters, namely, the start point and the end point. Therefore, it is possible to map a time interval to a point in a 2D space, using these two parameters as the coordinate. Given a time interval I in the time line, two straight lines (L_1 and L_2) are projected from I^- and I^+ (Figure 7-5). The angle between L_1 and the time line is α_1, while The angle between L_2 and the time line is α_2 , where $\alpha_1 = -\alpha_2 = \alpha$. The angle α is a constant that is identical for all intervals. Therefore, the intersection point of L_1 and L_2 is completely decided by I^- and I^+. In other words, the time interval I can be represented by this point in the 2D space. This representation of time intervals is called the Triangular Model (TM). Because $\alpha_1 = -\alpha_2$, it is straightforward to deduce that the horizontal location of the point indicates the middle point of the interval, i.e. $mid(I)$. In the vertical dimension, the height (h) of the point is proportional to the length of the linear interval (l), i.e. $h = \frac{\tan\alpha}{2} \cdot l$. Thus, the height of an interval point in the TM indicates the duration of the interval. Using this approach, every time interval can be represented as a unique point in the 2D space, and the characteristics of a time interval are completely expressed by the location of the point. Note that α can be different values for specific purposes. In this chapter, we set α = 45°, to be consistent with earlier work (Kulpa, 1997a; Qiang et al., 2010; Van de Weghe et al., 2007). Considering this setting, the eight directions in the 2D space correspond to eight changing directions of interval properties (Figure 7-6).

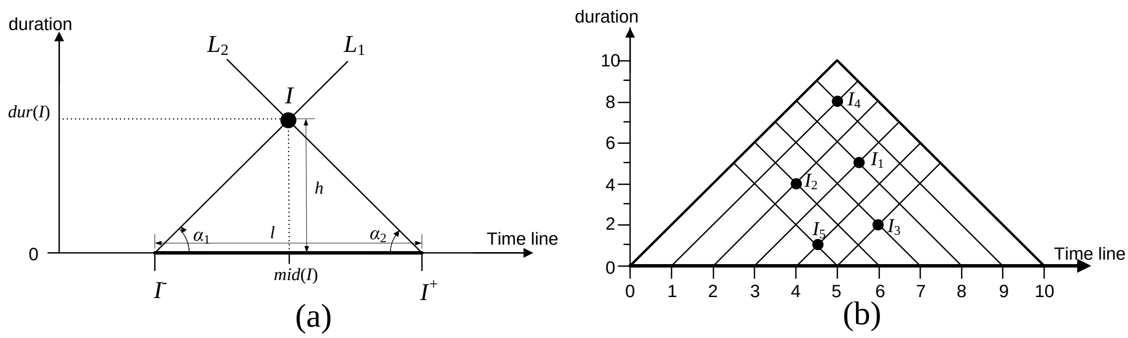


Figure 7-5: The Triangular Model. (a): The construction of an interval point. (b): The representation of the intervals in Figure 7-1(b) in the Triangular Model.

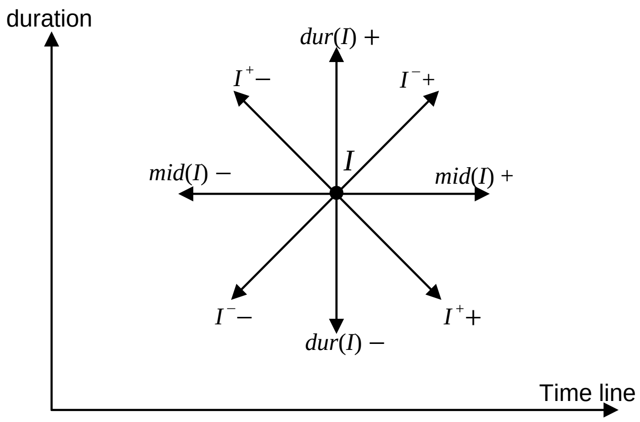


Figure 7-6: The meanings of the eight directions in the Triangular Model. The minus sign means a decrease, while the plus sign means an increase.

Since the TM represents time intervals as points in a 2D coordinate space, the relations between time intervals are expressed by the spatial relations between points. Let us consider a study interval that contains all points of examined intervals. Given a reference interval I_1 in the study area, all intervals in a certain temporal relation to I_1 are located in a specific zone. Figure 7-7 shows the zones of thirteen temporal relations with respect to I_1. For instance, all intervals that are *before* I_1 are located in the triangular zone in the left corner of the study area. In other words, each of these zones represents the set of intervals that are in a specific temporal relation to I_1. These zones are called relational zones. The boundaries of the relational zones are in α or $-\alpha$ angle to the horizontal axis. It is noteworthy that the zones of the open relations are 2D geometries (i.e. triangle or rectangle), whilst the zones of the semi-open relations are 1D geometries (i.e. line). The zone of the closed relation (i.e. the *equal* relation) is a point (0D). Moreover, the zones of a pair of inverse relations are centrosymmetric about the reference interval. In Figure 7-7, the reference interval I_1 is chosen in the centre of the study area in order to avoid visual bias. Of course the reference interval can be put in other locations. However, the topological configuration of its relational zones always remains the same (Figure 7-8). This way, the TM transfers a temporal topology to a spatial topology between points and zones. The intervals in a specific relation to the referenced interval can be found in a specific zone.

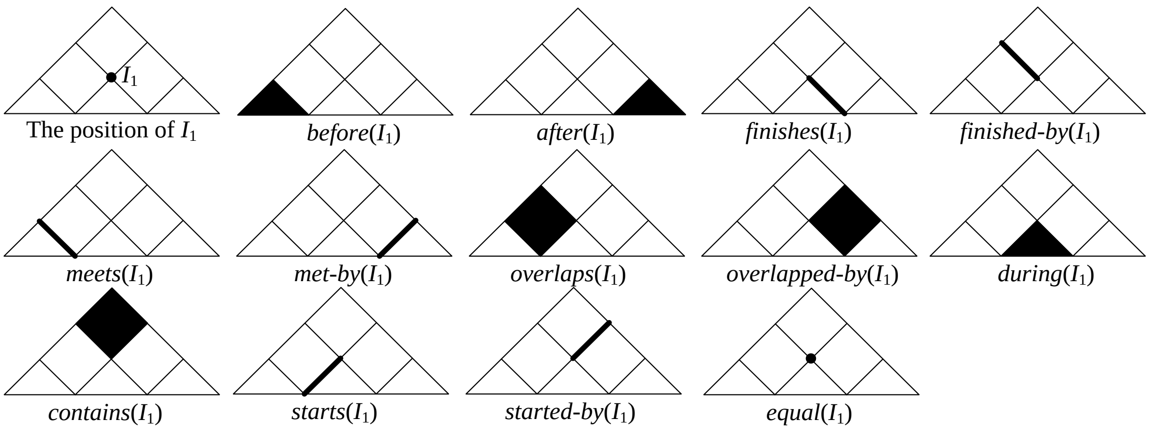


Figure 7-7: The zones of the temporal relations in the Triangular Model.

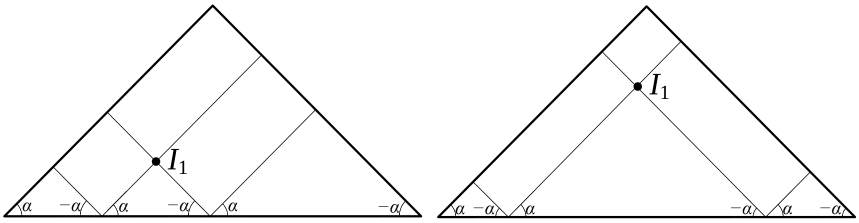


Figure 7-8: The consistent topological configuration of relational zones in the Triangular Model

7.3 Empirical Study

The potential and advantages of the TM in view of the analysis of time intervals has been elaborated in previous literatures (Qiang et al., 2010; 2012a; 2012b). Following the TM, every interval is represented by a point at a specific location. Therefore, a set of time intervals is represented as a fixed structure of points, facilitating the visual observation and analysis of interval distributions. This feature is particularly useful when comparing two sets of intervals. The pattern differences can be easily observed from the structural difference of the two point sets, which is not possible within the LM where the linear segments do not have an fixed location. Also, the TM transfers temporal topology to spatial topology between points and zones. The intervals that are in a specific relation to a reference interval can be found in a specific zone, which allows the visual query process in the interval visualization. In addition, the points in the TM are apparently more space-efficient than linear segments, making it capable of analytical tasks involving large datasets of intervals.

However, there is a lack of empirical evidence that the TM can be easily understood and manipulated by non-experts. To the best of our knowledge, till present, no empirical work has been done to evaluate any 2D representations of time intervals similar to the TM. To this end, we have conducted an empirical study that aims to assess the understandability and usability of the TM, as compared to the conventional LM as a reference. More specifically, this empirical study aims to address the following questions: whether 1) people can easily learn the TM; 2) people are able to correctly use it to answer question about time intervals; 3) the TM is more efficient to answer these questions; 4) people like to use it.

7.3.1 Learning the Two Models

Representations are central to the development and storage of conceptual knowledge. Established working memory models explicitly refer to the role of the visuo-spatial sketchpad to process information in view of storage in long term memory (Baddeley and Hitch, 1974; Miyake and Shah, 1999). As such, complex knowledge, such as manipulating time intervals, will be fostered when learners are presented with efficient representations. In addition, following cognitive load theory, presenting learners with well developed representations, will also help to reduce extraneous cognitive load that hinders the active processing and consecutive storage and retrieval of knowledge (Sweller, 1994; Sweller and Chandler, 1994). Lastly, the Cognitive Theory of Multimedia Learning (CTML) stresses the critical need to present learners with clear multimedia representations to develop the organization of complex information into mental models that can be linked to prior knowledge in long-term memory (Mayer and Moreno, 2003). CTML explicitly puts forward the potential of multimedia representations (such as a graphical representation of time intervals) to foster knowledge processing, resulting in better knowledge retrieval and application.

Putting the above theoretical frameworks in the context of learning about time representations, the question can be posed: to what extent there is a differential impact of learning with the support of the LM or the TM. Considering the fact that learners are acquainted with the LM, it can be assumed that they will perform in a better way when coping with time interval problems based on the LM representation. In contrast, the TM might invoke extraneous cognitive load due to the unknown nature of the representation that does not help to deal with the time interval problems. The actual situation greatly depends on the comprehension and usability of both time representation models. No empirical evidence is available to put forward conclusive guidelines. As such, the presented study aligns with the need to study alternative representations in complex knowledge domains. Also, the available theoretical base suggests a critical element to be considered prior to the comparison of the LM versus the TM. Namely, a learning phase has to be considered to introduce learners to the alternative time representation, i.e. the TM.

7.3.2 Research Instruments: Pretest and Posttest

A pretest-posttest intervention study was set up. The pretest and posttest aimed at measuring the comprehension and usage of the LM and the TM within a group of participants. In the tests, time intervals were represented in either of the two models.

Participants were requested to answer specific questions with respect to these time intervals. Every question was asked twice with the same number of intervals presented, but shifting the use of either the LM or the TM. When comparing participant comprehension and usage, the following corresponding parameters could be compared: correctness of the answer, time the participant spent on the question and the preference for a model by the participant. Parameters obtained from the two models were compared using appropriate statistical tests, building on the null hypothesis stating that the parameters from the two models are not significantly different.

Two categories of questions were included in the tests. The first type aimed to evaluate how well the participants could identify the properties of time intervals in the two models, i.e. the start point, the end point, the duration and the midpoint, which are considered as the elementary properties of time intervals in previous research about temporal reasoning (Freksa, 1992). In this type of questions, given a number of time intervals, the participants were requested to find intervals with the specific properties. The second category aimed at evaluating how well the participants could identify the temporal relation between time intervals with the two models. Participants were requested to find the intervals that satisfied certain relations to the reference intervals. Each question category consisted of two sub-types, which either asked about one item (i.e. property or relation) or two. Table 7-2 exemplifies the questions types. The asked items are highlighted with bold font in the example questions, which can be replaced by other items in other questions. In order to reduce the total number of questions, the questions including two temporal relations only include open relations introduced in Section 7.2.1.

Table 7-2: The examples of questions in the test

Item number / Type of question	One item	Two items
Interval property	Please find and select the intervals that have a start point between 20 and 40.	Please find and select the intervals that have a start point between 20 and 40 and an end point between 60 and 80.
Temporal relation	Please find and select the intervals before the reference interval I_1.	Please find and select the intervals that overlap I_1 and are during I_2.

Tests were presented in different ways at the time of the pretest and the posttest (Table 7-3). The pretest was presented as a paper-and-pencil test, containing a small number of questions with fewer intervals being presented in the two models. In the pretest, the participants were requested to answer the questions by circling the correct intervals. The posttest was presented via a website containing a set of questions (in total 37 questions) with a larger number of intervals. Due to the time limitation, every participant only needed to finish a random set of 12 questions out of the 37 questions. The 12 questions covered all types of questions illustrated in Table 7-2.

Table 7-3: Comparing the pretest and the posttest

	Pretest	Posttest
Media	Paper-based	Computer-based
Function	Training and test	Test
Number of questions	8	37
Numbers of intervals	Less (8 to 10)	More (50 to 200)
Linear segment in the Linear Model	Randomly sorted	Can be sorted by the participant
Preference of the model	Not asked	Asked
Tested parameters	1. Score of the answers	1. Score of the answers 2. Time used 3. Preference of the model

In the computer-based posttest, participants were able to select the correct intervals by clicking or dragging a selection box in the diagram. When the time intervals were displayed in the LM, the participants could sort the linear segments along the vertical dimension according to a certain interval property, including the start point, the end point, the midpoint and the duration. Every question was asked twice in both the LM and the TM in a random order with the same number of intervals presented. In order to exclude extra interference, both the LM and the TM were presented in their most basic forms. Equal-interval ticks and labels were placed on the axes to indicate the scales of the axes. Gridlines from these ticks were drawn on the background in order to help the participants to read the metrics in the space. Figure 7-9 and Figure 7-10 illustrate two example questions with intervals presented in the TM and the LM. Since the study involved Dutch-speaking participants, all information and tests were presented in Dutch. At the end of the question, participants were requested to evaluate the two models by selecting one from the five options: 1): The TM is much easier; 2): The TM is a little easier; 3) They are the same; 4): The LM is a little easier; 5): The LM is much easier.

The website automatically recorded participant responses, answer scores, preferences of a model and the time participants spent on each question.

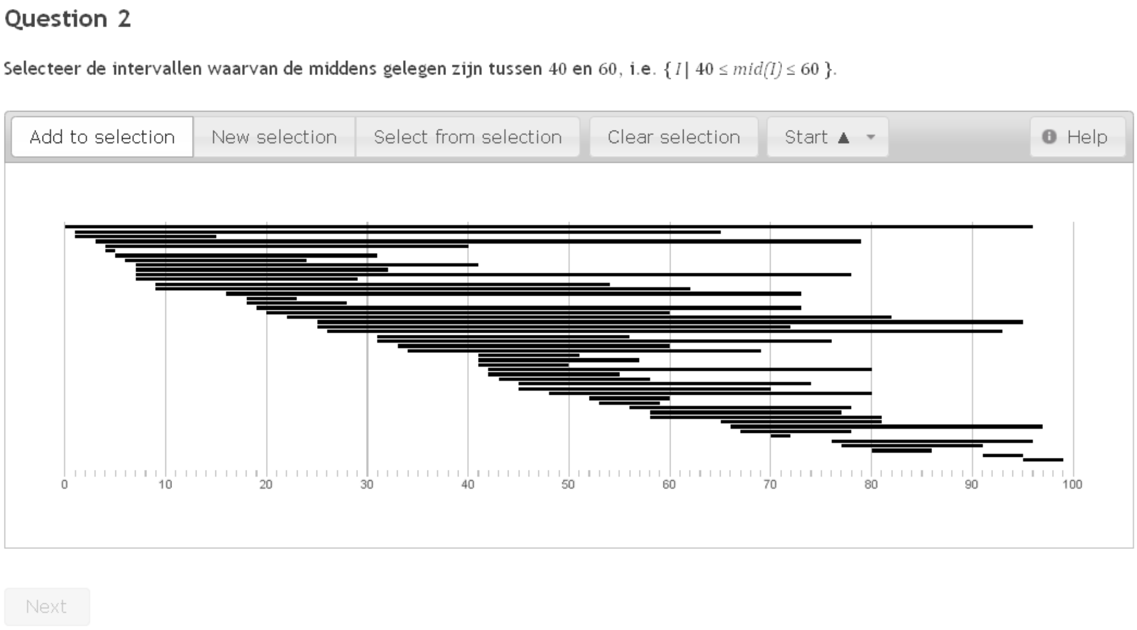


Figure 7-9: An example of posttest question with time intervals represented in the Linear Model. The time intervals are sorted by ascending start point along the vertical dimension.

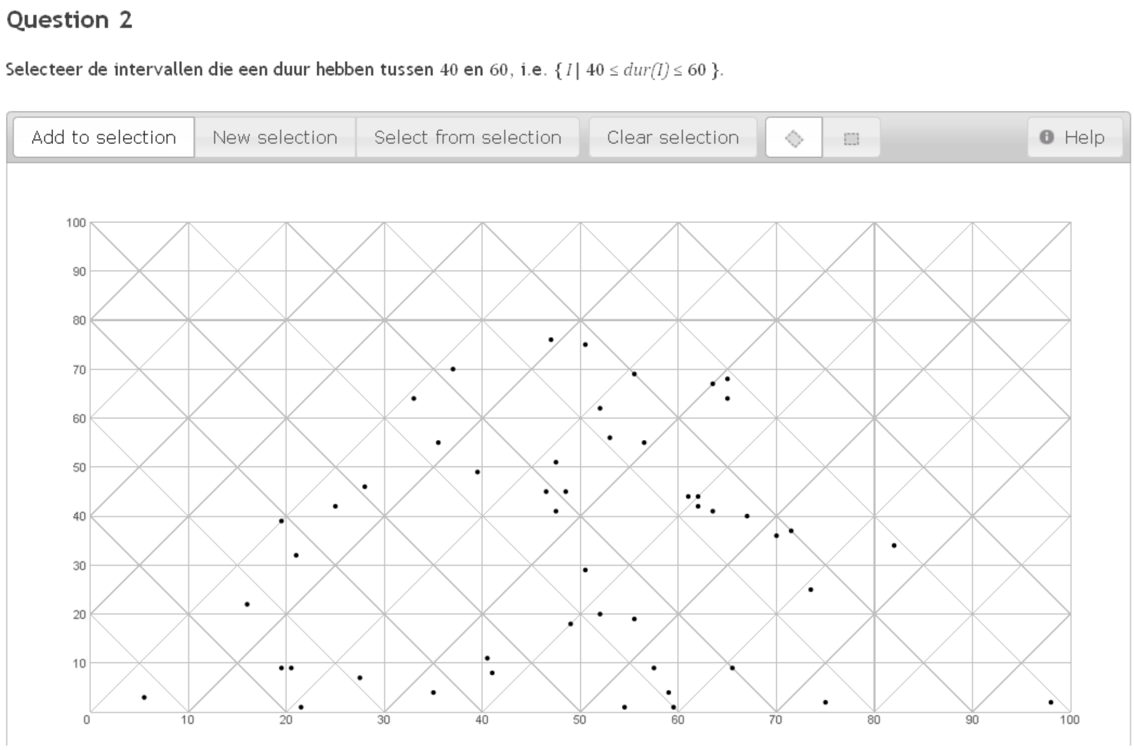


Figure 7-10: An example of posttest question with time intervals represented in the Triangular Model.

7.3.3 Research Procedure

The intervention study consisted of a training phase followed by the pretest. This training phase was considered to be crucial in two ways. First, it guaranteed the activation of the prior knowledge of the participants about the LM; and second, it guaranteed that a basic introduction to the TM could be implemented. The participants were all undergraduate students from Ghent University in the subject of educational sciences. In total 258 students participated in the study. Considering the group size, the study was repeated seven times within a one-week time frame. Students could plan their participation via an online calendar tool. The study took place in a computer lab in which every participant was allocated a desktop computer. To ensure that every group experienced exactly the same training process, the instructional intervention was consistently based on standardized video clips. All students watched these video-based instructions on a central screen in front of the computer classroom. Two video-based instructions were presented to the students. In the first video instruction, students were introduced to the conceptual base about time intervals and the representations of time intervals following the two models. In the second video instruction, students were introduced to the temporal relations between time intervals and the corresponding representations following the LM and TM. After each video instruction, the participants were requested to answer the corresponding part of the pretest, in which the questions were closely related to the knowledge taught in the video clip. The results from the pretest provided a preliminary evaluation about how well the students comprehended the two models. After students finished either pretest, a feedback video was played to demonstrate the correct answers of the questions and explain the underlying rationale. This feedback video was also considered as part of the instructional intervention since it offered further chances to learn or to correct their understanding about the two models. After studying the video instructions, solving the pretests, and studying the feedback videos, students entered to a website including the comprehensive posttest. The entire research procedure is illustrated in Figure 7-11.

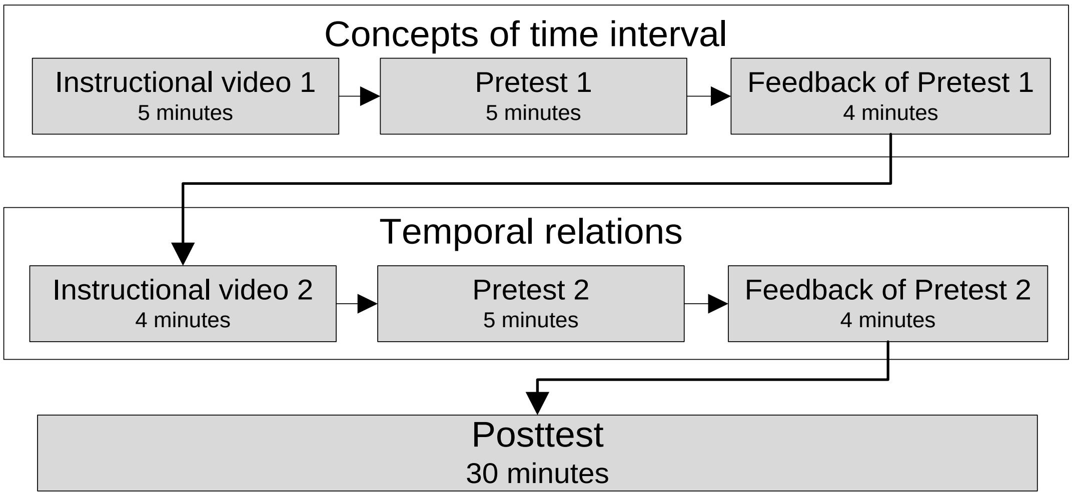


Figure 7-11: The flowchart of the research procedure.

7.4 Results

In this section, the result obtained from the test is presented and analysed. Since every question has been answered by participants in both the TM and the LM, it is feasible to compare the parameters obtained in these two models, including the answer score, the time used to solve the question, and the preference for a particular model. In general, the comparison has been made at two levels: the entire test and individual questions. T-test for paired samples is used to compare the scores obtained and time spent in the LM and TM. Wilcoxon signed-rank test is used to determine the preference for a particular model. The Pearson correlation coefficient is applied to indicate the correlation between the time consumption and the number of presented intervals.

7.4.1 Test Score

The correctness of the answer is described by a score between zero and one. In the pretest, the participants produced significantly higher average scores using the LM than the TM (Table 7-4). However, the situations reversed in the posttest, in which the participants produced significantly higher average score using the TM (Table 7-4).

In the posttest, the participants produced significantly higher scores with the LM in 3 questions, while in 17 questions they produced significantly higher scores with the TM. The participants produced higher scores with the LM in the questions about the start points and end points (Question 1 and 2 in Table 7-5). However, in the questions about the duration and midpoint, the participants (Question 3 and 4) produced higher score when using the TM. No significant differences have been detected in most of the questions concerning two interval properties, except the questions that ask about both the duration and midpoint (Question 10) in which the TM generated higher scores. Moreover, with the TM, participants produced higher scores in 7 questions with one temporal relation (Question 11 to 22), including the relation *before*, *meets*, *overlaps*, *started-by*, *during*, *contains*, *after*. Only in the question about the *met-by* relation, participants produced higher scores within the LM. In Figure 7-12, we depict the mean scores of these questions as grey scales in the conceptual neighbourhood structure of temporal relations. As can be derived from Figure 7-12, the mean scores obtained from the LM shows a symmetric distribution about the central point, i.e. the *equal* relation, which means that the participants produced similar scores in pairs of inverse relations. However, this centrosymmetry is less obvious in the TM. For instance, the participants produced much higher scores in the *meets* and *overlaps* relations than their inversions *met-by* and *overlapped-by*. Furthermore, in terms of the questions with two temporal relations, the scores obtained in the TM are higher than those obtained in the LM in 7 questions (Table 7-5). However, the opposite situation (i.e. the score obtained in LM is higher) has not been detected in any such questions.

7.4.2 Time and Preference

As stated earlier, in addition to the correctness of the question answers, the time participants spent and the preference of participants for a specific model were recorded in the posttest as well. With respect to the entire posttest, the participants generally spent significantly less time when the intervals were represented in the TM (Table 7-4). According to specific questions, the advantage of the TM in time consumption was also overwhelming (Table 7-5), except the two questions that asked about the start point or end point (Question 1 and 2 in Table 7-5) where no significant difference was detected. On the other hand, the correlation between the number of intervals and the time used has been detected significant in more than half of the questions (20 out of 37) when the interval is represented in the LM (Table 7-6). Namely, the more intervals are represented in the LM, the more time the participant spent on answering the question. In contrast, such correlation has only been detected in 6 questions in the TM (Table 7-6), which reflects that the efficiency of the TM is less influenced by the increasing number of intervals.

In terms of the model preference, most participants considered the TM as the easier representation mode with respect to the entire test (Table 7-4). At the question level, the preference of the TM is also predominant, with the exception of 7 questions in which the participants gave a neutral opinion (Table 7-5).

Table 7-4: The average scores, time consumption and preference in the pretest and posttest

Test	Mean score		Mean time		Preference
	LM	TM	LM	TM	
Pretest	0,79**	0,73	N/A	N/A	N/A
Posttest	0,61	0,69**	123**	63	TM**

** The underlying distribution is higher at the significance level of 0.01 (2-tailed)
N/A: The parameter is not available in this test

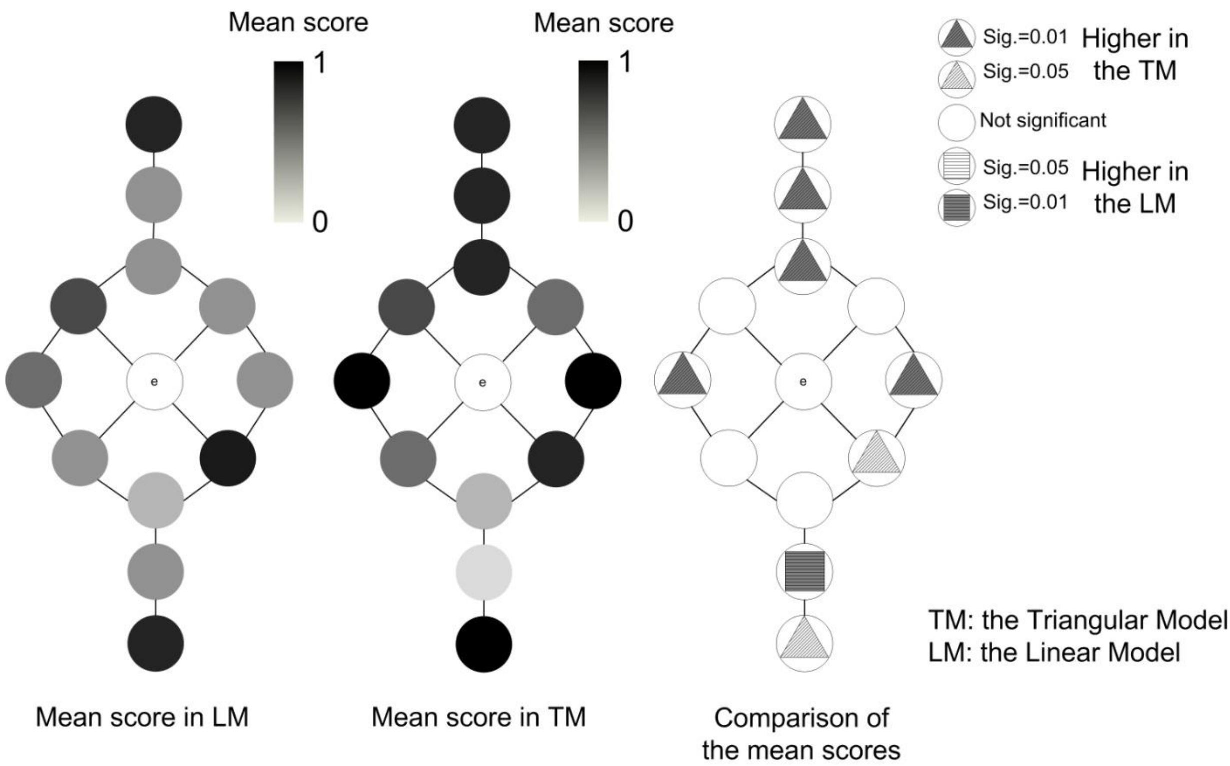


Figure 7-12: The mean scores of the questions about temporal relations represented as grey scales in the conceptual neighbourhood structure of the temporal relations.

Table 7-5: The scores, the time consumption and preference in the posttest, according to specific questions.

Type	No.	Asked concept	Mean score		Mean time (seconds)		Preference
			LM	TM	LM	TM	
Interval property	1	Start point	0.85**	0.70	95	77	----
	2	End point	0.85*	0.71	117	87	----
	3	Duration	0.46	0.70**	170**	91	TM**
	4	Midpoint	0.60	0.76**	206**	71	TM**
	5	Start point & end point	0.75	0.76	131**	88	----
	6	Start point & duration	0.60	0.57	166**	100	TM**
	7	Start point & midpoint	0.73	0.73	162**	80	TM**
	8	End point & duration	0.61	0.51	177**	93	TM**
	9	End point & midpoint	0.76	0.69	155**	87	TM**
	10	Duration & midpoint	0.45	0.72**	169**	57	TM**
Interval relation	11	Before	0.81	0.89**	80**	29	TM**
	12	Meets	0.55	0.81**	83**	43	TM**
	13	Overlaps	0.52	0.84**	109**	36	TM**
	14	Finishes	0.53	0.69	93**	34	----
	15	Starts	0.73	0.71	56**	38	----
	16	During	0.65	0.95**	107**	29	TM**
	17	Contains	0.58	0.90**	89**	35	TM**
	18	Started-by	0.83	0.80	53**	29	TM**
	19	Finished-by	0.53	0.66*	39**	70	TM*
	20	Overlapped-by	0.48	0.42	105**	36	TM**
	21	Met-by	0.57**	0.31	77**	44	TM**
	22	After	0.89	0.94*	75**	45	TM*
	23	Before & overlap	0.50	0.79**	151**	72	TM**
	24	During & before	0.62	0.67	114**	51	TM**
	25	Before & contain	0.44	0.70**	110**	49	TM**
	26	Overlapped-by & before	0.39	0.36	67**	123	----
	27	Before & after	0.62	0.73**	66**	38	TM**
	28	Overlap & during	0.49	0.79**	106**	52	TM**
	29	Contain & overlap	0.44	0.60	91**	48	TM**
	30	After & overlap	0.57	0.44	141**	80	TM**
	31	During & contain	0.40	0.78**	89**	35	TM**
	32	During & overlapped-by	0.47	0.65**	86**	48	TM**
	33	During & after	0.61	0.69	93**	48	TM**
	34	Contain & overlapped-by	0.43	0.52	94**	50	TM**
	35	Contain & after	0.52	0.61	119**	66	TM**
	36	Overlaps & overlapped-by	0.45	0.70**	135**	81	----
	37	Overlapped-by & after	0.41	0.39	120**	89	TM**

* The underlying distribution is higher at the significance level of 0.05 (2-tailed)
** The underlying distribution is higher at the significance level of 0.01(2-tailed)
---- No significant preference detected

Table 7-6: The correlation coefficients between the time consumption and the number of intervals, according to specific questions.

Type	No.	Asked concept	Correlation coefficient in LM	in TM
Interval property	1	Start point	0.18	0.27
	2	End point	0.15	-0.12
	3	Duration	0,13	0.21
	4	Midpoint	0.34*	0.22
	5	Start point & end point	0.30**	-0.01
	6	Start point & duration	0.28*	0.11
	7	Start point & midpoint	0.47**	0.24
	8	End point & duration	0.08	0.01
	9	End point & midpoint	0.26*	-0.08
	10	Duration & midpoint	0.25**	0.09
Interval relation	11	Before	0.42*	0.34**
	12	Meets	0.20	0.02
	13	Overlaps	0.31*	0.31*
	14	Finishes	0.14	0.40*
	15	Starts	0.47*	0.26
	16	During	0.26*	0.21
	17	Contains	0.24	0.12
	18	Started-by	-0.16	0.29
	19	Finished-by	0,39*	-0,31
	20	Overlapped-by	0.16	0.15
	21	Met-by	0.13	-0.08
	22	After	0.21	0.27*
	23	Before & overlap	0.33**	0.19
	24	During & before	0.39**	0.38**
	25	Before & contain	-0.01	-0.04
	26	Overlapped-by & before	0.42**	0.18
	27	Before & after	0.30**	0.02
	28	Overlap & during	0.10	-0.05
	29	Contain & overlap	0.20	0.14
	30	After & overlap	0.28*	0.06
	31	During & contain	0.19*	0.02
	32	During & overlapped-by	0.25*	0.04
	33	During & after	-0.05	-0.01
	34	Contain & Overlapped-by	0.38**	0.28*
	35	Contain & after	0.17	0.00
	36	Overlaps & overlapped-by	0.36**	0.03
	37	Overlapped-by & after	0.01	-0.12

* The correlation is at the significance level of 0.05 (2-tailed)
** The correlation is at the significance level of 0.01(2-tailed)

7.5 Discussion

In terms of the overall test scores, the reversed results have been observed in the pretest and the posttest. In the pretest, the participants produced higher scores using the LM than the TM, which could be explained as that the participants had more prior knowledge about the LM than the TM. This is in line with our expectations about the impact of the available prior knowledge that exclusively builds on the representation via the LM. However, in the posttest, the scores from the TM exceeded the scores from the LM. Based on the Working Memory Model, we can assume that the TM has played a more facilitating role when supporting the processing of new information and solving of temporal relation in working memory due to the better support of the visuo-spatial sketchpad. Consequently, we can assume that the TM has induced to a lesser extent extraneous cognitive load as compared to the LM. An additional catalyst can be found in the additional video clips that were presented after the pretest, which gave explicit feedback about the problems being solved. This could have fostered in particular the processing and storage of knowledge about the TM.

The scores obtained in the posttests show that the TM is generally a more effective graphical representation of time intervals with respect to these questions. Moreover, the participants spent less time on answering the majority of the questions when time intervals were represented in the TM than in the LM. This reflects a higher efficiency of the TM for visual observation of intervals, as compared to the LM. Furthermore, the efficiency of the TM was less influenced by an increase in the number of intervals, as the correlation between the number of intervals and time consumption was significant in only 6 of the 37 questions. In contrast, this correlation was significant in the majority of questions where time intervals were presented in the LM.

The LM showed advantages in questions that asked about the start point or the end point of intervals. This can be explained as that these properties can be directly observed from the two boundary points of the linear segments. However, the LM was not easy to use in the questions asking about duration and midpoint, which were expressed by the extent and location of the linear segments. It is hard to visually compare the lengths of intervals distributed at different locations. Also, it is not straightforward to compare the midpoints of intervals of different length. The visual identification of these two properties has to take account of both the start point and end point of the linear segments, which may invoke extra visual processing and as such extraneous cognitive load as suggested above. The same applies in the questions about the temporal relations. The identification of the temporal relations in the LM always requires the visual comparison of the two boundary points between the linear segments. Therefore, the search for the intervals in a certain temporal relation to a referenced interval is difficult and error-prone. For example, one has to compare a reference interval with other interval scattered in the diagram in order to find the ones that *overlap* it. This task becomes even more complex when the questions involve two temporal relations. Although the sorting tool in the website might have helped the search process to some extent, finding the correct intervals from the sorted linear segments is still not a straightforward task. Moreover, with increasing number of intervals, such visual observation or comparison has to be done increasing times, which makes the questions with a large number of intervals rather time-consuming. This has been reflected in the positive correlation between the number of intervals and the time that the participant spent on the question, which has been detected in most questions asked in the LM.

In contrast, every interval has a fixed location in the TM. The intervals having a specific property or satisfying a specific temporal relation are located within a specific zone. Therefore, in the questions, the search for the correct intervals is transferred to the identification of specific zones. The question scores indicate that the participants could correctly identify most of the relational zones in the TM. Once the zone has been located, the selection of the intervals within the zone is rather simple and straightforward, which is, moreover, not affected much by the increasing number of intervals presented in the model. This is probably the reason that the correlation between the number of intervals and the time consumption has been detected in fewer questions built on the TM. This finding reveals the potential of the TM in handling a large amount of intervals, which might be difficult for the LM.

The positive learning results of the TM were generated by a group of undergraduate-level participants without prior knowledge about the TM, and after a 20-minute training process. The increase in the knowledge about the TM is reflected in the differences in scores between pretest and posttest. In spite of the biased understanding about the *met-by* and *overlapped-by* relations in the TM being detected in the posttest, the participants have been quickly trained to understand and manipulate the TM at a satisfactory level. This provides empirical evidence that the TM and its implementation can be easily comprehended and used by non-expert users. Moreover, the participants overwhelmingly preferred the usage of the TM, which shows that the above-mentioned merits of the TM are perceived and appreciated by the participants.

7.6 Conclusion and Direction for Future Work

This chapter elaborated an empirical study of an alternative representation of time intervals, i.e. the TM. The study used a set of questions about time intervals to evaluate the understandability and usability of the TM, as compared to the traditional LM. The results showed that the participants scored better and spent less time in questions building on the TM. Thus, we contend that the TM is more effective and efficient than the LM in terms of interval visualization. Moreover, the TM showed the potential of representing a large amount of time intervals, which seemed to cause extra difficulties when being dealt with in the LM. Since these positive results of the TM were obtained by non-expert participants, after a 20-minutes training process, we believe that the TM can be easily understood and used by non-expert users. While this empirical study is based on the comparison between the TM and the LM, its aim is not proposing an entire substitution of the LM with the TM in all application areas. The LM remains an intuitive and founding representation of linear time, and thus remains advantageous in representing simple linear processes and a small number of time intervals. The findings from this study, together with findings from our previous research, provide strong support to consider the adoption of the TM in special tasks of interval visualization and analysis. It also supports our plans of developing tools based on the TM and releasing them to a broader research and practical contexts.

In future work, further evaluations of the TM will be conducted in different age groups and education levels in order to determine the appropriate user group it can be applied. It would be very interesting to evaluate the TM within younger groups, in which the linear time concept is probably less predominant. In addition to the parameters measured in this study, new technologies, such as eye-tracking, can be applied to obtain extra information in research process. Moreover, we plan to study more applications of the TM and further evaluate it in solving domain-specific problems. One the one hand, the implementation of the TM will be provided to the users that need to analyse massive interval-based data in the field of, for example, computer science, archaeology and geographical information science. Different visualization variables and metaphors can be applied in the TM to represent contextual information. In addition to interval visualization, this implementation will also allow users to formulate temporal queries by creating 2D zones on top of the interval visualization (Qiang et al., 2012b). On the other hand, attempts will be made to apply the TM in education of history and chronological knowledge. The initial idea is to represent the interval-based knowledge in history textbooks via an interactive website implementing the TM. For example, when the user clicks on a certain chapter or a certain category of events, the website will automatically display the intervals of events in the selected chapters or categories in an interactive TM diagram. It is also interesting that, users can select intervals in the TM using the graphic query devices, to study and analyse the knowledge linked to these intervals. The special merits of the TM combined with interactive visualization techniques can potentially benefit the perception and memorization of the temporal structure of interval-based knowledge.

References

Allen, J. F. (1983). Maintaining Knowledge about Temporal Intervals. Communications of the ACM, 26(11), pp.832-843.

Baddeley, A. and Hitch, G. J. (1974). Working memory. ed. by Bower, G., in Recent Advances in Learning and Motivation, Academic Press, New York.

Bhatt, M., Guesgen, H., Wolfl, S. and Hazarika, S. (2011). Qualitative Spatial and Temporal Reasoning: Emerging Applications, Trends, and Directions. Spatial Cognition and Computation, 11(1), pp.1-14.

Bittner, T. (2002). Approximate Qualitative Temporal Reasoning. Annals of Mathematics and Artificial Intelligence, 36(1), pp.39-80.

Cohn, A. G., Bennett, B., Gooday, J. and Gotts, N. M. (1997). Qualitative Spatial Representation and Reasoning with the Region Connection Calculus. GeoInformatica, 1(3), pp.275-316.

De Tré, G., Van de Weghe, N., De Caluwe, R. and De Maeyer, P. (2006). Towards a Flexible Visualization Tool for Dealing with Temporal Data. ed. by Larsen, H., Pasi, G., Ortiz-Arroyo, D., Andreasen, T. and Christiansen, H., in Flexible Query Answering Systems, pp.109-120, Springer Berlin, Heidelberg.

Freksa, C. (1992). Temporal Reasoning Based on Semi-Intervals. Artificial Intelligence, 54(1-2), pp.199-227.

Galton, A. (1990). A Critical Examination of Allen's Theory of Action and Time. Artificial Intelligence, 42(2-3), pp.159-188.

Gottfried, B. (2008). Global Feature Schemes in Spatial and Temporal Reasoning. Spatial Cognition and Computation, 8(1-2), pp.27-46.

Hochheiser, H. and Shneiderman, B. (2004). Dynamic query tools for time series data sets: timebox widgets for interactive exploration. Information Visualization, 3(1), pp.1-18.

Knauff, M. (1999). The cognitive adequacy of Allen's interval calculus for qualitative spatial representation and reasoning. Spatial Cognition and Computation, 1(3), pp.261-290.

Kulpa, Z. (1997a). Diagrammatic Representation for a Space of Intervals. Machine Graphics & Vision, 6, pp.5-24.

Kulpa, Z. (1997b). Diagrammatic Representation of Interval Space in Proving Theorems about Interval Relations. Reliable Computing, 3(3), pp.209-217.

Kulpa, Z. (2006). A Diagrammatic Approach to Investigate Interval Relations. Journal of Visual Languages and Computing, 17(5), pp.466-502.

Li, X. and Kraak, M.-J. (2008). The Time Wave. A New Method of Visual Exploration of Geo-data in Time–space. The Cartographic Journal, 45(3), pp.193-200.

Ligozat, G. (1994). Temporal reasoning made simpler. Proceedings of the 7th International Conference on IEA/AIE'94, Austin, Texas, pp.123-130, Gordon and Breach Science Publishers, Inc.

Ligozat, G. (1997). Figures for thought: temporal reasoning with pictures. Paper presented in Workshop on Spatial and Temporal Reasoning (at AAAI-97) Providence, RI, pp.31-36.

Mayer, R. E. and Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. Educational Psychologist, 38(1), pp.43-52.

Miyake, A. and Shah, P. (1999). Models of working memory. Mechanisms of active maintenance and executive control. Cambridge University Press, Cambridge.

Neutens, T., Witlox, F., Van de Weghe, N. and De Maeyer, P. (2007). Space-time opportunities for multiple agents: a constraint-based approach. International Journal of Geographical Information Science, 21(10), pp.1061-1076.

Qiang, Y., Delafontaine, M., Asmussen, K., Stichelbaut, B., De Tré, G., De Maeyer, P. and Van de Weghe, N. (2010). Modelling Imperfect Time Intervals in a Two-Dimensional Space. Control and Cybernetics, 39(4), pp.983-1010.

Qiang, Y., Delafontaine, M., Neutens, T., Stichelbaut, B., De Tré, G., De Maeyer, P. and Van de Weghe, N. (2012a). Analysing Imperfect Temporal Information in GIS using the Triangular Model. The Cartographic Journal, (in press).

Qiang, Y., Delafontaine, M., Versichele, M., De Maeyer, P. and Van de Weghe, N. (2012b). Interactive Analysis of Time Interval in a Two-Dimensional Space. Information Visualization, (in press).

Schockaert, S., De Cock, M. and Kerre, E. E. (2008). Fuzzifying Allen's Temporal Interval Relations. IEEE Transactions on Fuzzy Systems, 16(2), pp.517-533.

Stichelbaut, B. and Bourgeois, J. (2009). The Aerial Imagery of World War One: A Unique Source for Conflict and Landscape Archaeology. Photogrammetrie - Fernerkundung - Geoinformation, (3), pp.235-244.

Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. Learning and Instruction, 4, pp.295-312.

Sweller, J. and Chandler, P. (1994). Why some material is difficult to learn. Cognition and Instruction, 12(3), pp.185-233.

Van de Weghe, N., Docter, R., De Maeyer, P., Bechtold, B. and Ryckbosch, K. (2007). The Triangular Model as an Instrument for Visualising and Analysing Residuality. Journal of Archaeological Science, 34(4), pp.649-655.

Weaver, C., Fyfe, D., Robinson, A., Holdsworth, D., Peuquet, D. and MacEachren, A. M. (2006). Visual analysis of historic hotel visitation patterns. ed. by Wong, P. C., Proceedings of IEEE Symposium on Visual Analytics Science and Technology (Vast 2006), New York, pp.35-42, IEEE.

Weber, M., Alexa, M. and Muller, W. (2001). Visualizing time-series on spirals. Proceedings of IEEE Symposium on Information Visualization 2001, Washington, DC, U.S.A., pp.7-13.

8 General Discussion and Conclusion

This chapter first summarizes the contributions of this thesis through an elaboration of how the general research questions raised in the introduction have been addressed. Then, the contributions are discussed in view of the broader research agenda. Issues and challenges for future research will be discussed where appropriate. At the end, the general conclusion of this thesis is presented.

8.1 Summary

This thesis is concerned with the use of a two-dimensional representation, the Triangular Model (TM), in the perspective of temporal reasoning and visual analytics. First, this model has been treated as a diagrammatic tool to study the reasoning of imperfect time intervals. Second, it has been applied as a visual representation for the analysis of crisp time intervals and rough time intervals. These approaches have been implemented into software prototypes and applied to solve practical problems in real-world scenarios. Third, it has been extended to visualise linear data. Fourth, an empirical study has been conducted to evaluate the usability of this representation within a group of non-expert users. In this section, the contributions are summarised with respect to the major research questions that were proposed in the introduction section.

RQ 1: How can the TM facilitate the reasoning research of imperfect time intervals?

Rough set theory and fuzzy set theory are two often used approaches to model imperfect time intervals. Particularly, a lot of attention has been paid to the fuzzy approach (Nagypál and Motik, 2003; Ohlbach, 2004; Schockaert and De Cock, 2008; Schockaert et al., 2008), because in many cases the start and end of imperfect time intervals can be treated as gradual and continuous processes, either according to the fuzziness of the events or according to the assumptions of possibilistic or probabilistic density. In contrast, the existing research on the rough approach is incomplete and limited in certain aspects (Bassiri et al., 2009; Bittner, 2002). Chapter 3 has first introduced a common situation in which rough set theory is more suitable for the modelling of time intervals than fuzzy set theory. This situation derives from discrete data acquisition activities in which uncertainty exists in-between the observations at two discrete time stamps. This chapter has taken remote sensing as an example, which relies on images or photographs taken at discrete time stamps (i.e. snapshots). With the snapshots, the time intervals of observed features can approximated as linear rough sets, which leads to the concept of rough time intervals.

Second, Chapter 3 has used the TM as a diagrammatic tool to investigate the reasoning principles of rough time intervals. The usefulness of diagrammatic tool has been proved in much research in temporal reasoning and spatial reasoning (Kulpa, 1994). Due to the tripartite structure of rough time intervals (i.e. the upper, lower approximations and boundary regions), using the conventional linear representation to study the relations of rough time intervals can be very complicated and error-prone. In contrast, the polygonal representation of rough time intervals in the TM is much simpler and straightforward with respect to human perception. This chapter has shown how the temporal relations between rough time intervals can be visually recognised according to the topologies of their polygonal representations. It has also generalised a framework of methods to interpret the underlying semantics behind the polygonal configuration. Furthermore, it has demonstrated how temporal queries with rough constraints can be modelled as 2D rough zones and described the principles of the composition of multiple queries (e.g. intersection and union).

Third, Chapter 3 has preliminarily studied how fuzzy time intervals can be represented in the TM. Due to the continuous structure of fuzzy time intervals, the relations of fuzzy time intervals are expressed by the overlap of fuzzy areas, which is difficult for humans to perceive and process. Therefore, the reasoning of fuzzy time intervals has not been studied as thoroughly as that of rough time intervals. However, it has shown that temporal queries with fuzzy constraints can be represented as continuous fields, which can be potentially useful for the mapping of probabilistic or possibilistic density when the fuzzy time intervals are based on probabilistic or possibilistic models.

RQ 2: Can the TM bring added value to the analysis of time intervals?

Chapter 4 and 5 have respectively demonstrated the use of the TM in the analysis of crisp time intervals and rough time intervals. In the TM, crisp time intervals are plotted as points in a coordinated 2D space. Therefore, a set of intervals are represented by a fixed point structure. This feature facilitates the visual analysis of large amounts of time intervals. In the case study of Bluetooth tracking, several interesting clusters have been observed in the point structure, which possibly implies the major movements of the crowd. In the future, quantitative measurements or additional visualization techniques (e.g. density map or transparency adjustment) are needed to solve the overplotting problem. Also, the coordinate space enables the comparison between sets of intervals. The difference of the interval distribution can be directly observed from the difference of the point structures, which is impossible in the traditional linear representation. On the other hand, an innovative temporal query mechanism has been proposed, which relies on the manipulation of geometries in the interval display. In this mechanism, a variety of temporal queries can be formulated by creating specific zones in the 2D space. The combinations of queries (e.g. unions and intersections) are intuitively expressed by the spatial operations of the query zones. This query-in-display approach is similar to the brushing technique that is applied in geographical maps or scatter plots (Monmonier, 1990), which increases the interactivity of the visualization.

Chapter 5 is based on the research outcomes of Chapter 3 and 4, however, with an emphasis on the analysis of rough time intervals. Instead of a point structure, a number of rough time intervals are represented in a polygonal structure in which overlapping areas are colour-coded. This representation offers an explicit image of the distribution of rough time intervals, which can also be used for visual pattern detection. The query mechanism introduced in Chapter 4 can be applied to rough time intervals as well. However, there arise uncertainty issues when the query zone partially covers the polygon of the rough time interval. To this end, a probabilistic framework has been proposed. The basic idea is that the probability of a rough time interval satisfying a query is the extent (i.e. the ratio) to which the polygon of the rough time interval is covered by the query zone. With this framework, temporal queries for rough time intervals can either be performed with a probabilistic threshold or by giving multiple answers according to the probabilities.

The data visualization and query mechanism for crisp time intervals and rough time intervals have been implemented in software prototypes, in which the data display of the TM is incorporated within a GIS. In these prototypes, observations and queries can be interactively carried out from the spatial, temporal and attribute perspectives. Also, the functionalities based on the probabilistic framework have been implemented in the version for rough time intervals, enabling users to analyse rough intervals with account of the probabilities. These prototypes have been applied to the analysis of two real-world datasets, in which one can interactively examine the spatio-temporal data in both the map view and the interval view, and discover interesting patterns and relationships that are hard to discover using traditional analytical tools. According to these applications, it is believed that such an implementation of the TM can effectively support exploratory analysis of interval-based geographical data.

RQ 3: How can the TM be used for the analysis of linear data?

Previous chapters focus on discrete time intervals. Chapter 6 has described how continuous linear data (e.g. time series and traffic data along a road) can be represented in the TM. The basic idea is assigning the aggregation value of every sub-interval of linear data to the corresponding point in the interval space. This way, linear data can be represented as a continuous field of values, which is called the Continuous Triangular Model. On the one hand, this representation offers an explicit overview of linear data aggregated in all different intervals. Variations in different scales and the hierarchies of the scales can be displayed in a single diagram. On the other hand, because these continuous fields are based on a coordinate space of intervals, which is similar to the geographical space, the methods of spatial analysis (such as map algebra and cartographical modelling) can be applied to them. Using these methods, one can analyse multiple linear data and find the most suitable intervals in view of relevant factors and constraints. The analysis is not limited to any scales or aggregation partitions, but presents flexible answers and suggestions in all different intervals. Real-life datasets have been used to demonstrate the use of this representation in the analysis of traffic data and the selection of surfing sites and periods.

RQ 4: Can the TM be understood and used by novice users.

Chapters 3 to 5 have revealed great potential of the TM in reasoning and analysis of time intervals. However, there is still a lack of feedbacks from non-expert users in terms of its understandability and usability. Chapter 7 has described an empirical study that evaluates whether the participants with no prior knowledge about this model can easily learn it and make use of it. In this empirical study, a series of questions about interval relations and interval properties have been answered by the participants after a 20 minutes training process. The result showed that the participants have more correctly and efficiently answered the questions using the TM than using the linear representation, which reflected that the TM better supported visual observation of interval properties and relations. It also showed that the increasing number of intervals presented in the TM does not affect the correctness and time consumption of the questions, which implies that it can be suitable for the visualization of large amounts of intervals. In addition, most participants considered the TM was an easier option than the linear representation with respect to the tested questions. The results obtained from this empirical study have confirmed the merits of the TM in visualizing time intervals and given support to its usability in a broader range of non-expert users.

8.2 Discussion

This section discusses the contributions of this thesis in a general research background. It elaborates the value of this research and remaining challenges in a number of aspects. The avenues for future research are proposed where appropriate.

8.2.1 Interval Analysis

In artificial intelligence and information science, it becomes a common view that everything in the world is referenced to a time interval. This view has been reflected in most of the fundamental ontologies (Bittner et al., 2004; Gangemi et al., 2002; Grenon, 2003), which define all entities in the world as endurants and perdurants. Both perdurants and endurants exist in certain time intervals. The major difference is that an endurant wholly exists in every subinterval during its lifetime but a perdurant partially exists in the subintervals during its lifetime. Similar to the footprints in the geographical space, time intervals can be considered as footprints of entities in time. While considerable amount of contributions have been made in analysis of time-dependent data, the majority of them focused on changes of certain properties of entities through time, which leads to an enormous body of methods for time series analysis. However, the analytical methods for the temporal footprints (i.e. time intervals) of entities have received much less attention, particularly in the perspective of visualization and visual analytics. The approach presented in this thesis offered an additional option for visual analysis of time intervals.

Of course, the continuous changes of entities in every time stamp are very important in many analytical tasks. However, in some cases, time intervals generalised from the changing processes are already enough for certain analytical purposes. The typical examples are lifetimes of events or certain periods of processes that have special characteristics. The generalisation from time series to interval-based data can greatly reduce the data volume and release the workload for the analytical systems. Also, the analysis of time intervals can bring extra insights that cannot be easily obtained through the analysis of time series. As demonstrated in the case studies of Chapters 4 and 5, the special patterns observed in the interval distribution may imply interesting phenomena behind.

On the other hand, the acquisition of interval-based data is often less expensive than the acquisition of time series data that requires full-time monitoring. For instance, from discrete observations, one can already specify time intervals or at least approximations of time intervals of the underlying processes. However, obtaining the time series of the processes may require equal-interval observations of higher frequency, which could significantly increase the expense for the data collection and sometimes is even technically infeasible. In contrast, interval-based data sometimes can be collected by special low-cost acquisition systems. For instance, in the Bluetooth tracking system described in Chapter 4, a single sensor can generate time intervals of Bluetooth devices within its physical detection range. The monitoring range could be greatly enlarged if proper installation strategies are applied. For example, the installation of several sensors at the entrances and exits of a closed area (e.g. shopping centres and museums) can acquire the time intervals of visitors' stays. This kind of system is much cheaper and more effective in many aspects than the full-time monitoring system such as closed-circuit television (CCTV) (Delafontaine et al., 2011; Versichele et al., 2012).

The analysis of interval-based data is very important in many disciplines. However, the methodologies of pattern detection, interpretation, and trend prediction are still yet to be established. The approach presented in this thesis can be considered as a visualization platform for these directions of research. In the perspective of visual analytics, more research is needed to build a stronger link between the visual observation and the ground truth.

8.2.2 Vagueness and Uncertainties in Time Intervals

Since early 1980, the appearance of Allen's interval (Allen, 1983) algebra has triggered a variety of research within and beyond temporal reasoning (Freksa, 1992; Galton, 1990; Güsgen, 1989; Knauff, 1999; Schockaert et al., 2008). This theory assumes an ideal situation: the temporal references of the events are well-defined and perfectly known. However, in reality, there are many cases that this assumption cannot be satisfied. The temporal information and knowledge in the real-world often involves vagueness and uncertainties, which makes the temporal references of events and processes hard to perfectly specify. The tackling of such issues are complicated and dangerous. The optimal solution is to just fully use what is known and also avoid overusing. Therefore, appropriate modelling of imperfect temporal knowledge is of vital importance. Chapter 3 has discussed the kind of situation in which fuzzy set theory or rough set theory is more suitable for the modelling of imperfect time intervals. It claimed that rough set theory is well suited to model the time intervals that have been obtained from discrete data acquisition activities. In such cases, the start and end of the observed processes may take place everywhere in-between observations at two time stamps. If there is no abundant prior knowledge or established assumptions of the temporal distribution of the start and end times, representing the time intervals of these events with quantitative models can be expensive and dangerous. This situation typically exists in the exploratory phase of the data analysis when there is not much known about the dataset. An improperly fitted fuzzy model can easily lead to biased analysis result. Therefore, the triple classification of rough set theory (i.e. 'definitely yes', 'maybe yes or no', and 'definitely no') is just enough to represent the time intervals of such events, which leads to the concept of rough time intervals.

Chapters 3 and 5 show that the TM is an excellent visual representation of rough time intervals for both reasoning and analysis. The discrete polygonal representation is simple but also very informative. The position, shape and area of a polygon completely express the characteristics of the rough time interval. The topological relations between polygons directly imply certain temporal relations of rough time intervals, which is straightforward to the research of qualitative reasoning. In the visualization aspect, the polygonal structure is more perceivable than linear segments with tripartite structures. Users may observe the interval distribution through the overlaps of polygons. When temporal queries are applied to rough time intervals, we realise that it is necessary to impose a probabilistic framework to rough time intervals. The basic assumption is that the crisp time interval approximated by the rough time interval can be located anywhere within the polygonal area with equal probability. Thus, the probability that a rough time interval satisfies a temporal query can be intuitively expressed by the overlapping ratio between the polygon of the rough time interval and the query zone. Although the polygonal representation differs from the conventional linear concept of time, there are many advantages of the TM in terms of visual perception and intuition. The prototype introduced in Chapter 5 demonstrates the use of the TM in analysing rough time intervals. In the future, additional functionalities can be implemented to enhance its ability in handling probabilities in rough intervals. For instance, the colour-coding in the interval visualization can be manipulated to display different probabilities. The graphic query tool can also be dynamically linked with a histogram to display the probability distribution of the numbers of selected intervals. Compared with rough time intervals, the representation of fuzzy time intervals in the TM is more complicated, and therefore, is not straightforward for the research of qualitative reasoning. However, fuzzy temporal constraints (e.g. during a fuzzy time interval) can be represented as 2D continuous fields, which is more perceivable than mathematical formulas. An interesting application is the development of a visual calendar, in which users can define their time constraints with certain degrees of possibility. From the continuous field in the TM, the users can have an overview of their available time slots with the awareness of the possibility.

8.2.3 Queries in Data Display

Visual and dynamic query techniques have been developed to support interactive visualization and analysis. These techniques have liberated people from learning formal query languages by intuitive graphic controls, such as sliding bars, check boxes and menus. In visualization systems, queries are not limited to the manipulation of graphic controls. A well-designed data display can be a rich platform for many queries based on visual observations. For instance, in a geographical map, one can visually identify the places within a country by the country boundaries, or find the cities near the coast line or sandwich shops near his or her office. In a scatter plot, one can directly recognise the data that are close to the regression line as well as the outliers far from it. These queries are directly carried out in the data display without manipulation of any additional controls. However, the limitation of this kind of queries is the accuracy. For instance, one can hardly tell whether Beijing or Madrid is nearer to the North Pole only by visual observation on the map.

Incorporating query tools into the data display can effectively supplement such observation-based query processes. The most prominent example is the brushing technique that is widely applied in many visualization systems (Monmonier, 1989, 1990). This technique allows users to formulate queries by directly clicking items or dragging query boxes in the data display. The data in the display is usually well-arranged or spatially coordinated. The position, range and shape of the query boxes thus can be interpreted by formal query expressions. Users can flexibly create queries according to the patterns they find interesting and also have a direct sense of how the query range covers the dataset. The brushing technique offers an interactive integration between visual observation and query formalization, which has many advantages over the query controls that are independent from data displays (Martin and Ward, 1995; Monmonier, 1989, 1990). The query temporal mechanism proposed in Chapter 4 and 5 belongs to this category, which is based on a coordinate space of time intervals. A variety of temporal queries and their compositions can be formulated as geometric zones and composition of such zones in the interval display, which are hard to formulate in the conventional linear representation. The prototype implementation described in Chapter 4 and 5 shows its potential in assisting exploratory analysis of interval-based data. This prototype is built on top the ArcGIS® platform, thus is limited in functionalities that can be developed. In the future, another development environment should be considered, where this query mechanism can be implemented in a more flexible manner. Other than dragging query zones, it is preferable that the query zones can be flexibly sketched, adapted, moved, and deleted in the interval visualization. The responses to queries can also be more dynamic: when the query zone is changing, the other views instantaneously update according to the current selection. Of course, the potential of this query mechanism is based on the assumption that the user is familiar with it. When this prototype is released to a broader user community, a learning process is probably needed before users can correctly and efficiently transfer the temporal queries to the corresponding query zones. Implementing more interactive and user-friendly functionalities can probably shorten this process. Furthermore, the actual usability of this query mechanism is still needed to be assessed through systematic user studies.

8.2.4 Empirical Study of Visual Analytics

The empirical study in Chapter 6 has evaluated the usability of the TM for the observation-based queries that are discussed in the previous sub-section (i.e. 8.2.3). The result shows that the TM can better support such queries than the linear representation. However, more evidence is still needed to prove that the non-expert users can effectively use this representation in practical analytical tasks. The usability of the TM in visual analysis needs to be further assessed by more comprehensive empirical studies. A challenge in this further step is that the analytical power of a visualization approach is usually difficult to evaluate because the users' perception of the data display is too complicated to be measured with traditional questionnaires. Innovative methods and techniques can be considered for the solution to this issue. For instance, the eye-tracking technique has improved its accuracy to millimetres and becomes more and more affordable. It can bring more objective and direct insights of users' behaviour when they use the analytical tools (Fabrikant et al., 2008; Ooms et al., 2012a; Ooms et al., 2012b). On the other hand, visual analysis is a sophisticated process, which can be felicitously summarised by Keim's mantra about visual analytics: "Analyze first, Show the Important, Zoom, filter and analyze further, Details on demand"(Keim, 2005). This procedure may go several loops until the user find phenomena or relationships in the data of their interests. However, a comprehensive and reliable evaluation of the entire process could be an extremely hard task. The lack of evaluation methods for analytical usability broadly exists in the area of visual analytics. The analytical usability of visualization techniques are usually demonstrated through case studies. However, there is usually a lack of direct evidence that the users are able to use it in the same way as demonstrated in the case study when they are confronted with similar problems. A lot of further research is needed to generalise the guidelines for visualization designers to conduct empirical studies for the evaluation of their systems. The empirical studies should not only focus on how well users can read the information from the visualization, but should also concern whether these systems can inspire users' analytical interest and facilitate their analytical thinking.

8.2.5 Multi-scale Analysis

Multi-scale analysis is a common issue in data analysis. To decide an appropriate scale for analysis, people often have to use the 'trial and error' approach, which is inefficient and imprecise. In addition to searching for an appropriate scale, the hierarchy of the data in different scales is also important in some analytical tasks. An explicit overview of data in different scales may contribute to the solution of this issue. The continuous triangular model (CTM) introduced in Chapter 7 provides such an overview for linear data, in which linear data aggregated in all possible units can be displayed in a single diagram. This breaks through the limitation of the conventional representations built on a pre-set scale and aggregation partition. A more interesting feature is that the CTM is based on a coordinate interval space. Thus, CTM diagrams of multiple linear datasets can be combined via map algebra, which is useful for the analysis considering different factors and constraints. The CTM-based analysis may offer flexible answers or suggestions in all possible intervals within the frame of the linear dataset. Chapter 7 has demonstrated the use of the CTM with several practical datasets. Besides these, it is possible to apply the CTM into other scenarios dependent on multi-scale analysis of linear data. For example, football, as well as many other sports, considers timing as a crucial issue. Applying appropriate strategies within appropriate time intervals is one of the keys to win a game. Multi-scale analysis of players' physical conditions and performance during games may offer valuable insights to managers to make strategies and substitutions. Furthermore, it is promising to integrate the CTM with a GIS to enhance spatio-temporal analysis. An initial idea is linking the CTM with the space-time cube (Gatalsky et al., 2004; Hägerstrand, 1970; Kraak, 2003). This may allow users to observe space-time trajectories in the space-time cube, and observe multi-scale information of trajectories in CTM. Interactive brushing can be applied between them: when a user selects trajectories in the space-time cube, the information of the selected trajectories is displayed in CTM; when users click on an interval point in CTM, trajectories within this interval are highlighted in the space-time cube. Such an interactive integration may potentially benefit multi-scale analysis of space-time trajectories.

8.3 General Conclusion

This thesis has investigated the use of a 2D time representation (i.e. the TM) in reasoning and analysis of temporal information. Beyond the previous studies on reasoning of crisp time intervals (Kulpa, 1997a; Kulpa, 1997b; Kulpa, 2001; Kulpa, 2006; Ligozat, 1994, 1997), this thesis has further developed its use in representation and reasoning of imperfect time intervals (i.e. rough time intervals and fuzzy time intervals). According to the findings of this thesis, we contend that the TM is an excellent diagrammatic tool for the reasoning research of imperfect time intervals. Compared with the traditional linear representation, the TM can represent imperfect time intervals and relations in a more perceivable way, which helps people to understand and study abstract temporal concepts. Based on the TM, the reasoning mechanism of rough time intervals have been systematically studied, which was a blank area in temporal reasoning. It has also shown the possibility of representing relations of fuzzy time intervals, which offers extra visual impression over mathematic formulas.

In addition to temporal reasoning, this thesis has investigated a new application area of the TM, i.e., interval analysis. Some drawbacks of the traditional linear representation have hampered the development of visualization approaches that deals with time intervals. This thesis has suggested that the TM offers a compact and fixed visualization of time intervals, which facilitates visual pattern recognition. It has also proposed and implemented a graphic query mechanism, which support flexible formulation of temporal queries by creating zones. The advantage of these tools is that users can direct perceive and control the range of the query over the displayed interval. Furthermore, operations of temporal queries can be modelled in the similar way of Venn Diagrams, which has already shown intuitive merits in set theory studies. With the support of interactive controls in graphic computer programmes, the TM combines the explicit data display and flexible query tools in an effective way to facilitate the analysis of time intervals. Instead of implementing the TM into a standalone tool, we chose to integrate it into a GIS with dynamic linkages to a map view and an attribute table so that queries can be specified from the space, time and attribute perspectives. Through a practical use case, we have shown that this integration can enhance the GIS for exploratory analysis of interval-based geographic data.

This thesis also addresses an uncertainty issue in the analysis of imperfect time intervals, which is a difficult problem for existing approaches. It starts from the fundamental level, where a probabilistic framework has been proposed to calculate probabilities of queries about imperfect time intervals. It is noteworthy that this framework is inspired by the geometric configuration (i.e. the overlapping ratio of zones) in the TM, which again confirms the advantage of the TM in temporal reasoning study. At the application level, functionalities based on this framework have been incorporated in the graphic query tools, which allow analysis of imperfect intervals with consideration of the probabilities. The usefulness of this tool has been demonstrated by solving a real world problem.

Furthermore, this work introduced an extension of the TM (i.e. the CTM) to deals with a more general class of linear data. The CTM overcomes the restrictions of scales and aggregation partitions, which can be considered as an interesting solution to the problem of multi-scale analysis. On the one hand, it presents a compact visualization of a linear data segment aggregated during all sub-intervals. On the other hand, as the CTM is based on a coordinate space of intervals, map algebra can be applied to multiple CTM diagrams for comparison, selection and filtering. Based on such map algebra, a similar idea of cartographic modelling can be applied to CTM diagrams for decision making analysis that considers different criteria and constraints. The analysis result is not restricted to certain scales or partitions, but based on all possible intervals during the studied linear frame. The other interesting feature of this approach is that the entire process is visually performed and the final result is presented in a diagram, which offers explicit insights to the intermediate processes and the set of solutions. Since linear data is a common data type in many disciplines, the potential of the CTM in visualising different types of linear data and solving different domain-specific problems certainly deserves further investigation. Moreover, we believe that an interactive implementation can better release the analytical power of the CTM.

An empirical study has compared the TM and the linear model (LM) for the visualization of crisp time intervals. The result shows that the TM is more efficient than the LM for visual recognition of interval properties and relations. This study has confirmed the merits of the TM in visualizing time intervals and given support to its usability within non-expert users. This study does not intend to suggest an entire substitution of the LM with the TM in all applications. The LM might be still preferred in many situations. It has become clearer that the TM is more suitable for analytical tasks, particularly those involving large amounts of intervals.

This thesis has introduced prototype implementations of the TM, which are developed on top of ArcGIS™. This development platform includes a robust library of spatial data models and operators that can be directly called into the program. Due to the similarity of the TM and the geographical space (both are based on a 2D coordinate space), these spatial models and operators can also be used to implement the TM. This platform allows basic functionalities of the TM to be implemented in a budget and efficient way. It has also strengthened the existing GIS in analysing interval-based data. In general, this implementation takes advantages of a GIS to support the GIS. However, its drawback is that the development is limited to this library, which impedes more interactive controls to be implemented. To further exploit the use of the TM, a more flexible development environment needs to be considered, where more interactive and complicated tools can be applied to the TM. The burgeoning development of open source GIS has offered such possibilities.

A time interval basically consists of a two-tuple property (i.e. a start point and end point). Extending the representation of time from 1D to 2D leads to a better visualization for time intervals and many interval-based concepts. Although the TM at the first glance differs from the conventional linear concept of time, beyond this, it has many advantages in terms of visual perception and intuition. This thesis has revealed the use of the TM in solving several problems in temporal reasoning and analysis. However, the potential of the TM is far from being exhausted. Future research can target at fuzzy interval analysis, advanced implementations, more application areas and further usability studies.

References

Allen, J. F. (1983). Maintaining Knowledge about Temporal Intervals. Communications of the ACM, 26(11), pp.832-843.

Bassiri, A., Malek, M., Alesheikh, A. and Amirian, P. (2009). Temporal Relationships between Rough Time Intervals. ed. by Gervasi, O., Taniar, D., Murgante, B., Laganà, A., Mun, Y. and Gavrilova, M., in Computational Science and Its Applications – ICCSA 2009, pp.543-552, Springer Berlin / Heidelberg.

Bittner, T. (2002). Approximate Qualitative Temporal Reasoning. Annals of Mathematics and Artificial Intelligence, 36(1), pp.39-80.

Bittner, T., Donnelly, M. and Smith, B. (2004). Endurants and perdurants in directly depicting ontologies. AI Communications, 17(4), pp.247-258.

Delafontaine, M., Versichele, M., Neutens, T. and Van de Weghe, N. (2011). Analysing spatial and temporal sequences in Bluetooth tracking data. Paper presented in Mobile positioning, Spatial Mobility, Geography, LBS Tartu, Estonia.

Fabrikant, S. I., Rebich-Hespanha, S., Andrienko, N., Andrienko, G. and Montello, D. (2008). Novel Method to Measure Inference Affordance in Static Small-Multiple Map Displays Representing Dynamic Processes. The Cartographic Journal, 45(3), pp.201-215.

Freksa, C. (1992). Temporal Reasoning Based on Semi-Intervals. Artificial Intelligence, 54(1-2), pp.199-227.

Galton, A. (1990). A Critical Examination of Allen's Theory of Action and Time. Artificial Intelligence, 42(2-3), pp.159-188.

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A. and Schneider, L. (2002). Sweetening Ontologies with DOLCE. Proceedings of the 13th International

Conference on Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web, pp.166-181, Springer-Verlag.

Gatalsky, P., Andrienko, N. and Andrienko, G. (2004). Interactive analysis of event data using space-time cube. Proceedings of the 8th International Conference on Information Visualisation (IV 2004), London, UK, pp.145-152.

Grenon, P. (2003). The Formal Ontology of Spatio-Temporal Reality and its Formalization. ed. by Guesguen, H. W., Mitra, D. and Renz, J., in Foundations and applications of spatio-temporal reasoning (FASTR), pp.27-34, AAAI Press, Menlo Park, CA.

Güsgen, H. W. (1989). Spatial reasoning based on Allen's temporal logic. International Computer Science Institute Berkeley.

Hägerstrand, T. (1970). What about people in Regional Science? Papers in Regional Science, 24(1), pp.6-21.

Keim, D. A. (2005). Scaling visual analytics to very large data sets. Paper presented in Workshop on Visual Analytics Workshop on Visual Analytics.

Knauff, M. (1999). The cognitive adequacy of Allen's interval calculus for qualitative spatial representation and reasoning. Spatial Cognition and Computation, 1(3), pp.261-290.

Kraak, M.-J. (2003). The space-time cube revisited from a geovisualization perspective. Proceedings of the 21st International Cartographic Conference, Durban, South Africa, pp.1988-1996.

Kulpa, Z. (1994). Diagrammatic representation and reasoning. Machine GRAPHICS & VISION, 3(1/2), pp.77-103.

Kulpa, Z. (1997a). Diagrammatic Representation for a Space of Intervals. Machine Graphics & Vision, 6, pp.5-24.

Kulpa, Z. (1997b). Diagrammatic Representation of Interval Space in Proving Theorems about Interval Relations. Reliable Computing, 3(3), pp.209-217.

Kulpa, Z. (2001). Diagrammatic representation for Interval Arithmetic. Linear Algebra and Its Applications, 324(1-3), pp.55-80.

Kulpa, Z. (2006). A Diagrammatic Approach to Investigate Interval Relations. Journal of Visual Languages and Computing, 17(5), pp.466-502.

Ligozat, G. (1994). Temporal reasoning made simpler. Proceedings of the 7th International Conference on IEA/AIE'94, Austin, Texas, pp.123-130, Gordon and Breach Science Publishers, Inc.

Ligozat, G. (1997). Figures for thought: temporal reasoning with pictures. Paper presented in Workshop on Spatial and Temporal Reasoning (at AAAI-97) Providence, RI, pp.31-36.

Martin, A. R. and Ward, M. O. (1995). High Dimensional Brushing for Interactive Exploration of Multivariate Data. Proceedings of the 6th conference on Visualization '95, IEEE Computer Society.

Monmonier, M. (1989). Geographic Brushing: Enhancing Exploratory Analysis of the Scatterplot Matrix. Geographical Analysis, 21(1), pp.81-84.

Monmonier, M. (1990). Strategies for the Visualization of Geographic Time-Series Data. Cartographica: The International Journal for Geographic Information and Geovisualization, 27(1), pp.30-45.

Nagypál, G. and Motik, B. (2003). A Fuzzy Model for Representing Uncertain, Subjective, and Vague Temporal Knowledge in Ontologies. Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics, (ODBASE), Catania, Italy, pp.906-923.

Ohlbach, H. J. (2004). Relations between fuzzy time intervals. Proceedings of 11th International Symposium on Temporal Representation and Reasoning (TIME 2004), Montréal, Québec, pp.44-51.

Ooms, K., Andrienko, G., Andrienko, N., De Maeyer, P. and Fack, V. (2012a). Analysing the spatial dimension of eye movement data using a visual analytic approach. Expert Systems with Applications, 39(1), pp.1324-1332.

Ooms, K., De Maeyer, P., Fack, V., Van Assche, E. and Witlox, F. (2012b). Interpreting maps through the eyes of expert and novice users. International Journal of Geographical Information Science, pp.1-16.

Schockaert, S. and De Cock, M. (2008). Temporal Reasoning about Fuzzy Intervals. Artificial Intelligence, 172(8-9), pp.1158-1193.

Schockaert, S., De Cock, M. and Kerre, E. E. (2008). Fuzzifying Allen's Temporal Interval Relations. IEEE Transactions on Fuzzy Systems, 16(2), pp.517-533.

Versichele, M., Neutens, T., Delafontaine, M. and Van de Weghe, N. (2012). The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities. Applied Geography, 32(2), pp.208-220.

9 Appendix - The Implementation

This appendix specifically describes the implementation of the TM (i.e. GeoTM). Two versions of GeoTM have been introduced in this thesis. One deals with crisp time intervals in Chapter 3, which is named GeoTM^C in this appendix. The other one deals with rough time intervals in Chapter 5, and is named GeoTM^R.

9.1 Development Environment

Both versions of GeoTM are built on ArcMap 9.3, which is a desktop GIS produced by ESRI. ArcMap embeds a robust development environment (i.e. ArcObjects) that exposes the full range of functionality available in ArcMap to developers. The development of GeoTM is based on ArcObjects with the programming language Visual Basic for Applications (VBA). Thus, an installation of ArcMap 9.3 is the prerequisite for running GeoTM.

9.2 Data Model

GeoTM deals with discrete geographical entities that are referenced to time intervals. An entity has (1) a location in space, which is represented as 2D objects in the map view; (2) a time interval (can be crisp or rough), which is represented as a 2D objects in the TM view; and (3) other attributes, which are stored in the attribute table. The data model and implementation in GeoTM that applies to GeoTM has been described in Chapter 4.4 and Chapter 5.4.

9.3 GeoTM for Crisp Time Interval

9.3.1 Graphic User Interface

GeoTM^C is embedded in an ArcMap project (an mxd file). To launch the graphic user interface (GUI) of GeoTM^C, the user first needs to open the ArcMap project and then click a customized button on the ArcMap tool panel. The GUI of GeoTM^C consists of a map view, a TM view and a set of controls (numbered in Figure 9-1). The functionality of these controls is described in Table 9-1.

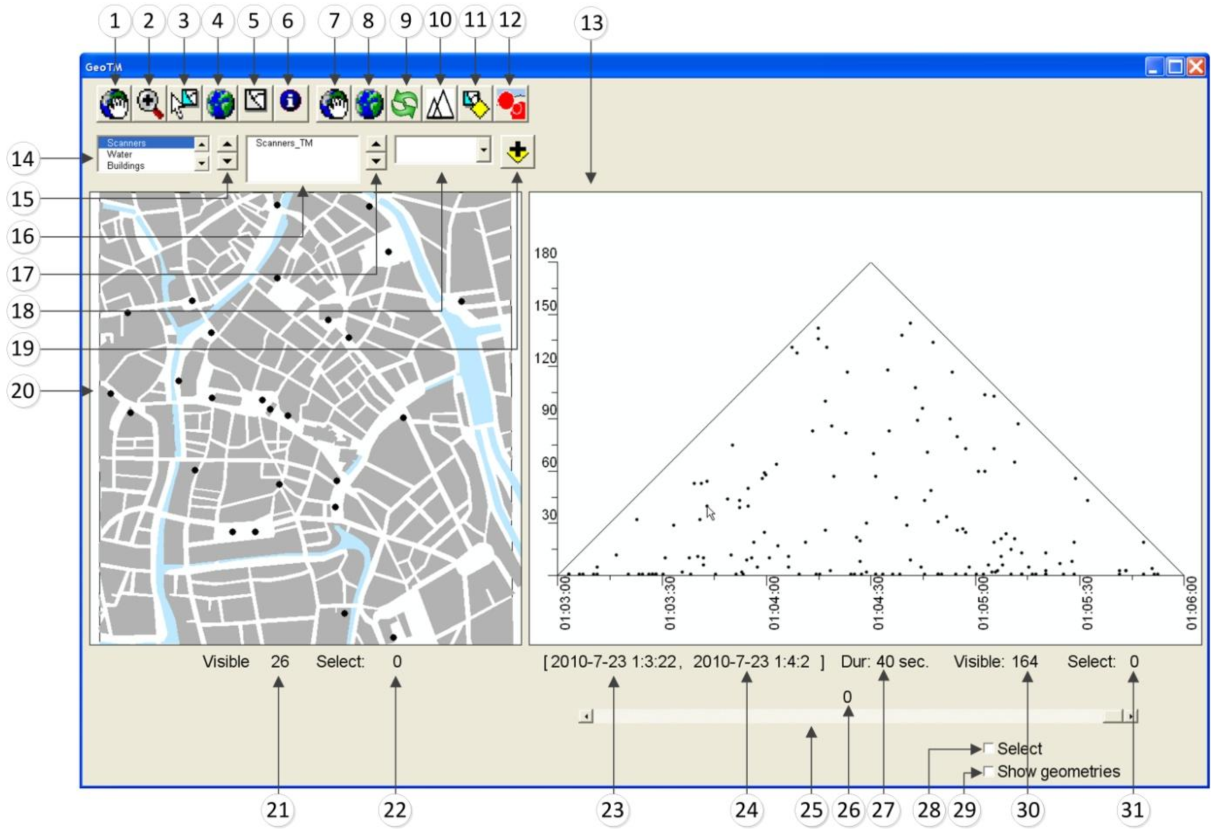


Figure 9-1: The GUI of GeoTM[C]

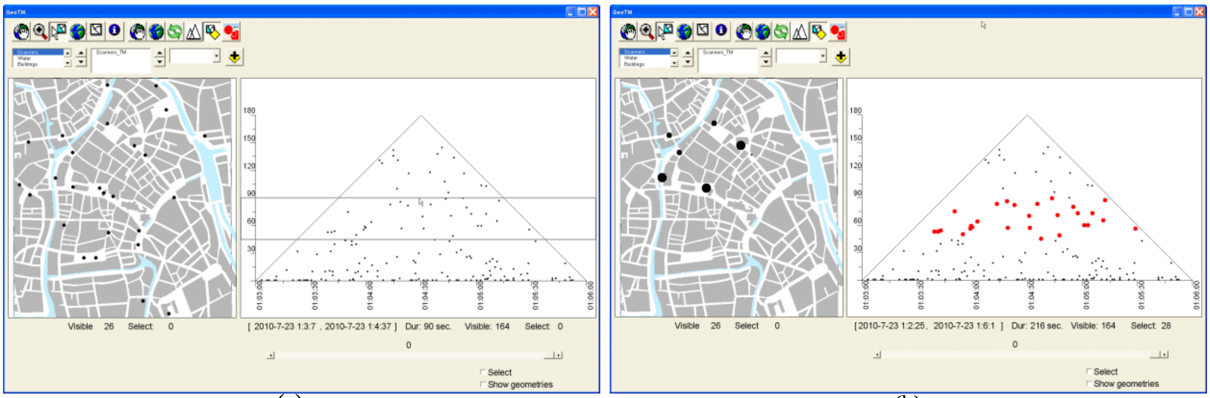
Table 9-1: Functionality of the controls numbered in Figure 9-1

#	Type	Functionality
1	Toggle button	When it is down, one can pan in the map view
2	Toggle button	When it is down, one can zoom in both views by dragging an orthogonal rectangle.
3	Toggle button	When it is down, one can select in both views by dragging an orthogonal rectangle.
4	Button	Return the map view to the full extent.
5	Button	Unselect all.
6	Toggle button	When it is down, one can click on an object to trigger a dialogue that displays its attributes, applying to both views.
7	Toggle button	When it is down, one can pan in the TM view.
8	Button	Return the TM view to the full extent.
9	Button	Refresh both views.
10	Toggle button	*Not applicable in this version.*
11	Toggle button	When it is down, special query tools are active in the TM view.

12	Toggle button	When it is down, sizes of objects in both views become proportional to a certain attribute.
13	Data display	The TM view.
14	List box	Layers in the map view.
15	Spin button	Move up and down the layers in the map view.
16	List box	Layers in the TM view.
17	Spin button	Move up and down the layers in the TM view.
18	Combo box	Select the data type of the layer to be added.
19	Button	Trigger a dialogue to add a layer.
20	Data display	The map view.
21	Text	Display the number of objects displayed in the working layer of the map view.
22	Text	Display the number of selected objects in the working layer of the map view.
23	Text	Display the start point of the interval pointed by the cursor.
24	Text	Display the end point of the interval pointed by the cursor.
25	Slider	Define a time point by moving the bar to a specific position.
26	Text	Display the time point specified by the slider (25).
27	Text	Display the duration of the interval pointed by the cursor.
28	Check box	Select intervals that contain the time point specified by the slider (25).
29	Check box	Display intervals that contain the time point specified by the slider (25).
30	Text	Display the number of objects displayed in the working layer of the TM view.
31	Text	Display the number of objects selected in the working layer of the TM view.

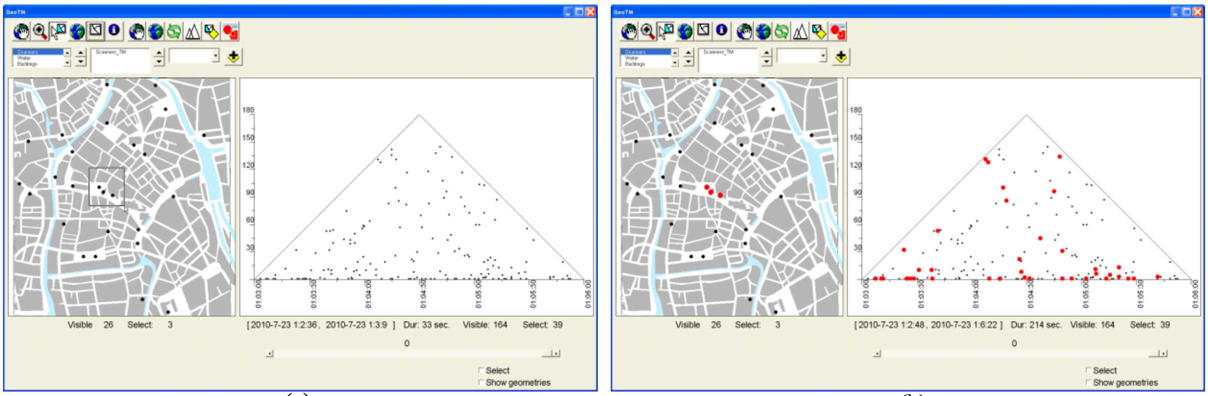
9.3.2 Dynamic link

The map view and the TM view are dynamically linked. As GeoTM^{C} is designed for a dataset from a Bluetooth tracking system (described in Chapter 4), many intervals are referenced to same sensor locations. Thus, the link between items in the map view and the TM view is one-to-many. When selecting intervals in the TM view, the size of locations in the map view updates to represent the number of intervals referenced to these locations (Figure 9-2). When selecting locations in the map view, intervals referenced to these locations become red in the TM view (Figure 9-3). The selection can also be made in the attribute table, and then the map view and the TM view update to display the corresponding selection (Figure 9-4).



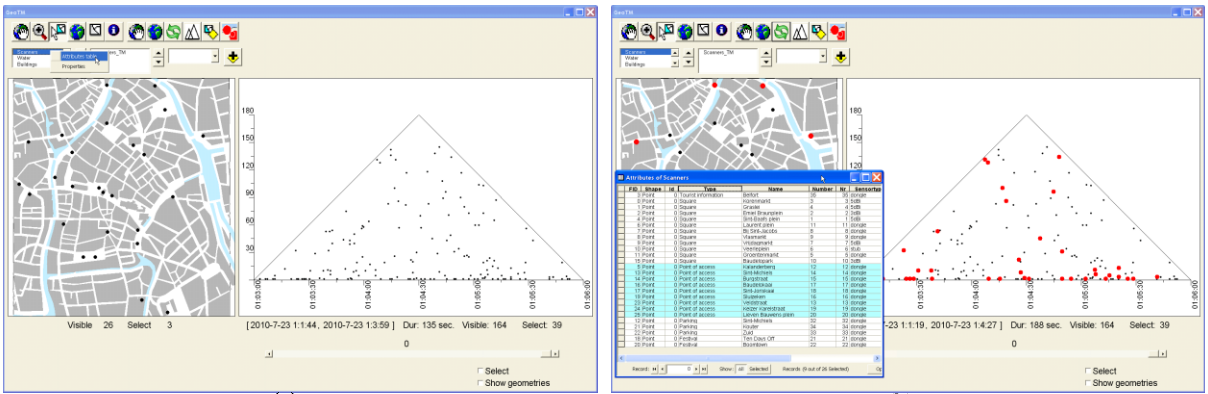
(a) (b)

Figure 9-2: Selection from the TM view. (a) Selecting intervals in the TM view. (b) The sizes of locations in the map view become proportional to the number of selected intervals.



(a) (b)

Figure 9-3: Selection from the map view. (a) Selecting locations in the map view. (b) The intervals referenced to these locations become red in the TM view.



(a) (b)

Figure 9-4: Selection from the attribute table. (a) Triggering the attribute table of a layer. (b) Select items in the attribute table. The map view and the TM view update according to the selection.

9.3.3 Temporal Query

The TM view implements a set of query tools that allow users to formulate temporal queries by creating zones. The fundamental of these query tools was described in Section 4.3 and the manipulation was described in Section 4.4. Note that in Figure 9-6 to 9-8, selected intervals are represented by larger black dots instead of red dots in order to reduce the number of colour page.

Query by clicking

Move cursor to a specific position in the TM, and right click to trigger the menu of Allen relations (Figure 9-5a). When a button of an Allen relation is clicked, intervals that satisfy the relation are selected (Figure 9-5b). When the button of 'Relational zone' is clicked, all relational zones are drawn in the TM (Figure 9-5c).

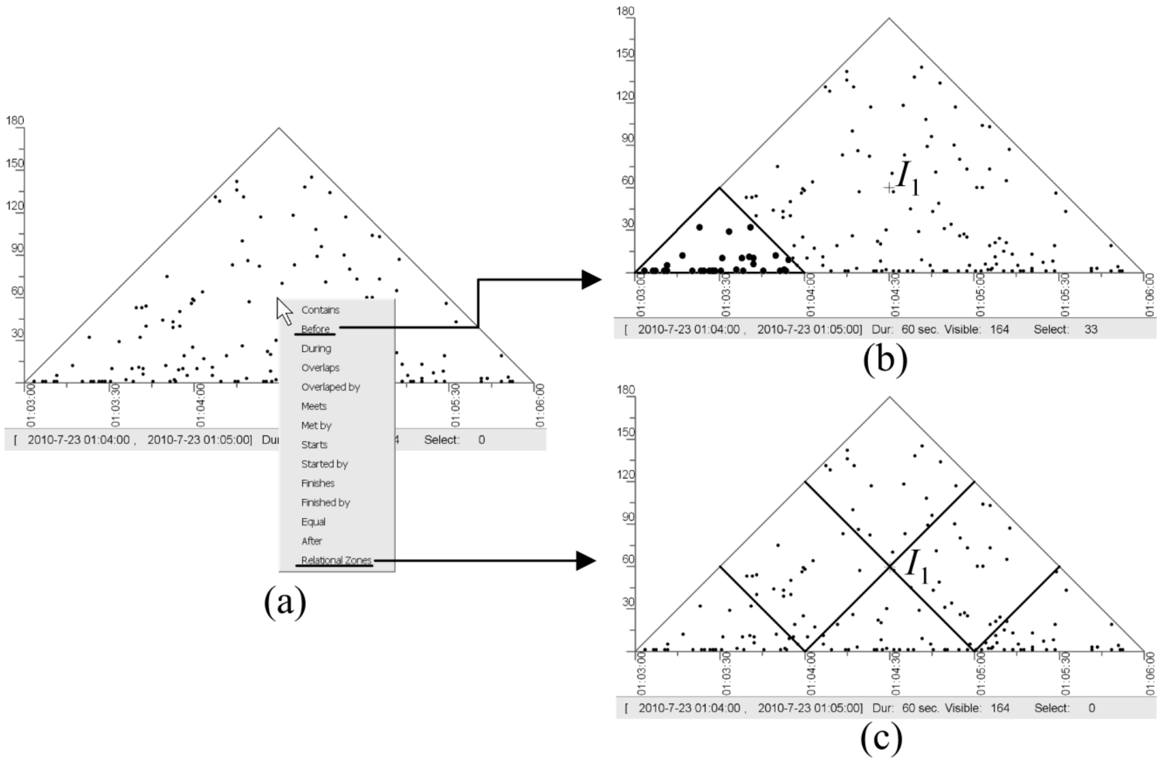


Figure 9-5: Query by clicking. (a) Click on a point in the TM to trigger the menu of Allen relations. (b) Click the 'Before' button of Allen relation to select intervals before the indicated interval. (c) Click the 'Relational zones' button to draw all relational zones in the TM.

Query by dragging zones

When the button (11) is down, special dragging tools are activated. The shape of dragged zone is determined by pressing down specific keys on the keyboard. The default dragging box is a non-orthogonal rectangle that selects intervals *in-between* two time intervals (Figure 4-10a). With the *I* key down, a triangle can be dragged on the horizontal axis to select intervals *in-between* two time instants (Figure 4-10b). With the *D* key down, a range along the vertical dimension can be dragged to select intervals with certain durations (Figure 4-10c). With the *S* key down, a range in 45° to the horizontal axis can be dragged to select intervals that *start-within* an interval. With the *E* key down, a range in -45° to select intervals that *end-within* an interval (Figure 4-10d). The formal definition of the relation *in-between*, *start-within* and *end-within* are described in Section 4.3.

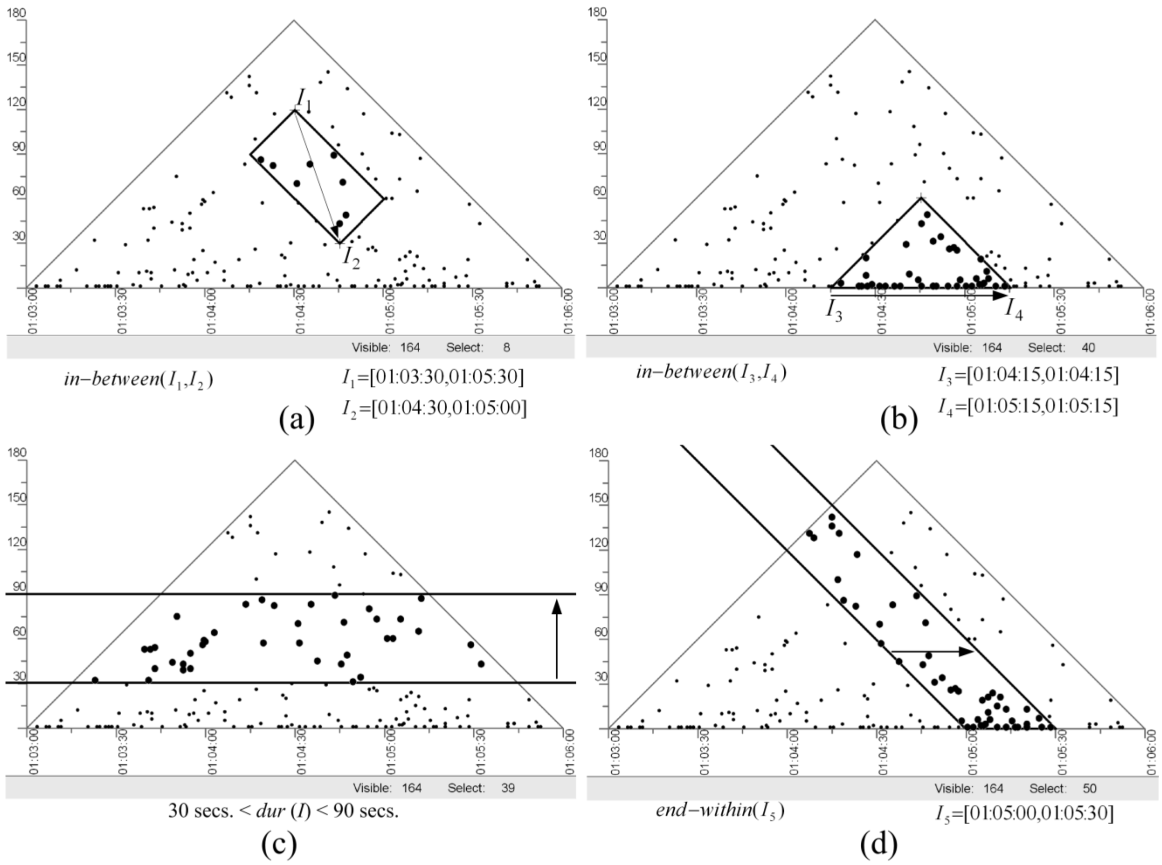
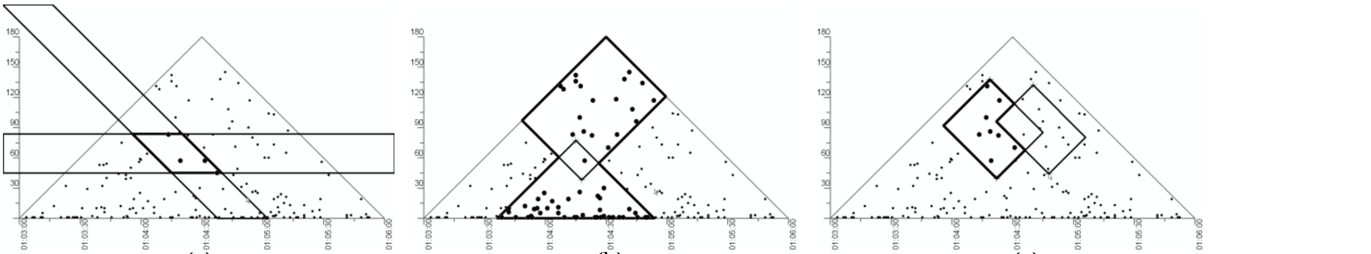


Figure 9-6: Queries by dragging zones in TM. (a): selecting intervals in a convex set *in-between* two time intervals. (b): selecting intervals in a convex set *in-between* two time instants. (c): selecting intervals with certain durations. (d): selecting intervals *end-within* an interval.

Composite query

Queries can be composed in the TM view. The operator that combines the query is decided by a pressing down a specific key on the keyboard. With the *Ctrl* key down, the latter query intersects the previous query (Figure 9-7a). With the *Shift* key down, the latter query unions the previous query (Figure 9-7b). With the *Alt* key down, the latter query subtract the previous query (Figure 9-7c). The combined query zone is highlighted by wider boundaries. These keys can be used in combination with the keys that specify the shape of dragged zones.

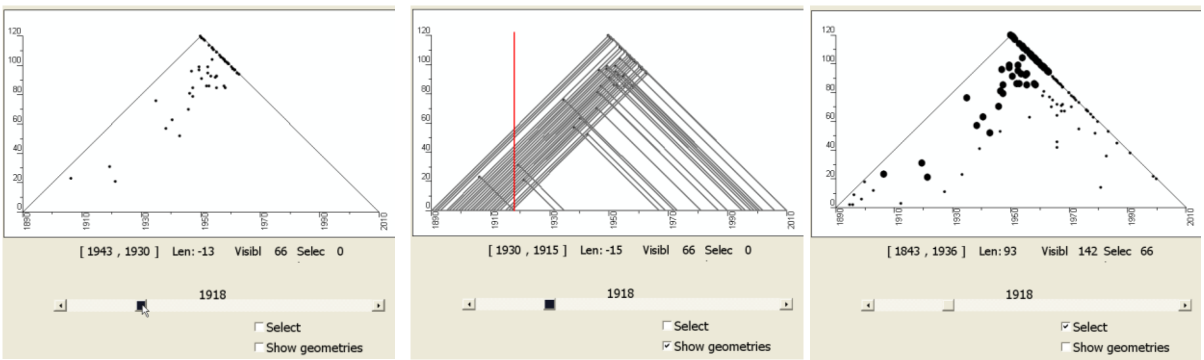


(a) (b) (c)

Figure 9-7: Composite queries. (a) Query intersection. (b) Query union. (c) Query subtraction.

Query by slider

By moving a slider (25), users can turn off the other intervals and only display the intervals that contain the time point specified by the slider (Figure 9-8a). If (27) is checked, assistant lines are drawn for the visible points (Figure 9-8b). If (28) is checked, all intervals are turned on and the intervals that contain the time point specified by (25) are selected (Figure 9-8c).



(a) (b) (c)

Figure 9-8: Specify a time point by the slider (25), (a): only intervals containing time point are visible; (b) assistant lines of the visible intervals are drawn; (c): intervals containing the time point are selected.

9.4 GeoTM for Rough Time Interval

9.4.1 Graphic User Interface

The GUI of GeoTM[R] is similar to that of GeoTM[C], except several additional controls that deal with the probability of temporal queries (Figure 9-9). The functionality of these controls is described in Table 9-2, where the controls in bold font tackle the query probability.

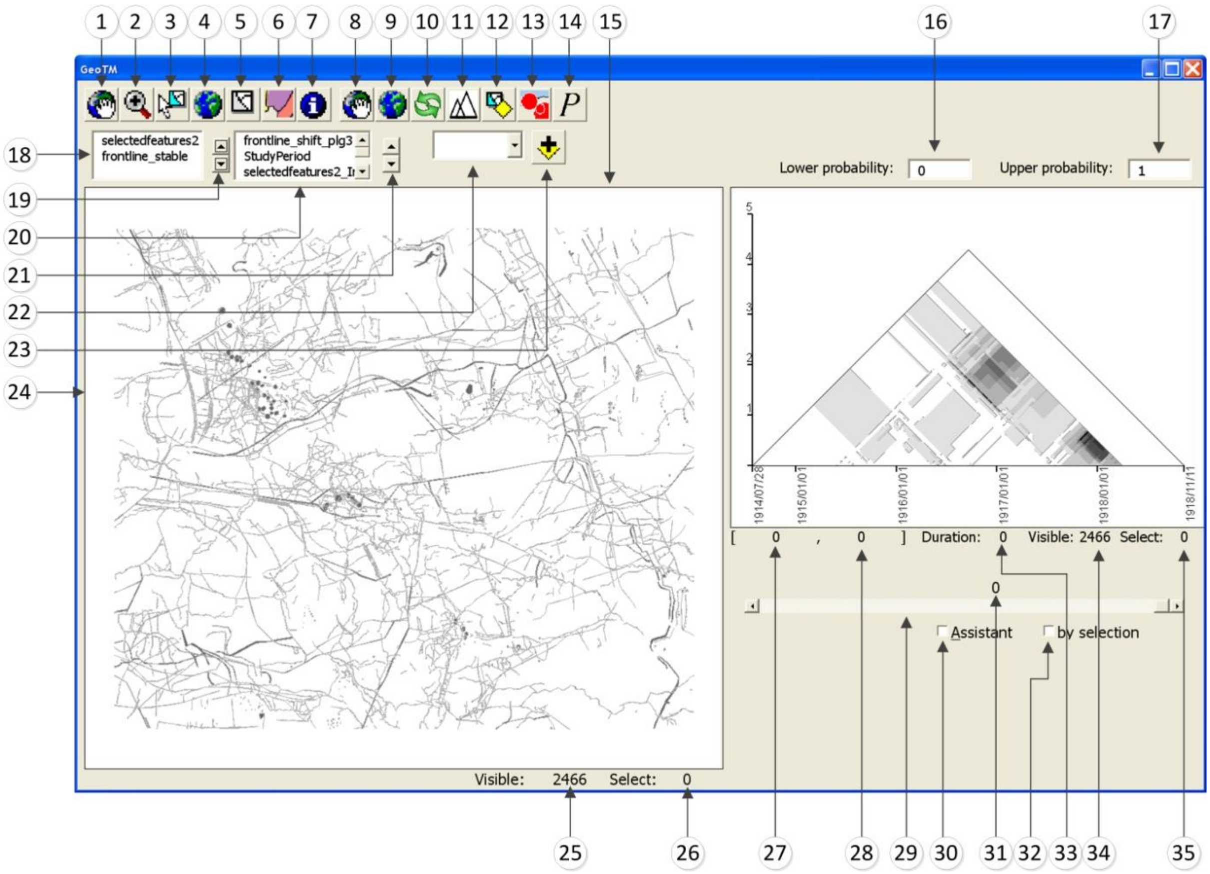


Figure 9-9: The GUI of the GeoTM[R].

Table 9-2: Functionality of the controls numbered in Figure 9-9.

#	Type	Functionality
1	Toggle button	When it is down, one can pan in the map view
2	Toggle button	When it is down, one can zoom in by dragging an orthogonal rectangle in both views.
3	Toggle button	When it is down, one can select by dragging an orthogonal rectangle in both views.
4	Button	Return the map view to the full extent.
5	Button	Unselect all.

6	Toggle button	*Not applicable in this version.*
7	Toggle button	When it is down, one can click on an object to trigger a dialogue that displays its attributes, applying to both views.
8	Toggle button	Pan in the TM view.
9	Button	Return the TM view to the full extent.
10	Button	Refresh both views
11	Toggle button	*Not applicable in this version.*
12	Toggle button	When it is down, special query tools are active in the TM view.
13	Toggle button	*Not applicable in this version.*
14	**Toggle button**	**Export the probability distribution of the query in the TM view to a txt file.**
15	Data display	The map view.
16	**Text box**	**The lower threshold of the probability of the query in the TM view.**
17	**Text box**	**The upper threshold of the probability of the query in the TM view.**
18	List box	Layers in the map view.
19	Spin button	Move up and down the layers in the map view.
20	List box	Layers in the TM view.
21	Spin button	Move up and down the layers in the TM view.
22	Combo box	Select the data type of the layer to be added.
23	Button	Trigger a dialogue to add a layer.
24	Data display	The map view.
25	Text	Display the number of objects displayed in the working layer of the map view.
26	Text	Display the number of selected objects in the working layer of the map view.
27	Text	Display the start point of the interval pointed by the cursor.
28	Text	Display the end point of the interval pointed by the cursor.
29	Slider	*Not applicable in this version.*
30	Check box	*Not applicable in this version.*
31	Text	*Not applicable in this version.*
32	Text	*Not applicable in this version.*
33	Text	Display the duration of the interval pointed by the cursor.
34	Text	Display the number of objects displayed in the working layer of the map view.
35	Text	Display the number of objects selected in the working layer of the map view.

9.4.2 Dynamic link

GeoTM^R is designed for a dataset of military features, which is described in Chapter 5. In this dataset, every feature has a unique rough time interval and geographical location. Thus, the dynamic link in GeoTM^R is one-to-one. In other words, if an item is selected in one view, a unique item is selected in the other view. In Figure 9-10, when a set of rough time intervals are selected in the TM view, a same number of corresponding items are selected in the map view and attribute table.

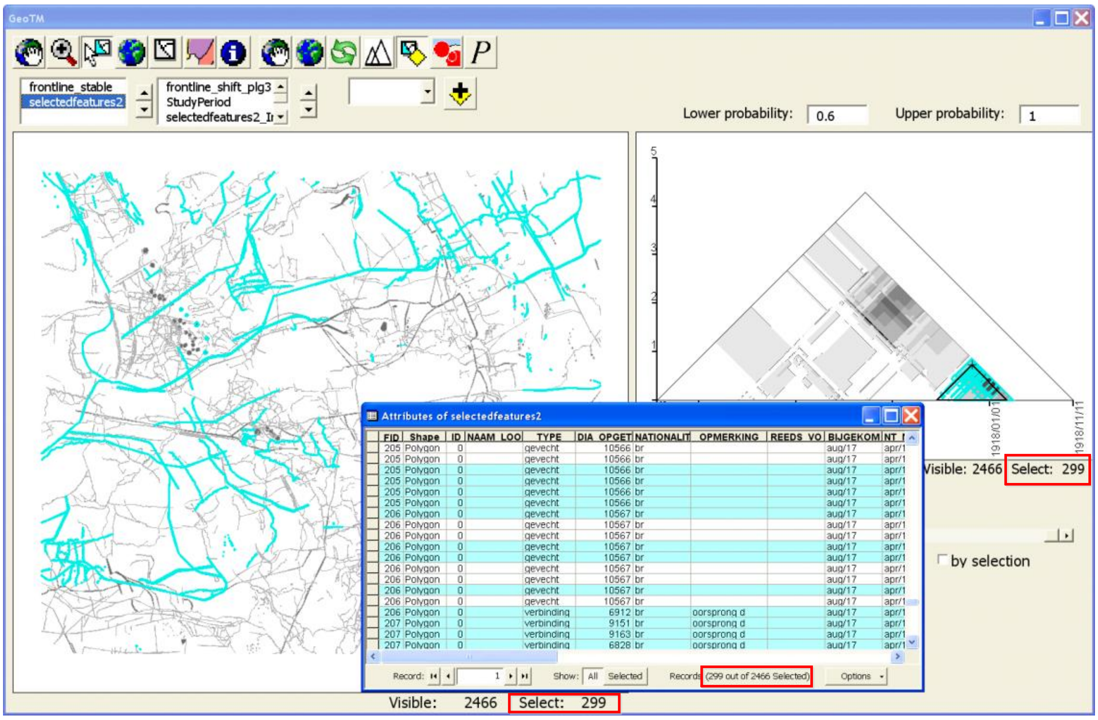
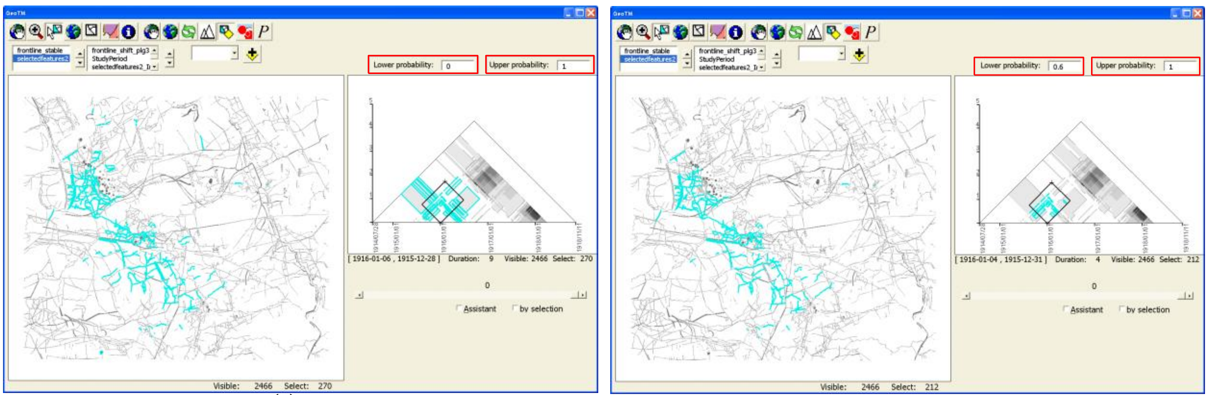


Figure 9-10: The one-to-one link in GeoTM^R



(a) (b)

Figure 9-11: A temporal query with different probability thresholds.

9.4.3 Temporal query with probability

The same set of graphic query tools in $GeoTM^C$ are implemented in $GeoTM^R$, except the queries by the slider bar (29). However, additional controls based on the probabilistic framework introduced in Section 5.3 are applied to these tools. On the one hand, the upper and lower probability thresholds can be set by inputting numbers between zero and one into the text boxes (16) and (17). This allows users to screen out intervals that do not meet the probability thresholds. Figure 9-11 illustrates a query with different probability thresholds applied. With the toggle button (14) down, $GeoTM^R$ can generate the probability distribution of a query, indicating the probabilities that specific numbers of intervals satisfy the query. This probability distribution is exported to a table (a txt file) containing two columns of numbers. One column indicates the number of intervals, while the other indicates the probability that this number of intervals satisfy the query.

9.5 Scalability

As GeoTM is developed in the same infrastructure as ArcGIS 9.3, the scalability of most functionalities of GeoTM can be referred to the scalability of ArcGIS 9.3. Most tools are running fluently with the datasets of the use cases. The scalability problem of $GeoTM^C$ is slow plotting of the around 150,000 interval points of the Bluetooth tracking dataset in the TM view. Plotting all these points takes around 10 seconds, which happens after refreshing the TM view and turning the TM view to the full extent. As $GeoTM^R$ implements algorithms to calculate the probability of temporal queries, selections in the TM view with non-default probability thresholds are slower than normal selections in ArcMap, which take 3 to 5 seconds extra processing time within the datasets introduced in Chapter 5 (containing 2466 rough time intervals). The processing time for generating the probability distribution table is dependent on several parameters, which takes 0 to 30 minutes within the dataset. These scalability facts are based on a desktop PC with following specifications:

Processors: Intel Pentium (R) Dual CPU E2160 @ 1.80GHz

Memory: 2.0 GB

Hard disk: 80 GB SATA

Graphic card: NVIDIA GeoForce 8400 GS

Operating system: Windows XP SP3

9.6 Limitation

The current versions of GeoTM have implemented the functionality that deal with crisp time interval and rough time interval. It can generally analyse geographical data with referenced to these two types of time intervals. Minor adaption may be needed for datasets with different inner structures (e.g. links between spatial and temporal data) from that of the use case datasets. Preferably a dialogue should be added in the future version to adapt these differences. Also, the queries tool based on probability needs to be further optimized to make the response time minimal. ArcObjects has allowed a wide range of ArcGIS operators and controls to be called in GeoTM, which has saved a lot of time and energy for bottom development. However, this platform prohibits functionality beyond ArcGIS to be developed. For example, the TM view is based on the map control in ArcObjects, which only support limited functionality for graphic visualization and manipulation. The transparency adjustment applies to layers rather than objects in layers. Also, it limits the flexibly of creating and adapting query zones. To further exploit the use of the TM, a more flexible development environment needs to be considered, where more interactive and complicated tools can be applied to the TM. The burgeoning development of open source GIS may provide feasible solutions.

SAMENVATTING (DUTCH SUMMARY)

Tijd is een alomtegenwoordig concept dat aanwezig is in elk aspect van ons dagelijks leven. Reeds enkele decennia, vanaf het moment dat computers een significante plaats innemen in het werk en in leven van mensen, blijkt dat het omgaan met tijd cruciaal is binnen de informatiewetenschappen en artificiële intelligentie. Enerzijds beseften computerwetenschappers het belang van het formaliseren van algemene kennis over tijd zodat het trekken van gevolgen en het redeneren over tijd kon uitgevoerd worden door computers, wat leidde tot het nieuwe onderzoeksdomein van temporeel redeneren. Anderzijds werden er veel visualisatiebenaderingen ontwikkeld om een interactieve analyse van temporele informatie mogelijk te maken. Het merendeel van dit onderzoek is gebaseerd op een lineair concept van de tijd, waarbij een tijdsinterval voorgesteld wordt als een lineair segment langs de tijdslijn. Dit concept brengt echter enkele inherente problemen met zich mee rond het voorstellen van vage tijdsintervallen en de analyse van tijdsintervallen. Om deze moeilijkheden te overkomen, werd een alternatieve voorstelling van tijdsintervallen in overweging genomen, welke in dit proefschrift het Triangulaire Model wordt genoemd. Het Triangulaire Model is gebaseerd op het werk van Ligozat en Kulpa (Kulpa, 2006; Ligozat, 1997) en stelt tijdsintervallen voor als punten in een tweedimensionale ruimte in plaats van lineaire segmenten. Dit proefschrift heeft als doel om de bruikbaarheid van het Triangulaire Model te onderzoeken, zowel voor het temporeel redeneren als voor de visuele analyse van tijdsintervallen. Volgende vier onderzoeksvragen (OV) werden onderzocht:

OV 1: Hoe kan het Triangulaire Model het onderzoek naar het redeneren van onvolmaakte tijdsintervallen vergemakkelijken?

OV 2: Welke meerwaarde heeft het Triangulaire Model voor de analyse van tijdsintervallen?

OV 3: Hoe kan het Triangulaire Model gebruikt worden voor de analyse van lineaire data?

OV 4: Kan het Triangulaire Model begrepen en gebruikt worden door niet-experten?

In de rest van deze samenvatting wordt besproken hoe deze onderzoeksvragen in dit proefschrift onderzocht werden.

Na de algemene inleiding in Hoofdstuk 1, wordt in Hoofdstuk 2 een overzicht gegeven van de belangrijkste literatuur in de domeinen van temporeel redeneren en temporele visualisatie, welke dicht aanleunen bij het onderzoek van dit proefschrift. Er komt een uitvoerige discussie aan bod over de huidige trends en ontwikkelingen en toekomstige uitdagingen in deze domeinen. Eveneens wordt de basis gelegd voor het besluit en wordt vermeldt hoe dit proefschrift past in het grotere geheel van eerdere onderzoeken.

In Hoofdstuk 3 wordt OV 1 behandeld. Twee veel gebruikte benaderingen om onvolmaakte tijdsintervallen te modelleren zijn de *rough set* theorie en de *fuzzy set* theorie. Vooral aan de *fuzzy* benadering werd reeds veel aandacht geschonken (Nagypál and Motik, 2003; Ohlbach, 2004; Schockaert and De Cock, 2008; Schockaert et al., 2008). Reeds uitgevoerd onderzoek naar de *ruwe* benadering is echter vaak onvolledig en in bepaalde aspecten vrij gelimiteerd (Bassiri et al., 2009; Bittner, 2002). In dit hoofdstuk wordt eerst een alledaagse situatie besproken waarbij de *rough set* theorie geschikter blijkt voor het modelleren van tijdsintervallen dan de *fuzzy set* theorie. Deze situatie vloeit voort uit een discrete data-acquisitie waarbij er een onzekerheid bestaat tussen de observaties op twee discrete tijdstippen. Teledetectie wordt in dit hoofdstuk als voorbeeld genomen, waarbij er beelden of foto's genomen worden op verschillende discrete tijdstippen (*snapshots*). Met deze snapshots kunnen de tijdsintervallen van geobserveerde entiteiten benaderd worden als lineaire *ruwe sets*, wat leidt tot het concept van *ruwe tijdsintervallen*.

Ten tweede wordt in Hoofstuk 3 het Triangulaire Model gebruikt als schematisch hulpmiddel om de redeneringseigenschappen van *ruwe tijdsintervallen* te onderzoeken. Het nut van zo'n schematisch hulpmiddel werd al in verscheidene onderzoeken rond temporeel en ruimtelijk redeneren aangetoond (Kulpa, 1994). Als gevolg van de tripartiete structuur van *ruwe tijdsintervallen* (de bovenste en onderste benaderingen, en de grensgebieden), kan het gebruik van de conventionele lineaire voorstelling om de relaties van *ruwe tijdsintervallen* te bestuderen heel ingewikkeld en foutgevoelig zijn. Daartegenover is de polygonale voorstelling van *ruwe tijdsintervallen* in het Triangulaire Model veel eenvoudiger voor de menselijke perceptie. In dit hoofdstuk wordt getoond hoe de temporele relaties tussen *ruwe tijdsintervallen* visueel herkend kunnen worden volgens de topologieën van meetkundige figuren. Daarnaast wordt ook een kader van methoden ontwikkeld om de semantiek achter de grafische voorstellingen te interpreteren. Ten slotte wordt er ook gedemonstreerd hoe temporele bevragingen met ruwe restricties gemodelleerd kunnen worden als tweedimensionale ruwe zones, en wordt ook de compositie van verschillende bevragingen (bv. intersectie en unie) beschreven.

Ten derde wordt in Hoofdstuk 3 onderzocht hoe *fuzzy tijdsintervallen* voorgesteld kunnen worden in het Triangulaire Model. Door de continue structuur van *fuzzy tijdsintervallen*, worden de relaties van deze tijdsintervallen uitgedrukt door de overlap van *fuzzy* gebieden, wat voor de mens moeilijk waarneembaar en te verwerken is. Daarom werd het redeneren rond *fuzzy tijdsintervallen* niet zo grondig bestudeerd als rond *ruwe tijdsintervallen*. Er wordt echter wel aangetoond dat temporele bevragingen met *fuzzy* restricties voorgesteld kunnen worden als continue velden, welke potentieel gebruikt kunnen worden voor het in kaart brengen van een probabilistische of possibilistische dichtheid wanneer de *fuzzy tijdsintervallen* gebaseerd zijn op probabilistische of possibilistische modellen.

Volgens de bevindingen in Hoofdstuk 3 kan er gesteld worden dat het Triangulaire Model een uitstekend schematisch hulpmiddel is bij het onderzoek naar kwalitatief redeneren rond intervallen. Het uitbreiden van de tijdsruimte van 1D naar 2D ondersteunt een groot aantal voorstellingen van temporele concepten. Het grafische karakter van het Triangulaire Model laat mensen toe om abstracte kennis op een meer waarneembare manier te begrijpen en te bestuderen

In Hoofdstukken 4 en 5 wordt respectievelijk getoond hoe het Triangulaire Model gebruikt kan worden bij de analyse van scherpe (*crisp*) en *ruwe tijdsintervallen* (OV 2). In het Triangulaire Model worden scherpe (*crisp*) tijdsintervallen uitgezet als punten in een 2D-ruimte. Bijgevolg wordt een reeks intervallen voorgesteld door een vast punt. Dit maakt het mogelijk om grote hoeveelheden tijdsintervallen visueel te analyseren. In de dataset van de casestudie werden enkele clusters geobserveerd in de puntstructuur, wat mogelijk grote bewegingen van de menigte betekent. Ook maakt de coördinaatruimte het mogelijk om vergelijkingen te maken tussen de reeksen intervallen. Het verschil van de intervalverdelingen kan rechtstreeks waargenomen worden uit de verschillen in de puntstructuren, wat niet mogelijk is met de traditionele lineaire voorstelling. Anderzijds wordt een vernieuwend temporeel bevragingsmechanisme voorgesteld, welke steunt op de manipulatie van geometrieën in de intervalvoorstelling. In dit mechanisme kunnen verschillende temporele bevragingen geformuleerd worden door specifieke zones te creëren in de 2D-ruimte en de combinatie van bevragingen (bv. unies en intersecties) wordt uitgedrukt door de ruimtelijke operaties van de bevragingszones. Deze *query-in-display* benadering is gelijkaardig aan de *brushing* techniek die gebruikt wordt in geografische kaarten of spreidingsplots (Monmonier, 1990), die de interactiviteit van de visualisatie verhoogt.

Hoofdstuk 5 is gebaseerd op de onderzoeksresultaten uit de Hoofdstukken 3 en 4, maar de nadruk ligt hier op de analyse van *ruwe tijdsintervallen*. In plaats van gebruik te maken van een puntstructuur worden *ruwe tijdsintervallen* voorgesteld als een polygonale structuur. Deze voorstelling geeft een expliciet beeld van de verdeling van *ruwe tijdsintervallen*, wat ook gebruikt kan worden bij visuele patroonherkenning. Het bevragingsmechanisme dat in Hoofdstuk 4 aan bod komt, kan eveneens toegepast worden op *ruwe tijdsintervallen*. Er ontstaan echter onzekerheden wanneer de bevragingszone de polygoon van het *ruwe tijdsinterval* deels bedekt. Daarom werd een probabilistisch kader ontwikkeld. Het basisidee hier is dat de probabiliteit waarmee een ruw tijdsinterval aan een bevraging voldoet gelijk is aan de grootte (concreet: de verhouding) van de polygoon van het *ruwe tijdsinterval* bedekt door de bevragingszone. Binnen dit kader kunnen temporele bevragingen ofwel uitgevoerd worden met een probabilistische limietwaarde ofwel door het geven van antwoorden rekening houdend met probabiliteiten.

De datavisualisatie en het bevragingsmechanisme voor scherpe (*crisp*) en *ruwe tijdsintervallen* werd geïmplementeerd in software prototypes waarbij de gegevensvoorstelling van het Triangulaire Model opgenomen is in een geografisch informatiesysteem (GIS). In deze prototypes kunnen observaties en bevragingen uitgevoerd worden voor zowel ruimtelijke als temporele aspecten. Ook werden de functionaliteiten van het probabilistisch kader verwerkt in de versie voor *ruwe tijdsintervallen*. De prototypes werden toegepast voor de analyse van twee realistische datasets, waarbij men op interactieve wijze de tijdruimtelijke data kan analyseren in zowel een kaart- als een tijdsbeeld. Hierbij kunnen interessante fenomenen ontdekt worden die met traditionele analytische hulpmiddelen moeilijk te vinden zijn. Volgens de bevindingen kan er besloten worden dat een dergelijke toepassing van het Triangulaire Model de analyse van interval-gebaseerde geografische data kan ondersteunen.

Bij de eerste hoofdstukken ligt de focus op discrete tijdsintervallen. Hoofdstuk 6 beschrijft hoe continue lineaire data (bv. tijdsreeksen en verkeersgegevens van een weg) in het Triangulaire Model voorgesteld kunnen worden (OV 3). Het basisidee is om de waarde van ieder deelinterval van de lineaire data toe te wijzen aan de 2D intervalruimte van het Triangulaire Model, welke een continu veld van nummers vormt. Enerzijds biedt deze voorstelling een overzicht van de lineaire data in alle verschillende schalen. Anderzijds kunnen de ruimtelijke analysemethoden (zoals kaartalgebra en cartografisch modelleren) erop toegepast worden omdat de 2D velden gebaseerd zijn op een coördinatenruimte van intervallen. Met deze methoden kan men meervoudige lineaire data vergelijken en de meest geschikte intervallen vinden inzake relevante factoren en restricties. Realistische datasets worden gebruikt om de toepassing van deze voorstelling bij de analyse van verkeersgegevens en de selectie van surflocaties en periodes te demonstreren.

Hoofdstukken 3 tot 5 veronderstellen een groot potentieel om het Triangulaire Model te gebruiken bij het redeneren over en de analyse van tijdsintervallen. Er is echter nog steeds een gebrek aan feedback van niet-experten inzake de begrijpelijkheid en bruikbaarheid. Hoofdstuk 7 beschrijft een empirisch onderzoek waarbij geëvalueerd werd of proefpersonen zonder voorkennis dit model gemakkelijk kunnen leren en er gebruik van kunnen maken (OV 4). In dit onderzoek wordt een reeks vragen over intervalrelaties en -proporties door de proefpersonen beantwoord nadat ze een training kregen van 20 minuten. Het resultaat toont dat de proefpersonen de vragen correcter en efficiënter beantwoorden als ze gebruik maken van het Triangulaire Model in plaats van met de lineaire voorstelling, wat aangeeft dat het Triangulaire Model de visuele observatie van intervalrelaties en -proporties beter ondersteunt. Ook beschouwen de meeste proefpersonen het Triangulaire Model als eenvoudiger dan de lineaire voorstelling met betrekking tot de gestelde vragen. Bovendien wordt aangetoond dat een toenemend aantal intervallen dat gepresenteerd wordt in het Triangulaire Model geen invloed heeft op de correctheid van de antwoorden en de tijd die de proefpersonen nodig hebben om de vragen op te lossen. Dit betekent dat het Triangulaire Model geschikt is om met grote hoeveelheden intervallen te werken. De resultaten van deze empirische studie bevestigen de verdiensten van het Triangulaire Model in het visualiseren van tijdsintervallen en tonen aan dat het door een breder publiek van niet-experten kan gebruikt worden.

In Hoofdstuk 8 worden de resultaten van dit proefschrift besproken in het licht van de bredere onderzoeksresultaten. Hierbij worden problemen en toekomstige uitdagingen duidelijk vermeld. Ten eerste wordt het belang van intervalanalyse besproken en komt er aan bod hoe de benaderingen uit dit proefschrift de huidige tekortkomingen ingevuld hebben. Ten tweede worden de algemene strategieën voor het modelleren van onvolmaakte tijdsintervallen uitgewerkt, waarbij vermeld wordt waarom deze voorstelling uitstekend geschikt is bij dit modelleren. Ten derde wordt de waarde van de combinatie van bevragingshulpmiddelen met datavisualisatie in de visuele analyse bediscussieerd. Ten slotte wordt erop gewezen dat verder onderzoek noodzakelijk is om een algemene richtlijn voor empirisch onderzoek te ontwikkelen om visuele analytische benaderingen te evalueren.

Referenties

Bassiri, A., Malek, M., Alesheikh, A. and Amirian, P. (2009). Temporal Relationships between Rough Time Intervals. ed. by Gervasi, O., Taniar, D., Murgante, B., Laganà, A., Mun, Y. and Gavrilova, M., in Computational Science and Its Applications – ICCSA 2009, pp.543-552, Springer Berlin / Heidelberg.

Bittner, T. (2002). Approximate Qualitative Temporal Reasoning. Annals of Mathematics and Artificial Intelligence, 36(1), pp.39-80.

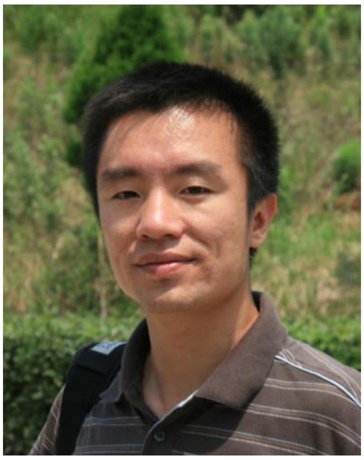
Nagypál, G. and Motik, B. (2003). A Fuzzy Model for Representing Uncertain, Subjective, and Vague Temporal Knowledge in Ontologies. Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics, (ODBASE), Catania, Italy, pp.906-923.

Ohlbach, H. J. (2004). Relations between fuzzy time intervals. Proceedings of 11th International Symposium on Temporal Representation and Reasoning (TIME 2004), Montréal, Québec, pp.44-51.

Schockaert, S. and De Cock, M. (2008). Temporal Reasoning about Fuzzy Intervals. Artificial Intelligence, 172(8-9), pp.1158-1193.

Schockaert, S., De Cock, M. and Kerre, E. E. (2008). Fuzzifying Allen's Temporal Interval Relations. IEEE Transactions on Fuzzy Systems, 16(2), pp.517-533.

Biographical sketch



Yi Qiang (°1983, China) started his education of geographical information science (GIS) in 2002 at Beijing Normal University, where he received a Bachelor degree in 2006. After this program, he has accomplished a Master program of GIS in the University of Edinburgh. Later, he has worked as a research assistant at the UK e-Science institute for a short project about Semantic Web and ontological modelling. In 2007, he received a grant from the Research Foundation – Flanders (FWO) to start a PhD in Ghent University. His PhD Research focuses on the use of a two-dimensional time representation in temporal reasoning, information visualization and spatio-temporal analysis. His research outcomes have been published in or submitted to a number of international journals and conference proceedings.