JYVÄSKYLÄ STUDIES IN COMPUTING 24

Vladimir Ryabov

# Handling Imperfect Temporal Relations

Esitetään Jyväskylän yliopiston informaatioteknologian tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston vanhassa juhlasalissa (S212)
joulukuun 20. päivänä 2002 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Information Technology of the University of Jyväskylä,
in Auditorium S212, on December 20, 2002 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2002

# Handling Imperfect Temporal Relations

JYVÄSKYLÄ STUDIES IN COMPUTING 24

Vladimir Ryabov

# Handling Imperfect Temporal Relations

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2002

Editors
Seppo Puuronen
Department of Computer Science and Information Systems, University of Jyväskylä
Pekka Olsbo, Marja-Leena Tynkkynen
Publishing Unit, University Library of Jyväskylä

URN:ISBN 9513913597
ISBN 951-39-1359-7 (PDF)

ISBN 951-39-1371-6 (Nid.)
ISSN 1456-5390

# ABSTRACT

Ryabov, Vladimir
Handling Imperfect Temporal Relations
Jyväskylä: University of Jyväskylä, 2002, 75 p. (+included articles)
(Jyväskylä Studies in Computing,
ISSN 1456-5390; 24)
ISBN 951-39-1359-7
Finnish summary
Diss.

In many areas of Artificial Intelligence (AI) there is a need to represent temporal information and reasoning about it. Temporal formalisms are applied in all areas where the time course of events plays an important role, for example, in temporal databases, process control, planning, natural language understanding, and in medical and reservation systems. Similar to almost all the information we have about the real world this temporal information is imperfect. In this thesis we propose a numerical formalism, based on probability theory, for handling imperfect temporal relations. The research problem includes three main issues: the representation of imperfect relations, the estimation of measures of imperfection, and reasoning about imperfect relations. An imperfect temporal relation between two primitives (points or intervals) is represented by the probabilities of the basic relations ("<", "=", and ">" for points, and thirteen Allen's relations for intervals) between these primitives. These probability values are calculated by the proposed formulas taking into account the information about the primitives. We further assume that the measurements of the temporal values of two primitives may include some measurement error, which needs to be taken into consideration during the estimation. Taking into account the maximum value of the measurement error, we derive the lower and upper probabilities of the basic relations between two primitives. The mechanism for reasoning about imperfect relations between temporal points includes four operations: inversion, composition, addition, and negation. The study presented in this thesis is constructive and includes an observation part and theory building. The main contribution of this work is a new formal technique for the representation of, estimation of, and reasoning about imperfect temporal relations. Finally, we proposed the application of the formalism to the area of temporal diagnostics (medical and industrial).

Keywords: temporal representation and reasoning, temporal relation, uncertainty, probability

**ACM Computing Review Categories**

**I.2.3.** Computing Methodologies: Artificial Intelligence, *Deduction and theorem proving, Uncertainty, ``fuzzy,'' and probabilistic reasoning*
**I.2.4.** Computing Methodologies: Artificial Intelligence, *Knowledge representation formalisms and methods, Temporal logic*
**J.3.** Computer Applications: Life and Medical Sciences, *Medical information systems*
**J.7.** Computer Applications: Computers in other systems, *Industrial control*


**Author** Vladimir Ryabov
University of Jyväskylä
Department of Computer Science and Information Systems
P.O. Box 35, FIN-40351, Jyväskylä, Finland
Email: vlad@it.jyu.fi

**Supervisors** Professor Seppo Puuronen
Department of Computer Science and Information Systems
University of Jyväskylä, Jyväskylä, Finland

Professor Vagan Terziyan
Department of Mathematical Information Technology
University of Jyväskylä, Jyväskylä, Finland

**Reviewers** Professor Peter van Beek
School of Computer Science
University of Waterloo, Waterloo, Ontario, Canada

Assistant Professor Curtis Dyreson
School of Electrical Engineering and Computer Science
Washington State University, Pullman, Washington, USA

**Opponent** Professor Ljudmil Golemanov
Department of Engineering Science
University of Vaasa, Vaasa, Finland

## ACKNOWLEDGEMENTS

I would like to thank all people who helped and supported me during making this research. Especially, I am thankful to my supervisors, Professor Seppo Puuronen (Department of Computer Science and Information Systems, University of Jyväskylä) and Professor Vagan Terziyan (Department of Mathematical Information Technology, University of Jyväskylä), for their guidance, encouragement, and valuable support. Particularly, I am thankful to Professor Terziyan for introducing me to the world of science, teaching me how to make research, spending plenty of his time in discussions during different stages of this work, and for his insightful comments and ideas. I am greatly indebted to Professor Puuronen for his valuable suggestions concerning all the aspects of this research project, for his critical and constructive comments on early drafts of many papers, and for our long discussions, which stimulated and inspired my further work. Moreover, I am thankful to both Professor Puuronen and Professor Terziyan for co-authoring the papers included in this thesis.

My special acknowledgment to COMAS graduate school (University of Jyväskylä), which financially supported this research during more than four years. I thank the grant (#55626) of the Academy of Finland for financial support during the last three months of finishing this dissertation. Finally, I am grateful to InBCT TEKES Project of Academy of Finland (Agora Center, University of Jyväskylä) for their financial support of conference traveling.

I would like to thank Professor Seppo Puuronen and Professor Pasi Tyrväinen for their valuable help with preparation of the Finnish summary. I am very thankful to the Department of Computer Science and Information Systems (University of Jyväskylä), where I started working at August 1997, for their friendly climate, help readiness, excellent working conditions, and their numerous financial support of my traveling to conferences. My especial sincere gratitude and appreciation to Professor Pekka Neittaanmäki, the Head of COMAS Graduate School, for his enthusiasm and anytime readiness to help with any organizational concerns.

I am thankful to Professor Andre Trudel (Acadia University, Canada) for sending me his numerous papers, for our discussions that introduced new vision of several problems, and for his favorable opinion about my work.

Many of my friends morally supported me during writing this dissertation. Especially, I would like to thank greatly my best friend Stephen Lord for our frank relationships, moral support, for helping me with language check of many papers, and for his genuine joy about all my small academic successes during the years of writing this dissertation.

Finally, I am very grateful to my dear parents Viktor Ryabov and Natalia Ryabova for their continuous moral support, for their anytime willingness to help me, and for their firm confidence in my scientific cleverness. I would like to thank sincerely my dearest and beloved wife Eugenia and little daughter Anastasia for their understanding, persistent and invaluable assistance, and for their open-hearted encouragement. I feel lucky to have such a supportive and loving family.

**CONTENTS**

## LIST OF INCLUDED ARTICLES

**I** Ryabov, V., Puuronen, S. & Terziyan, V. 1999. Representation and Reasoning with Uncertain Temporal Relations. In A. Kumar & I. Russel (Eds.) Proceedings of the 12-th International Florida AI Research Society Conference, Menlo Park, California: AAAI Press, 449-453.

**II** Ryabov, V. 2000. Uncertain relations between indeterminate temporal intervals, In R. Agrawal, K. Ramamritham & T. Vijayaraman (Eds.) Proceedings of the 10-th International Conference on Management of Data, New Delhi, India: Tata McGraw-Hill Publishing Company Limited, 87-95.

**III** Ryabov, V. 2001. Probabilistic estimation of uncertain temporal relations. Colombian Journal of Computation 2 (2), 61-77.

**IV** Ryabov, V. & Puuronen, S. 2000. Estimation of uncertain relations between indeterminate temporal points, In T. Yakhno (Ed.) Proceedings of the 1-st Biannual International Conference on Advances in Information Systems, Lecture Notes in Computer Science 1909, Heidelberg, Germany: Springer-Verlag, 108-116.

**V** Ryabov, V. 2001. Estimating uncertain relations between indeterminate temporal points and intervals. In C. Bettini & A. Montanari (Eds.) Proceedings of the 8-th International Symposium on Temporal Representation and Reasoning (TIME'01), Los Alamitos, California: IEEE Computer Society Press, 69-74.

**VI** Ryabov, V. & Puuronen, S. 2001. Probabilistic reasoning about uncertain relations between temporal points. In C. Bettini & A. Montanari (Eds.) Proceedings of the 8-th International Symposium on Temporal Representation and Reasoning (TIME'01), Los Alamitos, California: IEEE Computer Society Press, 35-40.

**VII** Ryabov, V. 2002. Handling uncertain interval relations, In M. Hamza (Ed.) Proceedings of the 2-nd IASTED International Conference on AI and Applications, Anaheim, Calgary, Zurich: ACTA Press, 291-296.

**VIII** Terziyan, V. & Ryabov, V. (manuscript, accepted), Abstract diagnostics based on uncertain temporal scenarios, International Conference on Computational Intelligence for Modeling Control and Automation (CIMCA 2003), Vienna, Austria.

**IX** Ryabov, V. & Terziyan, V. (manuscript, accepted), Industrial diagnostics using algebra of uncertain temporal relations. IASTED International Conference on Artificial Intelligence and Applications, Innsbruck, Austria.

# PART I:
# OVERVIEW AND SUMMARY

# 1 INTRODUCTION

Representation and reasoning about time is important in modeling dynamic aspects of the world. Time plays a critical role, and cannot be considered of secondary relevance, as in most natural language systems or hidden in a search process, as in most problem-solving systems (Allen & Fergusson, 1994). Indeed, there are many AI applications, where the time course of events plays a crucial role, for example, in temporal databases (Jensen & Snodgrass, 1999; Tsortas & Kumar, 1996), scheduling (Bouzid & Mouaddib, 1998), and in medical diagnostic systems (Combi & Shahar, 1997; Chountas & Petrounias, 2000). During the last several decades plenty of papers have been published in the area of temporal representation and reasoning, in which many formalisms have been proposed. Nevertheless, there still exist topics within this area that require and deserve further research attention. One such topic is providing temporal mechanisms with an ability to handle imperfect information.

Imperfect information surrounds us everywhere - almost all that we know about the real world is not fully certain, complete, precise, or consistent. This means that when we just study certain information, we only concentrate on a small part of a big and complex problem. In the area of temporal representation and reasoning, we also need to deal with imperfect information. This topic was underlined in recent surveys by Chittaro and Montanari (1996, 2000) as one of the important areas of temporal representation and reasoning which requires further research. Most of the present temporal reasoning techniques have been built on the assumption that precise and certain information is available, even though, in reality, this is false. Therefore, as was pointed out by Smithson (1989), all models that are built upon such idealizations, fail to describe the situations of the real world adequately. Other authors (e.g., Cohen, 1985; Motro, 1993; Parsons, 1996; Saffiotti *et al.*, 1994) also underline the necessity to be able to deal with imperfection in order to model the real world accurately.

In many situations there is a need to deal with temporal relations. A temporal relation is a relation between two temporal primitives (points, intervals, etc.), which includes the temporal meaning characterizing the difference between these primitives. Reasoning is a procedure, in which a conclusion can be inferred from known facts and from the relations between facts and conclusions. These relations are the core of reasoning. However, information about temporal relations might not be always perfect. One particular type of situation when this is the case is when there are two or more alternative values for a particular relation. To model this situation accurately, we need to be able to formalize this imperfect information, and take it into account during further reasoning.

The research problem of this thesis is situated at the convergence of the topic of temporal representation and reasoning with that of handling imperfect information. The main problem considered is the development of a formalism for handling imperfect temporal relations. Such a formalism needs to address three important issues: representation, estimation, and reasoning. An imperfect temporal relation needs to be represented along with the numerical measures of imperfection of this relation. The estimation of imperfection suggests how these measures can be obtained, because they are not readily available in many situations. The reasoning mechanism defines the operations for reasoning about imperfect temporal relations. These operations allow us to derive previously unknown imperfect temporal relations, taking as operands known imperfect relations.

The rest of the thesis is organized as follows. The research problem addressed by the thesis and the research methodology used are described in Chapter 2. In Chapter 3 we present three classifications of imperfect information, and define the concepts of imperfection used throughout this thesis. Chapter 4 presents a brief introduction to the main numerical formalisms for handling imperfect information, such as probability theory, possibility theory, and the Demspter-Shafer theory of evidence. After that, we address the main limitations of the numerical formalisms, and select the formalism used. In Chapter 5 we present the basic concepts of time ontology, including the representation of temporal primitives. In Chapter 6 we consider the two representations of imperfect relations between temporal points used in the thesis. Chapter 7 deals with the representation of imperfect relations between temporal intervals. The reasoning operations are presented in Chapter 8. Chapter 9 presents two possible application areas for the formalism proposed in the thesis. The organization of the study and the summary of the articles included in the thesis are discussed in Chapter 10. Finally, the main contribution of the thesis, limitations, and directions for further research are summarized in Chapter 11.

# 2 RESEARCH PROBLEM AND METHODOLOGY

In this chapter we present the main research problem considered in the thesis and the research methodology used.

In this study we are concerned with the topic of handling imperfect temporal information, and more specifically, we are interested in handling imperfect temporal relations. The main issues related to this topic are: representation and estimation of temporal relations, and reasoning with temporal relations. Figure 1 presents the conceptual schema showing the logical connection between these issues.

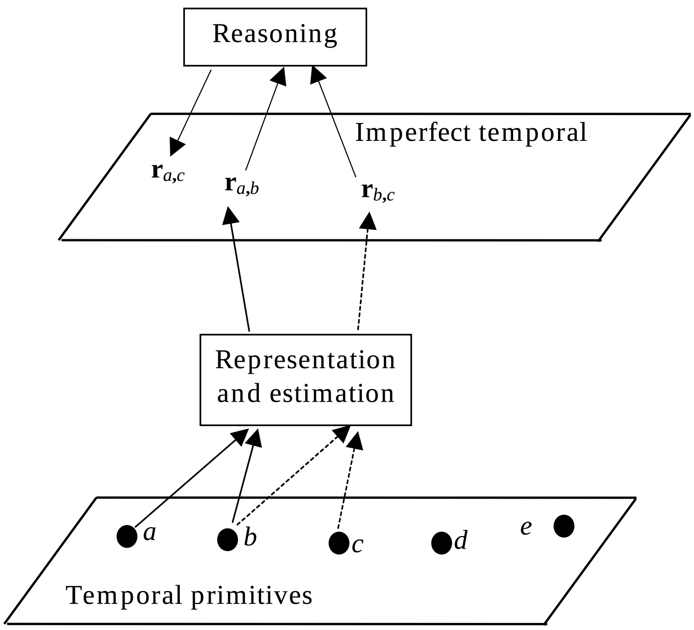



FIGURE 1 Conceptual schema for the basic notions related to handling imperfect temporal relations

In Figure 1 at the lower level there are temporal primitives. At the upper level there are imperfect temporal relations between these primitives. A temporal relation between two primitives is derived using the information about these primitives through representation and estimation procedure. With imperfect temporal relations we can perform reasoning using reasoning operations as inversion, composition, and addition.

Let us consider the main issues related to handling imperfect temporal relations in more detail.

- *Representation and estimation of imperfect temporal relations.* The representation should be able to formalize imperfect temporal relations, and should include the numerical measures of imperfection to distinguish between different imperfect temporal relations. Often the information about the temporal relation between two primitives is not readily available but needs to be derived, for example, by using the information about the temporal primitives. The information about the primitives can also be imperfect, and which leads to an imperfect relation between them. The estimation of an imperfect temporal relation is the process of deriving the measures of imperfection for this relation.
- *Reasoning with imperfect temporal relations.* Reasoning is a procedure for deriving new temporal relations in using reasoning operations. The operands and the derived relation might be imperfect. Reasoning operations need to calculate the measures of imperfection for the derived relation using the measures of imperfection for the operand relations.

Compounding the above main issues of handling imperfect temporal relations together with the probabilistic approach, the selection of which is discussed in Chapter 4, let us state the main research problem considered in this thesis as the following:

> The development of a formal approach to represent an imperfect temporal relation between two temporal primitives, to estimate the imperfection of this relation, and to reason with imperfect temporal relations using a probabilistic approach.

The study presented in this thesis is constructive and includes two main stages: observation and theory building. Observation is useful when little is known about the research area, providing a hypothesis for testing or enabling the further focusing of the research. Our study is based on the approaches to temporal representation and reasoning and the approaches to handling imperfect information as proposed in the literature. The overview of these revealed, that there exists a problem in handling imperfect temporal relations, which was stated as a goal of the present study.

Theory building includes the development of new ideas, concepts, conceptual frameworks, and models. Analysis of the current approaches to processing imperfect information (Chapter 4) and requirements with regard to handling imperfect temporal relations, lead to the selection of the probabilistic approach as a basis for the developed theory. The theory was developed along the directions: representation and estimation of imperfect temporal relations, and reasoning with such relations. Combining the representation and reasoning mechanisms for temporal relations, together with probabilistic measures of imperfection, we obtained the representation, estimation, and reasoning mechanism for imperfect temporal relations.

# 3 CLASSIFICATIONS OF IMPERFECT INFORMATION

In this chapter we present three classifications of imperfect information, and define the concepts and types of imperfection used throughout this thesis.

There is no general consensus among researchers about the terms used regarding information which is not perfect, certain, complete, or precise. Prior to considering the classifications of imperfect information, it is necessary to clarify the difference between the uses of the terms "uncertainty" and "imperfection". As has been underlined by Parsons (1996), the term "uncertainty" in the literature is overloaded, and is commonly used both as a generic term for imperfection in data, and as a term for a particular form of imperfect knowledge, that is, whether or not a statement is true. We support this point of view and aim to use the term "imperfection" in its generic sense, and the term "uncertainty" in the specific sense according to Parsons' classification (Parsons, 1996).

In the literature one can find many attempts to classify different types of imperfect information and to establish the relationships between them. None of these classifications however, has been widely recognized as better than the others, although several of them seem to be reasonable. In the following three sections we present the classifications of Bonnissone and Tong (1985), Bosc and Prade (1993), and of Parsons (1996). In the last section we summarize the concepts of imperfection that are used in this thesis.

## 3.1 Bonnissone and Tong's classification

One of the earliest classifications, proposed by Bonnissone and Tong (1985) includes three main types of imperfection: incompleteness, imprecision, and uncertainty (Figure 2).

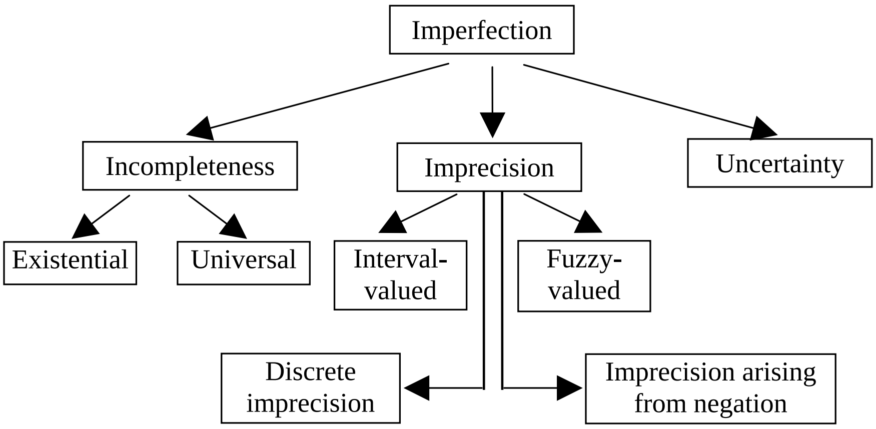



FIGURE 2 Types of imperfection according to Bonnissone and Tong's classification (1985)

According to this classification, incompleteness arises from the absence of a value, and can be of two types: existential and universal. Existential incompleteness occurs when a particular instance of the value of an attribute is unknown. Alternatively, universal incompleteness occurs when all instances of particular attributes are unknown. Imprecision arises from the existence of a value, which cannot be measured with suitable precision. Imprecision is interval-valued when the value of a variable is given within some interval, for example: "The meeting will be between 5 p.m. and 8 p.m.".

Fuzzy-valued imprecision occurs when some fuzzy concept is used, for example: "It is quite late". Discrete forms of imprecision arise from disjunctive information, for example: "The meeting will be either at 5, 6, or 7 p.m.". Finally, imprecision can arise from negation, as in the statement "The meeting will not be on Monday". This information is of little value for us, since the meeting can easily be on any other day of the week.

In Bonnissone and Tong's classification (Bonnissone & Tong, 1985), both imprecision and incompleteness are objective to some degree. Uncertainty is considered here as a subjective thing. It is supposed that some person makes his estimation of the truth of some fact. This estimation may be made, by using different formalism, e.g., probabilities, degrees of possibilities, etc.

## 3.2 Bosc and Prade's classification

Bosc and Prade (1993) proposed another classification, which is an extension of an earlier classification by Dubois and Prade (1988). It includes four types of imperfection: uncertainty, imprecision, vagueness, and inconsistency (Figure 3).

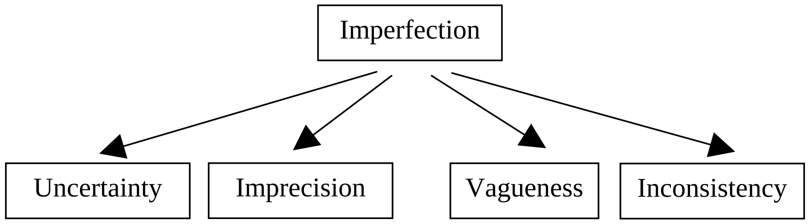



FIGURE 3 Types of imperfection according to Bosc and Prade's classification (1993)

The definitions for the concepts in this classification are slightly different from the definitions given by Bonnissone and Tong's classification. Particularly, uncertainty arises from the lack of information about the state of the world. In this case, we are unable to determine if certain statements about the world are true or false, but it is possible to estimate the tendency of a statement to be true or false. In order to do this, we again can use some numerical formalism or symbolic approach.

Imprecision can be considered arising from the granularity of the language that is used to make imprecise statements. So, the statement "it is 2 o'clock now" is considered imprecise when we are interested in an exact time given in hours and minutes. Bosc and Prade (1993) point out that uncertainty and imprecision can arise together in the same piece of information as it is, for example, in the statement "the meeting will be between 5 and 8 p.m.". This statement can be considered imprecise if we are interested in exact time of the meeting given in minutes and uncertain since the certainty of this statement might depend on the knowledge of the source of this information. It is interesting that very imprecise statements might be more correct than precise ones. Vagueness is considered similar to the fuzzy-valued type of imprecision according to the classification of Bonnissone and Tong. A vague statement includes a vague predicate, e.g., "it is quite late".

Inconsistency represents the situation when there are two or more conflicting values, and there is no a possibility to find a consensus among them. Inconsistency can arise from a number of information sources, for example, after merging databases. One approach to resolve inconsistency is to select the information from the most reliable source, although we might not always have information about the reliability of the information sources. Another solution includes assigning different pieces of information with credibility rates, and then selecting the most credible information. Finally, there is an approach (Roos, 1992), which aims to locate and eliminate the information source that makes information inconsistent. For example, when three information sources suggest non-conflicting values for a variable and the forth one suggests a different (conflicting) value, the latter one will not be taken into account.

## 3.3 Parsons' classification

Combining the classifications of Bonnissone and Tong and of Bosc and Prade, Parsons (1996) proposed to distinguish between the following types of imperfect information: uncertainty, imprecision, incompleteness, inconsistency, and ignorance (Figure 4).

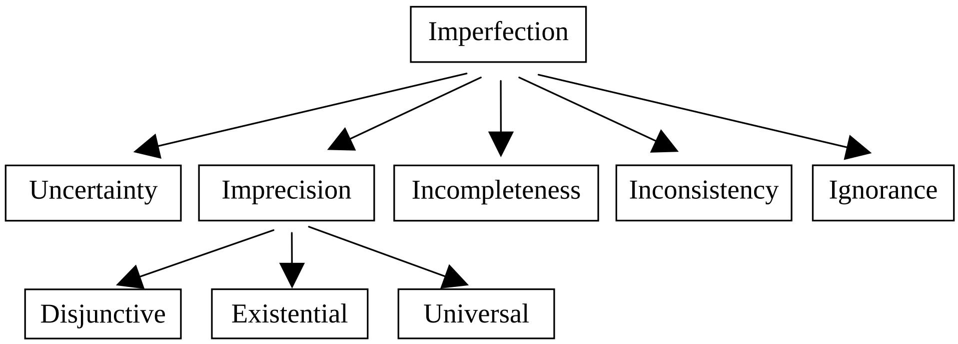



FIGURE 4 Types of imperfection according to Parsons' classification (1996)

Uncertainty is a situation when we do not have enough certain information about the state of the world. This can stem from subjectivity (or error) of the information source. Imprecision can be of three types: disjunctive, existential, and universal. Incompleteness is simply a lack of relevant information. Inconsistency is similar to the definition given in the previous section. Ignorance describes a lack of knowledge, particularly, a lack of knowledge about the relative certainty of a number of statements.

## 3.4 Concepts used in the thesis

In this section we summarize the concepts of imperfection used throughout this thesis and briefly discuss possible sources of imperfection.

We have selected the model of imperfection as outlined in Parsons' recent classification for use in this thesis, as it is self-consistent and intuitively acceptable, as well as having the advantage of combining concepts and definitions proposed by other researchers. In this thesis we will use the notions of uncertainty, inconsistency, and disjunctive form of imprecision. The latter one will be referred to as indeterminacy, as defined in the temporal database concepts glossary (Jensen & Dyreson, 1998). Disjunctive imprecision describes the situation, when a number of propositions or statements are present, but only one of them is true for the current state of the world. We will not consider in this thesis incompleteness, ignorance, and existential or universal forms of imprecision.

In general, different types of imperfect information can arise from different sources, and these ones might also depend on particular applications. For instance, different kinds of imperfection can result from unreliable sources, as was suggested by Motro (1993), such as faulty sensors and input errors, or from the inappropriate choice of representation. Secondly, if data is recorded statistically, it is inherently uncertain. Moreover, some information can be intentionally made uncertain for the reasons of security, as it was pointed out by (Kwan *et al.*, 1993). Measurement instruments might also introduce some imprecision in recorded data.

Particularly, the main origin of uncertainty is subjective error in the information provided by some information source. Imprecision stems from granularity mismatch, i.e. the situation when the particular value was obtained in one granularity but is recorded in a system with a finer granularity. Inconsistency often occurs after merging of a number of data sets, for example, databases, or as a result of combining opinions in expert or decision making systems.

In this chapter we have considered three classifications of imperfect information, and defined the concepts to be used later in the study. In the next chapter we overview the numerical methods for handling different kinds of imperfection.

# 4 NUMERICAL APPROACHES TO HANDLING IMPERFECTION

There are two general classes of techniques for handling imperfection: numerical and symbolic. In this thesis our interest is with the class of numerical techniques, which is motivated by the wish to use numerical measures for handling imperfect temporal relations. The reader interested in symbolic formalisms can find an excellent overview of the most essential ones (as well as numerical formalisms) in Parsons (1996), and in Parsons and Hunter (1998). These overviews present such symbolic formalisms as nonmonotonic logic, circumscription, default logic, autoepistemic logic, and the numerical formalisms presented later in this chapter. Usually, only one formalism from any one class (symbolic or numerical) is applied in a particular task, but there are also a number of proposals for combining numerical and symbolic formalisms. For example, Parsons (1996) suggested that in some applications, it might be reasonable to apply numerical and symbolic techniques sequentially.

In this chapter we present three main numerical techniques for handling imperfection: probability theory, possibility theory, and Dempster-Shafer theory of evidence, which are considered in the next three sections correspondingly. In Section 4 we discuss the main limitations of all the numerical approaches. In Section 5 we evaluate three presented techniques, based on the criteria by Walley (1996). Finally, in Section 6 we select an approach to be used in the thesis.

## 4.1 Probability theory

The oldest numerical technique for handling imperfection is definitely probability theory, which existed already in different forms several hundred years. During that time a number of definitions and formulations for this theory have been proposed. A brief overview of the basics of probability theory presented below is drawn from Parsons (1996), which in turn is based on the discussion of probability theory by Lindley (1975).

The classical definition of probability given below is drawn from Neapolitan (1990, 28), but is based on Laplace (1951):

> "The theory of chance consists in reducing all the events of some kind to a certain number of cases equally possible, that is to say, such as we may be equally undecided about in regard to their existence, and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of the probability".

Neapolitan (1990) was formally to define the notion of probability used within the framework of this classical approach, which we will also use as a definition of probability in this thesis:

> "For each event $E \subset F$, there corresponds a real number *P(E)*, called the probability of *E*. This number is obtained by dividing the number of equipossible alternatives favorable to *E* by the total number of equipossible alternatives". (Neapolitan, 1990, 31)

A probability measure is an estimate of the degree to which an uncertain event is likely to occur. According to Lindley (1975), probability theory is based on three axioms or laws that define the behavior of probability measures. These ones are convexity, addition, and multiplication laws.

The convexity law suggests that the probability measure for an event **A** given information **H** is such that:

$$0 \le \Pr(\mathbf{A}|\mathbf{H}) \le 1.$$

The addition law relates the probabilities of two events to the probability of their union. For two exclusive events **A** and **B**, that is two events that cannot both occur, we have:

$$\Pr(\mathbf{A} \cup \mathbf{B}|\mathbf{H}) = \Pr(\mathbf{A}|\mathbf{H}) + \Pr(\mathbf{B}|\mathbf{H}).$$

Often, explicit reference to the information **H** is omitted, since this information is the same in all cases. When the events are not exclusive we have:

$$\Pr(\mathbf{A} \cup \mathbf{B}) = \Pr(\mathbf{A}) + \Pr(\mathbf{A}) - \Pr(\mathbf{A} \cap \mathbf{B}).$$

The sum of the probabilities of a set of **n** mutually exclusive and exhaustive events $\mathbf{A_i}$ is equal to 1:

$$\sum_{\mathbf{i}=1}^{\mathbf{n}} \Pr(\mathbf{A_i}) = 1.$$

The multiplication law derives the probability of two events **A** and **B** occurring together:

$$\Pr(\mathbf{A} \cap \mathbf{B}) = \Pr(\mathbf{A})\Pr(\mathbf{B}|\mathbf{A}).$$

The probability measure Pr(**B**/**A**) is the conditional probability of **B** given **A**, that is, the probability that **B** will occur when **A** is known to have occurred. There is a well-known further rule concerning probability measures - the Bayes' theorem, which is crucial from the AI point of view. It can easily be derived from the fundamental laws of probability theory stated above:

$$\Pr(\mathbf{A}|\mathbf{B}) = \frac{\Pr(\mathbf{B}|\mathbf{A})\Pr(\mathbf{A})}{\Pr(\mathbf{B})}.$$

There have been several extensions of probability theory within artificial intelligence literature (e.g., Duda *et al.*, 1976; Quinlan, 1983; Tawfik & Neufeld, 1998). An important approach when using probability theory in computing, is that of a probabilistic network, also called a Bayesian network or a causal network (Heckerman & Wellman, 1995; Pearl, 1987; Pearl, 1988). Using conditional probabilities combined with further structural information these networks are intended to represent and describe more efficiently probabilistic information. Much attention has been given to the problem of propagating probabilities through probabilistic networks efficiently (e.g., Pearl, 1992). The interested reader can find a further discussion in Parsons (1996), Parsons and Hunter (1998), and in Hunter and Parsons (1998).

Clear interpretation of numerical measures is very important not only to justify the usefulness of the theory but also to guide applications of the theory. The issue of interpretation is distinguished from that of mathematical representation. There are many kinds of probability models, and any of these models can be given various interpretations. Similarly, any single interpretation of probability can be given various mathematical representations. The basic distinctions between different interpretations of probability were proposed by Walley (1991). According to that, the most fundamental distinction is between aleatory and epistemic probabilities. Aleatory probabilities model randomness in empirical phenomena. Epistemic probabilities model logical or psychological degrees of partial belief of a person or an intentional system.

Among epistemic probabilities, which depend on available evidence, we can distinguish between logical, personalist, and rationalist interpretation. In its logical interpretation the epistemic probability of some hypothesis, related to a particular evidence, is uniquely determined. In personalist interpretation probabilities are constrained only by axioms of coherence and not by evidence. Finally, rationalistic interpretation lies between logical and personalist interpretations and requires probabilities to be consistent in certain way with the evidence, without requirement to be uniquely determined.

Epistemic probabilities could also be classified as having behavioral and evidential interpretations. In behavioral interpretation probabilities are interpreted in terms of behavior, e.g., betting behavior. In evidential interpretation a probability measures a logical or linguistic relation between the hypothesis and the available evidence. The interpretation of probabilities can also be classified according to the issue of measurement. They could be measured either by observation of quantities that they influence, or by construction of probabilities from knowledge of the factors influencing them. Finally, interpretations are distinguished between operationalist and theoretical interpretations. In operationalist interpretation probabilities are identified with the observations obtained from specified procedures. In theoretical interpretation probabilities model the underlying theoretical quantities that are not directly observable, but influencing the observable quantities through interaction with them and other quantities.

## 4.2 Possibility theory

Another well-known numerical formalism for handling imperfection is possibility theory, which emerged from the notion of fuzzy sets (Zadeh, 1965), and was firstly introduced by Zadeh (1978). A fuzzy set is a set whose membership is not absolute, but a matter of degree, for example the set of young people. A fuzzy set **F** is characterized by a membership function $\sigma_{\mathbf{F}}$ which specifies the degree to which each object in the universe **U** is a member of **F**. Let **X** be a variable which takes values in **U**. The assignment of a value **u** to **X** has the form:

$$\mathbf{X} \leftarrow \mathbf{u} : \sigma_F(\mathbf{u}),$$

where $\sigma_{\mathbf{F}}(\mathbf{u})$ is the degree to which the constraint **F** is satisfied when **u** is assigned to **X**. To denote the fact that **F** is a fuzzy restriction on **X** we write:

$$R(\mathbf{X})=\mathbf{F}.$$

Now, the proposition "**X** is **F**" translates into "R(**X**)=**F**", and associates a possibility distribution $\Pi_x$ with **X,** and this distribution is taken to be equal to R(**X**):

$$\Pi_x = R(\mathbf{X}).$$

Along with this, we have a possibility distribution function $\phi_{\mathbf{X}}$ that is defined to be equal to the membership function of **F**:

$$\phi_{\mathbf{X}} = \sigma_{\mathbf{F}}.$$

Thus $\phi_{\mathbf{X}}(\mathbf{u})$, the possibility that **X**=**u**, is taken to be equal to $\sigma_{\mathbf{F}}(\mathbf{u})$. For example, let **U** be the set of different people's ages, and **F** be the fuzzy set of young people. This set is described by the following set of pairs, each of the form $(\mathbf{u}, \sigma_{\mathbf{X}}(\mathbf{u}))$:

$$\mathbf{F} = \{(15,1),(16,1),(17,1),...,(30,0.5),(31,0.48),...,(40,0.1)\}$$

Given this, the proposition "**X** is a young person" associates the possibility distribution $\Pi_x$ with **X** where $\Pi_x$ is written as a set of pairs $(\mathbf{u}, \phi_{\mathbf{X}}(\mathbf{u}))$:

$$\Pi_x = \{(15,1),(16,1),(17,1),...,(30,0.5),(31,0.48),...,(40,0.1)\}$$

According to the above distribution, the possibility, for example, that a person of age 16 is young is 1, and the possibility of the same proposition for a person of age 40 is only 0.1. Possibility distributions are used to define possibility measures. If **A** is a fuzzy subset of **U**, then the possibility measure $\Pi(\mathbf{A})$ of **A** is defined by

$$Poss(\mathbf{X} \text{ is } \mathbf{A}) = \Pi(\mathbf{A}) = \sup_{\mathbf{u} \in \mathbf{U}} \min\left(\sigma_{\mathbf{A}}(\mathbf{u}), \phi_{\mathbf{X}}(\mathbf{u})\right).$$

When **A** is a strict subset of **U** the above equation is written as:

$$\Pi(\mathbf{A}) = \sup_{\mathbf{u} \in \mathbf{U}} \phi_{\mathbf{X}}(\mathbf{u}).$$

The following equations define the rules for combining possibility measures:

$$\Pi(\mathbf{A} \cup \mathbf{B}) = \max(\Pi(\mathbf{A}), \Pi(\mathbf{B})) \text{ and } \Pi(\mathbf{A} \cap \mathbf{B}) = \min(\Pi(\mathbf{A}), \Pi(\mathbf{B})).$$

There is a heuristic connection between possibility and probability as noticed by Parsons (1996), since if something is impossible, it is likely to be improbable, on the other hand, a high degree of possibility does not imply a high degree of probability, nor does a low degree of probability reflect a low degree of possibility. At the same time, as was pointed out by Neapolitan (1990), possibility theory addresses a fundamentally different class of problems than that addressed either by probability theory or by the Dempster-Shafer theory. These two deal with propositions that are definitely true or false.

There have been attempts to combine possibility theory with classic logic to form a possibilistic logic (e.g., Bigham, 1998; Dubois *et al.*, 1991; Dubois & Prade, 1987; Dubois & Prade, 1989).

## 4.3 Dempster-Shafer theory

The third well-known numerical formalism for dealing with imperfection is the theory of evidence, which is also often referred to as the Dempster-Shafer theory (Shafer, 1976). The theory deals with the so-called frame of discernment, the set of base elements $N = \{\chi_1, ..., \chi_n\}$ in which we are interested, and its power set $2^{\chi}$, which is the set of all the subsets of the base elements. The basis of the measure of uncertainty is a probability mass function $m(\cdot)$ that assigns zero mass to the empty set, $m(\emptyset) = 0$, and a value within the interval $[0,1]$ to each element of $2^{\chi}$, the total mass distributed being 1 so that:

$$\sum_{A \subseteq N} m(A) = 1.$$

In the evidence theory we deal with all possible subsets of the set of propositions. Therefore, we can distribute the probability mass between particular subsets as we wish. The belief in a subset **A** of the set of all propositions, is defined as the sum of all the probability masses that support its parts:

$$\mathrm{Bel}(\mathbf{A}) = \sum_{\mathbf{B} \subseteq \mathbf{A}} \mathbf{m}(\mathbf{B})\,.$$

The plausibility of **A** is defined as the probability mass not supporting $\neg$**A**:

$$\mathrm{Pl}(\mathbf{A}) = \sum_{\mathbf{B} \cap \mathbf{A} \neq \emptyset} \mathbf{m}(\mathbf{B})\,.$$

The interval $[\mathrm{Bel}(\mathbf{A}), \mathrm{Pl}(\mathbf{A})]$ is considered as a measure of ignorance about **A**, and can vary from zero, when we have the same degree of belief in **A** as would be generated by probability theory, to 1 when **A** has belief 1 and plausibility 1. The latter means that no mass is assigned to **A** or any if its subsets, but equally no mass is assigned to $\neg$**A**.

Evidence is combined by Dempster's rule of combination. It calculates the probability mass assigned to $\mathbf{C} \subseteq N$ using the probability mass assigned to **A** and **B**, where **A** and **B** are also subsets of N. If we let the distribution function assigning probability mass to **A** be $m_1(\cdot)$ and the function distributing probability mass to **B** be $m_2(\cdot)$, then the probability mass assigned to **C** is calculated by the equation:

$$\mathbf{m}_{12}(\mathbf{C}) = \frac{\sum_{\mathbf{A} \cap \mathbf{B} = \mathbf{C}} \mathbf{m}_1(\mathbf{A})\mathbf{m}_2(\mathbf{B})}{1 - \sum_{\mathbf{A} \cap \mathbf{B} = \emptyset} \mathbf{m}_1(\mathbf{A})\mathbf{m}_2(\mathbf{B})}\,.$$

The division in the above equation normalizes the new distribution by reassigning any probability mass that is assigned to the empty set by the combination.

There is a potential problem with the computational complexity of Dempster's rule of combination, which has been discussed by several researchers. Barnett (1981) showed that the apparent exponential time requirement of the theory could be reduced to simple polynomial time if the theory was applied to a single hypothesis, rather than to sets of hypotheses, and if the evidence was combined in an orderly fashion. Barnett's approach was extended by Gordon and Shortliffe (1985), and further by Shafer and Logan (1987). Furthermore, Shenoy and Shafer (1990) introduced a method for the efficient propagation of belief functions in networks by means of local computations. Whereas, Wilson (1992) proposed a method in which the explicit use of Dempster's rule of combination is avoided and which allows us to perform belief function calculations in better than exponential time but worse than in polynomial time. Three examples of the applications of evidence theory, such as in the retrieving of documents, in the running of a radio communications system, and in automated inspection, can be found in (Hunter & Parsons, 1998).

## 4.4 Limitations of the numerical approaches

There are a number of other numerical formalisms for dealing with imperfect information, for example, certainty factors (Shortliffe, 1976). This approach assigns a numerical weight, called the certainty factor, to the consequent of every rule in a rule-based system. The value of the certainty factor belongs to the interval Ψ1,1β and is a difference between the degree of belief and the degree of disbelief assessed by the domain expert. However, several researchers have challenged the validity of the certainty factors model, e.g., Heckerman (1986). Generally, none of the proposed techniques for handling imperfection is blameless. There are three clear problems common to all numerical formalisms:

1) Obtaining the "numbers".
2) The interpretation of the results.
3) Computational expense.

The first limitation is considered obtaining the "numbers" needed to apply formalism, since sophisticated computational mechanisms are of little value without good numerical assessments (Parsons, 1996). For example, to apply probability theory in its "frequentist" interpretation we often need to obtain the kind of strong statistical data, which might not be possible in many application domains, as it was noted in Fox (1986). Other schools of probability theory, i.e. the personalist and necessarian schools (Shafer, 1988), argue that probabilities may always be obtained, either from rational human reasoning, or because they exist as a measure of the degree to which sets of propositions confirm one another. As it was pointed out by Parsons (1996), there is no clear better argument in the dispute about obtaining the numbers for numerical formalisms. It is apparent, that if there is a possibility to obtain numbers, then a particular numerical technique could be applied. If there is not, then perhaps a symbolic method for handling imperfection should be utilized.

The interpretation of the results produced by formalism might not be always obvious, and therefore is considered also as a limitation of numerical formalisms. All numerical techniques generate results as numerical values. However, these values can represent and denote different things in different techniques, and to interpret them correctly, it is perhaps necessary to label them with the type of belief that they measure.

Complexity of computations can be considered as the third limitation of numerical approaches, because sometimes computations become too expensive and might require a massive amount of time. To overcome this obstacle special techniques, for example, certainty factors, were proposed. In overall, the main problem of computational expense remains; there have been a number of attempts to identify efficient calculation methods in particular situations.

## 4.5 Evaluation of the numerical approaches

In this section we compare the three main numerical approaches (probability theory, possibility theory, and the Dempster-Shafer theory of evidence) for handling imperfect information. We assess them using the criteria proposed by Walley (1996, 3-4).

- *Interpretation.* The technique should have a clear interpretation that is sufficiently definite to be used to guide assessment, to understand the conclusions of the system and use them as a basis for action, and to support the rules for combining and updating measures.
- *Imprecision.* The technique should be able to model partial or complete ignorance, limited or conflicting information, and imprecise assessments of uncertainty.
- *Calculus.* There should be rules for combining measures of imperfection, updating them after receiving new information, and for drawing conclusions and making decisions.
- *Consistency.* There should be methods for checking the consistency of all imperfection assessments and default assumptions used by the system, and the rules of the calculus should ensure that the conclusions are consistent with these assessments.
- *Assessment.* It should be practicable for a user of the system to make all the imperfection assessments that are needed as input. The system should give some guidance on how to make the assessments.
- *Computation.* It should be computationally feasible for the system to derive inferences and conclusions from these assessments.

Walley (1996) does not claim that this list of criteria is exhaustive and sufficient for any situation, but he believes that they are important in the evaluation of numerical imperfection handling techniques. The first four criteria are theoretical in the sense that one would expect an adequate theory of imperfection to show that they can be satisfied, irrespective of the specific application. The last two criteria are practical in the sense that they will be satisfied in some applications but not in others, depending on different factors. Walley (1996) further underlines that the first criteria (“interpretation”) is the most fundamental, because an interpretation is needed to support the rules of the calculus, to guide assessment, and to understand conclusions. Therefore, the “interpretation” criterion is a prerequisite for “calculus”, “consistency”, and “assessment”.

Table 1 presents an evaluation of the three main numerical techniques for handling imperfection, based on the six criteria above. We can see that probability theory does well on the criteria of “interpretation”, “calculus”, and “consistency”, because probabilities have a simple behavioral interpretation (Walley, 1996). The rules of probability calculus can be justified through this interpretation, and these rules guarantee consistency.

There is no a consensus among the researchers about the marks for the criteria “imprecision” and “assessment” for probability theory. Many of them claim that probability theory does not do very well on “imprecision” and “assessment”, because in practice, it can be both difficult to make the many precise assessments of probabilities that are needed to determine the complete probability model, and to check that the assessments are consistent and that they determine a unique probability model. On the other hand, adherents of probability theory, people dealing with this formalism often, suggest that they do not face unsolvable problem with getting the numbers. Whatever mark we will put here for the probability theory, it might be criticized for subjectivity. Nevertheless, we propose “average” for “imprecision” and “assessment” criteria to reflect different opinions on this subject.

TABLE 1 Comparative evaluation of probability and possibility theory and the Dempster-Shafer theory of evidence, based on the criteria of Walley (1996)

| | Probability theory | Possibility theory | Theory of evidence |
|---|---|---|---|
| Interpretation | well | fair | poor |
| Imprecision | average | mainly well | mainly well |
| Calculus | well | poor | poor |
| Consistency | well | poor | poor |
| Assessment | average | mainly well | mainly well |
| Computation | mainly feasible | mainly feasible | average |

Lastly, “Computation” is feasible for some important types of models, for example, singly connected networks (Walley, 1996). Overall, probability theory is highly developed, especially for dealing with judgments of conditional independence, and useful for many practical problems.

Possibility theory translates natural-language expressions into a mathematical formalism of possibility measures. It is widely recognized that possibility is distinct from probability and that this distinction is central to all theories of probability (e.g., de Finetti, 1974, 1975). Some early attempts to provide possibility measures with a clear interpretation can be found, for example, in Zadeh (1978) and in Dubois and Prade (1988). The recent work by Dubois, Prade, and Smets (Dubois *et al.*, 2001) proposes new semantics for qualitative possibility theory giving a clear idea of how possibility measures should be interpreted.

Possibility measures can be used to model some types of imprecise or partial information. Especially if second-order possibility measures are allowed (although they are more complicated than first-order measures), possibility measures can model a wide variety of uncertainty judgements, including imprecise judgements in natural language (Walley, 1996). However, the calculus rules and methods for checking the consistency of possibility measures appear to be quite arbitrary. Assessment on the other hand is mainly feasible, and is easier for first-order distributions. There is some help in the literature on how to go about selecting the required functions (e.g., Dubois & Prade, 1988). Finally, the computation of inferences and decisions from possibility distributions requires in general, the solution of a nonlinear programming problem (Walley, 1996). Therefore, computations will often be difficult, despite the apparent simplicity of the calculus. These computations are generally easier for first-order possibility distributions than for second-order distributions.

Belief functions in the Dempster-Shafer theory of evidence lack a single clear interpretation. A survey of various interpretations can be found in Smets (1991). Belief functions can model partial ignorance and limited or conflicting evidence. However, due to the lack of a clear interpretation, the calculus rules and methods for checking consistency are often arbitrary. Belief functions can be assessed through multivalued mappings or in other ways, and the assessment strategies suggested by the theory are useful in many applications (Walley, 1996). As belief functions can be represented in terms of a probability mass function, they in many cases appear to be mathematically and computationally simple.

A further discussion regarding the assessments provided in Table 1 can be found in Walley (1996).

## 4.6 Selection of the theory to be used

After an overview of the main numerical techniques, a discussion of their limitations and their comparative evaluation, we can select the technique to be used in this thesis for modeling imperfect temporal relations.

The research problem considered in the thesis as discussed in Chapter 2 imposes some requirements on the technique to be used to model imperfect temporal relations.

1) *Alternative relations*. One type of an imperfect temporal relation (an uncertain relation) is defined as a number of alternative basic relations that can hold between two primitives. Only one of these basic relations definitely holds between the primitives, but we are simply uncertain about which one. The selected technique should accurately model this situation and provide numerical measures for the alternatives.
2) *Dependent values*. The values of the endpoints of two temporal intervals can be dependent. To model accurately the relation between these intervals we need to model accurately these dependencies.
3) *Calculus*. We need to have clear and explicit rules for combining the imperfection measures. These rules are an essential part of the combination of the imperfection measures, and are needed to define the reasoning operations. Clear and explicit calculus rules stem from a clear interpretation of the imperfection measures as shown in Table 1.

In this study as we are only dealing with precise assessments of imperfection measures, we are not interested in the ability of a formalism to model imprecision, and therefore, we can ignore the evaluations describing imprecision in Table 1.

Probability theory does well on the first criteria, because we can easily provide different probability values for different alternatives. Possibility theory however, does not do as well, because in a fuzzy set, a number of propositions within this set can be true. This does not correspond to the situation within an uncertain relation, where only one of the basic relations is the true relation, which also does not require us to consider all the possible subsets within the set of possible values as in the Demspter-Shafer theory. Probability theory includes a rich and highly developed mathematical means for modeling dependencies, although, dependencies in general can be modeled by all three formalisms. Probability theory has also very clear calculus rules, which stem from the clear interpretation of probability measures. In overall, it seems that probability theory more than the other techniques suits the three requirements above, and taking into account also assessments form Table 1, we select the probabilistic approach to be used in this thesis to model imperfect temporal relations.

In this chapter we have considered the three main numerical approaches to handling imperfection. We have discussed their limitations, undertaken a comparative evaluation and, finally, selected the probabilistic approach for use in the thesis. In the next four chapters we are going to overview the basic concepts within the field of temporal representation and reasoning that are used in this study.

# 5 TIME ONTOLOGY AND TEMPORAL PRIMITIVES

In the first section of this chapter we present ontology of time, including the selection of temporal primitives to be used and the structure of time. In the second section, we consider the representation of temporal primitives.

## 5.1 Ontology of time

Approaches to temporal representation and reasoning in AI have been situated in the context of philosophical theories of time (e.g. Hamblin, 1972; Newton-Smith, 1980; van Benthem, 1983). The modeling of the notion of time initially most important in AI, was within the areas of natural language understanding, medical decision making, and planning. However, it was only from the beginning of the 1980's that more general theories of time and action were proposed, such as McDermott's temporal logic (McDermott, 1982), Allen's theory of action and time (Allen, 1984), and Vilain's theory of time (Vilain, 1982). These theories established the basic representational issues and reasoning algorithms, and proposed two main temporal ontological primitives: temporal points (also called instants in the literature) and temporal intervals (also called periods). A possible third approach uses a combination of the above primitives. Each approach has its own adherents and opponents, and there is no standard agreement on which of these primitives is generally more applicable, this is rather decided according to the context of a particular application.

Early representations of time used points as temporal primitives such as in Situational Calculus (McCarthy & Hayes, 1969), in Time Specialist (Kahn & Gorry, 1977), and in McDermott's temporal logic (McDermott, 1982). Other authors on the other hand argued for using temporal intervals as ontological primitives. This direction was initiated by Allen (1983), who proposed thirteen basic relations between two temporal intervals and an Interval Calculus. A further discussion on this subject can be found in Vila (1994), Böhlen *et al.* (1998), and Steiner (1998).

In this thesis, we believe that both types of temporal primitives are important for the accurate modeling of time across different applications. We select temporal points as our main ontological primitives, whereas a temporal interval is constructed as a pair of temporal points denoting the start and the end of the interval.

After deciding on the ontological primitives, it is necessary to define the structure of time. This will determine the desired properties of time and the behavior of temporal primitives and relations. The question of the structure of time needs to be considered very carefully, because the complexity of a reasoning formalism strongly depends on this structure, almost independently of the defined formalism itself. The structure of time needs to be defined along three important dimensions, as is proposed in Vila (1994).

- *Discrete vs. Dense.* The time axis is either considered as a sequence of discrete temporal elements, or in the case of the dense model, it is assumed that between two temporal elements there is always another element.
- *Bounded vs. Unbounded.* A time axis can be finite or infinite in either or both directions.
- *Precedence.* Time can be linear, branching, parallel, or circular.

In this thesis, time is represented by a time axis which is unbounded in both directions, and linear. We assume that the time model is discrete, that is that the time line is considered as a sequence of indivisible temporal elements (chronons) of minimal duration. We further assume that a temporal point is located during one particular chronon.

## 5.2 Representation of temporal primitives

Information concerning temporal points can also sometimes be imperfect. One example of this is temporal indeterminacy (Dyreson & Snodgrass, 1993), which means that we do not know exactly when a particular event happened. This type of imperfection is similar to the disjunctive form of imprecision, according to our classification of imperfect information (Chapter 3, Figure 3). In terms of our own time ontology, this means that we do not know exactly during which particular chronon the temporal point is located, rather, that the point is located somewhere during a set or range of chronons. An indeterminate temporal point is described by a lower bound, an upper bound, and a probability mass function (Dyreson & Snodgrass, 1993; Voss, 1997). The bounds are the chronons that delimit where the point is located. The indeterminate point **a** cannot be earlier than the lower bound (denoted as $\mathbf{a}^{\mathbf{l}}$), or later than the upper bound (denoted as $\mathbf{a}^{\mathbf{u}}$), as is shown in Figure 5.

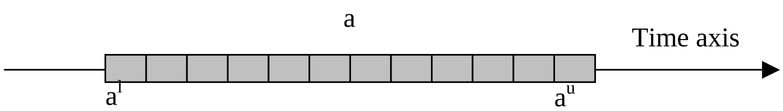



FIGURE 5 Indeterminate temporal point $\mathbf{a}\langle \mathbf{a}^{\mathbf{l}},\mathbf{a}^{\mathbf{u}}\rangle$

In many cases not all the possible chronons are equally probable. For example, it might be possible that the chronons located in the middle of the interval $[\mathbf{a}^l,\mathbf{a}^u]$ are more probable than the chronons near the endpoints $\mathbf{a}^l$ and $\mathbf{a}^u$. The probability mass function (p.m.f.) f(**a**) gives the probability of each chronon within $[\mathbf{a}^l,\mathbf{a}^u]$. The p.m.f. is defined so that $\sum_{\mathbf{a}=\mathbf{a}^l}^{\mathbf{a}^u} \mathbf{f}(\mathbf{a}) = 1$, and $\mathbf{f}(\mathbf{x})=0$ when $\mathbf{x}<\mathbf{a}^l$ or $\mathbf{x}>\mathbf{a}^u$. The requirement that the sum of all the probabilities of the chronons is equal to 1 results from the definition of our time ontology, according to which, a temporal point occurs exactly during one particular chronon. A determinate point is defined within the interval consisting of only one possible chronon with an underlying probability value equal to 1.

Often it is supposed, that the p.m.f. is supplied when an indeterminate point is created. The possible sources for this p.m.f. were suggested in Dyreson and Snodgrass (1993), these are:

- *Granularity mismatch*, which occurs when an event is known in one granularity but is recorded in a system with a finer granularity. In this case, we often have no reason to favor one possible chronon over another.
- *Dating techniques*. Some dating techniques are inherently imprecise and possibly might provide the typical distribution of the variable, for example, the normal distribution.
- *Analysis of past data* can sometimes provide a hint to defining the p.m.f.
- *Clock measurements* might assume the specific distribution. For details, see (Petley 1991).

Several other additional means of determining the p.m.f. were suggested in Dey and Sarkar (1996). Moreover, the distribution can be specified as missing if the underlying p.m.f. is unknown (Dyreson & Snodgrass, 1993), although partially known distributions can be allowed as in the Probabilistic Data Model (Barbara *et al.*, 1992). In this thesis however, we assume that the p.m.f.'s for the indeterminate points are totally known, do not allow partially known distributions, and make no provisions for joint or dependent probabilities.

An indeterminate temporal interval **A** is a temporal interval, the endpoints of which **a** (the starting point of the interval **A**) and **b** (the endpoint of **A**) are the indeterminate temporal points, defined according to the definition above. We also suppose, that $\mathbf{a}<\mathbf{b}$. We further assume, that the intervals within which the endpoints are defined, do not overlap, i.e. $\mathbf{a}^u<\mathbf{b}^l$. This assumption can be relaxed in a particular implementation by providing an additional mechanism for checking the consistency of the indeterminate interval. Since the indeterminate interval **A** starts during any chronon within the interval $[\mathbf{a}^l,\mathbf{a}^u]$ and ends up during any chronon within $[\mathbf{b}^l,\mathbf{b}^u]$ the duration of this indeterminate interval might be imprecise.

In the next chapter we consider the representation of imperfect relations between temporal points.

# 6 RELATIONS BETWEEN TEMPORAL POINTS

In many situations there is a need to know the temporal relation between two primitives. In the first section of this chapter we will present the general properties of, and types of temporal relations between points. In the second and third sections we consider two types of imperfect relations and their representations correspondingly: inconsistent and uncertain relations.

## 6.1 General properties

A *temporal relation* is a relation between two temporal primitives (points, intervals, etc.), which includes the temporal meaning characterizing the difference between these primitives. In the literature, one can often find the term “constraint” used regarding a temporal relation, which can be interpreted as a constraint on the temporal entities (primitives). Generally, temporal relations can be divided into two main classes: qualitative and quantitative (metric). Also, there are several temporal formalisms that use both classes of relations together (e.g., Kautz & Ladkin, 1991; Meiri, 1996). Throughout this thesis, we will consider only qualitative temporal relations (temporal relations for short) between temporal points and intervals.

A temporal relation takes its symbolic value over a finite domain of the basic temporal relations. For temporal points, taken as ontological primitives, there are three basic qualitative relations: “before” (<), “at the same time” (=), and “after” (>). It is supposed that a particular temporal relation can have only one definite temporal value. When there are several alternative values, the temporal relation is often represented as a disjunction of these values. For example, using the relations between temporal points, we can obtain four such disjunctions: “Ω’ (“before” or “at the same time”), “∅’ (“at the same time” or “after”), “∏’ (“before” or “after”), and “?” (“before” or “at the same time” or “after”).

## 6.2 Inconsistent relations

According to the classification of imperfect information by Parsons (1996), inconsistency is one type of imperfect information that describes the situation when there are two or more conflicting values for a variable. It is common to consider knowledge consistent if it supposes the absence of contradiction, and inconsistent when it contains contradiction (Roos, 1992). Generally, inconsistency arises from having too much information from too many sources. For example, through merging the opinions of experts we might derive a contradiction, and hence, inconsistent knowledge.

There are two general approaches to deal with inconsistency. The first approach identifies inconsistency with a kind of error, which needs to be corrected before a reasoning system can continue working properly. Usually, the consistency of information is restored by eliminating one or more of the information sources that have brought about inconsistency within the resulting knowledge. There are several means of selecting information that can be taken into account, such as estimating the reliability of the information source and estimating the credibility of information (e.g., Ekenberg *et al.*, 1997). For example, all information sources might be assigned with reliability degrees, and to restore consistency, the information from the most reliable source is selected.

This approach has been adopted in many reasoning systems dealing with temporal relations. For example, van Beek (1989) and van Beek and Cohen (1990), develop approximate algorithms for reasoning about temporal relations. One of the key concepts they use is the “consistent scenario” for a relational network, which means that no contradiction in the network exists. Generally, restoring consistency according to this approach often means that some information will not be taken into account and probably will be lost, which in some situations can lead to a misleading result if the ignored information was essential.

The second approach considers inconsistency more as a phenomenon of normality than as a kind of error. Supporting this point of view, Gabbay and Hunter (1991, 1993) have pointed out that inconsistency in information is the norm, and we therefore should delighted if we are able to formalize it. They also underline, that there is a fundamental difference between artificial and human intelligent behavior: a human often resolves inconsistencies not by “restoring” consistency as is so often done in AI, but by applying rules telling one how to act when an inconsistency arises. This point of view concords with ours, in suggesting that there are many applications in which inconsistency needs to be formalized, but not necessarily resolved immediately.

The area of temporal representation and reasoning is not an exception, for in many situations we can derive the conflicting values for a particular temporal relation. These situations include, for example, merging information from several temporal databases, and in combining experts’ opinions in an expert system dealing with temporal information. We believe however, that not in all of those situations when an inconsistency is derived, is it necessary to restore consistency immediately. This also might be impossible, for example when at a particular moment in time we might not have enough relevant information to select one of the conflicting values. In this situation, it might be possible to formalize the conflicting information but possibly leave any decision about restoring consistency to further reasoning.

Let us suppose, that an *inconsistent temporal relation* between two points includes the conflicting values of the temporal relations between these points, and is represented as a conjunction of these values. It is important to notice, that the basic relations within the inconsistent relation in our formulation are not alternatives, but rather separate parts of the whole inconsistent relation. For example, let us obtain information from a first data set, so that the relation $\mathbf{r}_{a,b}$ between the points **a** and **b** is “<”. At the same time, let us have information derived from a second data set where the value of this relation is “>”. If we know nothing else about this relation and the datasets, but can assume the hundred percent reliability of information in each data set, then the inconsistent relation can be derived, and we suggest to represent it as “< and >”.

Formally, let the inconsistent relation between two points be presented by three values $[\mathbf{d}^{<},\mathbf{d}^{=},\mathbf{d}^{>}]$, where the values $\mathbf{d}^{<}$, $\mathbf{d}^{=}$, $\mathbf{d}^{>}$ are the percentages, divided by 100, of each of the three basic relations within the inconsistent relation, and $\mathbf{d}^{<}+\mathbf{d}^{=}+\mathbf{d}^{>} = 1$ (Ryabov *et al.*, 1999). In the above example, we suppose that both of the basic relations “<” and “>” equally contribute to the inconsistent relation, and so we suggest that the inconsistent relation $\mathbf{r}_{a,b}$ includes 50% of the relation “<” and 50% of the relation “>”, and is represented as $[0.5,0,0.5]_{a,b}$. In some other situations, the percentages can be different, depending on, for example, the number of information sources supporting each of the basic relations, and the credibility of information provided.

## 6.3 Uncertain relations

Uncertainty is another type of imperfect information, according to the classification presented in Chapter 3. Uncertainty arises from the lack of information about the state of the world, which makes it impossible to determine if certain statements about it are true or false. We are able only to estimate the tendency of a statement to be true or false using, for example, a numerical measure of the degree of our belief.

Let the *uncertain temporal relation* between two points be represented as a disjunction of the basic relations that can hold between these points. This means, that the basic relations within the uncertain relation are alternatives, and only one of them definitely holds between the two points. To estimate the tendency of one of the basic relations to hold between the two points, we propose to provide the basic relations within the uncertain relation with their probabilities. Generally, let us distinguish between two cases. The first case is when we do not need to formalize inconsistent temporal relations, and we can assume that all the inconsistencies are resolved immediately after their occurrence by ordinary means. In the second case, we suppose that any derived inconsistent temporal relation needs to be formalized, as argued in the previous section. The principal difference between these two situations, is that in the first case, we are dealing with the three basic relations that can hold between two points: "<", "=", and ">". In the second case, we are dealing with four basic relations: "<", "=", ">", and an "inconsistent relation". For each situation we propose a separate representation of the imperfect relation between two points.

When we do not need to formalize inconsistency, we can formally represent the imperfect relation between temporal points **a** and **b** by a vector with the three probability values $(\mathbf{e}^{<}, \mathbf{e}^{=}, \mathbf{e}^{>})_{\mathbf{a,b}}$, where: $\mathbf{e}^{<}$ is the probability that **a** < **b**, $\mathbf{e}^{=}$ is the probability that **a** = **b**, and $\mathbf{e}^{>}$ is the probability that **a** > **b**. The sum of the probability values $\mathbf{e}^{<}$**,** $\mathbf{e}^{=}$, and $\mathbf{e}^{>}$ is equal to 1, because in this case they represent all the possible basic relations between two temporal points.

In the situation when the inconsistent temporal relation is one of the basic relations, we can represent the imperfect relation between two temporal points **a** and **b** by a vector $(\mathbf{e}^{<}, \mathbf{e}^{=}, \mathbf{e}^{>}, \mathbf{e}^{i}[\mathbf{d}^{<}, \mathbf{d}^{=}, \mathbf{d}^{>}])_{\mathbf{a,b}}$, where the value $\mathbf{e}^{<}$ is the probability that **a**<**b**, the value $\mathbf{e}^{=}$ is the probability that **a**=**b**, the value $\mathbf{e}^{>}$ is the probability that **a**>**b**, and $\mathbf{e}^{i}$ is the probability that the relation between **a** and **b** is inconsistent. The inconsistent relation is represented by three values $[\mathbf{d}^{<},\mathbf{d}^{=},\mathbf{d}^{>}]$, as it is in the previous section. We also suppose that the percentage values of $\mathbf{d}^{<}$, $\mathbf{d}^{=}$, and $\mathbf{d}^{>}$ are defined only when $\mathbf{e}^{i} \neq 0$. The sum of the probability values of $\mathbf{e}^{<}$, $\mathbf{e}^{=}$, $\mathbf{e}^{>}$, and $\mathbf{e}^{i}$ is equal to 1, because these values represent the probabilities of all the possible basic relations that can hold between the points **a** and **b**.

Let us consider two examples illustrating the two above representations. Let us suppose, that the information about the temporal relation between the points **a** and **b** is obtained from one information source. This source suggests that $\mathbf{a} \leq \mathbf{b}$. In this case, within the vector representing this relation, the probability needs to divided equally between the two alternatives, i.e. $\mathbf{r}_{\mathbf{a,b}}$= $(0.5, 0.5, 0)_{\mathbf{a,b}}$. Another situation is when we suppose, that the information about the temporal relation between two temporal points **a** and **b** is obtained from three information sources, as in the example discussed in the previous section. The first source suggests that **a**<**b**. The second source says, that **a**=**b**, and the third information source suggests that **a**>**b**. In this situation, we suggest that the relation $\mathbf{r}_{\mathbf{a,b}}$ is totally inconsistent and the probability $\mathbf{e}^{i}$ is equal to 1. In this example, the percentages of the three relations "<", "=", and ">" within the inconsistent relation are distributed equally $\mathbf{r}_{\mathbf{a,b}}= \left(0,0,0,1\left[\frac{1}{3},\frac{1}{3},\frac{1}{3}\right]\right)_{\mathbf{a,b}}$.

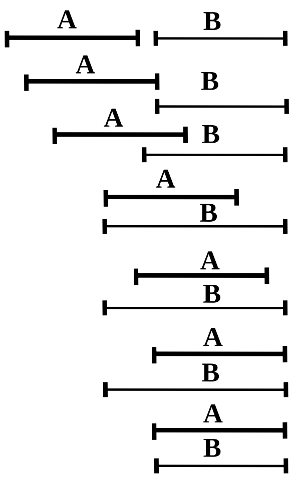
In this chapter we have considered the representation of, and different types of imperfect relations between two temporal points. In the next chapter we will consider the imperfect relations between temporal intervals, and after that we will review some relevant approaches to the representation of imperfect temporal relations as proposed in the literature.

# 7 RELATIONS BETWEEN TEMPORAL INTERVALS

In this chapter we consider the two main ways to represent the relation between two temporal intervals. The first approach, considered in the first section, uses Allen's thirteen interval relations. The second approach, presented in the second section, uses the four relations between the endpoints of the intervals. In the third section, we provide a discussion of related research.

## 7.1 Allen's interval relations

Allen (1983) developed an Interval Algebra based on thirteen interval relations, (Figure 6) corresponding to the simple definite mutually exclusive relations that may exist between two intervals.




| | |
|---|---|
| **A** before (b) **B** | **B** after (bi) **A** |
| **A** meets (m) **B** | **B** met-by (mi) **A** |
| **A** overlaps (o) **B** | **B** overlapped-by (oi) **A** |
| **A** starts (s) **B** | **B** started-by (si) **A** |
| **A** during (d) **B** | **B** contains (di) **A** |
| **A** finishes (f) **B** | **B** finished-by (fi) **A** |
| **A** equals (eq) **B** | **B** equals **A** |

FIGURE 6 Allen's interval relations (1983, 835)

Generally, to represent the imperfect relation between two intervals using Allen's thirteen relations, we need to allow for any arbitrary disjunctions on these relations. Allen for this task proposed using vectors of relations that are interpreted as the disjunction of the basic relations. For example, the vector (**A** b m o **B**) means that the interval **A** is "before", "meets", or "overlaps" the interval **B**. Allen (1983) proposed to maintain the temporal relations within the relational network, which is simply a directed graph. The nodes of the network represent the individual intervals. Each arc is labeled to indicate the possible relations between the two intervals represented by its nodes.

## 7.2 Relations between the endpoints of the intervals

The relation between two intervals can also be represented by the four relations between the endpoints of these intervals, as is shown in Figure 7.

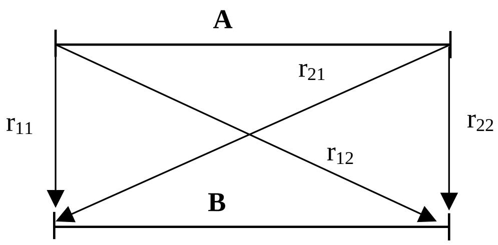



FIGURE 7 The four relations between the endpoints of the intervals **A** and **B**

This representation is quite obvious and has been used, for example, in van Beek (1992), Hirsch (1994, 1996), Vilain and Kautz (1986), and Wetprasit *et al.* (1997). It is convenient to represent the relation between two temporal intervals using the matrix $\vee = \begin{pmatrix} \mathbf{r}_{11} & \mathbf{r}_{12} \\ \mathbf{r}_{21} & \mathbf{r}_{22} \end{pmatrix}_{\mathbf{A},\mathbf{B}}$, where $\mathbf{r}_{11}$, $\mathbf{r}_{12}$, $\mathbf{r}_{21}$, and $\mathbf{r}_{22}$ are the relations between the endpoints of the intervals. The matrices corresponding to Allen's interval relations are represented in Figure 8.

| | | | | |
|---|---|---|---|---|
| $\begin{pmatrix} \vert & \lbrace \\ \rbrace & \vert \end{pmatrix}$ "equals" | $\begin{pmatrix} \lbrace & \lbrace \\ \lbrace & \lbrace \end{pmatrix}$ "before" | $\begin{pmatrix} \rbrace & \rbrace \\ \rbrace & \rbrace \end{pmatrix}$ "after" | $\begin{pmatrix} \rbrace & \lbrace \\ \rbrace & \lbrace \end{pmatrix}$ "during" | $\begin{pmatrix} \lbrace & \rbrace \\ \lbrace & \rbrace \end{pmatrix}$ "includes" |
| $\begin{pmatrix} \lbrace & \lbrace \\ \rbrace & \lbrace \end{pmatrix}$ "overlaps" | $\begin{pmatrix} \rbrace & \rbrace \\ \lbrace & \rbrace \end{pmatrix}$ "overlapped-by" | $\begin{pmatrix} \lbrace & \lbrace \\ \vert & \lbrace \end{pmatrix}$ "meets" | $\begin{pmatrix} \rbrace & \rbrace \\ \vert & \rbrace \end{pmatrix}$ "met-by" | $\begin{pmatrix} \vert & \lbrace \\ \rbrace & \lbrace \end{pmatrix}$ "starts" |
| | $\begin{pmatrix} \vert & \rbrace \\ \lbrace & \rbrace \end{pmatrix}$ "started-by" | $\begin{pmatrix} \rbrace & \lbrace \\ \rbrace & \vert \end{pmatrix}$ "finishes" | $\begin{pmatrix} \lbrace & \rbrace \\ \lbrace & \vert \end{pmatrix}$ "finished-by" | |

FIGURE 8 Matrices of Allen's interval relations (Hirsch, 1996, 279)

In many situations, as was discussed in the previous chapter, these four relations between the endpoints can be imperfect, and, in this case, the relation between the intervals is imperfect also. We suggest to represent the imperfect relation between two intervals using the modified matrix representation

$$\vee = \begin{pmatrix} (\mathbf{e}^{\{}, \mathbf{e}^{|}, \mathbf{e}^{\}})_{\mathbf{r}_{11}} & (\mathbf{e}^{\{}, \mathbf{e}^{|}, \mathbf{e}^{\}})_{\mathbf{r}_{12}} \\ (\mathbf{e}^{\{}, \mathbf{e}^{|}, \mathbf{e}^{\}})_{\mathbf{r}_{22}} & (\mathbf{e}^{\{}, \mathbf{e}^{|}, \mathbf{e}^{\}})_{\mathbf{r}_{22}} \end{pmatrix}_{\mathbf{A},\mathbf{B}},$$

where the relations between the endpoints are represented by the vectors of imperfect point relations as is proposed in the previous chapter.

Using the modified matrix $\vee$, it is possible to represent any disjunction of Allen's thirteen relations between intervals. However, this representation has some weak points. In many applications it is more natural to use representations based on interval relations rather than on points. We can use intervals instead of points as the basic entity, which does significantly increase the expressive power, but in general however, involves a loss of tractability (Hirsch, 1996). For example, when the probabilities within the vectors in the matrix $\vee$ are non-zero**,** it is difficult to guess which of Allen's relations are probable, and how probable they are, and this makes representation using the four relations between the endpoints less informative compared to representation using Allen's relations. However, in many practical situations this is the only way to estimate the imperfect relation between two indeterminate temporal intervals.

Therefore, in this thesis we propose a transition from representation using endpoints to representation using Allen's relations, when the relation between two intervals is imperfect. In a representation using the relations between endpoints we will use the matrix $\vee$, as shown in this section. In a representation using interval relations we propose to calculate the probabilities of Allen's relations between two intervals (Ryabov, 2000). This calculation uses the probability values within the vectors from the matrix $\vee$.

## 7.3 Related research

A number of formalisms for handling imperfect temporal information have already been proposed, but only a few of them consider the representation of imperfect temporal relations. Of the latter, almost all of them are built on the assumption that there is present several alternative values for the temporal relation.

Allen (1983) when proposing his representation of interval relations, assumed that in many practical situations we cannot be sure about which of the basic relations holds between two temporal primitives, as several alternative relations may be available. Allen represented the imperfect temporal relation between two intervals as a disjunction of the basic thirteen relations that are possible. He also discussed the importance of reasoning algorithms and analyzed their complexity. This representation however, did not include any numerical measures of imperfection.

Van Beek (1989, 1990, 1991) proposed a formalism for dealing with imperfect temporal relations (referred to as indefinite relations). He considered the representation of imperfect relations using a relational network approach, where the temporal entities are represented by nodes, and the temporal relations among these entities are represented by arcs. The main research problem considered by van Beek was to make explicit the strongest possible assertions about temporal relations when we are given a set of events, represented as intervals, and when we have some knowledge of the relations between some of the intervals. Van Beek in his formalism also did not specify any numerical measures of imperfection, although he did not consider them as a research goal.

Freksa (1992) presented a generalization of Allen's interval-based approach to temporal reasoning. The notion of a "conceptual neighborhood" of qualitative relations between events is central to the presented approach. Relations between semi-intervals rather than intervals are used as the basic units of knowledge. In addition to the logical constraints considered by Allen, Freksa takes into account neighborhood relationships between temporal relations. Two relations are conceptual neighbors if they can be derived from each other by shifting one of the four interval endpoints, leaving the other three fixed. A set of relations forms a conceptual neighborhood if each relation is a conceptual neighbor of at least one other relation in the set.

One important problem to be solved by temporal logic is interval representation problem. The core of this problem is in representing the relation which exists between the truth of an assertion over an interval and the truth of an assertion over internal points and/or subintervals of an interval. Several solutions of this problem proposed in the literature are observed and critically evaluated in Trudel (1991). That paper also proposes another approach to this issue based on first-order temporal logic called GCH. This logic uses different approach to represent information associated with an interval, based on the assumption that what is true at every point in an interval completely determines what is true over the interval.

An approach to represent temporal knowledge based on elementary calculus was proposed in Trudel (1997). Many traditional methods for representing temporal information in a first-order logic suppose associating the information with a temporal point or interval via a relation. Trudel (1997) proposed a new paradigm in which all point-based information is translated to real-valued functions in the Cartesian plane. In this way, information that is true/false at a point becomes a 0-1 function.

In the area of temporal databases the topic of dealing with imperfect relations has also received attention. Dyreson (1994), Dyreson and Snodgrass (1993, 1998), and Cowley and Plexousakis (2000) considered the problem of handling temporal indeterminacy in databases, by calculating the probabilities of the three basic relations between two temporal points. They stressed the need for the development of a query language for querying temporal information and have studied the complexity of processing different queries, which requires dealing with temporal indeterminacy. They did not however consider imperfect interval relations, nor the representation of imperfect relations between points.

Many other authors have attempted to incorporate the handling of imperfection in their reasoning about temporal relations. The probabilistic approach is a numerical technique most widely used for that purpose, and which is incorporated, for example, into probabilistic temporal networks (also called Bayesian networks), see for example (Heckerman & Wellman, 1995; Tawfik & Neufeld, 1994, 1996; Young, 1996; Young & Santos, 1996). In the research area of integrating time and probability for instance, there are a number of approaches that have introduced into temporal contexts mathematical formalisms such as Bayesian networks (e.g., Haddawy, 1994; Pearl, 1992) and Markov processes (e.g., Shafer, 1988). Moreover, a number of approaches to handling temporal uncertainty by integrating probabilistic reasoning into temporal logic have been proposed, e.g. (Goodwin *et al.*, 1994; Eberbach & Trudel, 1993).

Some research attention has also been paid to non-probabilistic methods for handling imperfect temporal relations. For example, Godo and Vila (1995) proposed a propositional temporal language based on fuzzy temporal relations. This language is provided with a natural possibilistic semantics to account for the imperfection issued by the fuzziness of temporal relations. The authors (Godo & Vila, 1995) also have presented an inference system, based on specific rules for dealing with temporal relations. Logical formalisms have also been applied for dealing with temporal intervals (e.g., Barber, 2000). Furthermore, Almeida (1999) proposed a new system of relations for reasoning about general intervals of time. This system is an integration of Allen's theory (Allen, 1983, 1984) and a one-dimensional version of Region Connection Calculus (Cohn *et al.*, 1997). And, finally, an extension of Allen's Interval Algebra based on fuzzy sets was proposed, for example, by Badaloni and Giacomin (2000).

In this chapter we have considered the representation of imperfect relations between temporal intervals. In the next chapter we will present the basics of the reasoning mechanism used in this thesis, and review some relevant reasoning techniques.

# 8 REASONING ABOUT TEMPORAL RELATIONS

Formalisms for reasoning about temporal relations are intended to handle temporal relations between temporal primitives (points, intervals, etc.), on the basis of the properties of the underlying time ontology. In this chapter, we overview the basics of the approaches intended for reasoning about temporal relations. In the first section, we present the reasoning operations that are used in the thesis. In the second section, we discuss the related temporal reasoning mechanisms.

## 8.1 Reasoning operations

In this thesis we consider three main reasoning operations: inversion, composition, and addition. Two of these operations, inversion (Figure 9a) and composition (Figure 9b), are widely used operations within different temporal reasoning formalisms.

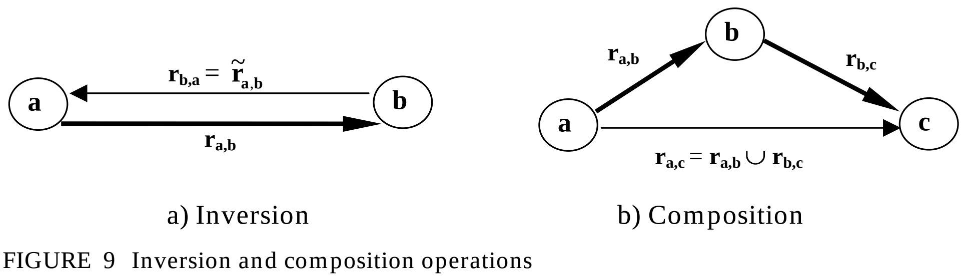



FIGURE 9 Inversion and composition operations

The operation of inversion allows us to derive the relation $\mathbf{r}_{b,a}$ between two temporal points **b** and **a**, when the relation $\mathbf{r}_{a,b}$ between **a** and **b** is known, and $\mathbf{r}_{b,a} = \tilde{\mathbf{r}}_{a,b}$, as is presented in Figure 9a. The operation of composition ($\cup$) derives the relation $\mathbf{r}_{a,c}$ between the temporal points **a** and **c**, when the relation $\mathbf{r}_{a,b}$ exists between the points **a** and **b** and the relation $\mathbf{r}_{\mathbf{b,c}}$ between the points **b** and **c**, as is presented in Figure 9b.

The rules for performing these operations using a symbolic representation of the basic relations have been proposed by Vilain and Kautz (1986). In this thesis we extend these proposals to the reasoning about imperfect temporal relations between points. We propose formulas to calculate the probabilities of the basic relations within the resulting relation, taking into account that the operands are imperfect relations.

In many situations, we need to deal with more than one possible imperfect relation between two temporal points. This happens when the information about a relation, for instance, is collected from a number of information sources or experts. For example, according to the first expert we might know that the relation between the points **a** and **b** is $\mathbf{r1}_{\mathbf{a,b}}$. At the same time, the second expert suggests that the relation between these points is $\mathbf{r2}_{\mathbf{a,b}}$. The binary operation of addition (∫ ) combines $\mathbf{r1}_{\mathbf{a,b}}$ and $\mathbf{r2}_{\mathbf{a,b}}$ into a single imperfect relation $\mathbf{r}_{\mathbf{a,b}}$ between the points **a** and **b**, as is illustrated in Figure 10.

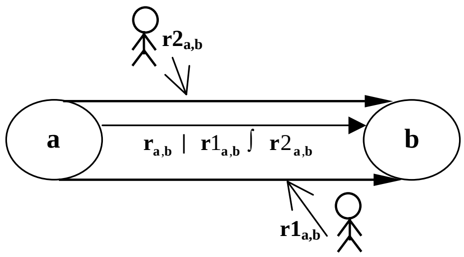



FIGURE 10 Operation of addition

The operation of addition is similar to the operation of intersection as defined, for example, in the Interval Algebra of Allen (Allen, 1983), in the Point Algebra of Vilain and Kautz (Vilain & Kautz, 1986), and further in the extensions of these by van Beek (van Beek, 1990, 1991, 1992), and Dey *et al.* (1996). We extend the operation of intersection, which uses a symbolic representation of temporal relations, with probabilistic measures of imperfection, and provide formulas to calculate the probability values within the resulting vector.

## 8.2 Related research

Allen's Interval Algebra (IA) (Allen, 1983) is definitely one of the most relevant formalisms for dealing with qualitative temporal relations. IA models the relation between two temporal intervals as a suitable subset of the set of the thirteen basic relations (Figure 6, Section 7). In such a way, $2^{12}$ relations can be specified between any pair of intervals, including the empty relation corresponding to the empty set of the basic relations. In IA the unary operation of inversion, and the binary operations of intersection and composition were defined. However, Allen’s formalism proved to be computationally intractable. Further research in this area therefore, paid significant attention to identifying the tractable fragments of IA, obtained by restricting the set of allowed relations. The three most important results from this perspective, Vilain and Kautz’s Point Algebra (Vilain & Kautz, 1986), van Beek’s Continuous Endpoint Algebra (van Beek, 1990), and Nebel and Bürckert’s ORD-Horn Algebra (Nebel, 1996; Nebel & Bürckert, 1995), are overviewed below.

Vilain and Kautz (1986) proposed a Point Algebra (PA), which models the relation between any two points as a subset (disjunctions) of the set of the three basic relations. In this way, they allow the specification of the following relations: “<”, “=”, “>”, “Ω’, “∅’, “∏’, “?”, plus the empty relation. To facilitate reasoning concerning these relations, Vilain and Kautz defined the unary operation of inversion, and the binary operations of intersection and composition. The relation between two temporal intervals is represented in PA, as it is in this thesis, as a conjunction of the four relations between the endpoints of the intervals. However, PA does not allow the representation of any disjunction of Allen’s relations between two intervals. So, for example, it allows the representation of the interval relation “before” or “meets” or “overlaps”, but does not allow the representation of “before” or “after”, because in the latter case a disjunction of point relations is required to represent it.

Van Beek (1989, 1990, 1992) identified a relevant proper subset of PA, called Continuous Endpoint Algebra (CEA), which only models continuous relations between time points. A relation $r_i$ between two points $P_1$ and $P_2$ is said to be continuous if, for any admissible value $x$ of $P_1$, from the fact that $x\ r_i\ y_1$ and $x\ r_i\ y_2$, where both $y_1$ and $y_2$ are admissible values for $P_2$ and $y_1 < y_2$, it follows that $x\ r_i\ y$, for any value y of $P_2$ such that $y_1 < y < y_2$, and visa versa (Chittaro & Montanari, 2000). It is possible to show that all the relations, except for “∏’, between two temporal points are continuous. This however means, that van Beek’s CEA is not able to model any relation between two intervals whose representations involve the relation “∏’ between the endpoints of the intervals. CEA covers approximately 1% of IA (Nebel & Bürckert, 1995).

Nebel and Bürckert’s ORD-Horn subclass of IA (OH) (Nebel, 1996; Nebel & Bürckert, 1995) is a strict superset of Simple Interval Algebra (SIA) (Tawfik & Neufeld, 1996), and it can be viewed as on the borderline between tractability and intractability, as was pointed out in Chittaro and Montanari (2000). OH is claimed to be the only maximal subset of IA which includes all the basic relations, and for which path-consistency is sufficient for deciding satisfiability (Nebel & Bürkert, 1995), and in which Ladkin and Reinefeld’s algorithm (Ladkin & Reinefeld, 1992) is sound and complete (Nebel, 1996). It has been shown that PA and CEA are proper subclasses of OH, which covers approximately 10 % of IA (Nebel, 1996).

There are a number of other approaches to reasoning about temporal relations, for example, Golumbic and Shamir’s Macro-Relation Algebra (Golumbic & Shamir, 1993), Ligozat’s Generalized Interval Calculus (Ligozat, 1991), Dechter, Meiri, and Pearl's Distance Algebra (Dechter *et al.*, 1991), the reasoning about repeating events by Morris and Khatib (1999), and the reasoning about qualitative and quantitative temporal information by Wetprasit and Sattar (1998). An overview of these and other approaches can be found in Chittaro and Montanari (2000).

## 8.3 Constraint Satisfaction Problem

One of the main research problems considered in the area of reasoning about temporal relations is the so-called Constraint Satisfaction Problem (CSP). In many applications it is supposed that we are given a set of temporal events, represented using temporal primitives, and a set of temporal relations between these events. Such kinds of situations are often represented as a directed graph with nodes representing temporal events, and arcs representing the temporal relations.

A general CSP can be represented by a finite set of variables (or nodes), their associated domains, and a set of constraints on these variables (Chittaro & Montanari, 1996, 2000). The domain of a variable is the set over which the variable takes its values. Each element of the domain is called a label. The association between a variable *v* and a label *l* is denoted by the pair (*v*,*l*). Relations can be distinguished between unary, binary, ternary, ..., n-ary relations over (*v*,*l*) pairs. The basic constraint satisfaction problem consists of finding one (or all) assignments of labels to variables which satisfy the given relations. The early formulations of the CSP were proposed by Mackworth (1977) and by Mohr and Henderson (1986).

The CSP can be divided into two sub-problems according to the two general types of temporal relations: qualitative CSP and quantitative CSP. In qualitative CSP variables take their values over a set of possible relations among temporal entities. In the framework of temporal points taken as ontological primitives, a variable takes its values over a set of three basic relations (before, at the same time, and after) that can hold between any two temporal points. In the framework of temporal intervals taken as ontological primitives, a variable takes its values from the set of the thirteen interval relations proposed by Allen (1983).

Two fundamental problems arise in this area: 1) deciding whether or not there exists an instant of the variables that satisfies the given set of relations; 2) determining for each pair of temporal entities, the strongest implied relation between them. In the sub-area of interval based formalisms the first problem is often referred to as the interval satisfiability problem (ISAT), and the second problem is called the minimal labeling problem (MLP), as in van Beek (1989, 1991) or the deductive closure problem (van Beek & Cohen, 1990).

Approaches for reasoning about temporal relations use a number of different reasoning algorithms for constraint satisfaction, which are based on the early works of Montanari (1974), Mackworth (1977), and Freuder (1978, 1982). These approaches use a network representation of relations by associating temporal entities to nodes, and labeling the arcs with the algebraic elements that denote the temporal relation between the nodes they connect. The problem of computing the minimal representation of a set of temporal relations, or proving its unsatisfiability, concerns the problem of computing the minimal consistent labeling of the correspondent network, a problem which was considered by van Beek (1989, 1992), by Kautz and Ladkin (1991), and by Vilain and Kautz (1986). The complexity of this problem depends on the expressiveness of the temporal formalism used.

Dechter *et al.* (1991) presented network-based method for constraint satisfaction that is able to handle continuous variables. The proposed by Dechter *et al.* framework allows to specify bounds on a temporal point, similarly as we do in this thesis, i.e. a point can be represented with lower and upper bounds of the interval, where the point is actually can be found. In that approach it is also possible to specify the bound on the difference between events, and this can be generally compared with the notion of maximum measurement difference in our work.

The temporal information might be relative (for example, A occurred before B) or metric (A had finished at most 5 minutes after B finished). The unique feature of the approach of Dechter *et al.* (1991) is in allowing the processing of metric information, i.e. assessments of time difference between events. Also in that paper the algorithms have been proposed for performing the following reasoning tasks: 1) finding all feasible times that a given event can occur; 2) finding all possible relationships between two given events; and 3) generating one or more scenarios consistent with the information provided.

# 9 APPLICATIONS

In this chapter we will discuss two possible application areas for the formalism proposed in this thesis. These are medical and industrial diagnostics, which are actually the two sub-areas of the big research field of automated diagnostics. Nevertheless the formalism to be applied is almost the same for both application areas; the motivation for the application and background of these areas and related research in those areas will be discussed separately in two sections below.

## 9.1 Medical diagnostics

It is almost inconceivable to try to represent clinical data and reason about them without a temporal dimension. As it was pointed out in many papers, i.e., (Nguyen *et al.*, 1999), dedicated to the problem of medical applications, the ability to reason with time-oriented data is central to the practice of medicine. Monitoring clinical variables over time often provides information that drives medical decision-making (e.g., clinical diagnosis and therapy planning). In some medical diagnostic applications, temporal information about the occurrence of the symptoms is vital for correct diagnostics and some medical expert systems, for example, Hamlet and Hunter (1987) tried to deal with this aspect. The crucial role of temporal representation and reasoning for modern medical information and decision support systems was also underlined by Shahar (2000), Dojat and Sayettat (1996), Combi and Shahar (1997), and Gamper (1996).

In Wainer and Rezende (1997) a temporal extension to the Parsimonious Covering Theory (PCT) was proposed to allow one to associate to a disease a temporal evolution of its symptoms. PCT is based on a model that associates to each disease a set of symptoms it may cause. Thus, PCT assumes that, at the moment of diagnostic, all symptoms are observable and that the order of occurrence of these symptoms is irrelevant for the diagnostic. The work of Wainer and Rezende extends the PCT model in such a way that to each disease one associates evolutions of symptoms (or sets of possible histories of symptoms). Thus, at diagnostic time, one will not just describe the symptoms present, as one would in a static diagnostic system, but describe the whole evolution of the symptoms.

One way, incorporating fuzzy temporal reasoning within diagnostic reasoning, was proposed by Wainer and Sandri (1999). Disorders are described as an evolving set of necessary and possible manifestations. Ill-known moments in time, e.g. when a manifestation should start or end, are modeled by fuzzy intervals, which are also used to model the elapsed time between events, e.g. the beginning of a manifestation and its end. Patient information about the intensity and times in which manifestations started and ended are also modeled using fuzzy sets. The paper discusses the ways to use that information to make predictions about future and past events in diagnosis.

A template system is described in Lowe *et al.* (1999) that uses fuzzy set theory to provide a consistent mechanism of accounting for uncertainty in the existence of events, as well as vagueness in their starting times and duration. Fuzzy set theory allows the creation of fuzzy templates from linguistic rules. The fuzzy template system that is introduced in this paper can accommodate multiple time signals, relative or absolute trends, and obviates the need to also design a regression formula for pattern matching. The target application for the fuzzy template system was anesthesia monitoring.

In Nejdl and Gamper (1994) a framework is described for model-based diagnosis of dynamic systems by using and expressing temporal uncertainty in the form of qualitative Allen's interval relations. Based on a logical framework extended by qualitative and quantitative temporal constraints it was shown how to describe behavioral models, how to use abstract observations and how to compute abstract temporal diagnoses. This yields an expressive framework, which allows the representation of complex temporal behavior with temporal uncertainty. An example of hepatitis diagnosis was considered. In later work (Gamper & Nejdl, 1997) the similar example of hepatitis B was considered to describe a model-based framework for complex temporal behavior. The concept of abstract observations was introduced as an abstraction from observations at time points into assumptions over time intervals. This leads to a more intuitive representation and makes diagnosis independent of the number of actual observations and the granularity of time.

All the mentioned above references underline the necessity to take into account temporal dimension in many areas of medicine, including diagnostics. Let us consider one possible situation. Imagine, that a patient has come to a therapist already when his unknown illness has quite neglected form. The therapist asks about dynamics of symptoms and usually the patient cannot answer precisely what particularly happened and when. Even more, the patient does not always remember which symptoms occurred first, i.e. temporal relations between symptoms are uncertain. From such a story the therapist derives quite uncertain scenario of the illness' dynamics but still needs to estimate possible diagnosis to decide about further investigation. To be able to solve such and similar diagnostic problems we propose to use the following approach.

We assume that there exist a certain number of symptoms (events) that are critical for particular diseases that we are able to classify. Examples of such symptoms are “body temperature exceeded 39°C” and “a qualitative characteristic of hemoglobin according to the blood analysis is lower than 90”. We believe that in many diagnostic situations it is easier to put the diagnosis and it will be more precise if we will take into account temporal relationships between the symptoms occurred.

The core of the approach to temporal diagnostics we propose is generation and further use of temporal scenarios of known illnesses. Such scenario is a relational network, where the nodes are the symptoms from the set of possible symptoms, and the arcs are the temporal relations between the symptoms. We propose to represent these relations using the formal mechanism proposed in this thesis, i.e. these relations are uncertain temporal relations. A particular course of illness for a particular patient is formalized using also a relational network with symptoms and relations between them. By this it is probable that the relations will be certain at this stage. When we observed a number of cases of this particular illness from a number of patients, we can generate a temporal scenario (pattern). At this stage we combine the networks for the illness into the one scenario. By this, it is very probable that the relations between the symptoms will become uncertain, since the same illness can have a different course for different patients.

Temporal scenarios for known illnesses are stored in a database and can be accessed when we try to put the diagnosis. In this situation, we can compare a relational network, describing the particular course of illness, with known temporal scenarios using the mechanism proposed in Article 8. To perform this comparison we use original measures of distance between different uncertain relations, and between network and scenario. In this way, we are able to provide a therapist with potential diagnoses (if appropriate) and their probabilities.

## 9.2 Industrial diagnostics

Industrial diagnostics is also one important application area for AI formalisms. In many industrial process control and monitoring tasks there is a need to identify and classify the situation occurred, and this need is crucial to enabling process improvements and the successful operation of industrial equipment. The examples of such situations are: automated inspection (flaw location) (Wilson *et al.*, 1998), automated diagnosis of failures occurred (Struss, 1992), fault localization in power transmission networks (Beschta *et al.*, 1993), and real-time diagnosis of the situation to prevent failures in future.

As it was pointed out by Struss (1997) the task of diagnosing technical systems becomes increasingly challenging. The complexity of these systems grows up rapidly, and therefore the complexity of the task of diagnostics does also. This relates to many types of industrial systems, such as production plants and factories, communications networks, and transportation systems.

In industrial, as well as in medical, diagnostics we use the approach of model-based diagnosis. The key idea of this approach is that we explicitly represent the knowledge about the device monitored, its structure, and operational behavior as a model. The diagnostics itself is organized as an inference process based on this model and the observed behavior (Dressler & Struss, 1996). This approach created the demand for rigorous theoretical foundations for automated diagnosis, and became one of the major fields of application and an important touchstone for the utility of many AI theories.

A number of techniques to temporal diagnostics have been proposed so far, but there still exist problems there require further research attention. One such problem is when the temporal information obtained from an industrial object is uncertain. Many researchers have paid their attention to this problem and a number of formalisms for handling imperfect information have been tried in those applications. For example, Markov chains were used in modeling temporal evolutions in model-based diagnosis by Portinale (1992). It was assumed in that paper, that probabilistic temporal knowledge is available for each component of the system.

The need for combination of uncertainty management techniques and temporal reasoning formalism for medical and industrial diagnosis and prediction was underlined in Arroyo-Figueroa and Sucar (1999). In that paper a novel representation called Temporal Nodes Bayesian Networks was proposed. In such a network each node represents an event or state change of a variable, and an arc corresponds to a causal-temporal relationship. Multiple granularity was also allowed based on the assumption that temporal intervals can differ in number and size for each temporal node. The proposed approach was applied to fault diagnosis and prediction as a subsystem for a fossil power plant.

The main advantage of temporal diagnostics is that it considers not only a static set of symptoms, but together with the time they were monitored, and which allows to have a broader view on the situation. Moreover, sometimes only considering temporal evolution of relations between different symptoms can give us a hint to precise diagnostics.

In general, the mechanism for industrial temporal diagnostics using generation and recognition of uncertain temporal scenarios is similar to the one already discussed in the previous section. The conceptual schemas for our approach to industrial temporal diagnostics can be found in Article 9 of this thesis.

# 10 ORGANIZATION OF THE THESIS

In this chapter we present the logical structure of the thesis and explain the organization of the articles included in the thesis.

## 10.1 Logical structure of the thesis

We presented above the introduction to the work, the classifications and methods used for dealing with imperfect information, and an overview of the current approaches to temporal representation and reasoning. We also summarized the research problem addressed by the thesis. The reminder of the study, which is also the main part of the thesis, includes nine articles. The classification of these articles and logical relationships between them are presented in Figure 11.

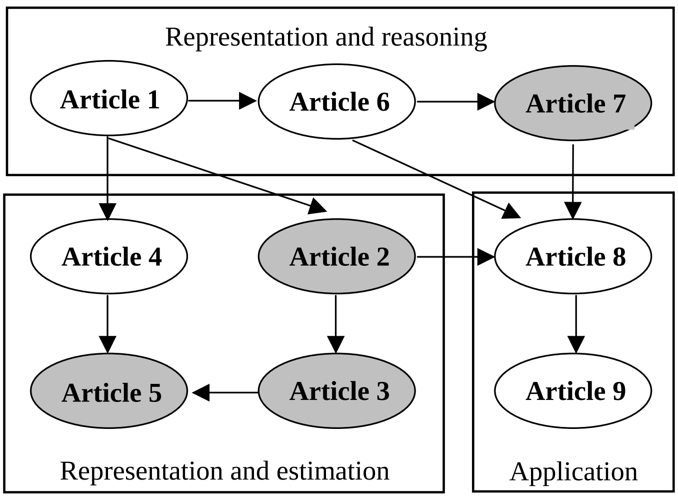



FIGURE 11 Articles included in the thesis

The nine articles included in the thesis can be classified in two ways. The first classification can be made based on the type of temporal primitives used in the articles. So, the first group (shown with a white background in Figure 11), includes articles 1, 4, and 6, and is concerned only with the representation of and reasoning about the relations between temporal points. The second group (shown with a gray background), includes articles 2, 3, and 5, and is concerned with the relations between temporal intervals. Finally, the third group, including articles 8 and 9, deals with both types of primitives.

The articles can also be divided into tree other groups using the criteria of research problem they deal with. The group of articles 2, 3, 4, and 5 deals with the representation of imperfect temporal relations and the estimation of these relations using the information about temporal primitives. Articles 1, 6 and 7 consider representation and reasoning with imperfect relations between temporal points and intervals. Finally, the group consisting of articles 8 and 9 discusses one possible application area for the formalism proposed in the thesis.

Arrows in Figure 11 stand for the logical connection between particular articles. For example, an arrow from Article 1 to Article 2 means that the latter one uses the results obtained in Article 1.

All the articles use the same representation of the imperfect temporal relations using the probabilities of the basic relations. The proposed approaches for estimating these uncertain relations show how this representation can be “filled” with the particular probability values. The reasoning operations show how the obtained imperfect relations can be used further to derive the previously unknown relations in a relational network combining the already known relations. We also show one possible application area (temporal diagnostics) for the formalism proposed in the thesis.

## 10.2 Contents of the thesis in brief

In the article, titled “Representation and Reasoning with Uncertain Temporal Relations”, we propose the representation of uncertain relations between temporal points, which includes explicit probability values of the consistent and the inconsistent parts of a temporal relation. The probability of the consistent part of the relation is divided between the basic temporal relations, i.e. “<” (before), “=” (at the same time), and “>” (after). The inconsistent part of the relation has one probability value, which is divided between the three basic relations presenting the percentage values of their support among knowledge sources. Both the probabilities and the percentage values are used in our reasoning mechanism, which consists of three operations: inversion, composition, and addition. These can be used to derive the probability and percentage values for a relation between any two temporal points.

In the article, titled “Uncertain Relations between Indeterminate Temporal Intervals”, we estimate the uncertain relation between two indeterminate intervals. The relation between two intervals is represented using the four relations between the endpoints of these intervals, as it is in Chapter 6. We first estimate these four relations between the endpoints, and then calculate the probabilities of Allen's relations between the intervals. To estimate the relation between two endpoints we compose three formulas for the three probabilities of the basic relations.

In the article, titled "Probabilistic Estimation of Uncertain Temporal Relations", we estimate the uncertain relation between two temporal intervals. The main problem discussed in this chapter is how to compose the probabilities of Allen's relations using the probabilities of the basic relations between the endpoints of the intervals. We formalize the dependencies between the values of the four relations between the endpoints of the intervals. After that, we compose the conditional probabilities for different combinations of these four relations. The probabilities of Allen's relations are composed as the joint conditional probabilities of the relations between the endpoints.

In the article, titled "Estimation of Uncertain Relations between Indeterminate Temporal Points", we are concerned with estimating uncertain relations between two indeterminate temporal points assuming that the temporal values of these points might be obtained with some measurement error, and this error needs to be taken into account when estimating the temporal relation between two indeterminate points (Ryabov & Puuronen, 2000). We assume that the probability mass functions of the values of these two temporal points, and the maximum measurement error are known. Using these assumptions we derive formulas for the lower and the upper bounds of the probabilities of the basic relations between two indeterminate temporal points.

In the article, titled "Estimating Uncertain Relations between Indeterminate Temporal Points and Intervals", we estimate the uncertain relation between two indeterminate temporal intervals, when the measurements of their temporal values may include some error. Combining the approaches proposed in Articles 3 and 4 we propose the formulas for the lower and upper probabilities of Allen's intervals relations between two indeterminate intervals.

In the article, titled "Probabilistic Reasoning about Uncertain Relations between Temporal Points", we present an extension of the approach proposed in the first article. The basic representation of the uncertain relation between two temporal points is proposed to be simpler for those situations where the formalization of inconsistent information is not required. The four reasoning operations (inversion, composition, addition, and negation) are redefined, taking into account the new representation.

In the article, titled "Handling Uncertain Interval Relations" we continue considering the topic of reasoning with imperfect temporal relations. The results achieved in (Ryabov & Puuronen, 2001) are extended to the representation using Allen's interval relations. The definitions for the probabilistic representation, as well as for the operations of inversion, composition, and addition, are given for interval relations. In this paper, we also show how to translate uncertain interval relations using Allen's representation into the representation using the relations between the endpoints of these intervals, which can simplify further reasoning in some practical situations.

In the article, titled “Abstract Diagnostics Based on Uncertain Temporal Scenarios” we showed one possible application area for the formalism proposed in the thesis. This area is temporal diagnostics. Particularly, in this article we considered the case of medical diagnostics, where identification of temporal patterns plays an important role. Temporal scenarios of different diseases, therapy protocols, and other temporal graphs provide important additional information for medical decisions. In this paper we propose to use the algebra of uncertain temporal relations in diagnostic problems. We represent uncertain temporal relations within a scenario graph using the probabilities of the basic relations that can hold between two temporal primitives. Also in the paper we show how: (a) to generate temporal scenarios by integrating appropriate relational networks from already diagnosed cases; and (b) to classify a new case using the measure of the distance between a network and a scenario.

In the article, titled “Industrial Diagnostics using Algebra of Uncertain Temporal Relations” we continue discussing possible applications for our formalism. Another application sub-area (within the area of automated diagnostics) is industrial diagnostics. Often, the information taken from an industrial object could be uncertain making the task of diagnostics more complex. We propose to use the algebra of uncertain temporal relations in solving this problem. We estimate temporal relations between the set of symptoms (crucial values of important variables) obtained from an industrial object to build the temporal relational network for this particular situation. After that, we compare the obtained network with known temporal scenarios (patterns) of failures, using the numerical measure of the distance between a network and a scenario. Using this approach we derive the probabilities of possible diagnoses for the particular situation. We also show how the learning for the database of scenarios can be performed, which will make diagnostics for future cases more precise.

## 10.3 About the joint articles

The first article introducing the new representation of imperfect temporal relations and basics of reasoning mechanism is co-authored with Seppo Puuronen and Vagan Terziyan. My particular contribution is in the definition of the representation of uncertain and inconsistent relations (Section 2), and partially in the definition of the reasoning operations (Section 3).

Article 4, presenting the formalism for estimation of uncertain relations between indeterminate temporal points is co-authored with Seppo Puuronen. My major contributions to that paper are related to the general idea of that approach, definitions of the basic concepts (Section 2), developing the formalisms for estimation of uncertain relations (Sections 3 and 4), and the motivating example (Section 5).

Article 6, presenting the approach to reasoning with uncertain temporal relations between points is co-authored with Seppo Puuronen. My contribution to that paper relates to the definition of the basic concepts (Section 2), the properties of reasoning operations (partially Section 3), reasoning mechanism (Section 3), and the example (Section 4).

In Article 8, the proposed in this thesis approach is applied to abstract temporal diagnostics, and the particular case of medical diagnostics is considered. My contribution to that paper is partially in the development of the general idea of that approach, the definition of the basic concepts (except for the measure of distance), definitions of the reasoning operations, and the algorithms for generation and recognition of temporal scenarios.

Article 9 proposes the industrial diagnostics as another application area for the mechanism described in this thesis. My contribution in that paper is partially in the conceptual schema for industrial diagnostics, the definition of the basic concepts and the algorithms for generation and recognition of temporal scenarios.

# 11 CONTRIBUTION, LIMITATIONS, AND FUTURE WORK

In this chapter we discuss the contribution of the thesis. We also address the major limitations of the study.

In this thesis we are dealing with the problem of handling imperfect temporal relations between temporal points and intervals. The work is situated in the area of temporal representation and reasoning and in the area of numerical techniques for handling imperfect information. We propose a formal numerical technique for the representation of, estimation of, and reasoning about imperfect temporal relations based on the probabilistic approach. The contribution of the thesis is described according to the three main issues related to handling imperfect temporal relations: representation, estimation, and reasoning. We also distinguish between two sub-areas: handling temporal relations between points and between intervals.

The main representational idea of the thesis is that an imperfect temporal relation between points is presented by providing the probability values of the basic relations that can hold between these points. For that purpose, we propose to use the uncertainty vector, which includes the probabilities of the basic relations. We define two types of representation: 1) the representation of uncertain and inconsistent temporal relations between points; and 2) the representation of uncertain temporal relations between points. The difference between these two representations was discussed in Chapter 6.

According to the classification presented in Section 1 of Chapter 4 we assume the epistemic interpretation for the probabilities. Also, our probabilities are rationalistic, theoretical, and behavioral. Finally, our probabilities are constructive since we allow assessing them from the available evidence, e.g., using the information about indeterminate temporal primitives we estimate the probabilities of the basic relations between them.

From the estimation of the imperfect temporal relation between two temporal points, we derive the probability values for the three basic relations, using the information about the points. We assume that the temporal points can be indeterminate, as was discussed in Chapter 5, which means that they are defined within the intervals of possible values, together with their probability mass functions. We propose formulas for the probabilities of the basic relations. Although this question has already been discussed in some papers in the area of temporal databases, we propose the formulas for the probabilities of the basic relations between two endpoints, which is a novel element. Moreover, we suppose that the measurements of the values of the temporal points may include some error, which needs to be taken into account during the estimation. We assume that the maximum value of this measurement error is known, and derive the formulas for the lower and upper probabilities of the three basic relations.

The proposed reasoning mechanism for temporal points, including the four operations, derives new imperfect relations in a relational network combining known imperfect relations. Our contribution here is, that we extend the standard definitions of the reasoning operations to ability to process imperfect temporal relations, which are represented as uncertainty vectors. The resulting relation is also represented as an uncertainty vector.

There are the two main ways to represent the temporal relation between two intervals. We either can use Allen's thirteen interval relations (Allen, 1983), or we might consider the four relations between the endpoints of these intervals, as was shown in Chapter 7. In the case when the intervals are indeterminate, it is in many situations not evident, which Allen's relation holds between the two indeterminate intervals. Moreover, a number of Allen's relations might be probable. We propose to estimate the imperfect relation between two indeterminate intervals by calculating the probabilities of Allen's relations, taking into account the information about the endpoints of the intervals. In that way, we show the transition between the two ways of representation of the imperfect relation between two indeterminate temporal intervals, which is also an important contribution of the thesis.

To compose the formulas for the probabilities of Allen's relations between two indeterminate intervals, we study and formalize the dependencies between the temporal values of the endpoints of these intervals. After that, we compose the probabilities of Allen's relations as the joint conditional probabilities of the relations between the endpoints, taking into account the dependencies derived.

We discuss the computational complexity of the estimation of Allen's relations. This discussion provides hints to the understanding of the behavior of the approach in real applications. In the thesis we also consider several examples from the area of temporal databases, which use synthetic but realistic settings, and help to illustrate the possible applications for the mechanism proposed in the thesis.

We proposed an application for the presented formalism in the area of temporal diagnostics (medical and industrial). Temporal diagnostics is based on analyzing temporal relations between values of crucial variables. We estimate temporal relations between the set of symptoms (crucial values of important variables) obtained from an industrial object (in industrial diagnostics) or a patient (in medicine) to build the temporal relational network for this particular situation. After that, we compare the obtained network with known temporal scenarios (patterns) of failures (diseases), using the numerical measure of the distance between a network and a scenario. Using this approach we derive the probabilities of possible diagnoses for the particular situation.

The presented work has several limitations. The study is mainly theoretical and we concentrated on the development of a formal approach to handle imperfect temporal relations. Nevertheless, we have proposed the application area where the formalism could be useful.

The approach proposed in this thesis is limited to the use of the discrete time model. This is motivated by the wide use of this model in many application areas of temporal reasoning, and particularly in the area of temporal databases. Moreover, the basic idea of the continuous time model can be modeled to some extent on the discrete time model by defining different time granularities (Bettini *et al.*, 1996; Bettini *et al.*, 2000; Wang *et al.*, 1997). The discrete time model is easier to implement compared to the continuous model, and its formalisms are simpler.

In the articles included in the thesis, where we discuss the estimation and representation of imperfect relations, we make the independence assumption when we derive the probabilities for the basic relations between two indeterminate points. Particularly, the points **a** and **b** are defined within two intervals of indeterminacy, and we consider the probabilities of pairs of values from these intervals. In this situation we assume that values of **a** and **b** are independent, and therefore we assume that the probability of the pair is a multiplication of the corresponding values of the p.m.f. functions for these points. This independence assumption made does not hold in some application areas, and which in its own turn limits the applicability of the proposed formalism.

Also, in a number of articles within this thesis defining temporal interval consisting of indeterminate points $\mathbf{s}[\mathbf{s}^l,\mathbf{s}^u]$ and $\mathbf{e}[\mathbf{e}^l,\mathbf{e}^u]$ denoting the start and the end of this interval accordingly, we made an assumption that $\mathbf{s}^u<\mathbf{e}^l$. This assumption was made to simplify the formulas for estimating the uncertain relations between indeterminate intervals, and easily can be relaxed in the implementation by providing additional consistency checking for indeterminate intervals.

In the thesis we have considered only two types of imperfection within temporal relations according to the classification by Parsons (1996), that is, uncertainty and inconsistency. The proposed formalism is limited to deal with these two types of imperfection. Although we also consider a disjunctive form of imprecision (indeterminacy) relative to temporal primitives, which is one source of imperfection within the relation between them, there still exists imprecision, ignorance, and incompleteness, which in some situations can also be found within temporal relations.

One of further research concerns is the implementation. A possible idea is to implement the technique within the context of the Semantic Web. In this case, the implementation may consist of a Java class to capture the computational aspects of the technique. This Java class could then be used to implement an ontology in RDF and e.g., OWL, in this way extending the Semantic Web.

Secondly, we plan to investigate further the application of the mechanism in industrial temporal diagnostics. Preliminary research has revealed that the model of branching time might be very relevant in this domain. Even perfect knowledge about temporal primitives in each branch can lead to imperfect temporal relations since the branches may not be synchronized, as it was discussed, for example, by Lamport (1998).

Another possible application for the formalism proposed in this thesis is natural language processing. In this domain we can find natural language concepts, which can help to estimate the initial probabilities of temporal relations necessary for reasoning. We might also perfectly know the temporal relations between two such concepts but the information about the primitives remains uncertain. It would be interesting to investigate this issue, and to try to find a heuristic connection between such temporal language concepts and numerical measures of imperfection of temporal relations.

In the situation when the endpoints of a temporal interval are indeterminate temporal points, we can consider also a set of probable intervals taking as their endpoints different combinations of the values from the intervals of indeterminacy. The representation and reasoning technique proposed in this thesis could be extended in the future to cover sets of probable intervals. Other possible directions for future research include: extending the formalism to deal with imprecision, incompleteness, and ignorance. Moreover, we think that the proposed approach can be extended to deal with the continuous time model, which might be useful in some applications.

# REFERENCES


Allen, J. 1983. Maintaining knowledge about temporal intervals. Communications of the ACM 26 (11), 832-843.

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

Allen, J. & Ferguson, G. 1994. Actions and events in interval temporal logic. Journal of Logic and Computation 4 (5), 531-579.

Almeida, M. 1999. A system for reasoning with nonconvex intervals. In C. Dixon & M. Fisher (Eds.) Proceedings of the 6-th International Workshop on Temporal Representation and Reasoning, Los Alamitos, California: IEEE Computer Society Press, 8-16.

Arroyo-Figueroa, G. & Sucar, L. 1999. A temporal Bayesian network for diagnosis and prediction. In K. Laskey & H. Prade (Eds.) Proceedings of the 15-th Conference on Uncertainty in Artificial Intelligence, San Francisco, California: Morgan Kaufmann Publishers, Inc., 13-20.

Badaloni, S. & Giacomin, M. 2000. A fuzzy extension of Allen's Interval Algebra, In E. Lamma & P. Mello (Eds.) Proceedings of the 6-th Congress of the Italian Association for AI, Lecture Notes in Artificial Intelligence 1792, Berlin, Heidelberg, New York: Springer, 155-165.

Barbara, D., Garcia-Molina, H. & Porter, D. 1992. The management of probabilistic data. IEEE Transactions on Knowledge and Data Engineering 4 (5), 487-502.

Barber, F. 2000. Reasoning on interval and point-based disjunctive metric constraints in temporal contexts. Journal of Artificial Intelligence Research 12, 35-86.

Barnett, J. 1981. Computational methods for a mathematical theory of evidence. In P. Hayes (Ed.) Proceedings of the 7-th International Joint Conference on Artificial Intelligence, Los Alamitos, California: William Kaufmann, 868-875.

Beschta, A., Dressler, O., Freitag, H. & Struss, P. 1993. A model-based approach to fault localization in power transmission networks, Insentient Systems Engineering 2 (1), 3-14.

Bettini, C., Jajodia, S. & Wang, S. 2000. Time granularities in databases, data mining, and temporal reasoning. Berlin, Heidelberg: Springer.

Bettini, C., Wang, X. & Jajodia, S. 1996. A general framework and reasoning models for time granularity. In L. Chittaro, S. Goodwin, H. Hamilton, & A. Montanari (Eds.) Proceedings of the Third International Workshop on Temporal Representation and Reasoning, Los Alamitos, California: IEEE Computer Society Press, 104-111.

Bigham, J. 1998. Correlation using uncertain and temporal information. In A. Hunter and S. Parsons (Eds.) Applications of uncertainty formalisms, Lecture Notes in Artificial Intelligence 1455, Berlin, Heidelberg: Springer-Verlag, 242-265.

Bonnissone, P. & Tong, R. 1985. Editorial: reasoning with uncertainty in expert systems. International Journal Man Machine Studies 22 (3), 241-250.

Bosc, P. & Prade, H. 1993. An introduction to fuzzy set and possibility theory-based treatment of flexible queries and uncertain or imprecise databases. In A. Motro & P. Smets (Eds.) Uncertainty management and information systems: from needs to solutions, Boston: Kluwer Academic Publishers, 285-324.

Bouzid, M. & Mouaddib, A.-I. 1998. Uncertain temporal reasoning for the distributed transportation scheduling problem. In R. Morris & L. Khatib (Eds.) Proceeding of the Fifth International Workshop on Temporal Representation and Reasoning, Los Alamitos, California: IEEE Computer Society Press, 21-28.

Böhlen, M., Busatto, R. & Jensen, C. 1998. Point- versus interval-based temporal data models, In S. Urban & E. Bertino (Eds.) Proceedings of the Fourteenth International Conference on Data Engineering, Los Alamitos, California: IEEE Computer Society Press, 192-200.

Chittaro, L. & Montanari, A. 1996. Trends in temporal representation and reasoning. The Knowledge Engineering Review 11 (3), 281-288.

Chittaro, L. & Montanari, A. 2000. Temporal representation and reasoning in artificial intelligence: issues and approaches. Annals of Mathematics and Artificial Intelligence 28 (1-4), 47-106.

Chountas, P. & Petronias, I. 2000. Factual and temporal imperfection. In T. Yakhno (Ed.) Proceedings of the First Biannual International Conference on Advances in Information Systems, Lecture Notes in Computer Science 1909, Berlin, Heidelberg: Springer-Verlag, 419-428.

Cohen, P. 1985. Heuristic reasoning about uncertainty: an artificial intelligence approach, London: Pitman.

Cohn, A., Bennet, B., Gooday, J. & Gotts, N. 1997. Representing and reasoning with qualitative spatial relations. In O. Stock (Ed.) Spatial and temporal reasoning, Dordrecht: Kluwer, 97-134.

Combi, C. & Shahar, Y. 1997. Temporal reasoning and temporal data maintenance in medicine: issues and challenges. In C. Combi & Y. Shahar (Eds.) Time-Oriented Systems in Medicine. Special Issue Computers in Biology and Medicine 27 (5), 353-368.

Cowley, W. & Plexousakis, D. 2000. Temporal integrity constraints with indeterminacy. In A. Abbadi, M. Brodie, S. Chkravarthy, U. Daya, N. Kamel, G. Schlageter & K.-Y. Whang (Eds.) Proceedings of the 26-th International Conference on Very Large Databases, San Francisco, California: Morgan Kaufmann Publishers, Inc., 441-450.

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

Dey, D., Barron, T. & Storey, V. 1996. A complete temporal relational algebra. The Very Large Databases Journal 5 (3), 167-180.

Dey, D. & Sarkar, S. 1996. A probabilistic relational model and algebra. ACM Transactions on Database Systems 21 (3), 339-369.

Dojat, M. & Sayettat, C. 1996. A realistic model for temporal reasoning in real-time patient monitoring, Applied Artificial Intelligence 10 (2), 121-143.

Dressler, O. & Struss, P. 1996. The consistency-based approach to automated diagnosis of technical devices. In G. Brewka (Ed.) Principles of Knowledge Representation, Stanford: CSLI Publications, 267-311.

Dubois, D., Lang, J. & Prade, H. 1991. Fuzzy sets in approximate reasoning, Part 2: Logical approaches. Fuzzy Sets and Systems 40 (1), 203-244.

Dubois, D. & Prade, H. 1987. Necessity measures and the resolution principle. IEEE Transactions on Systems, Man, and Cybernetics 17 (3), 474-478.

Dubois, D. & Prade, H. 1988. Possibility theory: an approach to the computerized processing of uncertainty, New York: Plenum Press.

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

Dubois, D., Prade, H., & Smets P. 2001. New semantics for qualitative possibility theory, In Proceedings of the 2-nd International symposium on Imprecise Probabilities and Their Applications, Ithaca, New York, (2.10.2002) http://decsai.ugr.es/~smc/isipta01/proceedings/index.html

Duda, R., Hart, P. & Nilsson, N. 1976. Subjective Bayesian methods for a rule-based inference system. In Proceedings of the National Computer Conference, AFIPS Conference Proceedings 45, Montvale, New York: AFIPS Press, 1075-1082.

Dyreson, C. 1994. Valid-time indeterminacy. University of Arizona, USA, Ph.D. thesis.

Dyreson, C. & Snodgrass, R. 1993. Valid-time indeterminacy. In Proceedings of the 9-th International Conference on Data Engineering, Los Alamitos, California: IEEE Computer Society Press, 335-343.

Dyreson, C. & Snodgrass, R. 1998. Supporting valid-time indeterminacy. ACM Transactions on Database Systems 23 (1), 1-57.

Eberbach, E. & Trudel, A. 1993. Representing spatial and temporal uncertainty, In B. Bouchon-Meunier, L. Valverde & R. R. Yager (Eds.) IPMU'92 - Advanced Methods in Artificial Intelligence, 4th International Conference on Processing and Management of Uncertainty in Knowledge-Based Systems, Lecture Notes in Computer Science 682, Heidelberg, Berlin: Springer, 129-138.

Ekenberg, L., Danielson, M. & Boman, M. 1997. Imposing security constraints on agent-based decision support. Decision Support Systems 20 (1), 3-15.

de Finetti, B. 1974. Theory of probability. Vol. 1, London: Wiley.

de Finetti, B. 1975. Theory of probability. Vol. 2, London: Wiley.

Fox, J. 1986. Three arguments for extending the framework of probability. In L. Kanal & J. Lemmer (Eds.) Uncertainty in artificial intelligence. Machine Intelligence and Pattern Recognition 4, Amsterdam, London, New York: Elsevier / North-Holland, 447-458.

Freksa, C. 1992. Temporal reasoning based on semi-intervals. Artificial Intelligence 54 (1-2), 199-227.

Freuder, E. 1978. Synthesizing constraint expressions. Communications of the ACM 21 (11), 958-966.

Freuder, E. 1982. A sufficient condition for backtrack-free search. Journal of the ACM 29 (1), 24-32.

Gabbay, D. & Hunter, A. 1991. Making inconsistency respectable 1: a logical framework for inconsistency in reasoning. In P. Jorrand & J. Kelemen (Eds.) Fundamentals of Artificial Intelligence Research, International Workshop (FAIR'91). Lecture Notes in Computer Science 535, Berlin, Heidelberg, New York: Springer, 19-32.

Gabbay, D. & Hunter, A. 1993. Making inconsistency respectable 2: meta-level handling of inconsistent data. In M. Clarke, R. Kruse & S. Moral (Eds.) Symbolic and qualitative approaches to reasoning and uncertainty. Lecture Notes in Computer Science 747, Berlin, Heidelberg, New York: Springer, 129-136.

Gamper, J. 1996. A temporal reasoning and abstraction framework for model-based diagnosis systems, Rheinisch-Westfälischen Technischen Hochschule, Aachen, Germany. Ph.D. thesis.

Gamper, J. & Nejdl, W. 1997. Abstract temporal diagnosis in medical domains, Artificial Intelligence in Medicine 10 (3), 209-234.

Godo, L. & Vila, L. 1995. Possibilistic temporal reasoning based on fuzzy temporal constraints. In C. Mellish (Ed.) Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, Vol. 2, San Mateo, California: Morgan Kaufmann Publishers, Inc., 1916-1923.

Golumbic, M. & Shamir, R. 1993. Complexity and algorithms for reasoning about time: a graph-theoretical approach. Journal of ACM 40 (5), 1108-1133.

Goodwin, S., Neufeld, E. & Trudel, A. 1994. Probabilistic temporal representation and reasoning, International Journal of Expert Systems 7 (3), 261-288.

Gordon, J. & Shortliffe, E. 1985. A method for managing evidential reasoning in a hierarchical hypothesis space. Artificial Intelligence 26 (3), 323-357.

Haddawy, P. 1994. Representing plans under uncertainty: a logic of time, change, and action, Lecture Notes in Computer Science 770, Berlin: Springer.

Hamblin, C. 1972. Instants and intervals. In J. Fraser (Ed.) The study of time. New York: Springer-Verlag, 325-331.

Hamlet, I. & Hunter, J. 1987. A representation of time for medical expert systems, In J. Fox, M. Fieschi & R. Engelbrecht (Eds.) Proceedings of the European Conference on Artificial Intelligence in Medicine (AIME’87), Lecture Notes in Medical Informatics 33, New York: Springer-Verlag, 112-119.

Heckerman, D. 1986. Probability interpretation for MYCIN's certainty factors. In L. Kanal & J. Lemmer (Eds.) Uncertainty in artificial intelligence. Machine Intelligence and Pattern Recognition 4, Amsterdam, London, New York: Elsevier / North-Holland, 167-196.

Heckerman, D. & Wellman, M. 1995. Bayesian networks. Communications of the ACM 38 (3), 27-30.

Hirsch, R. 1994. Relational algebra of intervals, Imperial College, London, UK. Ph.D. thesis.

Hirsch, R. 1996. Relational algebra of intervals. Artificial Intelligence 83 (2), 267-295.

Hunter, A. & Parsons, S. (Eds.) 1998. Applications of uncertainty formalisms. Lecture Notes in Computer Science 1455, Berlin, Heidelberg: Springer-Verlag.

Jensen, C. & Dyreson, C. (Eds.) 1998. The consensus glossary of temporal database concepts - February 1998 version. In O. Etzion, S. Jajodia & S. Sripada (Eds.) Temporal databases - research and practice. Lecture Notes in Computer Science 1399, Berlin, Heidelberg: Springer-Verlag, 367-405.

Jensen, C. & Snodgrass, R. 1999. Temporal data management. IEEE Transactions on Knowledge and Data Engineering 11 (1), 36-44.

Kahn, K. & Gorry, G. 1977. Mechanizing temporal knowledge. Artificial Intelligence 9 (1), 87-108.

Kautz, H. & Ladkin, P. 1991. Integrating metric and qualitative temporal reasoning. In T. Dean & K. McKeown (Eds.) Proceedings of the 9-th National Conference on Artificial Intelligence, Menlo Park, California: AAAI Press, 241-246.

Kwan, S. Olken, F. & Rotem, D. 1993. Uncertain, incomplete and inconsistent data in scientific and statistical databases. In A. Motro & P. Smets (Eds.) Uncertainty management and information systems: from needs to solutions, Boston: Kluwer Academic Publishers, 127-154.

Ladkin, P. & Reinefeld, A. 1992. Effective solution of qualitative interval constraint problems. Artificial Intelligence 57 (1), 105-124.

Lamport, L. 1998. Proving possibility properties. Theoretical Computer Science 206 (1-2), 341-352.

Laplace, P. 1951. A philosophical essay and probabilities, New York: Dover (originally published in 1820).

Ligozat, G. 1991. On generalized interval calculi. In T. Dean & K. McKeown (Eds.) Proceedings of the 9-th National Conference on Artificial Intelligence, Menlo Park, California: AAAI Press, 234-240.

Lindley, D. 1975. Making decisions, Chichester, UK: John Wiley and Sons.

Lowe, A., Jones, R., & Harrison, M. 1999. Temporal pattern matching using fuzzy templates, Journal of Intelligent Information Systems 13(1-2), 27-45.

Mackworth, A. 1977. Consistency in networks of relations. Artificial Intelligence 8 (1), 99-118.

McCarthy, J. & Hayes, P. 1969. Some philosophical problems from the standpoint of AI. In B. Meltzer & D. Michie (Eds.) Machine Intelligence 4, Edinburgh, UK: Edinburgh University Press, 463-502.

McDermott, D. 1982. A Temporal logic for reasoning about processes and plans. Cognitive Science 6, 101-155.

Meiri, I. 1996. Combining qualitative and quantitative constraints in temporal reasoning. Artificial Intelligence 87 (1-2), 343-385.

Mohr, R. & Henderson, T. 1986. Arc and path consistency revisited. Artificial Intelligence 28 (2), 225-233.

Montanari, U. 1974. Networks of constraints: fundamental properties and applications to picture processing. Information Sciences 7 (2), 95-132.

Morris, R. & Khatib, L. 1999. Optimization in constraint reasoning about repeating events. In C. Dixon & M. Fisher (Eds.) Proceedings of the 6-th International Workshop on Temporal Representation and Reasoning, Los Alamitos, California: IEEE Computer Society Press, 82-87.

Motro, A. 1993. Sources of uncertainty, imprecision, and inconsistency in information systems. In A. Motro & P. Smets (Eds.) Uncertainty management and information systems: from needs to solutions, Boston: Kluwer Academic Publishers, 9-34.

Neapolitan, R. 1990. Probabilistic reasoning in expert systems, New York, Chichester, Brisbane, Toronto, Singapore: John Willey and Sons, Inc.

Nebel, B. 1996. Solving hard qualitative temporal reasoning problems: evaluating the efficiency of using the ORD-Horn class. In W. Wahlster (Ed.) Proceedings of the 12-th European Conference on Artificial Intelligence, Chichester: John Wiley and Sons, 38-42.

Nebel, B. & Bürckert, H.-J. 1995. Reasoning about temporal relations: a maximal tractable subclass of Allen's interval algebra. Journal of the ACM 42 (1), 43-66.

Nejdl, W. & Gamper, J. 1994. Model-based diagnosis with qualitative temporal uncertainty, In R. Lopez de Mantaras & D. Poole (Eds.) Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence, San Francisco, California: Morgan Kaufmann Publishers, Inc., 432-439.

Newton-Smith, W. 1980. The structure of time, London: Routledge and Heagan Paul.

Nguyen, J., Shahar, Y., Tu, S., Das, A., & Musen, M. 1999. Integration of temporal reasoning and temporal-data maintenance into a reusable database mediator to answer abstract, time-oriented queries: the Tzolkin system, Journal of Intelligent Information Systems 13 (1-2), 121-145.

Parsons, S. 1996. Current approaches to handling imperfect information in data and knowledge bases. IEEE Transactions on Knowledge and Data Engineering 8 (3), 353-372.

Parsons, S. & Hunter, A. 1998. A review of uncertainty handling formalisms. In A. Hunter & S. Parsons (Eds.) Applications of uncertainty formalisms.

Lecture Notes in Artificial Intelligence 1455, Berlin, Heidelberg: Springer-Verlag, 8-37.

Pearl, J. 1987. Bayesian decision methods. In S. Shapiro (Ed.) Encyclopedia of Artificial Intelligence. New York: Wiley Interscience, Inc., 48-56.

Pearl, J. 1988. Probabilistic reasoning in intelligent systems: networks of plausible inference, San Mateo, California: Morgan Kaufmann Publishers, Inc.

Pearl, J. 1992. Probabilistic Reasoning in Intelligent Systems, 2-nd edition, San Francisco, California: Morgan Kaufmann Publishers, Inc.

Petley, B. 1991. Time and frequency in fundamental metrology. Proceedings of the IEEE 79 (9), 1070-1077.

Portinale, L. 1992. Modeling uncertain temporal evolutions in model based diagnosis. In D. Dubois, M. Wellman, B. D'Ambrosio, & P. Smets (Eds.) Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence, San Francisco, California: Morgan Kaufmann Publishers, Inc., 244-251.

Quinlan, J. 1983. INFERNO: a cautious approach to uncertain inference. Computer Journal 26 (3), 255-269.

Roos, N. 1992. A logic for reasoning with inconsistent knowledge. Artificial Intelligence 57 (1), 69-103.

Ryabov, V. 2000. Uncertain relations between indeterminate temporal intervals, In R. Agrawal, K. Ramamritham & T. Vijayaraman (Eds.) Proceedings of the Tenth International Conference on Management of Data, New Delhi, India: Tata McGraw-Hill Publishing Company Limited, 87-95.

Ryabov, V. & Puuronen, S. 2000. Estimation of uncertain relations between indeterminate temporal points, In T. Yakhno (Ed.) Proceedings of the First Biannual International Conference on Advances in Information Systems, Lecture Notes in Computer Science 1909, Heidelberg, Germany: Springer-Verlag, 108-116.

Ryabov, V. & Puuronen, S. 2001. Probabilistic reasoning about uncertain relations between temporal points, In C. Bettini & A. Montanari (Eds.) Proceedings of the 8-th International Symposium on Temporal Representation and Reasoning (TIME’01), Los Alamitos, California: IEEE Computer Society Press, 35-40.

Ryabov V., Puuronen S. & Terziyan V. 1999. Representation and reasoning with uncertain temporal relations. In A. Kumar & I. Russel (Eds.) Proceedings of the Twelfth International Florida AI Research Society Conference, Menlo Park, California: AAAI Press, 449-453.

Saffiotti, A., Parsons, S. & Umkehrer, E. 1994. Comparing uncertainty management techniques. Microcomputers in Civil Engineering - Special Issue on Uncertainty in Expert Systems 9, 367-380.

Shafer, G. 1976. A mathematical theory of evidence, Princeton: Princeton University Press.

Shafer, G. 1988. Comments on 'an inquiry into computer understanding' by Peter Cheeseman. Computational Intelligence 4, 121-124.

Shafer, G. & Logan, R. 1987. Implementing Dempster's rule for hierarchical evidence. Artificial Intelligence 33 (3), 271-298.

Shahar, Y. 2000. Dimension of time in illness: an objective view, Annals of Internal Medicine 132 (1), 45-53.

Shenoy, P. & Shafer, G. 1990. Axioms for probability and belief function propagation. In R. Shachter, T. Levitt, L. Kanal & J. Lemmer (Eds.) Uncertainty in Artificial Intelligence. Amsterdam: North-Holland, 169-198.

Shortliffe, E. 1976. Computer-based medical consultations: MYCIN, New York: Elsevier.

Smets, P. 1991. The transferable belief model and other interpretations of Dempster-Shafer's model. In P. Bonissone, M. Henrion, L. Kanal & J. Lemmer (Eds.) Uncertainty in artificial intelligence, Amsterdam: North-Holland, 375-383.

Smithson, M. 1989. Ignorance and uncertainty: emerging paradigms, New York: Springer-Verlag.

Steiner, A. 1998. A generalisation approach to temporal data models and their implementations, Swiss Federal Institute of Technology, Zurich, Switzerland. Ph.D. thesis.

Struss, P. 1992. Knowledge-based diagnosis - an important challenge and touchstone for AI, In B. Neumann (Ed.) Proceedings of the 10th European Conference on AI, Chichester: John Wiley & Sons, 863-874.

Struss, P. 1997. Model-based diagnosis for industrial applications. Colloquium - Applications of model based reasoning, Institute of Electrical Engineers (IEE), Savoy Place, London. (2.9.2002)
http://wwwradig.informatik.tu-muenchen.de/research/MQM/
/Publications/1997/Struss97d_Abstract.html

Tawfik, A. & Neufeld, E. 1994. Temporal Bayesian networks. In S. Goodwin & H. Hamilton (Eds.) Proceedings of the International Workshop on Temporal Reasoning (TIME-94), Regina, Canada: University of Regina, 85-92.

Tawfik, A. & Neufeld, E. 1996. Irrelevance in uncertain temporal reasoning. In L. Chittaro, S. Goodwin, H. Hamilton & A. Montanari (Eds.) Proceedings of the 3rd International Workshop on Temporal Representation and Reasoning (TIME-96), Los Alamitos, California: IEEE Computer Society Press, 182-187.

Tawfik, A. & Neufeld, E. 1998. Model-based diagnosis: a probabilistic extension. In A. Hunter & S. Parsons (Eds.) Applications of uncertainty formalisms, Lecture Notes in Computer Science 1455, Berlin, Heidelberg: Springer-Verlag, 379-396.

Trudel, A. 1991. The interval representation problem, International Journal of Intelligent Systems 6 (5), 509-547.

Trudel, A. 1997. A temporal knowledge representation approach based on elementary calculus, Computational Intelligence 13 (4), 465-485.

Tsortas, V. & Kumar, A. 1996. Temporal databases bibliography update. SIGMOD Record 25 (1), 41-51.

van Beek, P. 1989. Approximation algorithms for temporal reasoning. In N. Sridharan (Ed.) Proceedings of the 11-th International Joint Conference on

Artificial Intelligence, Detroit, Michigan: Morgan Kaufmann Publishers, Inc., 1291-1296.

van Beek, P. 1990. Exact and approximate reasoning about qualitative temporal relations. University of Waterloo, Canada. Ph.D. thesis.

van Beek, P. 1991. Temporal query processing with indefinite information. Artificial Intelligence in Medicine 3 (6), 325-339.

van Beek, P. 1992. Reasoning about qualitative temporal information. Artificial Intelligence 58 (1-3), 297-326.

van Beek, P. & Cohen, R. 1990. Exact and approximate reasoning about temporal relations. Computational Intelligence 6 (3), 132-144.

van Benthem, J. 1983. The logic of time, Dordrecht: Kluwer Academic Publishers.

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

Vilain, M. 1982. A system for reasoning about time. In D. Waltz (Ed.) Proceedings of the Second National Conference on Artificial Intelligence, Menlo Park, California: AAAI Press, 197-201.

Vilain, N. & Kautz, H. 1986. Constraint propagation algorithms for temporal reasoning. In T. Kehler & S. Rosenschein (Eds.) Proceedings of the Fifth National Conference of the American Association for Artificial Intelligence, Menlo Park, California: AAAI Press, 377-382.

Voss, D. 1997. Timestamps to support valid-time indeterminacy in temporal databases, Vanderbilt University, Nashville, Tennessee, USA. Ph.D. thesis.

Wainer, J. & Rezende, A. 1997. A temporal extension to the parsimonious covering theory, Artificial Intelligence in Medicine 10 (3), 235-255.

Wainer J. & Sandri, S. 1999. Fuzzy temporal/categorical information in diagnosis, Journal of Intelligent Information Systems 13 (1-2), 9-26.

Walley, P. 1991. Statistical reasoning with imprecise probabilities, Monographs on statistics and applied probability 42, London, New York, Tokyo, Melbourne, Madras: Chapman and Hall.

Walley, P. 1996. Measures of uncertainty in expert systems. Artificial Intelligence 83 (1), 1-58.

Wang, X., Bettini, C., Brodsky, A. & Jajodia, S. 1997. Logical design for temporal databases with multiple granularities. ACM Transactions on Database Systems 22 (2), 115-170.

Wetprasit, R. & Sattar, A. 1998. Temporal reasoning with qualitative and quantitative information about points and durations. In C. Rich & J. Mostow (Eds.) Proceedings of the Fifteenth National Conference on Artificial Intelligence, Menlo Park, California: AAAI Press, 656-663.

Wetprasit, R., Sattar, A. & Khatib, L. 1997. A generalised framework for reasoning with multi-point events. In R. Shyamasundar & K. Ueda (Eds.) Proceedings of the Third Asian Computing Science Conference (ASIAN'97), Lecture Notes in Computer Science 1345, Berlin, Heidelberg, New York: Springer, 121-135.

Wilson, N. 1992. Some theoretical aspects of the Dempster-Shafer theory. Oxford Polytechnic. Ph.D. thesis.

Wilson, D., Greig, A., Gilby, J. & Smith, R. 1998. Some problems in trying to implement uncertainty techniques in automated inspection, In A. Hunter & S. Parsons (Eds.) Applications of uncertainty formalisms, Lecture Notes in Artificial Intelligence 1455, Berlin, Heidelberg: Springer-Verlag, 225-241.

Young, J. 1996. On unifying time and uncertainty: the probabilistic temporal network. Air Force Institute of Technology, Air University, USA. M.Sc. thesis.

Young, J. & Santos, E. 1996. Introduction to temporal Bayesian networks. In M. Gasser (Ed.) Online Proceedings of the 7-th Midwest AI and Cognitive Science Conference. (15.8.2002). http://www.cs.indiana.edu/event/maics96/Proceedings/Port/port.html

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

Zadeh, L. 1978. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1, 3-28.

## YHTEENVETO (FINNISH SUMMARY)


Aikatiedon formaalia esittämistä ja siihen perustuvaa päättelyä tarvitaan useilla kohdealueilla todellisuuden dynaamisten piirteiden mallintamisessa. Tällaisia tapahtumien ajallisen järjestyksen hallintaa vaativia sovellusalueita ovat esimerkiksi prosessien säätö, toimenpiteiden suunnittelu, luonnollisen kielen käsittely ja diagnostiikka. Toisaalta lähes kaikki ympäröivään todellisuuteen liittyvät tiedot ovat puutteellisia – vain harvoin tieto on täysin varmaa, täydellistä, täsmällistä tai ristiriidatonta. Niinpä myös aikaa koskevien tietojen käsittelyssä, niin esittämisessä kuin päättelyssäkin, joudutaan toimimaan puutteellisen tiedon varassa.

Väitöskirja käsittelee epätäydellisen ajallisen tiedon esittämistä ja käsittelyä tavoitteenaan epätäydellisille aikarelaatioille sopivan formalismin kehittäminen. Työssä keskitytään kahteen seikkaan: 1) epätäydellisten aikarelaatioiden esittämiseen ja estimointiin sekä 2) niiden pohjalta tapahtuvaan päättelyyn. Kehitetyssä esitystavassa epätäydellinen aikarelaatio kuvataan numeerisin arvoin. Koska nämä numeeriset arvot eivät useinkaan ole välittömästi saatavissa, käytetään estimointia tarvittavien arvojen määrittämiseksi. Työssä kuvattuun päättelymekanismiin kuuluvilla operaatioilla johdetaan tunnettujen epätäydellisten aikarelaatioiden pohjalta uusia arvoltaan aikaisemmin tuntemattomia epätäydellisiä aikarelaatioita mielenkiintoisten aikaprimitiivien välille.

Työssä esitetään todennäköisyysteoriaan perustuva numeerinen formalismi epätäydellisille aikarelaatioille. Siinä tarkastellaan diskreetin aikakäsityksen aikaprimitiiveinä sekä yksittäisiä ajankohtia että aikavälejä. Kahden yksittäisen ajankohdan välinen epätäydellinen aikarelaatio esitetään antamalla todennäköisyysarvo kaikille kolmelle näiden väliselle mahdolliselle ajalliselle perusrelaatiolle (ennen, samaan aikaan ja jälkeen). Kahden aikavälin välinen epätäydellinen aikarelaatio esitetään vastaavasti antamalla todennäköisyysarvo kullekin kolmelletoista mahdolliselle, niiden väliselle ajalliselle perusrelaatiolle. Työssä on kehitetty laskukaavat näiden todennäköisyysarvojen laskemiseen. Tarkasteluun otetaan myös sellainen tilanne, jossa laskennassa mukana olevien aikaprimitiivien (ajankohta tai aikaväli) arvo voi sisältää mittausvirhettä. Tilanteeseen, jossa mittausvirheen maksimiarvo tunnetaan, esitetään kaavat todennäköisyysarvojen ala- ja ylärajojen laskemiseksi kullekin aikaprimitiivien väliselle perusrelaatiolle.

Työssä esitetään lähestymistavalle sopivana kaksi diagnostiikan automatisoinnin sovellusaluetta: terveydenhoito ja teollisuus. Esitetyn lähestymistavan soveltaminen taudin diagnosointiin perustuu aikaulottuvuuteen liitetyn skenaarion käyttöön. Soveltamisessa lähdetään siitä olettamuksesta, että on olemassa tietty joukko kunkin diagnosoitavissa olevan taudin kannalta oleellisia oireita. Itse skenaario esitetään relaatioverkkona, jonka solmut vastaavat oireita ja jonka kaaret vastaavat oireiden välisiä epätäydellisiä aikarelaatioita esitettynä työssä kehitetyllä esitystavalla. Kokoamalla riittävän laajan aineiston eri henkilöiden osalta taudin etenemisskenaarioista voidaan muodostaa yhdistetty skenaario, jota voidaan käyttää taudin diagnosoinnin apuna. Vastaavasti teolli-

suuden vikadiagnostiikkaan esitetään käytettäväksi menettelytapaa, joka perustuu laitteen rakenteen ja toiminnan eksplisiittisesti esitettyyn aikaskenaariomalliin, josta mallin ja havaitun toiminnan avulla päätellään diagnoosi (kuten terveydenhoidonkin sovellusalueella).

Tutkimuksen pääkontribuution muodostaa epätäydellisten aikarelaatioiden esittäminen ja estimointi sekä siihen pohjaava päättelymekanismi. Tuloksen käyttökelpoisuuden rajoitteena voidaan nähdä sen teoreettisuus ja keskittyminen formaalin lähestymistavan kehittämiseen pelkästään diskreetille aikamallille. Työssä on edelleen esitetty ratkaisu ainoastaan aikarelaatioiden epävarmuuden ja epäjohdonmukaisuuden käsittelemiseen jättäen muut epätäydellisyyden ilmenemismuodot tutkimuksen ulkopuolelle. Mielenkiintoinen jatkotutkimuskohde olisikin lähestymistavan laajentaminen käsittelemään myös muita epätäydellisyyden ilmenemismuotoja, kuten epätäsmällisyyttä, vaillinaisuutta ja epätietoisuutta. Suunnitteilla oleva lähestymistavan soveltaminen teollisuuden diagnosointiongelmiin tarjonnee työssä esitettyä vankemman pohjan lähestymistavan käyttökelpoisuuden arvioimiselle.

# J Y V Ä S K Y L Ä   S T U D I E S   I N   C O M P U T I N G

1 ROPPONEN, JANNE, Software risk management - foundations, principles and empirical findings. 273 p. Yhteenveto 1 p. 1999.

2 KUZMIN, DMITRI, Numerical simulation of reactive bubbly flows. 110 p. Yhteenveto 1 p. 1999.

3 KARSTEN, HELENA, Weaving tapestry: collaborative information technology and organisational change. 266 p. Yhteenveto 3 p. 2000.

4 KOSKINEN, JUSSI, Automated transient hypertext support for software maintenance. 98 p. (250 p.) Yhteenveto 1 p. 2000.

5 RISTANIEMI, TAPANI, Synchronization and blind signal processing in CDMA systems. - Synkronointi ja sokea signaalinkäsittely CDMA järjestelmässä. 112 p. Yhteenveto 1 p. 2000.

6 LAITINEN, MIKA, Mathematical modelling of conductive-radiative heat transfer. 20 p. (108 p.) Yhteenveto 1 p. 2000.

7 KOSKINEN, MINNA, Process metamodelling. Conceptual foundations and application. 213 p. Yhteenveto 1 p. 2000.

8 SMOLIANSKI, ANTON, Numerical modeling of two-fluid interfacial flows. 109 p. Yhteenveto 1 p. 2001.

9 NAHAR, NAZMUN, Information technology supported technology transfer process. A multi-site case study of high-tech enterprises. 377 p. Yhteenveto 3 p. 2001.

10 FOMIN, VLADISLAV V., The process of standard making. The case of cellular mobile telephony. - Standardin kehittämisen prosessi. Tapaustutkimus solukkoverkkoon perustuvasta matkapuhelintekniikasta. 107 p. (208 p.) Yhteenveto 1 p. 2001.

11 PÄIVÄRINTA, TERO, A genre-based approach to developing electronic document management in the organization. 190 p. Yhteenveto 1 p. 2001.

12 HÄKKINEN, ERKKI, Design, implementation and evaluation of neural data analysis environment. 229 p. Yhteenveto 1 p. 2001.

13 HIRVONEN, KULLERVO, Towards Better Employment Using Adaptive Control of Labour Costs of an Enterprise. 118 p. Yhteenveto 4 p. 2001.

14 MAJAVA, KIRSI, Optimization-based techniques for image restoration. 27 p. (142 p.) Yhteenveto 1 p. 2001.

15 SAARINEN, KARI, Near infra-red measurement based control system for thermo-mechanical refiners. 84 p. (186 p.) Yhteenveto 1 p. 2001.

16 FORSELL, MARKO, Improving Component Reuse in Software Development. 169 p. Yhteenveto 1 p. 2002.

17 VIRTANEN, PAULI, Neuro-fuzzy expert systems in financial and control engineering. 245 p. Yhteenveto 1 p. 2002.

18 KOVALAINEN, MIKKO, Computer mediated organizational memory for process control. Moving CSCW research from an idea to a product. 57 p. (146 p.) Yhteenveto 4 p. 2002.

19 HÄMÄLÄINEN, TIMO, Broadband network quality of service and pricing. 140 p. Yhteenveto 1 p. 2002.

20 MARTIKAINEN, JANNE, Efficient solvers for discretized elliptic vector-valued problems. 25 p. (109 p.) Yhteenveto 1 p. 2002.

21 MURSU, ANJA, Information systems development in developing countries. Risk management and sustainability analysis in Nigerian software companies. 296 p. Yhteenveto 3 p. 2002.

22 SELEZNYOV, ALEXANDR, An anomaly intrusion detection system based on intelligent user recognition. 186 p. Yhteenveto 3 p. 2002.

23 LENSU, ANSSI, Computationally intelligent methods for qualitative data analysis. 57 p. (180 p.) Yhteenveto 1 p. 2002.

24 RYABOV, VLADIMIR, Handling imperfect temporal relations. 75 p. (145 p.) Yhteenveto 2 p. 2002.