COMPUTER–SUPPORTED ANALYSIS OF SCIENTIFIC MEASUREMENTS

Illustration of a bend test used in Galileo's *Discorsi e dimostrazioni mathematiche intorno a due nuove scienze attenenti alla meccanica et i movimenti locali* ("Two new sciences"). Reproduced from *Le Opere di Galileo Galilei*, Barbera, Firenze, 1968.

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

De Jong, Hidde

Computer-Supported Analysis of Scientific Measurements
H. de Jong – [S.l. : s.n.]. – Ill.
Thesis Enschede – With ref. – With summary.
ISBN: 90-365-1143-7
ISSN: 1381-3617-98-23 (CTIT-Ph.D.-series No. 98-23)
Subject headings: knowledge-based systems / measurements / qualitative reasoning / experiments

CTIT Ph.D.-thesis series 98-23

P.O. Box 217, 7500 AE Enschede, the Netherlands
telephone +31-53-4893779/4892100, fax +31-53-4894524

Centre for
Telematics and
Information
Technology

COMPUTER–SUPPORTED ANALYSIS OF SCIENTIFIC MEASUREMENTS

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof.dr. F.A. van Vught,
volgens besluit van het College van Promoties
in het openbaar te verdedigen
op vrijdag 5 juni 1998 te 15.00 uur.

door
Hidde de Jong
geboren op 7 juli 1968
te Delft

Dit proefschrift is goedgekeurd door:

prof.dr.ir. N.J.I. Mars (promotor)
dr. P.E. van der Vet (assistent-promotor)

Duke. There is no composition in these news
That gives them credit.
First senator. Indeed, they are disproportion'd;
My letters say a hundred and seven galleys.
Duke. And mine, a hundred and forty.
Second senator. And mine, two hundred . . .

William Shakespeare, *Othello, the Moor of Venice* (I, 3).

Abstract

In the past decade, large-scale databases and knowledge bases have become available to researchers working in a range of scientific disciplines. In many cases these databases and knowledge bases contain measurements of properties of physical objects which have been obtained in experiments or at observation sites. As examples, one can think of crystallographic databases with molecular structures and property databases in materials science.

These large collections of measurements, which will be called measurement bases, form interesting resources for scientific research. By analyzing the contents of a measurement base, one may be able to find patterns that are of practical and theoretical importance. With the use of measurement bases as a resource for scientific inquiry questions arise about the quality of the data being analyzed. In particular, the occurrence of conflicts and systematic errors raises doubts about the reliability of a measurement base and compromises any patterns found in it. On the other hand, conflicts and systematic errors may be interesting patterns in themselves and warrant further investigation.

These considerations motivate the topic that will be addressed in this thesis: the development of systematic methods for detecting and resolving conflicts and identifying systematic errors in measurement bases. These measurement analysis (MA) methods are implemented in a computer system supporting the user of the measurement base.

Despite their obvious importance, MA methods for conflict resolution and error identification have been largely unexplored thus far. Statistical methods assist in detecting conflicts between measurements, but do not offer much help in resolving them. In addition, they focus on random errors and largely neglect the problem of systematic errors. In contrast with statistical methods, the methods developed in this thesis draw upon knowledge about the domain under study. More specifically, they are model-based MA methods since this knowledge takes the form of models of the physical systems investigated in the experiments.

Chapter 2 provides a framework for conflict detection, conflict resolution, and error identification by relating conflicts and errors to the experiments in which measurements are performed. An experiment is conceptualized as the activity of creating and sustaining a controlled physical system, referred to as an experimental system. Measurement of a certain property amounts to an empirical determination of the value of a quantity of the experimental system. A conflict between two property measurements can be explained by reference to structural differences between the experimental systems on which the measurements are performed and differences between the experimental conditions. A systematic error in a property measurement can be predicted from differences between the structure of the experimental system actually investigated and the structure of a hypothetical ideal experimental system, and from differences between the actual experimental conditions and the ideal experimental conditions.

Experimental systems are modeled by means of differential equations. More particularly, qualitative differential equations (QDEs) are used, since much of the knowledge about the systems will be qualitative in nature, especially when certain idealized experimental circumstances cannot be realized. Given this representation, the methods for

model-based conflict resolution and error identification can be elaborated by means of two techniques from qualitative reasoning (QR): qualitative simulation and comparative analysis.

Chapter 3 reviews the well-known qualitative simulation algorithm QSIM which is used to infer the possible qualitative behaviors of an experimental system from an initial state representing the experimental conditions. QSIM has a solid foundation in mathematics which allows one to prove certain properties of the algorithm. In particular, all genuine possible behaviors of an experimental system are inferred, but occasionally spurious behaviors as well.

Chapter 4 introduces the CEC* algorithm for comparative envisionment construction which allows one to compare a model and behavior of one experimental system with a model and behavior of another system. CEC* improves upon existing comparative analysis algorithms in that it is able to compare structurally different experimental systems. Starting from initial relative values for some of the quantities, CEC* derives possible comparative behaviors of the two systems. A comparative behavior describes the differential dynamics of the experimental systems being compared in a qualitative manner. An explanatory comparative analysis finds possible causes of differences in the response of two systems, whereas a predictive comparative analysis finds possible consequences of differences in the initial conditions. As in QSIM, certain guarantees can be given on the outcome of a comparative analysis. All genuine comparative behaviors of the systems will be inferred, but sometimes spurious comparative behaviors as well.

In chapter 5 the techniques for qualitative simulation and comparative analysis are combined to formalize the tasks of conflict resolution and comparative analysis. In addition, a standard statistical method for conflict detection is given. Conflict resolution is defined as an explanatory comparative analysis, where the resulting comparative behaviors represent possible explanations of the conflict between two measurements. Error identification is defined as a predictive comparative analysis with the comparative behaviors pointing at possible systematic errors in the measurement. From the formal properties of QSIM and CEC* it can be proven that all genuine explanations of a conflict and all genuine predictions of a systematic error are found, but that the occurrence of spurious explanations and predictions cannot be excluded.

The conflict detection, conflict resolution, and error identification algorithms have been implemented in Common Lisp and together form the KIMA system for Knowledge-Intensive Measurement Analysis. KIMA is built on top of the implementations of QSIM and CEC* and repeatedly calls the main functions of these programs.

KIMA has been successfully applied in a case-study on a realistic though simplified problem: the analysis of measurements of the fracture strength of brittle materials obtained in tension tests and four-point bend tests. Chapter 6 reviews basic theories on brittle fracture and fracture testing which underlie the models required by the conflict resolution and error identification algorithms. The results of the case-study are presented in chapter 7. KIMA is shown to be able to reproduce a number of interesting phenomena reported in the literature on tension tests and four-point bend tests.

In chapter 8 the MA methods and their application are discussed in the context of related work. In particular, attention is given to different forms of knowledge-based measurement analysis, the use of qualitative knowledge, the relationship between model-based measurement analysis and model-based diagnosis, the (computer-supported) construction and revision of models of experimental systems, the generality of the methods for conflict resolution and error identification, and the practical use of the methods.

The concluding chapter 9 summarizes the two main contributions of the thesis: first, the development of methods for model-based conflict resolution and error identification which supplement conventional statistical analyses; and second, the development of a general technique for comparative analysis which improves upon existing approaches and may prove useful for design, diagnosis, and discovery problems as well. A few directions for further research are indicated and the chapter concludes with a speculative outlook by viewing model-based measurement analysis as a part of future computer-supported discovery environments.

Samenvatting

In de afgelopen jaren hebben wetenschappers in diverse vakgebieden grootschalige databanken en kennisbanken tot hun beschikking gekregen. In veel gevallen bevatten deze databanken en kennisbanken metingen van eigenschappen van fysische objecten die zijn verkregen in experimenten of uit observatiestudies. Voorbeelden hiervan zijn kristallografische databanken met moleculaire strukturen en databanken met mechanische eigenschappen van materialen.

Deze grote verzamelingen metingen, die hier metingenbanken worden genoemd, vormen interessante hulpmiddelen voor wetenschappelijk onderzoek. Door de inhoud van een metingenbank te analyseren kunnen patronen van praktisch en theoretisch belang worden gevonden. Het gebruik van metingenbanken als een hulpmiddel voor wetenschappelijk onderzoek roept vragen op over de kwaliteit van de geanalyseerde data. In het bijzonder zijn het conflicten tussen metingen en systematische fouten die twijfel doen rijzen aan de betrouwbaarheid van de metingenbank en de patronen die erin zijn gevonden. Vanuit een ander invalshoek bekeken kunnen conflicten en systematische fouten echter *zelf* interessante patronen vormen die een nader onderzoek verdienen.

Deze overwegingen motiveren het onderwerp van dit proefschrift: de ontwikkeling van systematische methoden voor het opsporen en oplossen van conflicten en het identificeren van systematische fouten in metingenbanken. Deze methoden voor het analyseren van metingen, ofwel MA-methoden, zijn geïmplementeerd in een computersysteem dat de gebruiker van de metingenbank ondersteunt.

Ondanks hun vanzelfsprekende belang zijn MA-methoden voor het oplossen van conflicten en het identificeren van systematische fouten tot dusver nog nauwelijks onderzocht. Statistische methoden helpen bij het opsporen van conflicten tussen metingen, maar bieden weinig ondersteuning bij het oplossen van die conflicten. Bovendien zijn ze voornamelijk gericht op toevallige fouten en besteden ze nauwelijks aandacht aan systematische fouten. In tegenstelling tot statistische methoden maken de methoden die worden ontwikkeld in dit proefschrift gebruik van kennis over het bestudeerde domein. Meer in het bijzonder zijn de methoden modelgebaseerd, omdat de domeinkennis de vorm aanneemt van modellen van fysische systemen die in de experimenten worden onderzocht.

Hoofdstuk 2 verschaft een raamwerk voor conflictdetectie, conflictresolutie en foutdentificatie door conflicten en fouten te relateren aan de experimenten waarin de metingen zijn uitgevoerd. Een experiment wordt geconceptualiseerd als het creëren en instandhouden van een gecontroleerd fysisch systeem, ook wel een experimenteel systeem genoemd. Metingen van een eigenschap zijn empirische bepalingen van de waarde van een specifieke grootheid in het experimentele systeem. Een conflict tussen twee metingen kan worden verklaard door verschillen in de struktuur van de experimentele systemen waarop de metingen zijn uitgevoerd, of door verschillen in de experimentele omstandigheden. Een systematische fout in een meting van een eigenschap kan worden voorspeld uit verschillen in de struktuur van het daadwerkelijk onderzochte experimentele systeem en het ideale experimentele systeem, en tussen verschillen in de geobserveerde experimentele omstandigheden en de ideale experimentele omstandigheden.

Experimentele systemen worden gemodelleerd met behulp van differentiaalvergelijkingen. Meer in het bijzonder worden kwalitatieve differentiaalvergelijkingen gebruikt aangezien veel van de kennis over de systemen kwalitatief van aard is, vooral wanneer bepaalde geïdealiseerde experimentele omstandigheden niet kunnen worden gerealiseerd. Gegeven deze representatie kunnen de methoden voor het oplossen van conflicten en het identificeren van fouten worden uitgewerkt met behulp van twee technieken uit het vakgebied van kwalitatief redeneren: kwalitatieve simulatie en comparatieve analyse.

Hoofdstuk 3 geeft een overzicht van het bekende QSIM-algoritme voor kwalitatieve simulatie dat wordt gebruikt om mogelijke kwalitatieve gedragingen van een experimenteel systeem af te leiden uit een begintoestand die de experimentele omstandigheden representeert. QSIM heeft een goede wiskundige basis wat het mogelijk maakt om bepaalde eigenschappen van het algoritme te bewijzen. Met name worden alle echt mogelijke gedragingen van het systeem afgeleid, maar soms ook onechte gedragingen.

Hoofdstuk 4 introduceert het CEC*-algoritme waarmee een model en een gedraging van een experimenteel systeem kunnen worden vergeleken met een model en een gedraging van een ander experimenteel systeem. CEC* is een verbetering ten opzichte van bestaande algoritmen voor comparatieve analyse in de zin dat het in staat is structureel verschillende experimentele systemen te vergelijken. Het algoritme leidt mogelijke comparatieve gedragingen van twee systemen af uit een verzameling initiële relatieve waarden voor systeemgrootheden. Een comparatieve gedraging beschrijft het verschil in gedrag van de vergeleken experimentele systemen op een kwalitatieve manier. Een verklarende comparatieve analyse vindt mogelijke oorzaken van geobserveerde verschillen in de respons van twee systemen, terwijl een voorspellende comparatieve analyse mogelijke gevolgen vindt van geobserveerde verschillen in de beginvoorwaarden. Net als bij QSIM kunnen bepaalde garanties worden gegeven voor de uitkomst van een comparatieve analyse. Alle echte comparatieve gedragingen van de systemen zullen worden afgeleid, maar soms ook onechte comparatieve gedragingen.

In hoofdstuk 5 worden de technieken voor kwalitatieve simulatie en comparatieve analyse gecombineerd om de taken van conflictresolutie en foutidentificatie te formaliseren. Verder wordt een standaardmethode uit de statistiek gebruikt voor het detecteren van conflicten. Conflictresolutie is gedefinieerd als een verklarende comparatieve analyse waarin de resulterende comparatieve gedragingen mogelijke verklaringen van het conflict tussen de twee metingen representeren. Foutidentificatie is gedefinieerd als een voorspellende comparatieve analyse waarin de comparatieve gedragingen verwijzen naar mogelijke systematische fouten in de meting. Met behulp van de formele eigenschappen van QSIM en CEC* kan worden bewezen dat alle echte verklaringen van een conflict en alle echte voorspellingen van een systematische fout zullen worden gevonden, maar dat dat het genereren van onechte verklaringen en voorspellingen niet kan worden uitgesloten.

De algoritmen voor conflictdetectie, conflictresolutie en foutidentificatie zijn geïmplementeerd in Common Lisp en vormen samen het KIMA-systeem voor Knowledge-Intensive Measurement Analysis. KIMA is gebouwd als een laag bovenop de implementaties van QSIM en CEC* en roept herhaaldelijk de hoofdfuncties van deze programma's aan.

KIMA is succesvol toegepast in een case-study over een realistisch maar vereenvoudigd probleem: het analyseren van metingen van de breuksterkte van brosse materialen die zijn verkregen in spanningstesten en vier-punts buigtesten. Hoofdstuk 6 geeft een overzicht van theorieën over brosse breuk en breuktesten die ten gronslag liggen aan de modellen die noodzakelijk zijn voor conflictresolutie en foutidentificatie. De resultaten van de case-study worden gepresenteerd in hoofdstuk 7. KIMA blijkt in staat te zijn om een aantal interessante, in de literatuur over spanningstesten en vier-punts buigtesten gerapporteerde fenomenen te reproduceren.

In hoofdstuk 8 worden de MA-methoden en hun toepassing besproken in samenhang met verwant onderzoek. In het bijzonder wordt aandacht geschonken aan verschillende vormen van kennisgebaseerde analyse van metingen, het gebruik van kwalitatieve kennis, de relatie tussen modelgebaseerde analyse van metingen en modelgebaseerde diagnose, de (computerondersteunde) constructie en revisie van modellen van experimentele systemen, de generaliseerbaarheid van de methoden voor conflictresolutie en foutidentificatie, en het praktische gebruik van de methoden.

Het afsluitende hoofdstuk 9 vat de twee voornaamste bijdragen van het proefschrift samen: ten eerste, de ontwikkeling van methoden voor modelgebaseerde conflictresolutie en foutidentificatie waarmee conventionele statistische analyses kunnen worden uitgebreid; en, ten tweede, de ontwikkeling van een algemene techniek voor comparatieve analyse die een verbetering vormt ten opzichte van bestaande aanpakken en die ook van waarde kan zijn voor ontwerp-, diagnose-, en ontdekkingsproblemen. Enkele mogelijkheden voor verder onderzoek worden belicht, en het hoofdstuk sluit af met een speculatieve vooruitblik waarin modelgebaseerde analyse van metingen wordt gezien als een onderdeel van toekomstige computerondersteunde ontdekkingsomgevingen.

Contents

Preface

The work described in this thesis has been carried out while I was a Ph.D. student in the Knowledge-Based Systems group of the faculty of Computer Science at the University of Twente.

I am grateful to my promotor Koos Mars for allowing me the freedom to choose a difficult topic. He kept a friendly eye on the progress of the work, while emphasizing the main lines and insisting on a clear and distinct formulation of my ideas. Further, I would like to thank my daily supervisor Paul van der Vet for his careful reading of drafts of the notes, memoranda, papers, and chapters that I produced through the years. I am indebted to both of them for supporting me in the period when disease prevented me from working in a normal way.

In computer science the proof of the pudding is in the implementation. I have been very lucky with the aid of Frank van Raalte and Hans de Wit, who have not only shown remarkable skill in transforming my sometimes vague specifications into working computer programs, but who have also contributed to the ideas underlying the specifications.

Apart from computer science, the work reported in this thesis involves fields as diverse as philosophy and sociology of science, mathematics, and materials science. I have profited from the help of experts in these fields who, though not directly involved in my Ph.D. research, were so kind to show the way when I felt lost during my explorations. In particular, I would like to thank Bart de Jong and Hans Zwart of the faculty of Applied Mathematics and Louis Winnubst of the faculty of Chemical Technology. A special word of thank for Arie Rip of the faculty of Philosophy and Social Sciences. The ideas on computer-supported discovery environments were developed in a M.Sc. thesis in Philosophy of Science, Technology, and Society which I completed under his inspiring supervision.

Good colleagues are a blessing. I have been quite fortunate with my room mates Jeroen Nijhuis, Erik Oltmans, Piet-Hein Speel, Wilco ter Stal, and Ivayla Vatcheva, and the other members of the Knowledge-Based Systems group, including Els Bosch and Sandra Westhoff. I will especially remember the many soccer and tennis games, the crazy bets and meta-bets, the lively discussions during the coffeebreaks, and the Friday-afternoon drinks. I have much enjoyed my stay in their company.

In the period in which I suffered from disease, I found relief in the medical treatments and the encouraging words of, especially, Janny Bos-Veldman, H. Kampman, and D. Meier. Also, I would like to express my gratitude for the support obtained from E. van Leer and Yvonne Jansma of the Arbo- en Milieudienst of the University of Twente and Eddy ter Heijne and Jessica Tonneyck of the faculty of Computer Science.

My sister and my friends have invariably displayed a sincere interest in the progress of my work and many other things that are dear to me. My thanks go out to them. Most of all I am indebted to my parents Bart and Johanna, who have shown by their example the joy that can be found in literature and science. Moreover, in the past few years they have helped me in every possible way to continue my work and cheered me up when my spirits were low. It is to them that I dedicate this thesis.

Part I

Overview

Chapter 1

Introduction

1.1 Problem description

In the past decade, large-scale databases and knowledge bases have become available to researchers working in a range of scientific disciplines. In many cases these databases and knowledge bases contain *measurements* of properties of physical objects which have been obtained in experiments or at observation sites. As examples, one can think of the sequence databases and knowledge bases in molecular biology, databases with molecular structures in crystallography, and property databases in materials science.

These large collections of property measurements, henceforth called *measurement bases*, form interesting resources for scientific research. By analyzing the contents of a measurement base, one may be able to find interesting patterns that are of practical and theoretical importance. In the medical sciences an analysis of a large number of clinical studies can provide conclusive evidence for the therapeutical effect of a particular medicine, even if only a few of the individual studies reveal a statistically significant effect. In molecular biology a search of a database with human DNA sequences may lead to the identification of genes coding for proteins with a critical role in a cell's metabolism. In even moderately large measurement bases the search for interesting patterns will quickly become so computationally expensive as to defy human information-processing capacities. Therefore, such tasks have increasingly been delegated to computer programs (e.g., Fayyad, Haussler & Stolorz [1996]).

With the use of measurement bases as a resource for scientific inquiry questions arise about the quality of the data being analyzed. In particular, the occurrence of *conflicts* and *systematic errors* raises doubts about the reliability of a measurement base and compromises any patterns found in it. When a measurement contains a systematic error it gives an inaccurate value for the quantity of interest. The values specified by two conflicting measurements are in contradiction, so that at least one of them must be inaccurate. Although problematic for some purposes, conflicts and systematic errors may be interesting patterns in themselves when viewed from a different perspective. For example, conflicts between measuments of a property obtained in one experiment and measurements of the same property obtained in another may suggest the presence of a physical process in the first experiment that is absent in the second, and thereby guide further theoretical investigations.

The above considerations motivate the topic that will be addressed in this thesis: the development of systematic methods for *detecting and resolving conflicts* and *identifying systematic errors* in measurement bases.

Despite their obvious importance, these *measurement analysis (MA) methods* have been largely unexplored thus far. Statistical methods assist in detecting conflicts between measurements, but do not offer much help in resolving them. For the resolution of conflicts we need to take recourse to knowledge about the experiments in which the measurements were conducted and knowledge about the processes occurring in the physical systems investigated. Also, statistical treatments of measurement focus on random errors and usually brush aside the problem of systematic errors with general recommendations like: "The detection of ... systematic errors ... depends on the observer's alertness and knowledge of the natural, instrumental, and personal factors that can influence his procedures." (Barry [1978], p. 13).

It seems that in order to resolve conflicts and identify systematic errors in a property measurement base we need to go beyond statistical methods and develop methods which draw upon knowledge about the domain under study. That is, the methods for conflict resolution and error identification will have to be *knowledge-based* MA methods. Although in some situations it will suffice to quickly dispatch conflicts and systematic errors by means of heuristic rules acquired through experience, in others it is necessary to gain a deeper understanding of the cause of a conflict or systematic error by using appropriate models of the physical systems investigated in the experiments. The latter situations will interest us here. That is, we are concerned with a specific form of knowledge-based measurement analysis, namely the *model-based* resolution of measurement conflicts and the *model-based* identification of systematic errors.

The methods for model-based conflict resolution and error identification can be implemented in a computer system supporting the user of the measurement base. This has a number of advantages. First, the programs based on the methods are able to systematically search the potentially large space of explanations of a conflict or predictions of systematic errors. Due to their limited information-processing capacities, humans are easily led to overlook possible causes of a conflict or possible systematic errors when they are dealing with large search spaces. Second, a computerized MA system is better able to deal with the large and increasing scale of measurement bases. The sheer number of conflicts and systematic errors to be expected tightens the human information-processing bottleneck even further, up to the point that a thorough analysis of the measurements becomes too large a time-investment. A third advantage lies in the explicit form of the reasoning steps leading to the explanation of a conflict or the prediction of a systematic error. Explicit reasoning steps facilitate the critical evaluation of the results obtained with the MA system, and may engender new research questions.

The introduction of knowledge-based MA methods as a means to deal with issues not addressed by statistical methods should not be interpreted as an argument against the use of the latter. Nothing like this is intended. I will assume that an MA system for the resolution of conflicts between property measurements and the identification of systematic errors finds a place alongside conventional statistical tools in *computer-supported discovery environments* (de Jong & Rip [1997]). The tools will interact with each other, and with other tools in the discovery environment, in a variety of ways. For instance, a statistical tool may provide the conflicts to be analyzed by the MA system, whereas model-based error identification presents a way to determine the weights for measurements that are required by a statistical integration technique.

The above discussion can be summarized in the following problem description:

> **Problem description** The development of model-based methods for the resolution of conflicts between measurements contained in large scientific measurement bases and the identification of systematic errors in these measurements. The methods are to be implemented in an MA system.

1.2 Research requirements

As a guideline to the elaboration of the problem description into a workable research plan, and the evaluation of the results in the final part of my thesis, three requirements for the knowledge-based MA methods and the system in which they are implemented can be formulated.

In the first place, the methods for conflict resolution and error identification should be developed bearing in mind that they are intended to form the basis of a computer tool for scientists and engineers. This implies that a number of aspects of the MA methods should be given special attention, such as their formal specification, the correctness of the answers produced by them, their efficiency, and their capability of generating explanations.

Functionality requirement The MA methods should be developed with a view to their use in a computerized MA system.

The development of model-based MA methods is a difficult problem which has hitherto hardly received any attention. It would therefore not be realistic to expect a professional, ready-to-use MA system as the output of this thesis. But as a lower bound on acceptable results, one may demand that the proposed methods have been implemented in a prototype MA system and applied in a case-study in a realistic scientific domain.

Application requirement The MA methods should be implemented in a prototype MA system and be of demonstrated effectiveness in a case-study in a realistic scientific domain.

Applying the implemented methods in a case-study is fine, but it should not be achieved at the cost of introducing *ad hoc* solutions for fundamental problems. The methods should be sufficiently general to tackle conflict resolution and error identification problems in a range of situations and domains, and so should be the system that is based on these methods. The initial focus of this thesis, however, will be on natural science domains in which property measurements are obtained in experiments.

Generality requirement The MA methods should be general enough to be applied to different conflict resolution and error identification problems in various domains.

1.3 Research approach

Activities like the resolution of conflicts and the identification of errors have been characterized as knowledge-based because domain knowledge is required for their accomplishment, in particular knowledge about physical systems and the way these systems are controlled in experiments. What needs to be determined first is how exactly the physical systems investigated in a scientific experiment are related to conflicts and errors, in other words we need to start with an account of the practices of producing and evaluating measurements in science. It is clear that such an account can profit much from the reconstructions of experimental science given by historically-oriented philosophers and sociologists of science.

Having thus provided a foundation for understanding MA activities, it is possible to specify how these activities could be performed in a systematic way, so as to open up the road to computer support. Methods for conflict resolution and error identification will be outlined in the form of a sequence of steps in which models of the physical systems investigated in experiments are simulated and compared. Statistical methods for conflict detection will be added as a preliminary to the method for conflict resolution. These three methods are then implemented and integrated in a system for knowledge-based measurement analysis.

The elaboration of the methods presupposes the availability of techniques for representing knowledge about the experiments in which measurements are performed and techniques for reasoning with this knowledge. In particular, one will need techniques to formally specify models of the physical systems whose quantities are measured and techniques to simulate these models and find possible behaviors of the systems. Further, one will need techniques to compare the models and behaviors of the physical systems that are created in different experiments, in order to explain conflicts between measured values of a quantity or to estimate the deviation of a measured value from an expected value.

Knowledge about quantities and relations between quantities in a physical system will often be qualitative in nature, which poses an additional requirement for the techniques. Especially when certain idealized circumstances cannot be materialized in an experiment, only qualitative or at best semi-quantitative instead of quantitative information can be given. Consider for instance the influence of friction forces and twisting on the strength of a material sample determined in a bend test. These influences are quite hard to quantify, although the direction of the influence and probably an order-of-magnitude estimate can be established. Besides this fundamental reason, there is also a pragmatic reason for using qualitative knowledge about the experiments. In many cases a qualitative understanding of a conflict or an error will be enough to allow one to repeat the experiment and obtain a new and better measurement.

Techniques developed in artificial intelligence (AI) for modeling and qualitative reasoning about physical systems (Weld & de Kleer [1990]; Kuipers [1994]) seem excellently suitable for dealing with the qualitative nature of much knowledge about the physical systems investigated in scientific experiments. In addition, these techniques lend themselves for giving explanations to the user, because they build upon a qualitative understanding of physical systems conveyed by textbooks and handbooks. A further advantage of the qualitative reasoning (QR) framework is its strong roots in mathematics and the engineering sciences, which contributes to the precise determination of the correctness and efficiency of the techniques.

I will exploit the capabilities of QR techniques for realizing model-based support in measurement analysis, but it should be added immediately that this approach does not rule out the use of (semi-)quantitative knowledge when it is available. Recently, much attention has been given to the integration of techniques for qualitative reasoning with techniques for (semi-)quantitative reasoning. Although the focus will be on providing computer support based on qualitative techniques, I will point to possible (semi-)quantitative extensions of the MA methods. Such extensions would allow a qualitative analysis for identifying a systematic error to be followed up by a quantitative analysis for estimating the magnitude of this error.

QR resources particularly relevant for my purpose are formal languages for mathematically describing physical systems, algorithms for performing a qualitative simulation of physical systems, and algorithms for doing a comparative analysis of two physical systems. All of these techniques have received considerable attention in the QR literature, although the capabilities of some techniques fall short of what is required for the MA methods. Whereas existing modeling languages and simulation techniques can be adopted with minor adaptations, comparative analysis techniques reported in the literature are deficient in that they cannot deal with structurally different physical systems. This is a major shortcoming since conflicts and errors are often explained by differences in the structure of the physical systems on which measurements are performed. Therefore, the development of a more general comparative analysis technique, a difficult and time-consuming problem, will take up a sizable portion of this thesis.

After having formalized the methods for conflict resolution and error identification in terms of the above techniques, a prototype of an MA system will be developed and used in a case-study. The system is called *KIMA*, after *Knowledge-Intensive Measurement Analysis*. Although all of the algorithms introduced in this thesis have been implemented, and all of the examples in the case-study have been worked out using the implementation, I will focus on a conceptual level of description. For details on the implementation of the KIMA system, the reader is referred to the appendices.

The case-study is concerned with measurement analysis problems arising in the context of the Plinius project. Plinius is a knowledge extraction project aiming at the construction of a large knowledge base on mechanical properties of ceramic materials.

The case-study will show how model-based computer support of measurement analysis could answer such questions as: Why are measurements of the fracture strength obtained with polished specimens higher than measurements obtained woth specimens not so treated (conflict resolution)? Will eccentric loading tend to increase or decrease the measured fracture strength in a tension test (error analysis)?

Much care will be taken to ensure that particularities of the case-study are not compiled into the formulation of the MA methods, so that their capabilities and limitations can be assessed independently from their application in the case-study. The formal character of the techniques by which the methods have been defined helps in precisely specifying their performance and their range of applicability.

As mentioned in passing in the introductory section, measurement analysis is not isolated from other concerns in science and engineering. Some attention will therefore be given to relationships with activities like model construction and model revision. I will conclude the thesis with some reasoned speculations on future computer-supported discovery practices in which these activities are connected through networks of computer tools.

1.4 Contributions

The research described in this thesis aims at making two main contributions, each contribution being oriented at a different body of literature.

In the first place, the MA methods and the implemented KIMA system are interesting for fields concerned with the analysis of measurements, such as meta-analysis in statistics and computer-supported discovery in AI. It shows how QR techniques can be used to supplement conventional statistical analyses of measurement bases by allowing MA methods to exploit domain knowledge about the controlled physical systems investigated in scientific experiments. The capabilities of the methods for model-based measurement analysis have been illustrated in a case-study in a realistic domain.

The second contribution is not concerned with the MA methods proper, but rather with the techniques underlying the methods. The QR technique for comparing structurally different dynamical systems has a wider import than its application in conflict resolution and error identification. It presents a theoretical contribution to the field of qualitative reasoning and is an improvement upon existing methods for the comparative analysis of dynamical systems. The technique has potential applications to diagnosis, design, discovery, and modeling problems.

1.5 Thesis organization

The material of this thesis has been organized into four parts. The introductory part, comprising this chapter and the next, gives an outline of the study. Chapter 2 will describe experimental practices and introduce methods for conflict detection, conflict resolution, and error identification. The methods employ models of the physical systems created and sustained in these practices.

Part II focuses on QR techniques for modeling the physical systems investigated in scientific experiments and for reasoning about their behavior. Further, it shows how the MA methods outlined in the previous part can be formalized in terms of these techniques. This part contains a chapter on the simulation of physical systems (chapter 3), a chapter on the comparative analysis of physical systems (chapter 4), and a chapter on the formalization of conflict detection, conflict resolution, and error identification (chapter 5).

The contents of part II are rather technical and may not always be easy to grasp for readers not familiar with the basic concepts of qualitative reasoning. To a large extent, the results presented in part III can be read without a thorough understanding of the simulation and comparative analysis techniques. Chapter 6 gives an introduction to the Plinius domain, mechanical properties of ceramic materials, while chapter 7 discusses the application of the KIMA system to examples from this domain, in particular the analysis of measurements of the fracture strength of alumina.

Part IV consists of two chapters. Chapter 8 discusses the MA methods, and the results obtained with them in the case-study, and places the methods in the context of related work. Chapter 9, the final chapter, returns to the problem description and the research requirements to see what has been achieved. In addition, it offers prospects for further work and a speculative outlook.

Chapter 2

Knowledge-Based Analysis of Measurements

The two measurement analysis activities central to this thesis, conflict resolution and error identification, are knowledge-based in the sense that they require domain knowledge for their accomplishment. In this chapter, I will describe in some detail which knowledge should be available and relate it to conflict resolution and error identification. Chapters 3 to 5, the bulk of this thesis, can be regarded as an elaboration of the framework presented here.

In section 2.1, I will develop a perspective on the production and analysis of property measurements. Measurements are interpreted as empirical determinations of the value of quantities of physical systems being created and controlled in experiments. Conflicts and systematic errors can be related to specific details of the creation and control of these physical systems. With this perspective in mind, systematic methods for supporting the MA tasks of conflict detection, conflict resolution, and error identification will be proposed in section 2.2. The methods employ models of the physical systems investigated in the experiments.

2.1 Measurements and measurement analysis

Even though measurements are basic to science and engineering, it is surprisingly difficult to clarify the notion of measurement and the issues surrounding its reliability. Statistical measurement theory (e.g., Barry [1978]) provides some help, but abstracts from the actual production of measurements. For a thorough understanding of measurements, conflicts, and errors, the account of statistical measurement theory should be supplemented by accounts of the practices in which measurements are performed. Such studies have been undertaken by practicing experimental scientists (e.g., Squires [1985]; Polak [1979]) and philosophers, sociologists, and historians of science (e.g., Hacking [1983]; Franklin [1986]; Gooding, Pinch & Schaffer [1989]; Galison [1987]; Collins [1992]; (ed.) [1981]). It is by connecting this practice view on measurements with the statistical view that the basic concepts of this thesis will be elaborated.

2.1.1 Measurements and errors

A *measurement* is here understood as the empirical determination of the value of a quantity. A measurement is performed by comparing the quantity with a constant quantity of some sort which has been adopted as a standard. Thus, by declaring the length of a piece of metal to be 3.57 meter we mean that its length equals 3.57 times the unit of length, the meter. I will also use the term measurement in a derivative sense to denote the product of performing a measurement: the value of the quantity.

A distinction can be made between *direct* and *indirect* measurements, that is, direct or indirect comparisons with a standard. An example of the former is the measurement of a force applied to a material by means of a load cell. An example of the latter is the measurement of the elasticity of a material by determining the strain of the material at various stress levels and calculating the slope of a curve fitted through the stress-strain coordinates. Notice that a direct measurement is often realized by standardizing an indirect measurement and embedding it into an instrument, like the measurement of a force by the measurement of the elongation of a stiff, calibrated spring (see the contributions in Wise [1995] for other examples).

The reported measurement is often not a single measurement, but rather a summary of a set of individual measurements obtained from repeating the experiment a number of times. Such a measurement will be called an *aggregated* measurement.

Each individual measurement underlying the aggregated measurement can be viewed as drawn from a hypothetical population of determinations of the quantity in the experiment. The summary commonly has the form of the mean value of the set of individual measurements supplemented by the standard error of the mean. It represents an estimate of the mean of the population of quantity values. Given the assumption that the population is normally distributed, one can calculate a confidence interval and confidence level allowing a more precise statement to be made on the value of the population mean. For instance, from a temperature measurement with a mean value of 2400 ^{0}C and a standard error of 10 ^{0}C it is possible to conclude that there is a 95% certitude that the mean of the hypothetical population lies in the interval $[2380\ ^0\mathrm{C}, 2420\ ^0\mathrm{C}]$.

In order to evaluate a measurement, one can consider its accuracy and its precision. The *precision* of a measurement refers to the care and refinement by which it was obtained. At a given sample size, a larger spread in the individual measurements implies a higher standard error of the mean and thus a lower precision of the aggregated measurement. For instance, a reported value for the melting temperature of a material having a standard error of 10 ^{0}C is more precise than a value having a standard error of 30 ^{0}C. The first measurement is more precise in that it yields a tighter confidence interval.

The *accuracy* of a measurement refers to its correctness. The aggregated measurement makes an assertion on the mean of the hypothetical population of all possible determinations of the quantity in the experiment, but is this really the value that we want to measure? If a mistake has been made in calibrating the thermometer, then the temperature values are drawn from a population with a mean differing from the intended value, often called the *true value* of the quantity. The problem with the true value is that it is unknown; if it was not, there would be no reason to perform the measurement in the first place! However, by analyzing the circumstances of the measurement, it may be possible to establish whether the measured value deviates from the true value. This idea will be further developed in section 2.1.5.

The precision and accuracy of a measurement are related with different kinds of error. Following Barry [1978], an *error* is defined as the deviation of a measurement from the true value. A *systematic error* is an error that invariably has the same magnitude and the same sign in the experiment.[1] It affects the accuracy of a measured value, such as the error introduced by the faulty thermometer calibration above. Besides errors in instruments, systematic errors encompass errors arising from neglecting certain natural phenomena and personal errors in the performance of an experiment. The occurrence of *random errors* affects the precision of a measurement. The sign and magnitude of random errors are not known to be related to the experimental circumstances and have a tendency to be mutually compensating. Usually, systematic errors and random errors occur in combination.

This account of the precision and accuracy of a measurement, and the random and systematic errors related to them, presupposes the reported measurement to be an aggregated measurement. However, when it is a *single* measurement obtained by following a standard procedure, possibly embodied in a measurement instrument, a similar evaluation of its quality can be given. The measurement then inherits the precision and accuracy of the standard procedure. This precision and accuracy have been established by making a large number of measurements of the same kind and by carefully analyzing the experimental circumstances prescribed by the procedure (see Barry [1978], ch. 9).

2.1.2 Experiments and experimental systems

Assuming that a measurement is carried out in an experiment, the accuracy and precision of the measurement are determined by the way in which the experiment is performed. An *experiment* will here be viewed as the activity of creating and sustaining a controlled physical system, also referred to as an *experimental system*. Measurements are determinations of the values of quantities of an experimental system. An example of an experimental system is a specimen of a material which is controlled in a hardness test, torsion test, or tension test. Quantities of this system include the temperature of the specimen and the applied stress.

In order to measure a certain quantity, experimenters will attempt to *control* the physical system of interest. That is, they will actively create and maintain the structure of this system and regulate its behavior by imposing certain *experimental conditions*. For instance, in a tension test they fixate a sample with a certain percentage porosity in the test machine and slowly extend it under a given elongation rate and temperature. Failures to control the system may cause errors in the measured value of a quantity, both random errors and systematic errors (section 2.1.5).

Control over a physical system does not usually mean that the system is isolated from

[1] This definition is actually too restrictive, since systematic errors may not be constant; it suffices for the discussion here, however.

its environment, that is, turned into a *closed* system (see Radder [1984] for a discussion of closed experimental systems). Rather, it refers to the regulation of interactions between the experimental system and its environment. In a melting temperature experiment, for example, we try to avoid chemical reactions between the sample to be melted and the holder in which it is contained. However, even if we would succeed in this, the system is not closed, because heat is deliberately supplied to it in order to activate the melting process. This causes the system to be open, though controlled if we manage to regulate the heat flow.

Tacitly presupposed in the above discussion is the role of scientists in achieving control over an experimental system. It is through their actions, possibly mediated by instruments, that a physical system of interest is created and experimental conditions are imposed upon the system. They have to design the experiment, prepare the specimen, calibrate the measurement instruments, regulate the temperature, pressure, humidity, and other relevant parameters, fixate the specimen in the test machine, and interpret the resulting data. In order to realize an experimental system, scientists need to have a variety of skills, ranging from the ability to manage an experimental work place to the sometimes subtle use of instruments (see, e.g., Ravetz [1996]; Collins [1992]; deSolla Price [1984]). The amount of work going into the control of a physical system may be quite large.[2]

2.1.3 Property measurements

Measurements are not usually made in an arbitrary way, but they are supposed to be directed at the determination of *properties* of a physical object, such as the fracture strength and melting temperature of a material. A property measurement can be seen as the determination of the value of a certain quantity when an experimental system has been brought into a certain state. The requirement that the system has been brought into a certain state is important. When measuring the strength of a material we are not interested in the value of the applied stress at the beginning of the elongation process, but we want to know this value when the specimen breaks.

A property can be determined in different ways, by performing measurements on different experimental systems. For instance, the strength of a material can be determined in a tension test, a three-point bend test, a four-point bend test, or a pressurized-ring test. In these experiments physical systems with a different structure are created and maintained. Different physical processes occur in the systems and are regulated by the imposition of different experimental conditions. Whereas in a tension test the elongation rate of a sample is controlled, in a bend test experimenters fix the deflection rate. Incidentally, this raises the philosophical question whether experiments of a different type really measure the *same* property.[3]

[2] A beautiful example in the field of biology is given by Griesemer [1992], who discusses the transformation of laboratory organisms into laboratory systems or 'organic machines'.

[3] Similar questions featured in philosophical debates on the operational definition of scientific concepts (see, e.g., Hempel [1966], ch. 7).

Notice that, just as the same property can be measured in different types of experiment, such as the fracture strength above, different properties can be measured in the same type of experiment. Tension tests, for example, are employed to determine the fracture strength as well as elasticity constants of a material.

I have thus far abstracted from the actual determination of the value of a quantity, like the measurement of the temperature in a melting point experiment. A fruitful way to look at this activity is to see it as the performance of an experiment within the experiment. The physical system created in such a subsidiary experiment will be called a *measurement system*. Often, a measurement system is judged to be so well under the control that no questions arise as to the reliability of the results produced by it. The task of creating and sustaining the measurement system can then be delegated to a material device, a measurement instrument which functions as a black box in the laboratory (Latour [1987]). As observed above, the delegation of a measurement to an instrument often marks the transition from an indirect to a direct measurement.

2.1.4 Experimental systems and conflicts

Property measurements are usually obtained in a variety of experiments performed at different places, under different conditions, by different methods, and by different experimenters. A review of the experimental determination of the melting temperature of refractory oxides provides a clear example of the heterogeneity found in measurements of the same property (Hlaváč [1982]). The six values for the melting temperature of zirconium oxide that have been published in the literature since 1963 were obtained in different environments, with samples of different degrees of purity, using different furnaces and containers.

Property measurements obtained in different experiments may not agree, that is, the measurements may be in conflict. A *conflict* between two measurements arises when it is unlikely that the population means estimated by the measurements are the same. The difference between the reported means should be significant, in the sense that it appears too large to be solely ascribed to random errors. The task of *conflict detection* establishes whether a pair of measurements forms a conflict.

Conflict resolution aims at resolving a conflict by giving possible explanations of the significant difference between the two measured values of a property. The cause of a conflict will be sought in the way the experiments underlying the measurements were carried out. In particular, the cause of a conflict is assumed to lie in differences in the structure of the experimental systems and differences in the experimental conditions imposed upon the systems. For instance, a possible cause of a conflict between two strength measurements could be the fact that one of the measurements was performed at a higher temperature, on a sample with a high stress concentration near the grip. Above a certain level, the stress concentration may induce microcracks and thus systematically lower the strength of the specimen. A higher temperature implies a lower theoretical strength which reduces the strength even further.

The identification of possible causes of a conflict can lead to qualitative or (semi-)

quantitative explanations of the conflict. In the former case, one merely explains that one measurement will be higher than another under the given circumstances, whereas in the latter case one also accounts for the observed magnitude of the difference. This thesis will focus on qualitative explanations of a conflict, although in the last part some attention will be given to (semi-)quantitative explanations as well.

2.1.5 Ideal experimental systems and systematic errors

Given a conflict, which of the two measurements provides the best value of a property? That is, which value is most accurate and most precise? The precision of two measurements can be easily compared by looking at the reported standard errors: a lower standard error implies a higher precision. For the accuracy of a measurement no such straightforward statistical criterion exists. A value is more accurate when it is closer to the true value, as we have seen in section 2.1.1. Without a further specification of the notion of true value, however, this merely begs the question.

One way to give an operational definition of the true value of a quantity would be to say that it is the mean of the population measured in an *ideal experiment*, that is, an experiment in which an experimental system is created with an ideal structure and an ideal behavior exhibited under ideal experimental circumstances. Such a system will be called an *ideal experimental system*. From this point of view, a measurement is considered more accurate than another when it is a closer approximation of the mean of the population determined in the ideal experiment.

What counts as 'ideal' is dependent upon the aim of the measurement, and may refer to such criteria as 'approximating an idealized theoretical situation' or 'conforming to a practical situation of interest'. An example of a situation in which the former criterion is intended can be found in the review of melting temperature measurements by Coutures & Rand [1989]. The material sample constituting the ideal experimental system is supposed not to react with the container nor with the atmosphere, since the ensuing changes in composition of the sample would alter the theoretically defined melting point. For the same reason, experimenters strive at the use of samples that are as pure as possible.

Note that it will often be impossible to realize an ideal experimental system in an experiment. In other words, the ideal experimental system is a *hypothetical* system and the true value of the property is a *hypothetical* value. The assessment of a measurement is concerned with the question how closely the measured property value approximates this hypothetical value, the value that would be obtained if the measurement had been performed upon the ideal experimental system.

An experimental system deviating from the ideal system is a *disturbed experimental system*. From the definition of experimental systems we conclude that disturbances are failures to realize the desired structure of the system or failures to impose the desired experimental conditions. Notice that a disturbed system is not bound to be out of control. We can create and maintain an experimental system with a structure deviating from the ideal case, for instance a sample with a high concentration of impurities in a melting temperature experiment. Even though the experimental system is not ideal, the temperature measurement conducted in this experiment yields a value for the property we are after, the melting temperature of the material. This value is not the true value, however, even though it may be the best value available.

The latter remark brings us to the central point that disturbances cause random and systematic errors in measurements. An example of a disturbance leading to random errors is a large spread in the percentage porosity of a batch of material samples used in a tension test. The individual measured values of the fracture strength will likewise show a large spread, thus yielding a low precision of the aggregated measurement. The accuracy of a measurement is affected by disturbances introducing systematic errors in the measurements. An example of a source of systematic errors in a tension test is the development of a high stress concentration near the grips attached to a specimen.

In the activity of *error identification* one looks for possible systematic errors in a measurement by analyzing the experiment in which it has been obtained. In particular, the structure of the physical system created in the experiment and the experimental conditions are compared with their counterparts in the ideal experiment. This comparison should result in an estimate of how the measured value of the property deviates from the true value. Notice the difference between the tasks of conflict resolution and error analysis: in the former case two actual measurements are involved, whereas in the latter case an actual measurement is compared with a value that would have been obtained in an ideal experiment.

Error identification results in a qualitative or quantitative estimate of a systematic error in a measurement. Measurements are corrected for these errors in the task of *error correction*. In the case of a quantitative estimate the correction for a systematic error is trivial, but a qualitative estimate leads to difficulties. In the latter case only the direction, and not the magnitude, of the error (and hence of the correction) is available, although sometimes the direction of the error can be supplemented by an order-of-magnitude estimate. When a systematic error can only be determined in a qualitative way, correction for a systematic error is not possible. In such a situation, one could choose to dismiss the measurement altogether and advice to do a new measurement in an experiment which is a better approximation of the ideal experiment. This strategy is not uncommon in measurement analysis (see, e.g., Hlaváč [1982]; Coutures & Rand [1989]).

By correcting measurements for systematic errors, it is possible to obtain a good approximation of the true value of a property in a non-ideal (i.e., disturbed) experiment. This is common practice in particle physics, where evidence that a certain interaction has occurred is obtained by subtracting from the measured number of events the expected number of events arising from all other interactions that are known to occur under the circumstances of the experiment (Galison [1987], ch. 4). The measured number of events is thus corrected for unintended interactions.

2.1.6 Conflicts, systematic errors, and models of experimental systems

In order to resolve a conflict between two property measurements we need to know which physical systems have been created, under which conditions the experiments have been carried out, and how structural differences and differences in experimental conditions can cause a conflict. This knowledge is captured in models of the experimental systems. Generally spoken, a *model* of a system is a symbolic representation of the system in which a particular view on the system is expressed. In order to be useful, the models of the experimental systems must be *adequate*, that is, they must be able to provide an explanation of the conflict and they must be consistent with what is known, theoretically or practically, about the systems. Similarly, we need adequate models of the actual and ideal experimental system when trying to find a systematic error in a measurement.

There can be different models of a particular physical system. In the first place, the available information about the system may not be sufficient to determine which of a number of alternative models is an adequate description of the system. For instance, from what we know about a certain strength measurement, the specimen being loaded in the tension test might have stress concentrations near the grip, surface cracks, stress concentrations at the shoulder, or a combination of these characteristics. For each of these cases we obtain a different model of the experimental system. In the absence of further information, all of these models need to be taken into consideration. In the second place, a system may be adequately described by a number of different models. The physical system in a tension test, for example, could be modeled as having a boundary which coincides with the physical dimensions of the specimen or as including the test apparatus and the measurement instruments as well. Both ways of framing the experimental systems may be adequate, depending on the question at hand.

The adequacy of a model can change as theoretical insights develop. For instance, the conceptual framework underlying the current view of a system could be adapted or extended to account for phenomena that are currently not well understood.[4] Such a change in conceptual framework may render an existing model inadequate, for instance because it does not take into account a newly distinguished physical process. Interestingly, the persistent failure to explain a conflict by means of existing models of an experimental system can lie at the root of such an extension of the conceptual framework. Attributing the conflict to the occurrence of the new process in the first experimental system and its absence in the second system might enable one to give an explanation where none was found before.

A direct consequence of the fact that experimental systems can be described by different and changing models is the relativity of explanations of conflicts and predictions of systematic errors to a theoretical and conceptual background. Consider the identification of a systematic error by comparing the actual experimental system with an ideal

[4] In their most dramatic form, these changes amount to the complete overthrow of a conceptual framework as occurring in a scientific revolution (Kuhn [1970]). The changes intended here, however, may be more modest.

experimental system. Since the true value of the property is defined as the mean of the population of quantity values sampled in the ideal experiment, a different conception of what counts as an ideal experimental system will generally change the true value. This (implicit) dependence of the true value of a property on a theoretical and conceptual background could be emphasized by speaking about the 'true' value of a property, as some authors do (e.g., Hlaváč [1982], p. 683).

2.2 Methods for model-based measurement analysis

The models that will be considered in this thesis are qualitative mathematical models of the structure of a dynamical system. They are employed by the methods for conflict resolution and error identification introduced below. The methods will be further elaborated in chapter 5, after basic techniques for specifying models of dynamical systems and techniques for reasoning with these models have been discussed. Chapter 5 will also describe how model-based conflict resolution and error identification have been integrated into an implemented MA system.

2.2.1 Conflict detection

As a preliminary to conflict resolution, I will briefly discuss a method for the detection of conflicts between measurements. The method operates upon a measurement base from which measurements of a certain property are retrieved. This retrieval process will not concern us here. The conflict detection method searches the measurements for conflicts and for each pair of conflicting measurements an entry is added to a *conflict matrix* (figure 2.1).

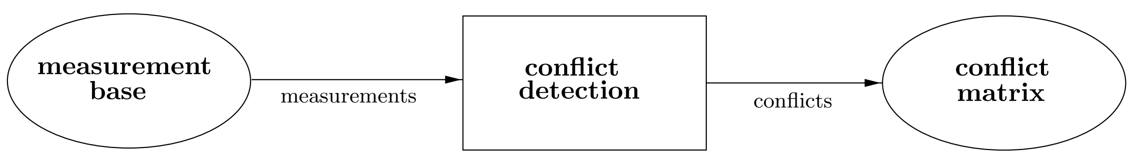


Figure 2.1: *Schematic overview of the method for conflict detection.*

The main problem is of course to fill in the details of the conflict detection algorithm by choosing an appropriate conflict criterion. Conflicts have been introduced within a statistical framework in the previous section: a conflict between two property measurements arises when it is unlikely that the population means estimated by the measurements are the same.

Barry [1978] describes a statistical technique for calculating the probability that two measurements summarize samples drawn from populations with a different mean. The applicability of the statistical technique rests on certain assumptions about the measurements to be compared, in particular that the standard errors of the mean are known and that the measured quantity has a normal distribution. These assumptions may not always be satisfied in practice. A heuristic criterion, such as the rule that two measurements are in conflict if their confidence intervals do not overlap, could be used as an alternative.

2.2.2 Conflict resolution

Conflict resolution starts with pairs of conflicting property measurements from the measurement base. Each property measurement is assumed to be accompanied by a set of candidate models of the experimental system investigated in the experiment, candidate descriptions of the experimental conditions, and a sequence of measurements of the state of the system. The result of conflict resolution is a set of possible explanations of the conflict. The explanations refer to differences in the structure and differences in the experimental conditions of the systems which can account for the observed difference between the measured property values. Figure 2.2 provides a schematic overview of model-based conflict resolution.

The models of the experimental system are qualitative mathematical models of the structure of the system, more specifically qualitative abstractions of differential equations. The variables in the models correspond with the quantities of the experimental system and the functional and differential relations between the variables capture relations between quantities of the system. The models are qualitative in the sense that functional relations between variables as well as values of variables are specified in a qualitative way. They are appropriate modeling devices when one does not know, or does not care about, the exact form of functional relations or the exact numerical value of system parameters (section 1.3).

Due to a lack of information about the experiment, it is often not possible to unambiguously determine the structure of the experimental system investigated in the experiment. Therefore, the property measurement specifies a set of *candidate models* of the experimental system, where each candidate model represents a different possible structure of the system. The candidate models of an experimental system are drawn from what is called a model space. The *model space* for a particular type of experiment is a set of alternative qualitative models of the experimental system. The models in the model space are mutually incompatible, so that at most one of the models is an adequate description of the experimental system under investigation.

A qualitative model of an experimental system can be simulated to obtain possible qualitative behaviors of the experimental system. A qualitative behavior is a qualitative description of the behavior of the experimental system, that is, of the development in time of the value of the quantities of the system. It is a *qualitative* behavior in the sense that it abstracts from the precise numerical value of the quantities. The qualitative behaviors are inferred from a qualitative model and initial qualitative values by means of a *qualitative simulation* algorithm. The initial qualitative values represent the experimental conditions.

The simulation process may result in several possible qualitative behaviors. Although at most one of these is an adequate description of the behavior of the physical system

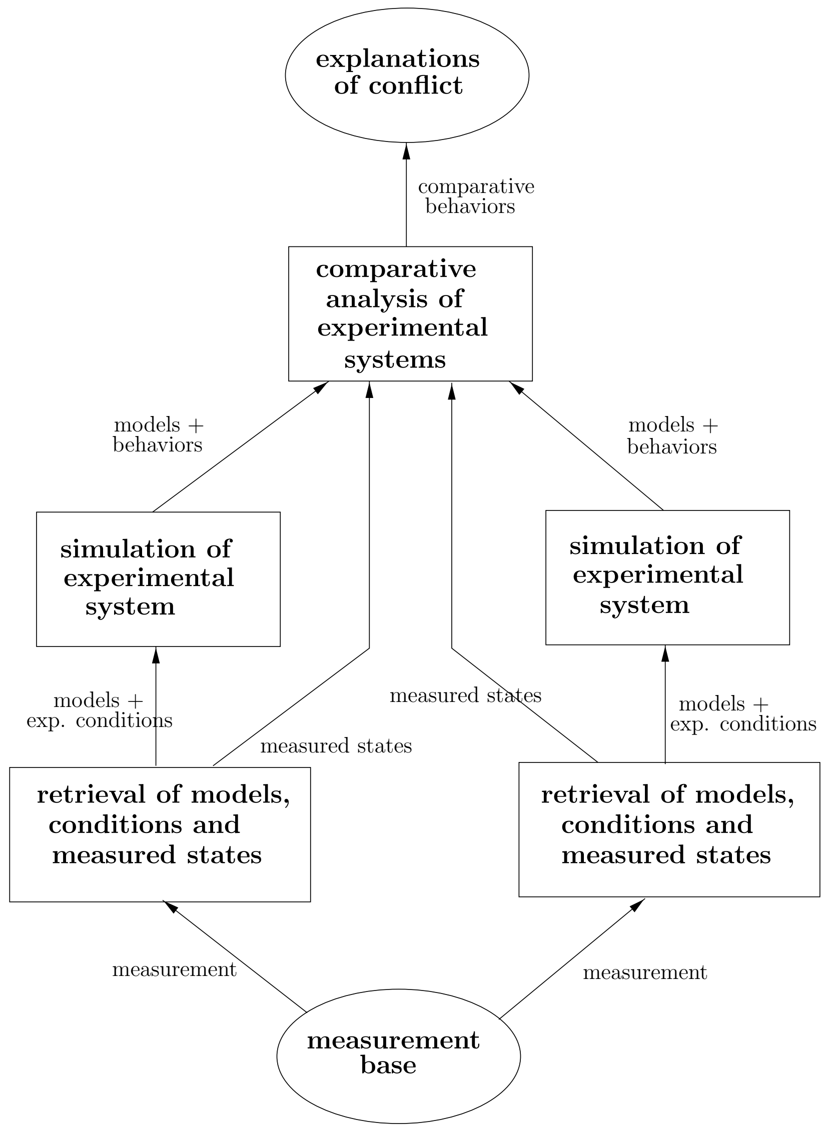


Figure 2.2: *Schematic overview of the method for model-based conflict resolution.*

having been created and controlled in the experiment, the available information about the experiment may not be sufficient to single out this behavior (one can eliminate behaviors inconsistent with the measured states of the system). Hence, we are left with a set of *candidate behaviors* of the system.

Remember that the value of a property was defined to be the value of a certain quantity when the system has been brought into a certain state, e.g., the value of the applied stress when the sample starts to break (section 2.1.3). As a consequence, the conflict between the two property measurements can be seen as a differential response of the two experimental systems.

In order to find explanations of the conflict, every combination of a candidate model and candidate behavior of the first experimental system is compared with every combination of a candidate model and candidate behavior of the second system. The comparison of a qualitative model and behavior of the first system with a qualitative model and behavior of the second system is effected by a technique for the *comparative analysis* of dynamical systems. Starting from initial relative values for some of the quantities, which are provided by the user and by measurements of the state of the systems, the comparative analysis algorithm derives possible comparative behaviors. A comparative behavior is a qualitative description of the differential dynamics of the experimental systems being compared. The comparative behaviors produced in model-based conflict resolution specify how differences in the structure of the experimental systems and differences in the experimental conditions can explain the observed difference in outcome. Each comparative behavior thus represents a possible explanation of the conflict.

2.2.3 Error identification

Whereas conflict resolution focuses on conflicting measurements of a property, error identification is concerned with individual measurements in the measurement base. It compares an actual experiment, in which the property has been measured, with an ideal experiment in order to find possible systematic errors affecting the accuracy of the value. In statistical treatments of measurement, the task of error identification is left to " the observer's alertness and knowledge of the natural, instrumental, and personal factors that can influence his procedures," as cited in the introductory section (see Squires [1985], p. 11, for an almost identical remark). The model-based method for error identification developed in this thesis attempts to approach the problem in a more systematic fashion. It is based on the same set of techniques used for conflict resolution (figure 2.3).

The input of the error identification method consists of a property measurement accompanied by a set of candidate models, candidate experimental conditions, and a sequence of measurements of the state of the experimental system. As described in section 2.2.2, a set of candidate behaviors of the experimental system can be derived by performing a qualitative simulation. For the purpose of error identification, the actual measurement of a property is supplemented by an ideal property measurement. The ideal property measurement specifies a model representing the structure of the ideal experimental system and ideal experimental conditions. Simulation of the ideal model

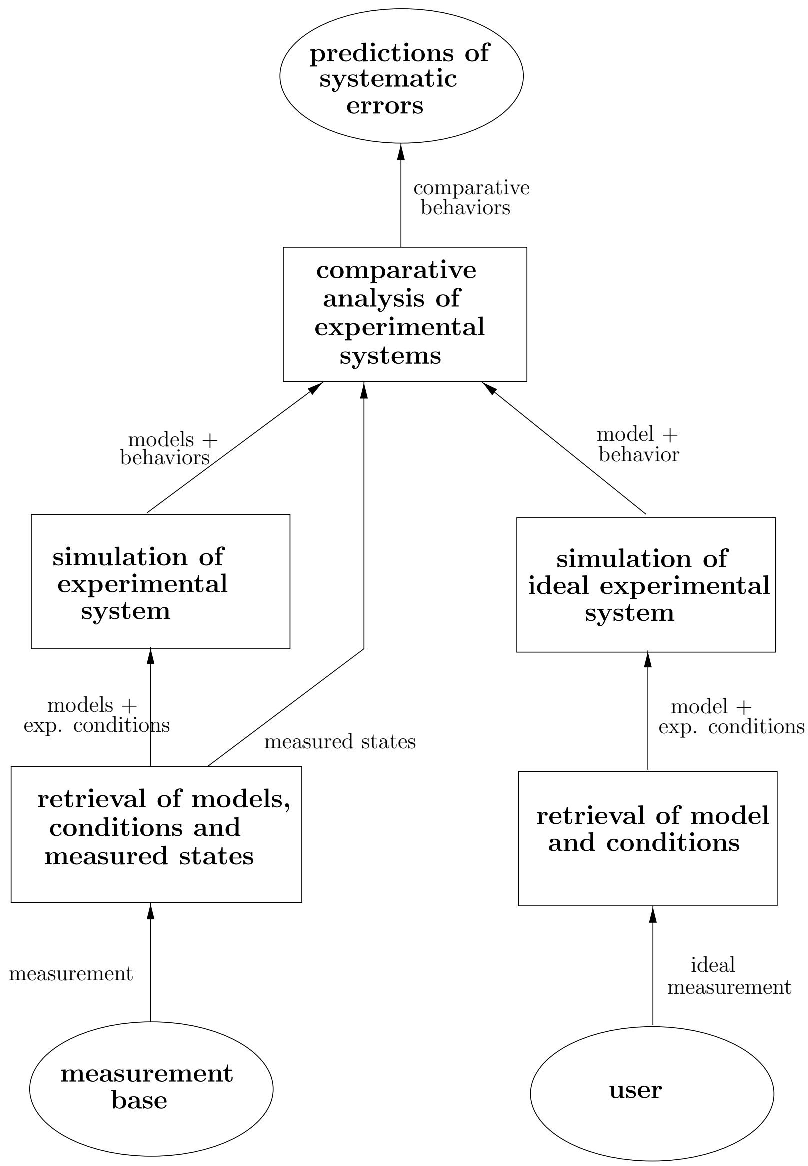


Figure 2.3: *Schematic overview of the method for model-based error identification.*

under the ideal conditions yields an ideal behavior that would have occurred if the ideal experiment had actually been carried out.

Performance of the ideal experiment would result in an estimation of the true value of the property (section 2.1.5). The measured value of the property in the actual experiment contains a systematic error when it deviates from this true value. A systematic error can consequently be interpreted as a differential response of the two experimental systems, i.e. the actual system and the (hypothetical) ideal system.

Possible systematic errors in a property measurement are identified by comparing the model and behavior of the ideal experimental system with every combination of a candidate model and candidate behavior of the actual experimental system. The technique for the comparative analysis of dynamical systems referred to above is used for this purpose. The comparative behaviors produced by the algorithm predict possible differences between the actually measured value and the true value of the property given certain structural differences between the actual and ideal experimental system and differences in the experimental conditions. Hence each comparative behavior predicting a deviation of the measured value from the true value points at a systematic error.

The correction of a systematic error identified by this method is not possible without further information. The predicted deviations from the true value are qualitative deviations, so that we can only conclude that the actually measured value is higher or lower than the true value.

Part II

Methods and Techniques

Chapter 3

Qualitative Simulation of Experimental Systems

Experimental systems are described by qualitative models to account for the fact that in many situations the available information about the structure of a system and the values of its quantities is incomplete. In order to determine the dynamics, the behavior in time, of an experimental system one can simulate its qualitative model. In this chapter I will discuss QSIM, a technique which provides a language to specify qualitative models and an algorithm to perform qualitative simulation. QSIM also forms the basis for the comparative analysis technique CEC*, which will be introduced in the next chapter.

The first section provides an overview of qualitative simulation and related approaches in the field of qualitative reasoning. The basic concepts of QSIM will be introduced in section 3.2. As a preliminary to the description of the algorithm in section 3.4, section 3.3 looks in some detail at the constraints employed in the simulation process. Formal properties of the algorithm, in particular its soundness, incompleteness, and computational complexity, are presented in section 3.5. The final section evaluates the use of QSIM as a means to model and simulate the physical systems investigated in scientific experiments. Various instances of cascaded-tanks systems and a mass-spring system are used as running examples in this chapter and the next. The systems are particularly suitable for illustrating the strengths and weaknesses of qualitative simulation and comparative analysis.

Although I have sometimes slightly modified the presentation and put different emphases to suit my purposes, the description of QSIM closely follows the original article (Kuipers [1986]) and a recent overview of fifteen years of work on the technique and its extensions (Kuipers [1994]). For more details the reader is referred to these sources.

3.1 Qualitative simulation

Qualitative reasoning (QR) is concerned with predicting and explaining the behavior of dynamical systems when there is only incomplete, qualitative information about these systems or when qualitative information is sufficient to answer the questions one would like to pose.[1] A qualitative description of the structure and behavior of a system abstracts from a complete, quantitative account in that it captures distinctions that are thought to be important or interesting, and ignores others. For instance, in order to explain how a material fractures in response to an external stress, we could describe how applying a tensile force to a specimen leads to a stress concentration at the tip of an internal crack. When this stress reaches the theoretical strength of the material the crack starts to propagate until the sample is broken.

In order to engage in qualitative reasoning one has to (1) construct an adequate model describing the structure of the system under study and (2) simulate or otherwise analyze this model to determine the behavior of the system from certain initial conditions. In this chapter the second problem will be addressed by introducing the *qualitative simulation* technique QSIM. The first problem of *model building* is much less explored, and much less constrained, than the second one and will only be given attention in passing. In chapter 8, I will return to model building when discussing how a qualitative model of an experimental system could be automatically inferred from a description of the experiment in which it is investigated.

Qualitative simulation is an important part of three seminal approaches in qualitative reasoning: ENVISION (de Kleer & Brown [1984]), QPT (Forbus [1984]), and QSIM (Kuipers [1986]). In *ENVISION* a qualitative model of a system is constructed from the models of its individual components. This model takes the form of a set of confluences, i.e. qualitative constraints on the signs of system variables. The confluences are used to infer a description of the behavior of the system through simulation. This results in an envisionment, a graph of all qualitative states of the system satisfying the confluences, and all possible transitions between these states. *Qualitative Process Theory (QPT)* provides a process-centered ontology of the physical world which allows one to express common-sense knowledge about physical systems. Given a scenario of a particular situation and a knowledge base of abstract model fragments describing objects, quantities, relations between objects, and processes in the domain, QPT generates a model of the physical system in question. The model is basically a set of constraints on the qualitative values of variables, called influences. As in the ENVISION approach, qualitative simulation is used to construct an envisionment describing the possible behaviors of the system. Unlike ENVISION and QPT, *QSIM* originally focussed on qualitative simulation and ignored the model-building aspect of qualitative reasoning. The qualitative model of a system is a qualitative differential equation, an abstraction of a class of ordinary differential equations. Qualitative simulation exploits continuity properties of the variables and constraints on their qualitative value implied by the qualitative differential equation. It produces a tree of possible qualitative behaviors of the system.

Of these three approaches, QSIM seems to have become the dominant way of performing qualitative simulation. For a large part this is due to its sound mathematical foundation, which has facilitated the adaptation and integration of results from branches of mathematics dealing with ordinary differential equations. In addition, it allows one to

[1] Good and concise introductions to QR can be found in Forbus [1988] and de Kleer [1992]. Somewhat longer, but by far the best introduction I know of, is Kuipers' textbook (Kuipers [1994]).

give guarantees on the validity of the outcome of the simulation process. These were two important motivations for adopting the QSIM approach towards qualitative simulation in this thesis.

ENVISION and QPT have exerted a considerable influence on automated model-building. The component-centered approach of ENVISION was inspired by work in system dynamics, and several other model-building approaches have also found inspiration in that field, for instance approaches based on bond graphs. The process-centered ontology of QPT has formed the basis for a number of model composition techniques (e.g., Forbus [1984]; Low & Iwasaki [1992] Farquhar [1994]). Model-building and qualitative simulation are complementary problems, and accordingly techniques for solving either of them have been used in combination, as for instance QPC and QSIM (Farquhar [1994]).

Sacks [1990a] has shown how the problem of qualitative simulation can be recast in terms of the phase space representation of dynamical systems theory. Basically, qualitative simulation consists of a transition analysis determining the sequence of qualitative states that a system traverses in the phase space and a global analysis deriving its long-term behavior. Although techniques developed in dynamical systems theory are sometimes presented as an alternative to qualitative simulation (e.g., Sacks [1990a]; Kalagnanam, Simon & Iwasaki [1991]), they have also been integrated into qualitative simulation to improve the results of the latter. Examples of work exploring the connections between qualitative simulation and dynamical systems theory will be discussed in section 3.5.

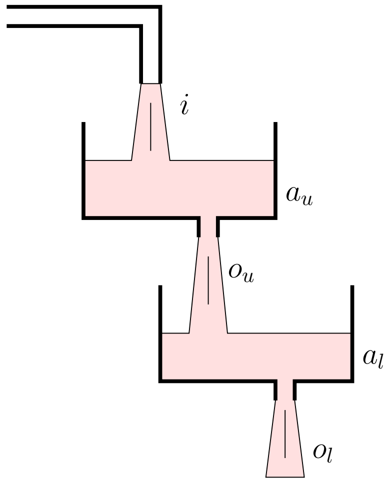
3.2 Basic concepts and outline of QSIM

3.2.1 Qualitative differential equations

Differential equations are arguably the most powerful tool to model dynamical systems in science and engineering. They are *models* of a physical system in the sense that they specify variables and relations between variables which form an adequate characterization of the structure of the system. The structural description of the differential equation can be used to predict the behavior of the system, that is, the evolution of the quantities over time starting from a given initial state (section 3.2.4).

An example of a differential equation, or more accurately a differential equation written as a system of equations, is shown in figure 3.1(b). The differential equation models a system of two cascaded tanks. The (constant) input variables are the inflow i and the sizes r_u, r_l of the orifices of the upper and lower tank. The amounts of water a_u, a_l in the tanks form the state variables.

If numerical values for the constant system parameters i, r_u, r_l and the initial state $a_u(t_0)$, $a_l(t_0)$ are supplied, and the functions f and g have been specified, then the amounts of water in the tanks at equilibrium can be predicted by analytically solving or numerically simulating the differential equation. Sometimes, however, we want to make a statement on the amounts of water at equilibrium *without* knowing or caring about



(a)

$$
\begin{aligned}
\frac{da_u}{dt} &= n_u \\
n_u &= i - o_u \\
o_u &= f(a_u, r_u), \;\; f \in M^{++} \\
\frac{da_l}{dt} &= n_l \\
n_l &= o_u - o_l \\
o_l &= g(a_l, r_l), \;\; g \in M^{++}
\end{aligned}
$$

(b)

Figure 3.1: *(a) Cascaded-tanks system and (b) the qualitative differential equation describing its structure. The variable names have the following interpretation: a water amount, r size of orifice, o outflow, i inflow, and n netflow. The subscripts* $\cdot_u$ *and* $\cdot_l$ *refer to the upper tank and lower tank, respectively.* r_l, r_u, *and* i *are constants*

the exact quantitative value of the system parameters and the initial state. Further, we may not know the exact form of functional relations. In these cases we abstract from a specific differential equation and initial conditions and obtain a more general model. The resulting model will be called a *qualitative differential equation (QDE)*, in contradistinction to the *ordinary differential equation (ODE)* from which it has been abstracted.

Definition 1 (Qualitative differential equation) A qualitative differential equation is an abstraction of an ordinary differential equation

$$\frac{d\boldsymbol{v}}{dt} = \boldsymbol{f}(\boldsymbol{v}),$$

in that:

1. The functions $\boldsymbol{f}$ are decomposed into additions, multiplications, negations, derivatives, and incompletely known functions specified only as being in a certain monotonicity class.

2. The values $\boldsymbol{v}$ are described in terms of a finite number of qualitatively significant landmark values and the intervals between them.

Figure 3.2 summarizes the representation and abstraction step leading from a real

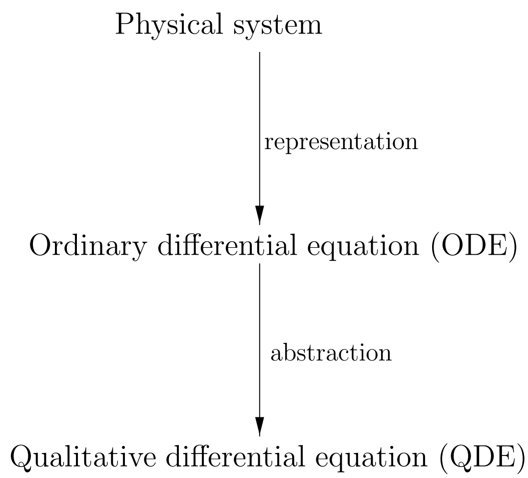


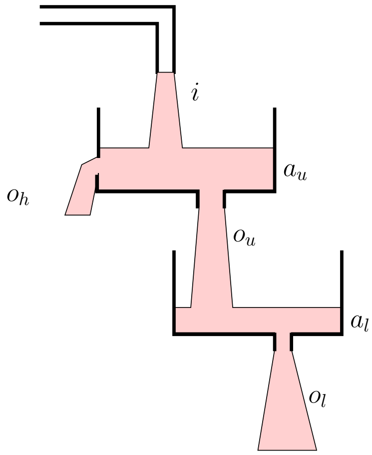
Figure 3.2: *Abstraction from a real physical system to an ordinary differential equation (ODE) and from the latter to a qualitative differential equation (QDE).*

physical system to a QDE.[2] Henceforth, we will equate the *qualitative model* of a dynamical system with a qualitative differential equation.

Kuipers [1994] lists a number of monotonicity classes. M^+ is the class of monotonically increasing functions, that is, for every $f \in M^+$ it holds that $f' > 0$ over the domain of the function. Similarly, M^- is the class of monotonically decreasing functions. M_0^+ and M_0^- are the classes of monotonically increasing and decreasing functions, respectively, such that for every element f we have $f(0) = 0$. The monotonicity classes can be generalized to multivariate functions in an obvious way. For instance, M^{+-} is the class of functions $y = f(x_1, x_2)$, such that $\partial f/\partial x_1 > 0$ and $\partial f/\partial x_2 < 0$ over the domain of the function f.

Qualitative differential equations are valid within a certain operating region. The boundaries of this region are described in terms of constraints on the (qualitative) value of system variables. For instance, the upper tank of a cascaded-tanks system may have a leak at its side (figure 3.3). Until the water in the tank has reached the level of the leak, the system is described by the qualitative differential equation in figure 3.1(b). This model ceases to be adequate, however, as soon as an additional outflow o_h from the leak commences. The system is then described by the QDE in figure 3.3(b). The condition triggering the region transition can be formulated as $a_h = 0$ and $da_h/dt > 0$, that is, the amount of water above the leak is zero but increasing.

[2] In contrast to Kuipers [1994], ch. 1, I will use the term representation instead of abstraction for the modeling relation between physical systems and differential equations. The representational relation between a model and a physical system is different in kind from the formal abstraction of a QDE from an ODE, and these distinctions should be reflected in terminology.



(a)

$$
\begin{aligned}
\frac{da_u}{dt} &= n_u \\
n_u &= i - o_u - o_h \\
o_u &= f(a_u, r_u),\ \ f \in M_0^{++} \\
o_h &= h(a_h),\ \ h \in M^{+} \\
a_h &= a_u - a_{lk} \\
\frac{da_l}{dt} &= n_l \\
n_l &= o_u - o_l \\
o_l &= g(a_l, r_l),\ \ g \in M^{++}
\end{aligned}
$$

(b)

Figure 3.3: *(a) Cascaded-tanks system with a leak at the side and (b) its QDE after the water has reached the height of the leak. Before this level has been reached, the system is described by the model in figure 3.1(b). The names of the variables have the same interpretation as in figure 3.1, with additional variables a_{lk} for the maximum amount of water below the leak, a_h for the amount of water above the leak, and o_h for the outflow from the leak.*

3.2.2 Quantity spaces

The variables in a qualitative differential equation represent time-varying quantities of a physical system. In QSIM these variables are assumed to be *reasonable* functions of time:

Definition 2 (Reasonable function) Given $[a, b] \subseteq \mathbb{R}^*$,[3] the function $v : [a, b] \to \mathbb{R}^*$ is a reasonable function over $[a, b]$ if

1. v is continuous on $[a, b]$,
2. v is continuously differentiable on $]a, b[$,
3. v has only finitely many critical points in any bounded interval, and
4. the one-sided limits $\lim_{t \downarrow a} v'(t)$ and $\lim_{t \uparrow b} v'(t)$ exist in $\mathbb{R}^*$ and are defined to be equal to $v'(a)$ and $v'(b)$, respectively.

At the boundaries of regions in which a QDE is valid the variables may exhibit discontinuities.

The possible qualitative values that a variable can have are determined by its *quantity space*. The quantity space contains *landmark values* capturing qualitatively important distinctions for the variable.

[3] $\mathbb{R}^*$ represents the extended set of real numbers, which includes ∞ and $-\infty$.

Definition 3 (Quantity space and landmark values) A quantity space is a finite, totally ordered set of symbols, the landmark values $l_1 < \ldots < l_k$.

A landmark value, or landmark for short, represents a value in $\mathbb{R}^*$ whose actual value is usually unknown. For instance, the quantity space for the size r_u of the orifice in the cascaded-tanks example can be defined to contain the landmarks $0 < \mathit{open} < \infty$ (figure 3.4). All we know about *open* is that it lies between 0 and ∞. QSIM may create new landmarks during the simulation process. Landmarks which by definition map to known values on the extended real line are $-\infty$, 0, and ∞. The *sign quantity space* consists of these three landmarks, with $-$ or < 0 corresponding with $[-\infty, 0[$; 0 with 0; and $+$ or > 0 with $]0, \infty]$.

$$\begin{aligned}
i: &\quad 0 < \infty \\
a_l, a_u: &\quad 0 < \infty \\
o_l, o_u: &\quad 0 < \infty \\
n_l, n_u: &\quad -\infty < 0 < \infty \\
r_l, r_u: &\quad 0 < \mathit{open} < \infty
\end{aligned}$$

Figure 3.4: *Quantity spaces of variables in figure 3.1.*

Time is itself a variable with the quantity space $t_1 < \ldots < t_n < \infty$. A time-point is a *distinguished* time-point if a variable changes from or to a landmark value at that instant.

Definition 4 (Distinguished time-point) A time-point $t \in [a, b]$ is a distinguished time-point of the variable v, if t is a boundary element of the set $\{t \in [a, b] | v(t) = x$, where $x \in \mathbb{R}^*$ is represented by a landmark value of $v\}$.

The time-point must be a *boundary* of the set to account for the fact that a variable may be constant during the interval. It can be proven that a reasonable function of time defined over a bounded interval has only a finite number of distinguished time-points in its domain and a finite number of landmark values in its range. This property directly carries over to variables in a QDE, since these are assumed to be reasonable functions.

3.2.3 Qualitative values, states and behaviors

The *qualitative value of a variable v at time-point t* is expressed in terms of the landmarks in its quantity space and the direction of change.

Definition 5 (Qualitative value at a time-point) The qualitative value of $v(t)$, $QV(v, t)$, with respect to the quantity space $l_1 < \ldots < l_k$, is the tuple $\langle \mathit{qmag}, \mathit{qdir} \rangle$, where

$$qmag = \begin{cases} l_j & \text{if } v(t) = l_j \\ ]l_j, l_{j+1}[& \text{if } v(t) \in]l_j, l_{j+1}[\end{cases}$$

$$qdir = \begin{cases} inc & \text{if } v'(t) > 0 \\ std & \text{if } v'(t) = 0 \\ dec & \text{if } v'(t) < 0 \end{cases}$$

For example, the size r_u of the upper tank at t_0 is represented by the qualitative value $QV(r_u, t_0) = \langle open, std \rangle$. Since r_u is a constant variable, it will have the same qualitative value at a later time-point: $QV(r_u, t_1) = \langle open, std \rangle$. If the orifice is only partially opened, then the qualitative value of its size $QV(r_u, t_0)$ equals $\langle]0, open[, std \rangle$.

In between two adjacent distinguished time-points the qualitative value of a variable does not change. This motivates the following definition of the *qualitative value of a variable v on a time-interval* $]t_i, t_{i+1}[$:

Definition 6 (Qualitative value on a time-interval) For adjacent distinguished time-points t_i and t_{i+1}, the qualitative value of v on $]t_i, t_{i+1}[$, $QV(v, t_i, t_{i+1})$, is defined to be equal to $QV(v, t)$ for any $t \in]t_i, t_{i+1}[$.

If the tanks are filled from empty, then between the initial time-point t_0 and the next distinguished time-point t_1 the tanks will contain a positive amount of water, more particularly $QV(a_u, t_0, t_1) = \langle]0, \infty[, inc \rangle$ and $QV(a_l, t_0, t_1) = \langle]0, \infty[, inc \rangle$. The amounts of water in the tanks are increasing.

A *qualitative state* of a dynamical system at a distinguished time-point or on an interval between two adjacent distinguished time-points is a tuple of qualitative values, one for each variable in the vector $\boldsymbol{v}$ of system variables.

Definition 7 (Qualitative state) Assume that each v_i in $\boldsymbol{v}$ is a reasonable function defined over the same domain $[a, b]$ and has its own set of landmarks and distinguished time-points. The distinguished time-points of $\boldsymbol{v}$ are now obtained from taking the union of the distinguished time-points of the individual functions v_i. The qualitative state of a dynamical system with a vector $\boldsymbol{v}$ of m variables is the m-tuple of individual qualitative values:

$$QS(\boldsymbol{v}, t_i) = \langle QV(v_1, t_i), \ldots, QV(v_m, t_i) \rangle$$
$$QS(\boldsymbol{v}, t_i, t_{i+1}) = \langle QV(v_1, t_i, t_{i+1}), \ldots, QV(v_m, t_i, t_{i+1}) \rangle$$

Notice that t_i and/or t_{i+1} are not necessarily distinguished time-points of *each* v_j; they are only distinguished time-points of *some* v_k. If t_i and t_{i+1} are distinguished time-points of $\boldsymbol{v}$ but *not* of a particular v_j, then $QV(v_j, t_i) = QV(v_j, t_i, t_{i+1}) = QV(v_j, t_{i+1})$.

Two further consequences of the definition deserve to be mentioned. First, for every variable v_j it holds that its qualitative value does not change during a particular qualitative state $QS(\boldsymbol{v}, t_i, t_{i+1})$. Second, for any pair of adjacent distinguished time-points t_i and t_{i+1} and qualitative states $QS(\boldsymbol{v}, t_i)$, $QS(\boldsymbol{v}, t_i, t_{i+1})$, $QS(\boldsymbol{v}, t_{i+1})$, there is some variable v_k which changes its qualitative value when entering $]t_i, t_{i+1}[$ and some variable v_l which changes its value when reaching t_{i+1}. These properties are inherited by the *qualitative behavior* of a dynamical system.

Definition 8 (Qualitative behavior) The qualitative behavior of a dynamical system with variables $\boldsymbol{v}$ on $[a, b]$ is the sequence of qualitative states

$$QB(\boldsymbol{v}) = \langle QS(\boldsymbol{v}, t_0), QS(\boldsymbol{v}, t_0, t_1), QS(\boldsymbol{v}, t_1), \ldots, QS(\boldsymbol{v}, t_{n-1}, t_n), QS(\boldsymbol{v}, t_n)\rangle,$$

with $t_0 = a$ and $t_n = b$.

A qualitative behavior alternates between qualitative states at distinguished time-points and qualitative states between distinguished time-points. Time is thus represented in QSIM as an alternating sequence of time-points and time-intervals. The term qualitative behavior is sometimes also used in a derivative sense for the sequence of qualitative values of individual variables v_j of the system,

$$QV(v_j, t_0), QV(v_j, t_0, t_1), QV(v_j, t_1), \ldots, QV(v_j, t_{n-1}, t_n), QV(v_j, t_n).$$

3.2.4 Simulation of qualitative differential equations

Differential equations function as tools for predicting or explaining the actual behavior of physical systems. Given certain initial conditions, a differential equation can be analytically solved or numerically simulated to obtain continuous functions $\boldsymbol{v}$ specifying the behavior of the system variables over time (figure 3.5). If the differential equation is an adequate model of the physical system, then the behavior specified by $\boldsymbol{v}$ will correspond with the actual behavior of the system, in the sense that it gives an adequate description of certain aspects of the latter. A QDE can be used as a tool for prediction and explanation in situations where one has to deal with incomplete knowledge about functional relations and parameter values. QDEs are abstractions of ODEs and, correspondingly, the qualitative behaviors generated by QSIM are intended to be abstractions of solutions $\boldsymbol{v}$ of the ODEs. In section 3.5, I will precisely define the abstraction relations and show how they are used to establish formal properties of QSIM.

First, however, I will explain how QSIM generates possible qualitative behaviors from a QDE and initial qualitative state information. Basically, the qualitative simulation process proceeds by repeatedly generating successor states from the qualitative

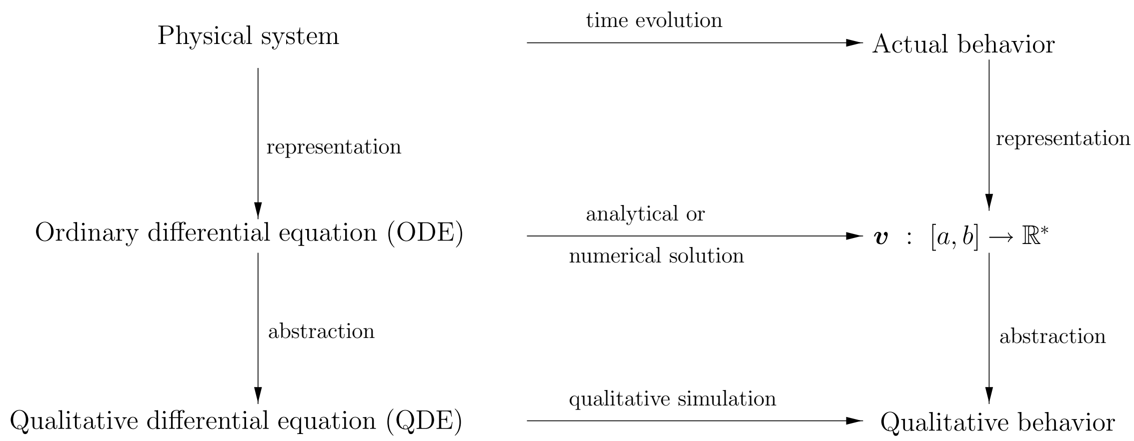


Figure 3.5: *Differential equations and their solutions are representations of the structure and behavior of a physical system. Qualitative differential equations and qualitative behaviors are abstractions of ordinary differential equations and their solutions.*

state at the current time-point or on the current time-interval, and forming qualitative behaviors from the sequences of qualitative states. Which successor states of a particular qualitative state are permitted is determined by constraints on the qualitative values of variables. QSIM distinguishes between three types of *qualitative constraints*: (1) constraints on the completion of a single qualitative state, (2) constraints on the transitions from a qualitative state to successor qualitative states, and (3) constraints on sequences of qualitative states. The next section introduces the qualitative constraints and section 3.4 describes the simulation algorithm in which the constraints are employed.

3.3 Constraints on qualitative values

3.3.1 State constraints

The first type of constraints puts restrictions on the qualitative values of variables in a qualitative state $QS(t_i)$ or $QS(t_i, t_{i+1})$. Given an incompletely specified qualitative state, consisting of qualitative values for some variables, the *state constraints* dictate which alternative completions are consistent with the QDE describing the system. As will be recalled from section 3.2.1, the QDEs considered by QSIM can be formulated as systems of equations containing only additions, multiplications, negations, derivatives, and incompletely known functions specified as belonging to a certain monotonicity class. Each type of equation constrains the qualitative magnitude and direction of the variables involved. The constraints implied by the equations are summarized in table 3.1.[4]

[4]In QSIM the equations in a QDE are often written in the equivalent constraint notation. I will depart from this practice in order to emphasize the similarity between qualitative and ordinary differential equations (compare also Kuipers [1994], ch. 2).

QDE equation	**Qualitative constraint**
$x + y = z$	$ADD(x, y, z)$
$x \cdot y = z$	$MULT(x, y, z)$
$y = -x$	$MINUS(x, y)$
$\frac{d}{dt}x = y$	$D/DT(x, y)$
$\frac{d}{dt}x = 0$	$CONSTANT(x)$
$y = f(x),\ f \in M^+$	$M{+}(x, y)$
$y = f(x),\ f \in M^-$	$M{-}(x, y)$

Table 3.1: *Equation types in a QDE and the corresponding qualitative constraints. The monotonicity classes have been restricted to M^+ and M^-.*

I will give two examples of constraints on qualitative values, in which it is assumed that the variables take their value from the sign quantity space discussed above.

The equation $x+y = z$ corresponds with the constraint $ADD(x, y, z)$. This constraint determines which qualitative values $\langle qmag_x, qdir_x \rangle$, $\langle qmag_y, qdir_y \rangle$, and $\langle qmag_z, qdir_z \rangle$ of the variables x, y, and z are permitted. The following pair of tables defines the *ADD* constraint:

		$qmag_y$		
	ADD	> 0	0	< 0
	> 0	> 0	> 0	?
$qmag_x$	0	> 0	0	< 0
	< 0	?	< 0	< 0

		$qdir_y$		
	ADD	*inc*	*std*	*dec*
	inc	*inc*	*inc*	?
$qdir_x$	*std*	*inc*	*std*	*dec*
	dec	?	*dec*	*dec*

A question mark means that the qualitative magnitude (direction) of z can be > 0, 0, or < 0 (*inc*, *std*, or *dec*).

The qualitative constraint $M{+}(x, y)$ implied by $y = f(x),\ f \in M^+$, is defined as $qdir_x = qdir_y$. There are no constraints on the qualitative magnitudes, since the point at which f crosses the x-axis is not known. If f were an element of M_0^+, then we would additionally have $qmag_x = qmag_y$.

Notice that the use of qualitative information may lead to ambiguities. Given that $ADD(x, y, z)$ and $x > 0,\ y < 0$, the qualitative magnitude of z can take any sign. The *ADD* constraint does not rule out any of these alternatives, since all of them are consistent with the qualitative information. During state completion the occurrence of ambiguities causes the simulation algorithm to produce several qualitative states consistent with the constraints (section 3.4).

The variables in the system may have other quantity spaces than the sign quantity space. The above constraints remain valid in the general case, but could be refined if *corresponding values* are taken into account. Corresponding values are tuples of landmark values for which a certain equation is satisfied. For instance, $\langle x_1, y_6, z_3 \rangle$ are corresponding values for the addition $ADD(x, y, z)$ when $ADD(x_1, y_6, z_3)$. If we now know that $x > x_1$ and $y = y_6$, then we can immediately conclude that $z > z_3$. In order to formalize qualitative constraints in an appropriate and general way, Kuipers [1994] uses a sign-real algebra.

3.3.2 Transition constraints

A qualitative behavior consists of a sequence of qualitative states, alternating between states coinciding with a time-point and states lasting over a time-interval. The *transitions* from a particular qualitative state $QS(t_i)$ to successor states $QS(t_i, t_{i+1})_1, \ldots, QS(t_i, t_{i+1})_m$ are restricted by so-called *transition constraints* in QSIM. These constraints are based on the continuous differentiability of the system variables (definition 2). For continuously differentiable functions the well-known intermediate value and mean value theorems from the differential calculus allow only a few possible transitions from one qualitative value to the next. For instance, the qualitative magnitude of a variable cannot change from positive to negative without passing through 0. The transition constraints rule out candidate successor states whose variables have qualitative values that are not consistent with these restrictions.

Two kinds of transition are distinguished in QSIM: transitions from a time-point to a time-interval (*point transitions* or *P-transitions*) and transitions from a time-interval to a time-point (*interval transitions* or *I-transitions*). The following pair of tables summarizes the possible successor relations from one qualitative value to the next for a continuously differentiable function $v : [a, b] \rightarrow \mathbb{R}^*$, where $l_{j-1} < l_j < l_{j+1}$ are three adjacent landmarks in the quantity space of v.

P-transitions

$QV(v, t_i)$	$\Rightarrow$	$QV(v, t_i, t_{i+1})$
$\langle l_j, std \rangle$		$\langle l_j, std \rangle$
$\langle l_j, std \rangle$		$\langle]l_j, l_{j+1}[, inc \rangle$
$\langle l_j, std \rangle$		$\langle]l_{j-1}, l_j[, dec \rangle$
$\langle l_j, inc \rangle$		$\langle]l_j, l_{j+1}[, inc \rangle$
$\langle l_j, dec \rangle$		$\langle]l_{j-1}, l_j[, dec \rangle$
$\langle]l_j, l_{j+1}[, inc \rangle$		$\langle]l_j, l_{j+1}[, inc \rangle$
$\langle]l_j, l_{j+1}[, dec \rangle$		$\langle]l_j, l_{j+1}[, dec \rangle$
$\langle]l_j, l_{j+1}[, std \rangle$		$\langle]l_j, l_{j+1}[, std \rangle$
$\langle]l_j, l_{j+1}[, std \rangle$		$\langle]l_j, l_{j+1}[, inc \rangle$
$\langle]l_j, l_{j+1}[, std \rangle$		$\langle]l_j, l_{j+1}[, dec \rangle$

I-transitions

$QV(v, t_i, t_{i+1})$	$\Rightarrow$	$QV(v, t_{i+1})$
$\langle l_j, std \rangle$		$\langle l_j, std \rangle$
$\langle]l_j, l_{j+1}[, inc \rangle$		$\langle l_{j+1}, std \rangle$
$\langle]l_j, l_{j+1}[, inc \rangle$		$\langle l_{j+1}, inc \rangle$
$\langle]l_j, l_{j+1}[, inc \rangle$		$\langle]l_j, l_{j+1}[, inc \rangle$
$\langle]l_j, l_{j+1}[, inc \rangle$		$\langle]l_j, l_{j+1}[, std \rangle$
$\langle]l_j, l_{j+1}[, dec \rangle$		$\langle l_j, std \rangle$
$\langle]l_j, l_{j+1}[, dec \rangle$		$\langle l_j, dec \rangle$
$\langle]l_j, l_{j+1}[, dec \rangle$		$\langle]l_j, l_{j+1}[, dec \rangle$
$\langle]l_j, l_{j+1}[, dec \rangle$		$\langle]l_j, l_{j+1}[, std \rangle$
$\langle]l_j, l_{j+1}[, std \rangle$		$\langle]l_j, l_{j+1}[, std \rangle$

3.3.3 Global constraints

The qualitative constraints discussed thus far are local, in the sense that they pertain to a single qualitative state or to a pair of successive qualitative states. The locality of the analysis may lead QSIM to generate qualitative states which are not valid in the light of the sequence of qualitative states preceding it. In order to filter out these qualitative states, *global constraints* global constraint can be formulated which put restrictions on sequences of qualitative states. In section 3.5.3, I will discuss these constraints in more detail.

3.4 The QSIM algorithm

3.4.1 Description of the algorithm

The input of the basic QSIM algorithm consists of a QDE and possibly incomplete initial state information. The algorithm starts by completing the initial state information to all possible qualitative states consistent with the state constraints. Next, it links each initial state to its possible successors determined by the transition constraints and global constraints. This process of state completion followed by successor generation is recursively applied to successor states until a state

1. is found to be inconsistent;
2. is a quiescent state, that is, a state in which all variables have a direction of change equal to *std*;
3. satisfies the conditions for a region transition;
4. is identical with a previous state, signalizing a cyclic behavior;
5. occurs at $t = \infty$.

For each completed initial state,the simulation algorithm thus produces a *qualitative behavior tree*. The qualitative behaviors are the paths from the root(s) to the leaves of a tree, that is, from the initial state(s) to states without successors.

The following algorithm formally describes the generation of qualitative behavior trees in QSIM.

Algorithm 1 (QSIM) Given a QDE with variables $\boldsymbol{v}$ and initial qualitative state information $QS(init)$.

Step 1 Complete $QS(init)$ to initial states $QS(t_0)_1, \ldots, QS(t_0)_n$, which are the roots of the behavior trees. Put the completed initial states on the agenda.

Step 2 If the agenda is empty, then stop. The paths from the root(s) to the leaves in the behavior tree(s) are the qualitative behaviors $QB_1, \ldots, QB_m$. Otherwise, pop a state $QS(t_i)$ or $QS(t_i, t_{i+1})$ from the agenda.

Step 3 For each variable v_i of the system, use the transition constraints to determine the possible successors to $QS(t_i)$ or $QS(t_i, t_{i+1})$.

Step 4 Determine the successor states consistent with the state constraints and global constraints.

Step 5 Add the consistent successor states to the behavior tree by linking them to $QS(t_i)$ or $QS(t_i, t_{i+1})$.

Step 6 Add each eligible successor state to the agenda. A successor state is eligible, unless it is inconsistent, quiescent, identical with a previous state, satisfies conditions for a region transition, or occurs at $t = \infty$. Continue with step 2.

This description of the QSIM algorithm concentrates on its basic features; it leaves out details and attempts a rational reconstruction of some aspects. For instance, in the source references one will find how state completion in steps 1 and 4 of the algorithm is achieved by formalizing it as a constraint satisfaction problem (CSP). QSIM uses the algorithm Cfilter for the solution of this CSP (Kuipers [1994], ch. 4). Most global constraints are applied only after the states have been completed by Cfilter and added to the behavior tree.

At a *transition state* the boundary of an operating region is reached. If there is a model for the system in the region which QSIM is about to enter, qualitative simulation may be resumed in the new region with an initial state defined by a *transition function*. The transition function maps the transition state to the new initial state. The variables may be discontinuous across region transitions (section 3.2.2).

3.4.2 Implementation of the algorithm

The QSIM algorithm has been implemented and embedded within a computer program for the specification and simulation of qualitative differential equations (Farquhar et al. [1993]). The qualitative behaviors resulting from the simulation process can be analyzed and graphically displayed. The user has the opportunity to control the simulation by switching on and off filters which check consistency with various types of qualitative constraints. A trace facility allows one to observe the impact of different constraints in the simulation process and debug the models. The QSIM program has been written in Common Lisp.

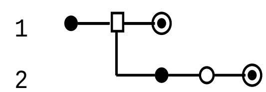
All the examples of QDEs discussed in this thesis have been simulated by means of the QSIM program and the behavior trees shown in the figures have been generated by the program. Appendix E shows examples of QDEs specified in the format required by the QSIM program.

3.4.3 Examples of behavior trees

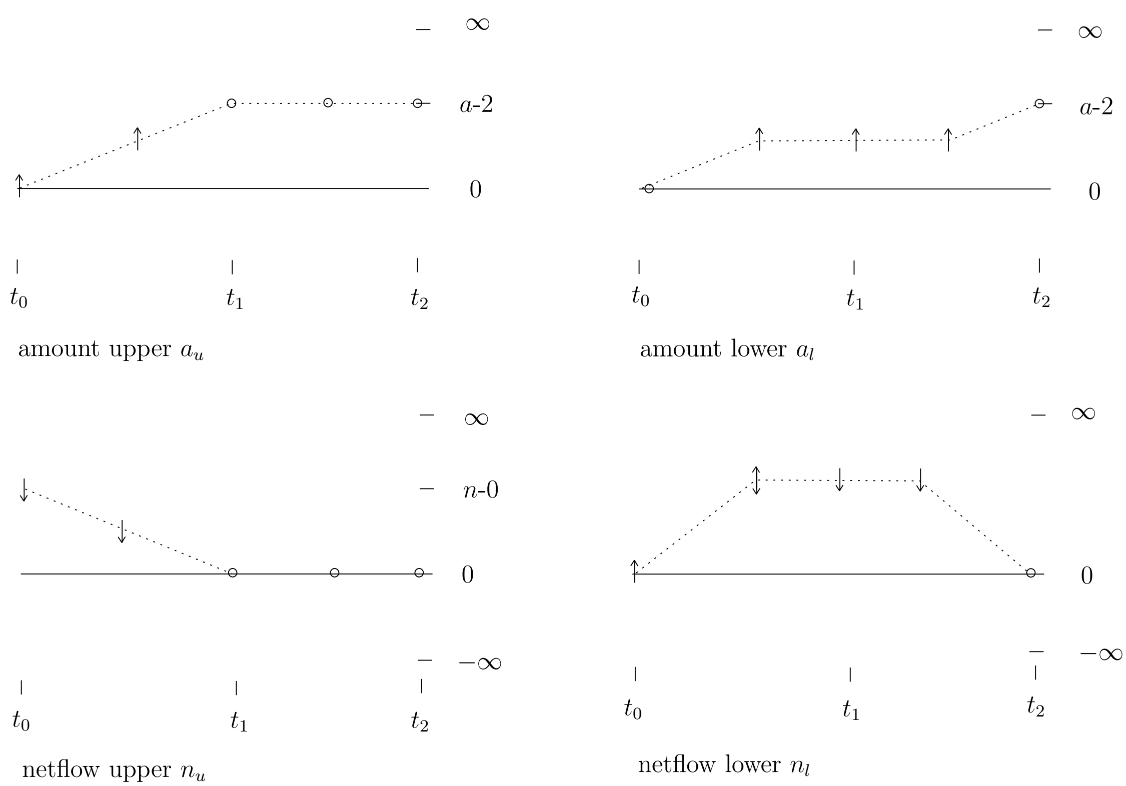
Suppose we simulate the QDE in figure 3.1 with the following initial state information:

$$QS(init) = \langle QV(a_u, t_0) = \langle 0, \_\rangle, QV(a_l, t_0) = \langle 0, \_\rangle, QV(i, t_0) = \langle\langle 0, \infty\rangle, std\rangle,$$
$$QV(r_u, t_0) = \langle\langle 0, open\rangle, std\rangle, QV(r_l, t_0) = \langle\langle 0, open\rangle, std\rangle\rangle$$

That is, the two tanks are filled from empty. This yields the qualitative behavior tree of figure 3.6(a) consisting of two behaviors, one of which is shown in part (b) of the figure. In the first behavior the upper and lower tank simultaneously reach equilibrium and in the second behavior the upper tank reaches equilibrium before the lower.



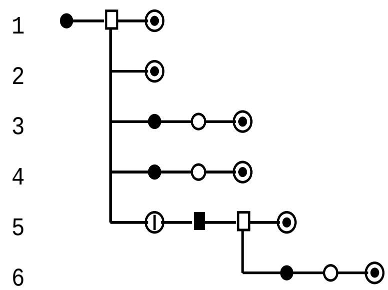
(a)



(b)

Figure 3.6: *(a) Qualitative behavior tree for the model of two cascaded tanks shown in figure 3.1. (b) The time behavior of a few distinctive variables according to one of the qualitative behaviors generated by QSIM. The notation is as follows: • denotes a qualitative state at a time-point, ○ a qualitative state over a time-interval, ◉ a quiescent state,⦶ a transition state, and ◎ the end of a period. The symbols ↑, ↓, ○ in the graphs refer to increasing, decreasing, and steady directions (Farquhar et al. [1993]).*

A qualitative simulation of the leaky cascaded-tanks system of figure 3.3, with the same initial state information, results in the behavior tree shown in figure 3.7. Six qualitative behaviors are generated, accounting for different ordinal relationships between the equilibrium water level in the upper tank and the level of the leak. The system may reach a quiescent state when the water has not yet reached the leak, when it has just reached the leak, or when water is flowing from the leak. Figure 3.7 shows a behavior for the latter case, with the upper tank reaching equilibrium before the lower one. Notice that a region transition occurs at t_1.



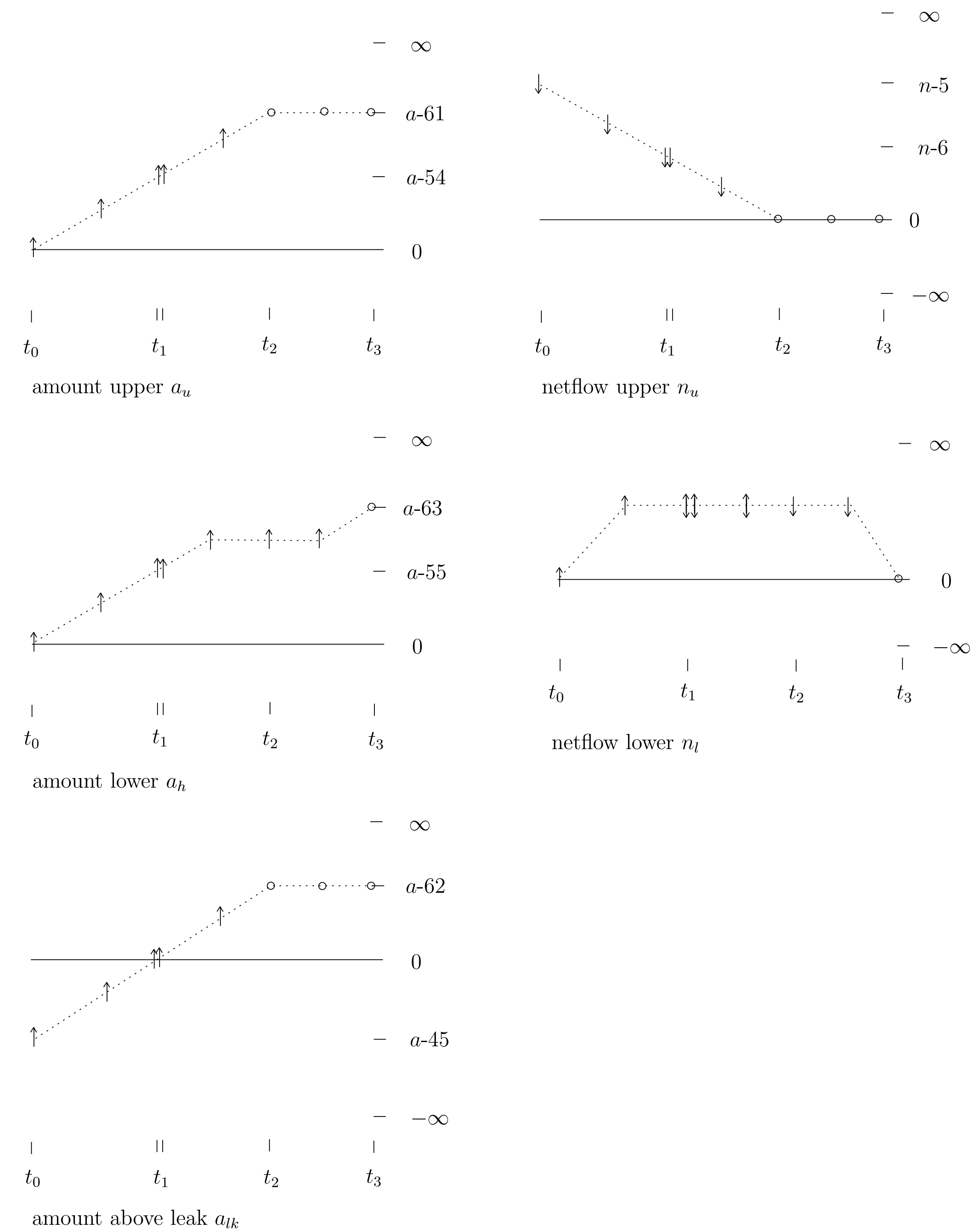
(a)

Figure 3.7: *(a) Qualitative behavior tree for the leaky cascaded-tanks system shown in figure 3.3. (b) The time behavior of a few distinctive variables according to the sixth qualitative behavior generated by QSIM.*

3.5 Properties of the QSIM algorithm

In preparation of the evaluation of qualitative simulation in the next section, I will now review results pertaining to the correctness (i.e., soundness and completeness) of the algorithm, its ability to deal with inconsistent input, and its computational complexity. The formal apparatus introduced to this end will be seen to provide a good starting-point for a similar discussion of the properties of comparative analysis in the next chapter.

Soundness, completeness, and inconsistency are rather broad concepts; before they can be used they should receive a precise definition. This will be achieved in the sections 3.5.1 and 3.5.2 by exploiting the abstraction relations between qualitative differential equations and ordinary differential equations. In its basic formulation the QSIM algorithm has been proven sound and incomplete (section 3.5.3). In many situations, however, global constraints can be employed to take away the consequences of QSIM's incompleteness. The ambiguities attendant upon the use of qualitative models and initial state information may lead QSIM to generate large, unwieldy behavior trees. Section 3.5.4 briefly discusses the computational complexity of QSIM in theory and in practice, and mentions approaches in the QR literature directed at preventing combinatorial



(b)

Figure 3.7: *(Continued)*

explosions.

3.5.1 Behavioral and structural abstraction

The QDE and qualitative behaviors have been interpreted as abstractions of a class of ODEs and a class of vectors $\boldsymbol{v}$ of continuously differential functions v_i, respectively (figure 3.2). These abstraction relations can be formally defined by means of the concepts in section 3.2. The results are summarized in the following pair of theorems (Kuipers [1994], ch. 3).

Theorem 1 (Qualitative behavior abstraction) Let $\boldsymbol{v}$ be a vector of reasonable functions $v_i : [a,b] \rightarrow \mathbb{R}^*$ which represents the time behavior of the quantities of a dynamical system on the bounded interval $[a,b]$. Then a qualitative behavior QB can be defined, consisting of the sequence of qualitative states

$$QS(\boldsymbol{v},t_0),\, QS(\boldsymbol{v},t_0,t_1),\, QS(\boldsymbol{v},t_1),\, \ldots\, ,\, QS(\boldsymbol{v},t_{n-1},t_n),\, QS(\boldsymbol{v},t_n),$$

with $a = t_0$ and $b = t_n$. QB is the qualitative behavior abstraction of $\boldsymbol{v}$.

Theorem 2 (Structural abstraction) Let ODE be an ordinary differential equation $d\boldsymbol{v}/dt = \boldsymbol{f}(\boldsymbol{v})$ written as a system of equations containing only additions, multiplications, negations, derivatives, and functions belonging to the monotonicity classes defined in section 3.2.1. Then a qualitative differential equation QDE corresponding to ODE can be defined, such that any solution $\boldsymbol{v} : [a,b] \rightarrow \mathbb{R}^*$ of ODE is consistent with the qualitative constraints of QDE. QDE is the structural abstraction of ODE.

Theorem 1 is a straightforward consequence of the definitions of qualitative values, states, and behaviors in section 3.2.3. The proof of theorem 2 immediately follows from the observation that the equations of QDE are mathematically equivalent to those of ODE, except for the occasional abstraction from fully specified functions to incompletely specified functions belonging to the monotonicity classes. The structural abstraction from the ODE to the QDE thus weakens the constraints upon the solution, so that any solution of the former must satisfy the constraints implied by the latter.

I will say that an ordinary differential equation ODE and a vector of reasonable functions $\boldsymbol{v}$ *satisfy* or *are consistent with* a qualitative differential equation QDE and a qualitative behavior QB, respectively. On the other hand, QDE and QB *abstract from* or *describe* ODE and $\boldsymbol{v}$. Notice that a whole class of ODEs may satisfy a given QDE.

3.5.2 Defining soundness, completeness, and inconsistency

A careful look at figure 3.2 reveals that it implies suggestions about the way in which the structural and qualitative behavior abstractions hang together. On the one hand, the qualitative abstraction of the solution to the ODE should be found among the behaviors generated by QSIM from the QDE abstracted from the ODE. On the other hand, a qualitative behavior generated by QSIM should describe a solution to some ODE described by the QDE. These suggestions provide the key for defining what is to be understood by soundness and completeness in the context of qualitative simulation.

A productive way to look at QSIM is to view it as a theorem prover deriving theorems of the following form:

$$QSIM \vdash QDE \wedge QS(init) \rightarrow QB_1 \vee \ldots \vee QB_m.$$

QSIM proves from a QDE and initial state information $QS(init)$ a disjunction of qualitative behaviors $QB_1, \ldots, QB_m$.

We can distinguish between *genuine* and *spurious* qualitative behaviors in the disjunction $QB_1 \vee \ldots \vee QB_m$.

Definition 9 (Spurious qualitative behavior) A spurious qualitative behavior QB_i in a QSIM prediction

$$QDE \wedge QS(init) \rightarrow QB_1 \vee \ldots \vee QB_m$$

is a qualitative behavior which describes no solution to any initial value problem $ODE \wedge \boldsymbol{v}(t_0) = \boldsymbol{v_0}$ satisfying $QDE \wedge QS(init)$. A genuine qualitative behavior is a behavior that is not spurious.

Whereas a genuine qualitative behavior is a behavior that can actually occur, because it describes the solution to some $ODE \wedge \boldsymbol{v}(t_0) = \boldsymbol{v_0}$ satisfying $QDE \wedge QS(init)$, a spurious behavior *cannot* occur. It does not describe the solution to any ordinary differential equation and initial state satisfying the qualitative differential equation and initial qualitative state information from which the spurious behavior has been derived. In the case of a spurious qualitative behavior, the behavioral abstraction relation in figure 3.2 breaks down.

These preliminaries directly lead us to a straightforward (though not undisputed, see Struss [1988b]) definition of the *soundness* and *completeness* of QSIM and other qualitative simulation algorithms.

Definition 10 (Soundness) QSIM is *sound*, iff all genuine qualitative behaviors are generated by the algorithm.

Definition 11 (Completeness) QSIM is *complete*, iff every qualitative behavior generated by the algorithm is genuine.

Generally speaking, an algorithm is sound, if it does not draw an invalid conclusion from valid input. An algorithm is complete, if it is able to derive every valid conclusion following from the input. These intuitive notions of soundness and completeness are captured by the definitions above. Consider a solution $\boldsymbol{v}$ to $ODE \wedge \boldsymbol{v}(t_0) = \boldsymbol{v_0}$. Soundness means that the qualitative behavior describing $\boldsymbol{v}$ is among the behaviors generated by QSIM. The algorithm is called sound, because a disjunction is always true when one of its disjuncts is true. If a qualitative behavior in the disjunction is a spurious behavior, however, then QSIM fails to prove the stronger theorem *without* this behavior. As a consequence, it is called incomplete.

The input of the QSIM algorithm is called *inconsistent* when there is no solution to any ODE and initial state satisfying the QDE and initial qualitative state information. consistency

Definition 12 (Inconsistent input) The input $QDE \wedge QS(init)$ of the QSIM algorithm is *inconsistent*, iff no conceivable qualitative behavior for this input is a genuine qualitative behavior.

Defining an inconsistency is one thing, detecting an inconsistency another. There is an interesting relationship between QSIM's ability to detect an inconsistent input and its completeness.

Theorem 3 If QSIM is complete, then it will detect every inconsistency in the input.

Obviously, an inconsistency in the input should lead to a trivial qualitative behavior tree without states and state transitions. Since a complete algorithm does not generate spurious behaviors, it is guaranteed to produce such a null tree.

3.5.3 QSIM is sound and incomplete

QSIM has been proven sound and incomplete (Kuipers [1994], ch. 5). soundness

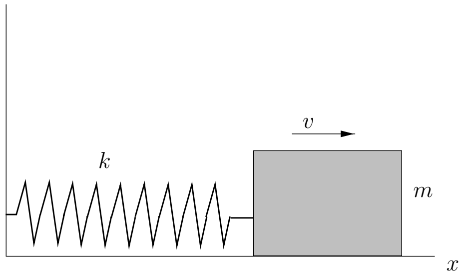
Theorem 4 QSIM is sound.

Theorem 5 QSIM is incomplete when it uses only state constraints and transition constraints.

As a consequence of this incompleteness property, QSIM may generate spurious qualitative behaviors. Only when a single qualitative behavior has been produced by QSIM, and the QDE and initial state information are known to be consistent, is it possible to conclude without further analysis that we are dealing with a genuine behavior. Further, inconsistencies in the input may be obscured by the generation of spurious behaviors. What can be guaranteed, however, is that all genuine behaviors are included in the behavior tree.

An example in which spurious behaviors manifest themselves is a simulation of the well-known ideal mass-spring system of figure 3.8. When the spring is stretched, it exerts a force on the block that causes it to move back. Apart from the position x, velocity v, and acceleration a, the relevant variables are the spring constant k and the mass constant m. If friction is ignored, as in figure 3.8, the motion of the mass is an undamped vibration. (Otherwise, the motion is either a damped vibration or an overdamped return to the equilibrium position.)

Figure 3.9 shows the behavior tree generated by QSIM for this model. The tree branches at t_3 to produce three qualitative behavior fragments at t_4. The first one



(a)

$$\frac{dx}{dt} = v$$
$$\frac{dv}{dt} = a$$
$$a = -\frac{k}{m} \cdot x$$

(b)

Figure 3.8: *(a) Ideal (undamped) mass-spring system and (b) the QDE by which it is defined. The variables refer to the position x, velocity v, acceleration a, and mass m of the block, and the spring constant k. The mass and the spring constant are time-invariant.*

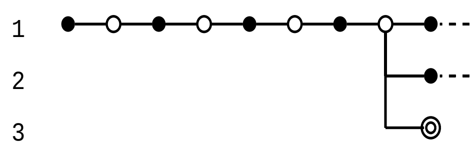
represents an increasing oscillation, the second one a decreasing oscillation, and the third one a steady oscillation. Although all three behaviors are consistent with the state and transition constraints, only the third behavior is a genuine possibility. One can easily show that the other two behaviors violate an implicit invariant: the sum of the potential and kinetic energy of the system. Since the total energy at t_4 should equal that at t_0, decreasing and increasing oscillations are impossible.

The energy constraint is an example of a global constraint on the behavior of a system, a constraint which puts restrictions on the possible sequences of qualitative states up and above state constraints and transition constraints. Global constraints can be made explicit and integrated into the qualitative simulation process to prune spurious qualitative behaviors from the behavior tree. Efforts in this direction have been guided to a large extent by the insights and results of the mathematical theory of dynamical systems. Examples of global constraints include the use of energy constraints, or more generally Lyapunov functions (Fouché & Kuipers [1992]), non-intersection constraints on trajectories in the phase space (Struss [1988a]; Lee & Kuipers [1988]), qualitative stability analysis (Ishida [1989]), and analytical function constraints (Kuipers [1994], ch. 10).

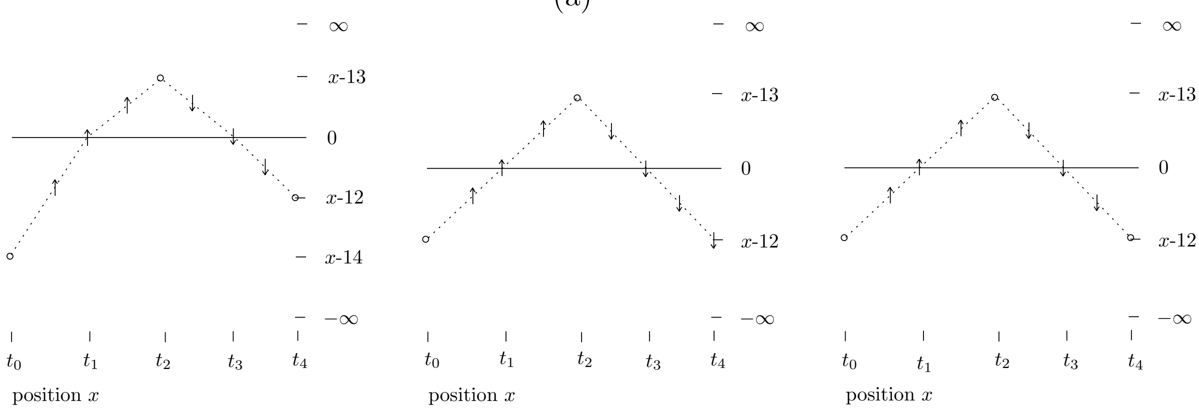
The inclusion of the explicit energy function

$$E = \frac{1}{2}mv^2 + \frac{1}{2}kx^2$$

in the QDE of the ideal mass-spring system, with E constant to reflect energy conservation, filters out the increasing and decreasing oscillations in figure 3.9. Notwithstanding the successful elimination of some spurious qualitative behaviors, one should bear in mind that the global constraints are not guaranteed to remove *all* spurious behaviors. In addition, incompleteness does not just arise from the failure to take into account constraints on sequences of qualitative behaviors. As Struss [1988b] has shown, *spurious states* may already be generated during state completion, due to the weakness of the qualitative calculus combined with the local nature of the CSP algorithm.



(a)



(b)

Figure 3.9: *(a) Qualitative behavior tree for the model of the undamped mass-spring system shown in figure 3.8. (b) The time behavior of the position of the spring in the three behaviors.*

Whether qualitative simulation can be made complete or is inherently incomplete thus remains an open problem. To put this incompleteness in perspective, though, it should be mentioned that problems due to the local character of the analysis also arise in numerical simulation. For example, round-off errors may lead numeric algorithms to miss global conservation constraints and careless simulation around singular points may produce non-sensical solutions (Boyce & DiPrima [1992], ch. 8).[5]

3.5.4 Computational complexity of QSIM

An estimate of the complexity of the QSIM algorithm is the number of qualitative states that need to be generated. In the worst case, the algorithm is exponential in the number of variables of the system and the length of the longest behavior (Kuipers [1986]). What is more, the algorithm need not halt and can continue producing ever more and ever longer qualitative behaviors.

The average-case behavior of the basic algorithm is usually much less than exponential. In some instances, however, the number of qualitative behaviors may be overwhelming and render QSIM useless as a tool for practical purposes (Makarovic [1991]). Assuming for the moment that all behaviors are genuine, the cause of intractable branching of the behavior tree must be sought in the weakness of the qualitative information contained in the model and initial state information. Qualitative information may give rise to ambiguities, sometimes quite severe, as the case of *chattering variables* illustrates. When a variable chatters, its qualitative value is totally unconstrained except by continuity, so that the algorithm is forced to repeatedly branch on every possible magnitude and direction.

Notwithstanding these complications, the conclusion that qualitative simulation is useless is certainly not warranted in general. It overlooks the possibility that additional information can be added to avoid intractable branching. One could include equations representing hidden assumptions into the QDE, incrementally grow the behavior tree by requesting new observations whenever the number of successor states threatens to explode, employ special techniques directed at the elimination of chatter (Kuipers [1994], ch. 10; Clancy & Kuipers [1997]), explore the behavior tree depth-first by prioritizing the most likely successor states (Leitch & Shen [1993]), or engage in order-of-magnitude reasoning (Raiman [1986]).

A particularly promising strategy is the extension of qualitative models with semi-quantitative information, i.e., interval bounds on the value of landmarks and numerical envelopes around monotonic functions (Kuipers [1994], ch. 9; Kay & Kuipers [1993]; Berleant & Kuipers [1997]; see Forbus & Falkenhainer [1992] for another approach towards the integration of qualitative and numerical information). The more sophisticated semi-quantitative simulation algorithms guarantee that the numeric interval in which the value of a landmark lies converges to a single point as the uncertainty in the initial conditions and the maximum step-size converges to zero.

[5]This point was suggested to me in discussion with Hans Akkermans.

3.6 Evaluation

In this thesis I will model experimental systems by means of QDEs and derive the behavior of an experimental system through qualitative simulation of its QDE. By this approach it is possible to represent and reason with incomplete information about the systems, in particular the value of quantities and the functional relations connecting them. Further, the qualitative behavior tree generated by QSIM provides a suitable starting-point for giving explanations of the possible behaviors of an experimental system. The solid foundation of qualitative simulation in the theory of differential equations, through the abstraction relations spelled out in the theorems 1 and 2, allows one to determine the capabilities of the QSIM algorithm in a precise and formal way.

The soundness and incompleteness of QSIM have consequences for the validity of the predictions in qualitative simulation, and thus for the validity of the answers produced in the resolution of conflicts and the identification of systematic errors. Although one can be sure that every possible qualitative behavior will be included in the behavior tree (soundness), some of these behaviors may not describe a possible behavior of the real system (incompleteness). Fortunately, the problem of incompleteness does not seem to be critical in the practical application of qualitative simulation in an MA system. Often one can determine on independent grounds (e.g., by taking additional measurements into account) which behavior or behaviors have occurred and select these behaviors from the QSIM output.

Several strategies to counter the problem of intractable branching were discussed, including the use of semi-quantitative information. Although I will focus on qualitative simulation, it should be kept in mind that the qualitative approach can be generalized to semi-quantitative simulation. When quantitative information is available and deemed useful, it would be possible to extend the models of experimental systems and refine the predictions of their behavior (chapter 8).

Chapter 4

Comparative Analysis of Experimental Systems

Different experiments may create and sustain different experimental systems or investigate experimental systems under different conditions. In order to find possible causes of deviations between two measurements of a property, or to estimate the deviation of a measurement from the value that would have been obtained in an ideal experiment, a comparative analysis of experimental systems can be carried out. In a comparative analysis, one propagates observed differences in the value of variables through the qualitative behaviors of two dynamical systems described by qualitative differential equations. Unfortunately, existing QR techniques for comparative analysis are not general and powerful enough to meet the requirements imposed by the problem at hand. Therefore, a major part of the work on this thesis has been directed at the development of CEC*, a general technique for the comparative analysis of dynamical systems. This chapter reports upon the results.

After a brief introduction to comparative analysis, the exposition starts with the definition of some basic concepts and the formulation of CEC* as a propagation process (section 4.2). The constraints employed in the propagation process form the topic of section 4.3. Section 4.4 provides a formal description of the algorithm by which the propagation process is realized and shows how CEC* deals with an example of a predictive and a diagnostic CA question. Properties of the algorithm – particularly its soundness, completeness, and computational complexity – are discussed in detail in section 4.5. Section 4.6 recapitulates the discussion by placing CEC* in the context of other work on comparative analysis. An evaluation of the applicability of the technique for the comparative analysis of experimental systems concludes the chapter.

4.1 Comparative analysis

An important problem in qualitative reasoning concerns the comparison of behaviors of two systems, that is, an assessment of the relative value of system variables at chosen time-points. More specifically, this *comparative analysis (CA)* task is directed at (1) predicting possible consequences of differential initial conditions and (2) finding possible causes of differential responses of the two systems. When fully-specified quantitative models of the systems are available, comparative analysis reduces to a trivial problem; one simply compares the numerical solutions of the ODEs at the time-points of interest. In the case of qualitative models and behaviors, however, a direct comparison of the qualitative values of variables may be quite uninformative. Nothing is known about the relative value of a variable when it has the qualitative magnitude $]0, \infty[$ in both the first and the second system. The QDEs describing the systems need to be taken into account in order to arrive at a stronger conclusion.

With few exceptions, QR techniques for comparative analysis have focussed on the comparison of systems with the same structure: *intra-model* CA (Weld [1990]; Chiu & Kuipers [1992]; Neitzke & Neumann [1994]; de Jong, Mars & van der Vet [1996]). The comparison of dynamical systems with a different structure, *inter-model* CA, is considered to be a "terribly difficult" problem in general (Weld [1990]). The limitation to intra-model CA is a serious shortcoming, since many interesting questions can be mapped onto inter-model CA problems, including the comparison of experimental systems. When experimental scientists fail to prevent the occurrence of a disturbing process in one experiment and effectively counter it in another, as explained in chapter 2, the experimental systems under investigation may have a different structure. Similar problems arise in other contexts, such as an engineer diagnosing an electrical circuit by comparing the observed faulty behavior with a reference behavior.

In this chapter I will introduce CEC*, a general technique for comparative analysis which treats intra-model CA *and* inter-model CA within a single framework (de Jong & van Raalte [1997a]; de Jong & van Raalte [1997b]). More specifically, intra-model problems are handled as a special case of inter-model problems. In CEC*, comparative analysis is reformulated as a propagation problem and elaborated by means of basic concepts from calculus and linear systems theory. From a triple input of a pair of QDEs, a pair of qualitative behaviors, and selected relative values, CEC* generates a comparative envisionment representing the differential dynamics of the two systems. The clear mathematical foundation of the technique allows one to provide guarantees on the correctness of the results produced by it. CEC* has been implemented and tested on a range of CA questions concerning simple and more advanced QR systems.

4.2 Basic concepts and outline of CEC*

4.2.1 Pairs of comparison

In a comparative analysis, a qualitative behavior of a first system is compared with a behavior of a second system. I will assume that the qualitative behaviors have been generated by the QSIM algorithm, although the principles underlying our approach do not depend on this choice.

Consider the example of the two cascaded tanks introduced in the previous chapter (figure 3.1). In one of the two qualitative behaviors produced by QSIM for this system, the upper tank was seen to reach equilibrium before the lower tank (figure 3.6). Now suppose we compare this system with a second cascaded-tanks system which has an upper tank with a leak at the bottom (figure 4.1). The leaky system is structurally different from the watertight one: the netflow into the upper tank is given by $n_u = i - o_u - o_h$ instead of $n_u = i - o_u$. An extra term o_h accounting for the flow out of the leak has been added. The qualitative behavior in which the upper tank reaches equilibrium before the lower one is also a behavior of the leaky system.

The behaviors originating from the qualitative simulation of two dynamical systems are compared at so-called *pairs of comparison*.

Definition 13 (Pair of comparison) A pair of comparison pc is a pair of distinguished time-points $\langle t, \hat{t} \rangle$ from the first and the second behavior at which the values of shared variables are compared.

Throughout this chapter, the hat accent $\hat{}$ will be used to denote time-points in the behavior of the second system.

Figure 4.2 depictures the comparison of behaviors of the normal and the leaky cascaded-tanks system in which the upper and the lower tank consecutively reach equilibrium, as in the figures 3.6(b) and 4.1(c). Three intuitively meaningful pairs of comparison are chosen: pc_0 when the upper and lower tank of both systems are empty, pc_1 when the upper tank reaches equilibrium in both systems, and pc_2 when the lower tank reaches equilibrium in both systems. In order to identify meaningful pairs of comparison in a systematic way, one would like to articulate the intuition underlying the choice of pairs of comparison. This will be achieved by definition 15 below.

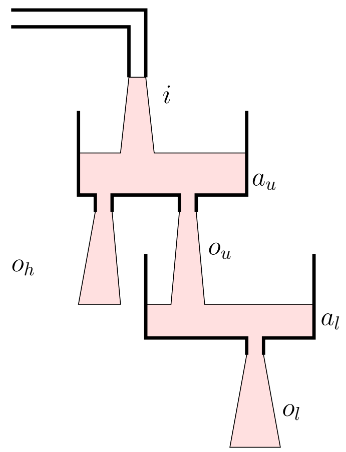
The choice to limit the comparison of two behaviors to selected pairs of *time-points* implies that nothing can be said about how the values of corresponding variables relate during *intervals*. This contrasts with, for example, Weld's DQ approach (Weld [1990]), but has the important advantage that it provides a guarantee on the soundness of CEC*'s output (section 4.5 and 4.6).

4.2.2 Ordering of pairs of comparison

A partial ordering relation can be defined on pairs of comparison which formalizes the notion that one pair of comparison occurs before another.

Definition 14 (Ordering relation $\preceq$) For two pairs of comparison $pc_1 = \langle t_1, \hat{t}_1 \rangle$ and $pc_2 = \langle t_2, \hat{t}_2 \rangle$ it holds that $pc_1 \preceq pc_2$, iff $t_1 \leq t_2$ and $\hat{t}_1 \leq \hat{t}_2$.

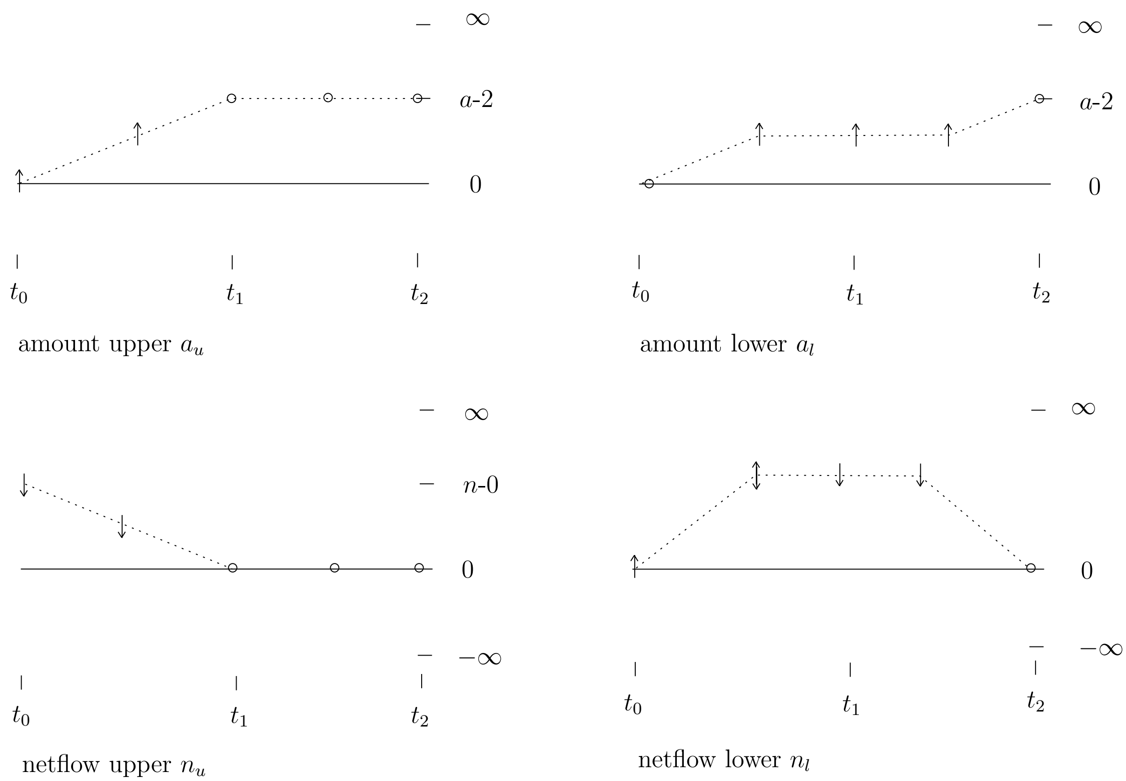
It is not difficult to verify that $\preceq$ has the reflexivity, antisymmetry, and transitivity properties of a partial ordering relation. If pc_1 and pc_2 are different pairs of comparison and $pc_1 \preceq pc_2$, then pc_1 is said to be a *predecessor* of pc_2 (and pc_2 a *successor* of pc_1). If pc_1 is a predecessor of pc_2, and there is no pc_3 such that $pc_1 \preceq pc_3 \preceq pc_2$, then pc_1 is called a *direct* predecessor of pc_2. Analogously, we can define direct successor pairs of comparison. Since we will be concerned with direct predecessors and successors most of the time, the qualification 'direct' is often omitted.



(a)

$$
\begin{aligned}
\frac{da_u}{dt} &= n_u \\
n_u &= i - o_u - o_h \\
o_u &= f(a_u, r_u),\ \ f \in M^{++} \\
o_h &= h(a_u),\ \ h \in M_0^{+} \\
\frac{da_l}{dt} &= n_l \\
n_l &= o_u - o_l \\
o_l &= g(a_l, r_l),\ \ g \in M^{++}
\end{aligned}
$$

(b)



(c)

Figure 4.1: *(a) Leaky cascaded-tanks system and (b) the qualitative differential equation describing its structure. The names of the variables have the same interpretation as in figure 3.1, with the additional variable o_h standing for the outflow from the leak. (c) One of the two qualitative behaviors produced by QSIM for this system.*

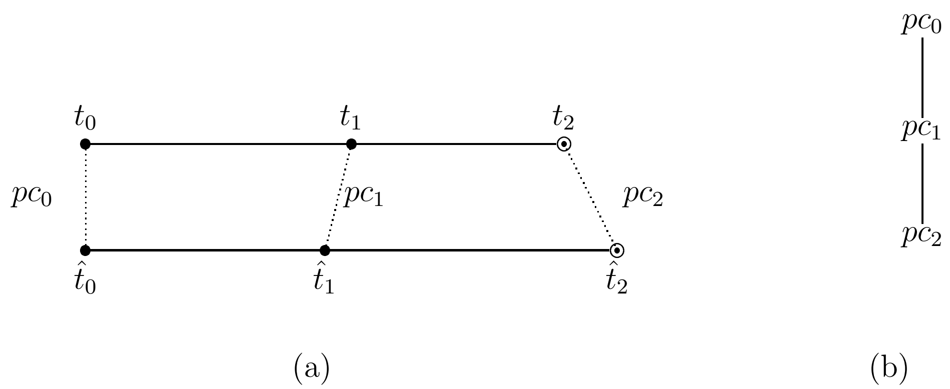


Figure 4.2: *(a) The comparison of qualitative behaviors of cascaded-tanks systems with a watertight upper tank and a leaky upper tank. In both behaviors the upper tank reaches equilibrium before the lower tank (figure 3.1(c)). The notation of the time-points follows the QSIM conventions (Farquhar et al. [1993]). Pairs of comparison are indicated by dotted lines. (b) The set of meaningful pairs of comparison ordered by the* $\preceq$ *relation.*

Figure 4.2(b) shows a graph with the ordering relation $\preceq$ imposed upon the set of pairs of comparison in that example: $pc_0 \preceq pc_1 \preceq pc_2$. A set of pairs of comparison ordered according to the $\preceq$-relation will be called an *ordered pairs of comparison (OPC) structure*. This structure always has a top element, the first pair of comparison, and a bottom element, the last pair of comparison. More formally, the first pair of comparison is the least upper bound and the last pair of comparison the greatest lower bound of the structure.

The behaviors in figure 4.2 are *topologically equal* (Weld [1990]) with respect to the shared variables, that is, when attention is restricted to the shared variables of the two systems, the behaviors show the same sequence of transitions between qualitative states and the variables have the same qualitative value in the corresponding states in this sequence.[1] If two behaviors are not topologically equal, they are *topologically different*. More generally, one can prove that CA problems with topologically equal behaviors, of which figure 4.2 is an instance, will lead to OPC structures with a single path from top to bottom. For topologically different behaviors such a guarantee cannot be given, and a pair of comparison may have more than one predecessor, leading to several paths in the OPC structure (see de Jong [1996] for examples).

Two successive pairs of comparison in an OPC structure mark out what we will call a pair of *behavior fragments*, one for each system. The behavior fragments coming with the successive pairs of comparison $pc_1 = \langle t_1, \hat{t}_1 \rangle$ and $pc_2 = \langle t_2, \hat{t}_2 \rangle$ are defined over the closed intervals $[t_1, t_2]$ and $[\hat{t}_1, \hat{t}_2]$, where $t_1 \leq t_2$ and $\hat{t}_1 \leq \hat{t}_2$. In the limiting case, when

[1] The restriction to shared variables is essential, for if the comparison of qualitative states of behaviors is extended to non-shared variables as well, behaviors arising from different models with different variables can never be topologically equal. The definition of topological equality given here is a generalization of Weld's original definition (Weld [1990]) and reflects the generalization from intra-model to inter-model CA undertaken here.

$t_1 = t_2$ or $\hat{t}_1 = \hat{t}_2$, an interval reduces to a single time-point.

If the interval of a behavior fragment does not contain any distinguished time-points except for its boundaries, we will say it is a *primitive* behavior fragment, otherwise a *composite* behavior fragment. During a primitive behavior fragment the qualitative models of the systems do not change, that is, the systems are described by the same variables and relations. In addition, the qualitative values of the variables only change, if they do, at the boundaries of the interval. In composite behavior fragments this is not necessarily the case, since changes in the qualitative value of variables, and possibly region transitions as well, will occur at the enclosed distinguished time-points.

Clearly, in figure 4.2 there are only primitive behavior fragments. In order to illustrate composite behavior fragments, consider the leaky cascaded-tanks system of the previous chapter, in which the leak is located at the side instead of at the bottom of the upper tank (figure 3.3). One of the behaviors of this system was shown in figure 3.7. In this behavior a region transition occurs at t_1 when the upper tank contains an amount a_{lk} of water reaching up to the level of the leak. A comparison of the behavior of the watertight system with the behavior of the leaky system, two topologically different behaviors, yields three interesting pairs of comparison (figure 4.3(b)). At pc_0 the upper and lower tanks are empty, at pc_1 the upper tanks reach equilibrium, and at pc_2 the lower tanks do. Time-point $\hat{t}_1$, when the water level reaches the leak in the second system, is not included in any of the selected pairs of comparison, because there is no comparable event in the behavior of the first system. The behavior fragment $[\hat{t}_0, \hat{t}_2]$ between pc_0 and pc_1 is an example of a composite behavior fragment.

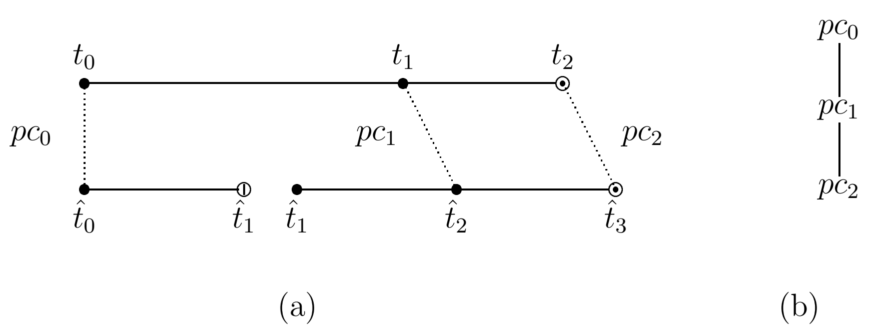


Figure 4.3: *(a) The comparison of a qualitative behavior of a cascaded-tanks system with a watertight upper tank (figure 3.1(c)) and that of a cascaded-tanks system with a leaky upper tank, where the leak is situated at the side of the tank (figure 3.3(c)). (b) The set of meaningful pairs of comparison ordered by the* $\preceq$*-relation.*

4.2.3 Meaningful pairs of comparison

What are meaningful pairs of comparison when performing a comparative analysis? The following definition presents an articulation of the intuitive notion of meaningfulness:

Definition 15 (Meaningful pairs of comparison) Two behaviors are given with sequences of distinguished time-points $t_0, \ldots, t_{n_1}$ and $\hat{t}_0, \ldots, \hat{t}_{n_2}$. The pair of comparison

$pc = \langle t, \hat{t} \rangle$ is a *meaningful* pair of comparison, if it satisfies at least one of the following criteria:

1. t and $\hat{t}$ are the initial time-points t_0 and $\hat{t}_0$;
2. t and $\hat{t}$ are the final time-points t_{n_1} and $\hat{t}_{n_2}$;
3. t and $\hat{t}$ are time-points at which a variable reaches the same basic landmark value 0, ∞, or $-\infty$ in both systems;
4. t and $\hat{t}$ are transition time-points at which the same region transition condition is fulfilled,

and, additionally, pc is not covered by a predecessor pair of comparison.

A pair of comparison pc_2 is *covered* by a predecessor pair of comparison pc_1, if pc_2 has a time-point in common with pc_1 and every criterion for choosing pc_2 as a meaningful pair of comparison is also a criterion for choosing pc_1.

A pair of comparison at which a variable has the same value in both systems is a natural point for assessing differences between the behaviors of the systems, as stipulated by criterion 3.[2] Criterion 4 states that a pair of region transitions brought about by the same event, by satisfaction of the same transition condition, also presents a suitable occasion for comparing the two systems.

The rationale for the requirement that a pair of comparison should not be covered by another pair of comparison is that it filters out superfluous candidates. If one considers a system *for the second time* at a particular time-point, that is, if the time-point is *again* included in a pair of comparison, there should be new reasons for doing so to make the comparison interesting. In the comparative analysis of two mass-spring systems, for instance, one would like to make a comparison at the time-points at which the masses cross the rest position for the first time in both systems; there is not much sense in comparing the time-point at which the first mass crosses the rest position for the first time with the time-point at which the second mass crosses the rest position for the seventh time after it has already been compared with the time-point at which the second mass does so for the first time.

Pairs of comparison pc_0 and pc_2 in figure 4.2 are chosen because of criteria 1 and 2, respectively, in the definition. At pc_1 the netflow n_u of the upper tank reaches 0 in both systems, so criterion 3 is satisfied. There are no coverage relations between these pairs of comparison. Time-point $\hat{t}_1$ in figure 4.3 is omitted from meaningful pairs of comparison, in conformity with our intuition, since no criterion in definition 15 is satisfied for any of the time-points in the first behavior.

[2]The basic landmarks of interest are restricted to 0, ∞, $-\infty$, because these represent the *same numerical value* in the quantity spaces of a particular variable in the first and second system. If one makes the guarantee that some other landmark also has this property, it could be taken into account by criterion 3 as well.

4.2.4 Relative values

Pairs of comparison have been introduced to specify when, at which time-points, the variables of the two systems can be meaningfully compared. Now that the necessary concepts to achieve this have been established, I continue to render the notion of comparison more precise.

A variable may occur in either one of the systems or in both. In the latter case, the variable is said to be *shared* by the systems. As with time-points, the hat accent $\hat{.}$ will be used to denote a shared variable as it occurs in the second system, so a_u and $\hat{a}_u$ represent the amount of water in the upper tank of the first and second cascaded-tanks system. Notice that we compare values of the *same* variable occuring in the first and second system. A more general approach, not pursued here, would be to define a mapping that relates possibly different variables in the two systems (cf. the notion of reformulation function in Weld [1989]).

In order to express how the values of a shared variable p in the first and second system relate at the instants included in a pair of comparison, the concept of *relative value (RV)* of p is introduced. The relative value of p is simply the difference between its values in the first and the second system.

Definition 16 (Relative value) Given a variable p and a pair of comparison $pc = \langle t, \hat{t} \rangle$. The relative value (RV) at pc is defined as:

$$\begin{aligned} p \Uparrow_{pc}, &\ \text{iff } \hat{p}(\hat{t}) > p(t), \\ p \,\|_{pc}, &\ \text{iff } \hat{p}(\hat{t}) = p(t), \\ p \Downarrow_{pc}, &\ \text{iff } \hat{p}(\hat{t}) < p(t). \end{aligned}$$

Often the RV of a variable at a pair of comparison $pc = \langle t, \hat{t} \rangle$ is denoted by the function $RV(p, pc)$, with $RV(p, pc) = \Uparrow$ being equivalent to $p \Uparrow_{pc}$. Remark that a relative value is only defined for variables that are shared by the two systems. A special kind of RV is the *relative duration* of the behavior fragments defined by two successive pairs of comparison. Written as $RV(T, pc_1, pc_2)$, this RV expresses whether the intervening behavior fragment of the second system has a longer, shorter or equal duration compared to that of the first system.

An important difference between this definition and the corresponding definitions in Weld [1990] and de Jong, Mars & van der Vet [1996] is that the (signed) values instead of the absolute values of p and $\hat{p}$ are used. Magnitudes are not well-suited for expressing differences between values of different sign (e.g., $p > 0$ and $\hat{p} < 0$), a situation not occurring in the problems addressed by CEC and Weld's DQ analysis, but certainly expected in inter-model comparative analysis.[3] Neitzke introduces a relative

[3] However, in some situations we would yet like to compare absolute values of variables, for instance when we are interested in speed instead of velocity, or distance instead of change in position. *Relative magnitudes* can be easily accommodated for in definition 16 by writing $|p| \Uparrow_{pc}$ or $RV(|p|)_{pc} = \Uparrow$.

description which normalizes the differences between values of variables with respect to a basic landmark (Neitzke [1992]). These relative descriptions play a central role in his simulation-based approach towards comparative analysis (Neitzke & Neumann [1994]).

4.2.5 A simple algebra for relative values

A few basic (in)equality relations and algebraic operations for RVs will now be introduced. This simple algebra is used in the definition and evaluation of propagation constraints for comparative analysis that are discussed in later sections. The algebra is predicated upon the relation of RVs to real numbers expressed in definition 16. With these semantics the validity of the relations and operations can be easily verified.

Definition 17 (Equality and inequality) For two relative values the equality $RV(p) = RV(q)$ and inequalities $RV(p) < RV(q)$ and $RV(p) > RV(q)$ are defined as follows:

		$RV(q)$		
	$=$	$\Uparrow$	$\|$	$\Downarrow$
	$\Uparrow$	T	F	F
$RV(p)$	$\|$	F	T	F
	$\Downarrow$	F	F	T

		$RV(q)$		
	$<$	$\Uparrow$	$\|$	$\Downarrow$
	$\Uparrow$	T	F	F
$RV(p)$	$\|$	T	F	F
	$\Downarrow$	T	T	T

		$RV(q)$		
	$>$	$\Uparrow$	$\|$	$\Downarrow$
	$\Uparrow$	T	T	T
$RV(p)$	$\|$	F	F	T
	$\Downarrow$	F	F	T

Similar definitions can be given for the relations $\geq$, $\leq$, and $\gtrless$.

Definition 18 (Basic algebraic operations) For two relative values the addition $RV(p) + RV(q)$ and subtraction $RV(p) - RV(q)$ operations are defined as follows:

		$RV(q)$		
	$+$	$\Uparrow$	$\|$	$\Downarrow$
	$\Uparrow$	$\Uparrow$	$\Uparrow$	?
$RV(p)$	$\|$	$\Uparrow$	$\|$	$\Downarrow$
	$\Downarrow$	?	$\Downarrow$	$\Downarrow$

		$RV(q)$		
	-	$\Uparrow$	$\|$	$\Downarrow$
	$\Uparrow$	?	$\Uparrow$	$\Uparrow$
$RV(p)$	$\|$	$\Downarrow$	$\|$	$\Uparrow$
	$\Downarrow$	$\Downarrow$	$\Downarrow$	?

A question mark means that the RV can be $\Uparrow$, $\|$, or $\Downarrow$.

4.2.6 Comparative states and behaviors

Suppose that $\boldsymbol{q}$ represents the vector of variables shared by two systems to be compared. The relative value of a variable in $\boldsymbol{q}$ at *pc* gives a qualitative description of the difference in value of the variable at that pair of comparison. The tuple of relative values of *all* variables in $\boldsymbol{q}$ at *pc* is the *comparative state* of the two systems.

Definition 19 (Comparative state) The shared variables $\boldsymbol{q}$ of two dynamical systems are reasonable functions of time over the domain $[a, b]$ in the first system and $[\hat{a}, \hat{b}]$ in the second system. Given that $\boldsymbol{q}$ has m elements, the comparative state of the systems at pair of comparison $pc = \langle t, \hat{t} \rangle$ is defined by the m-tuple of relative values:

$$CS(\boldsymbol{q}, pc) = \langle RV(q_1, pc), \ldots, RV(q_m, pc) \rangle.$$

Now assume that the shared variables of the systems are compared at the meaningful pairs of comparison $pc_0, \ldots, pc_n$ deduced from definition 15. The *comparative behavior* is then given by a tuple consisting of the comparative states at the pairs of comparison and the relative durations between direct successor pairs of comparison.

Definition 20 (Comparative behavior) The shared variables $\boldsymbol{q}$ of two dynamical systems are reasonable functions of time over $[a, b]$ and $[\hat{a}, \hat{b}]$. The systems are compared at pairs of comparison $pc_0, \ldots, pc_n$, with k direct successor relations implied by the OPC structure. The comparative behavior of the systems is the $n + k$-tuple of comparative states and relative durations:

$$CB(\boldsymbol{q}) = \langle CS(\boldsymbol{q}, pc_0), \ldots, CS(\boldsymbol{q}, pc_n), RV(T, pc_{i_1}, pc_{j_1}), \ldots, RV(T, pc_{i_k}, pc_{j_k}) \rangle,$$

and $pc_{i_1} \preceq pc_{j_1}, \ldots, pc_{i_k} \preceq pc_{j_k}$.

On the face of it, the definitions of relative value, comparative state, and comparative behavior have much in common with the corresponding definitions of qualitative value, qualitative state, and qualitative behavior in QSIM (section 3.2.3). Both QSIM and CEC* view values, states, and behaviors as qualitative abstractions of the reasonable functions representing system variables. Below, and in section 4.5, I will take advantage of this similarity, but it should immediately be added that there is also a fundamental difference in meaning between the corresponding concepts in QSIM and CEC*. This can be illustrated by the notions of qualitative state and comparative state; whereas a qualitative state refers to the state of a physical system, possibly persisting over an interval, a comparative state refers to a *comparison* of the states of *two* systems *at a pair of time-points*.

4.2.7 Comparative analysis as a propagation process

The basic idea underlying qualitative simulation was schematically summarized in the previous chapter by means of figure 3.5. A QDE is an abstraction of an ordinary differential equation (ODE) describing a dynamical system, and qualitative simulation is intended to yield a corresponding abstraction of the solution $\boldsymbol{v} : [a, b] \to \mathbb{R}^*$ of the ODE. In a similar vein, the comparative analysis task can be clarified (figure 4.4). The QDEs and qualitative behaviors are abstractions of the ODEs and the solutions describing the dynamical systems to be compared, and comparative analysis is intended to yield an abstraction of the difference between the solutions at the pairs of comparison. Section 4.5 discusses the abstraction relations in more detail.

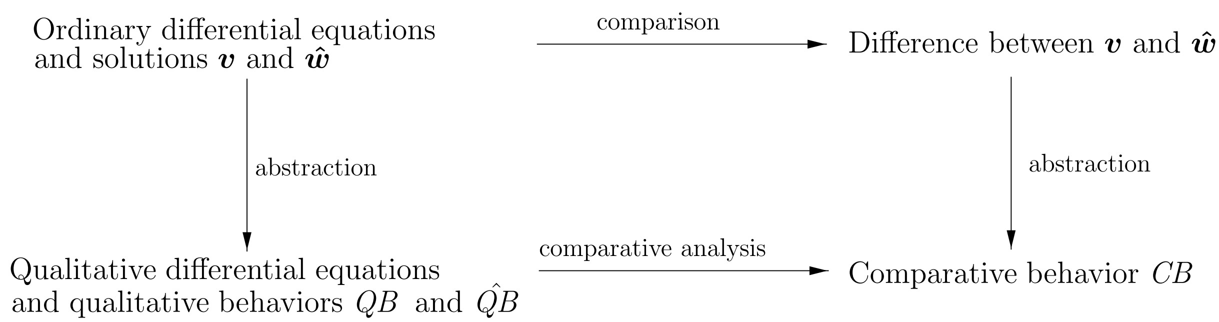


Figure 4.4: *Abstraction relations in comparative analysis.*

By means of the concepts introduced in this section, comparative analysis can be formalized as a propagation process. The input of CEC* consists of a pair of qualitative behaviors, the QDEs giving rise to them, and a set of relative values at either the first or the last pair of comparison, that is, the top or the bottom of the OPC structure. The analysis proceeds by propagating the initial RVs through the OPC structure, from a pair of comparison to its successor(s) or predecessor(s). At each pair of comparison, CEC* generates comparative states consistent with the *propagation* or *RV constraints* derived from the qualitative behaviors and QDEs of the systems to be compared. Analogously to the state constraints and transition constraints in qualitative simulation, CEC* uses propagation constraints relating RVs at the same pair of comparison and propagation constraints relating RVs at successive pairs of comparison. The next section introduces general theorems from which RV constraints can be derived in a particular CA problem under consideration. The output of CEC* consists of a set of possible comparative behaviors connecting the comparative states. The comparative behaviors are contained in a comparative envisionment (section 4.4), and describe in a qualitative manner the differential dynamics of the systems to be compared.

When the direction of propagation is from top to bottom, the algorithm is predicting the consequences of differential initial conditions. The reverse direction corresponds with the explanation of a differential response of the two systems. I will employ the terms *predictive* and *explanatory* comparative analysis, respectively, to refer to these two uses of CEC*. Although convenient and intuitively clear, one should bear in mind that the forward and backward directions of propagation on the one hand, and the tasks of making predictions and giving explanations on the other, do not match completely. An explanation of an observed difference in response of two systems could also be explained by exhaustively generating differential initial conditions and performing a predictive CA for each of them to see whether they account for the differential response.

4.3 Propagation constraints

4.3.1 Variables at a pair of comparison

A direct way to infer the RV of a variable p at a pair of comparison $pc = \langle t, \hat{t} \rangle$ is to examine the state information in the qualitative behaviors of the first and second system. In the QSIM framework, the state information provides one with the qualitative value of the continuously differentiable functions of time p and $\hat{p}$. The qualitative value of a variable p at t is a tuple of the qualitative abstractions of $p(t)$ and $p'(t)$ (definition 5), but since the latter is of less interest to us here, we will often speak of a qualitative value when its qualitative magnitude is intended.

By definition the qualitative value of $p(t)$ is either a landmark value l_j or an open interval between two adjacent landmark values $]l_j, l_{j+1}[$ in which $p(t)$ lies. If we extend the applicability of the relations $>$, $=$, and $<$ to intervals, the following theorem can be formulated:

Theorem 6 (Qualitative values) Let pv and $\hat{pv}$ be qualitative values of p and $\hat{p}$, respectively, at pair of comparison $pc = \langle t, \hat{t} \rangle$. The sign of the difference $\hat{p}(\hat{t}) - p(t)$ is given by the following table:

	$\hat{p}(\hat{t}) - p(t)$	pv: > 0	pv: $= 0$	pv: < 0
$\hat{pv}$	> 0	?	> 0	> 0
	$= 0$	< 0	$= 0$	> 0
	< 0	< 0	< 0	?

Proof. Consider the situation that $pv > 0$ and $\hat{pv} < 0$. pv is either a landmark l_j or an open interval between two landmark values $]l_j, l_{j+1}[$. Similarly, $\hat{pv}$ equals $\hat{l}_k$ or $]\hat{l}_k, \hat{l}_{k+1}[$. $\hat{p}(\hat{t}) - p(t)$ has the value or lies in the interval determined by $\hat{pv} - pv$. If $\hat{pv} < 0$, then $\hat{l}_k$ or the upper bound of $]\hat{l}_k, \hat{l}_{k+1}[$ is < 0. Also, $pv > 0$ implies that l_j or the lower bound of $]l_j, l_{j+1}[$ is > 0. Combining this information, it is easy to see that $\hat{pv} - pv$ is a value < 0 or an interval with an upper bound < 0, and consequently $\hat{p}(\hat{t}) - p(t) < 0$. The proof goes analogously for other values pv and $\hat{pv}$. □

The interest of theorem 6 derives from the fact that knowing the sign of the difference $\hat{p}(\hat{t}) - p(t)$ immediately provides a constraints on the relative value of p at pc, $RV(p, pc)$. For example, in figure 3.1(c) we see that at pc_0 $a_u = 0$ and $\hat{a}_u = 0$, so that $RV(a_u, pc_0) = \|$.

Theorem 6 does not cover all possible combinations of values for p and $\hat{p}$. In particular, it does not help in establishing $RV(p)$ when the values of p and $\hat{p}$ lie 'on the same side' of the basic landmark 0. A consideration of the functional relations in the QDEs provides additional constraints on the value of $\hat{p} - p$ at a pair of comparison.

4.3.2 Functional relations at a pair of comparison

Suppose that a variable p shared by the two systems is defined by a continuously differentiable function of a vector of variables $\boldsymbol{r}$ in the first system and $\boldsymbol{s}$ in the second system:

$$p = f(\boldsymbol{r}) \quad \text{and} \quad \hat{p} = g(\hat{\boldsymbol{s}}). \tag{4.1}$$

The problem is that at a pair of comparison $\langle t, \hat{t} \rangle$, p and $\hat{p}$ cannot be straightforwardly compared, because in general the vectors $\boldsymbol{r}$ and $\boldsymbol{s}$ and the functions f and g will be different. The expressions for p and $\hat{p}$ have to be made comparable. To begin with, define $\boldsymbol{q}$ as the vector of variables occurring both in $\boldsymbol{r}$ and in $\boldsymbol{s}$, and $\boldsymbol{a}$ as the vector of variables occurring either in $\boldsymbol{r}$ or in $\boldsymbol{s}$, but not in both. We can now give a criterion for the comparability of p and $\hat{p}$:

Definition 21 (Comparable functions) Suppose p and $\hat{p}$ are defined by the continuously differentiable functions f and g, as in (4.1). The functions f and g are called comparable with respect to a continuously differentiable function h and vectors $\boldsymbol{a}^{\mathbf{0}}$ and $\hat{\boldsymbol{a}}^{\mathbf{0}}$, if

$$p = h(\boldsymbol{q}, \boldsymbol{a}^{\mathbf{0}}) \quad \text{and} \quad \hat{p} = h(\hat{\boldsymbol{q}}, \hat{\boldsymbol{a}}^{\mathbf{0}}).$$

The vectors $\boldsymbol{a}^{\mathbf{0}}$ and $\hat{\boldsymbol{a}}^{\mathbf{0}}$ are obtained from $\boldsymbol{a}$ and $\hat{\boldsymbol{a}}$ by assignment of specific constant *comparison values* to variables. If a_i does not occur in $\boldsymbol{r}$, then a_i^0 is defined to have a comparison value c_i. In the same way, if $\hat{a}_j$ does not occur in $\hat{\boldsymbol{s}}$, then $\hat{a}_j^0$ has a comparison value $\hat{c}_j$.

An example will clarify this definition. In figure 4.2 two cascaded-tanks behaviors are compared and from the models of the systems we know that:

$$n = i - o_u \quad \text{and} \quad \hat{n} = \hat{i} - \hat{o}_u - \hat{o}_h.$$

Take $\boldsymbol{q} = [i \ o_u]'$ and $\boldsymbol{a} = o_h$, and define a function h as follows: $h([x_1 \ x_2]', x_3) = x_1 - x_2 - x_3$. Further, set $\boldsymbol{a}^{\mathbf{0}} = 0$ and $\hat{\boldsymbol{a}}^{\mathbf{0}} = \hat{o}_h$, that is, assign the comparison value 0 to $\boldsymbol{a}^{\mathbf{0}}$. Now, $n = h([i \ o_u]', 0) = i - o_u$ and $\hat{n} = h([\hat{i} \ \hat{o}_u]', \hat{o}_h) = \hat{i} - \hat{o}_u - \hat{o}_h$, as required by the definition. For standard functions f and g, the function h and the comparison values are usually easy to find.

The comparability of the functions defining p and $\hat{p}$ allows one to formulate the following theorem:

Theorem 7 (Functional relations) The variables p and $\hat{p}$ are defined by the continuously differentiable functions f and g. If f and g are comparable with respect to a continuously differentiable function h, then at pair of comparison $pc = \langle t, \hat{t} \rangle$

$$\hat{p}(\hat{t}) - p(t) = \boldsymbol{d}_{\boldsymbol{q}}(\hat{\boldsymbol{q}}(\hat{t}) - \boldsymbol{q}(t)) + \boldsymbol{d}_{\boldsymbol{a}}(\hat{\boldsymbol{a}}^{\mathbf{0}}(\hat{t}) - \boldsymbol{a}^{\mathbf{0}}(t)). \tag{4.2}$$

$\boldsymbol{d_q}$ and $\boldsymbol{d_a}$ are vectors with

$$d_{q,i} = \frac{\partial}{\partial q_i} h(\bar{\boldsymbol{q}}, \bar{\boldsymbol{a}}^{\mathbf{0}}) \quad \text{and} \quad d_{a,i} = \frac{\partial}{\partial a_i} h(\bar{\boldsymbol{q}}, \bar{\boldsymbol{a}}^{\mathbf{0}}), \tag{4.3}$$

where each $\bar{q}_j$ lies between $q_j(t)$ and $\hat{q}_j(\hat{t})$, and each $\bar{a}_k^0$ between $a_k^0(t)$ and $\hat{a}_k^0(\hat{t})$.

Proof. The comparability of f and g implies, by definition 21:

$$\hat{p}(\hat{t}) - p(t) = h(\hat{\boldsymbol{q}}(\hat{t}), \hat{\boldsymbol{a}}^{\mathbf{0}}(\hat{t})) - h(\boldsymbol{q}(t), \boldsymbol{a}^{\mathbf{0}}(t)),$$

with h continuously differentiable. Suppose that $\boldsymbol{q}$ has n elements and $\boldsymbol{a}$ has m elements. By means of the generalized mean value theorem for differential calculus (Courant [1959], vol. II, ch. 2), the above equation can be written as:

$$\begin{aligned} \hat{p}(\hat{t}) - p(t) = {} & \frac{\partial}{\partial q_1} h(\bar{\boldsymbol{q}}, \bar{\boldsymbol{a}}^{\mathbf{0}})(\hat{q}_1(\hat{t}) - q_1(t)) + \ldots + \frac{\partial}{\partial q_n} h(\bar{\boldsymbol{q}}, \bar{\boldsymbol{a}}^{\mathbf{0}})(\hat{q}_n(\hat{t}) - q_n(t)) \\ & + \frac{\partial}{\partial a_1} h(\bar{\boldsymbol{q}}, \bar{\boldsymbol{a}}^{\mathbf{0}})(\hat{a}_1^0(\hat{t}) - a_1^0(t)) + \ldots + \frac{\partial}{\partial a_m} h(\bar{\boldsymbol{q}}, \bar{\boldsymbol{a}}^{\mathbf{0}})(\hat{a}_m^0(\hat{t}) - a_m^0(t)), \end{aligned}$$

where each $\bar{q}_j$ lies between $q_j(t)$ and $\hat{q}_j(\hat{t})$ and each $\bar{a}_k^0$ between $a_k^0(t)$ and $\hat{a}_k^0(\hat{t})$. As can be readily verified, this is equivalent to equations (4.2) and (4.3). □

By this theorem, the functional relation in the cascaded-tanks example at $pc_1 = \langle t_1, \hat{t}_1 \rangle$ can be written as:

$$\hat{n}(\hat{t}_1) - n(t_1) = (\hat{i}(\hat{t}_1) - i(t_1)) - (\hat{o}_u(\hat{t}_1) - o_u(t_1)) - (\hat{o}_h(\hat{t}_1) - 0),$$

with all partial derivatives equal to 1. Since we know from the qualitative behavior that $\hat{o}_h(\hat{t}_1) > 0$ the equality can be reformulated as an inequality:

$$\hat{n}(\hat{t}_1) - n(t_1) < (\hat{i}(\hat{t}_1) - i(t_1)) - (\hat{o}_u(\hat{t}_1) - o_u(t_1)).$$

The advantage of this reformulation is clear. One can now express the difference $\hat{n} - n$ at pc_1 in terms of the differences $\hat{i} - i$ and $\hat{o}_u - o_u$ at the same pair of comparison. In other words,

$$RV(n, pc_1) < RV(i, pc_1) - RV(o_u, pc_1).$$

By means of the simple algebra in section 4.2.5, this RV constraint can be evaluated. For instance, $i \,\|_{pc_1}$ and $o_u \Uparrow_{pc_1}$ make $n \Downarrow_{pc_1}$.

In a similar way, the model equations

$$o_u = f(a_u, r_u) \quad \text{and} \quad \hat{o}_u = f(\hat{a}_u, \hat{r}_u)$$

lead to the following RV constraint at pc_0:

$$RV(o_u, pc_0) = RV(a_u, pc_0) + RV(r_u, pc_0).$$

An equal amount of water in the upper tank ($a_u \|_{pc_0}$) and larger size of the orifice ($r_u \Uparrow_{pc_0}$) lead to a higher outflow ($o_u \Uparrow_{pc_0}$).

Theorem 7 provides the key to the formulation of propagation constraints from functional relations by enabling one to express the difference in the dependent variable p in terms of the difference in the shared independent variables $\boldsymbol{q}$. If the sign of each $\hat{a}_j^0(\hat{t}) - a_j^0(t)$ is known, equation (4.2) reduces to an (in)equality involving only the differences $\hat{p}(\hat{t}) - p(t)$ and $\hat{\boldsymbol{q}}(\hat{t}) - \boldsymbol{q}(t)$. Since either $\hat{a}_j^0(\hat{t})$ or $a_j^0(t)$ has a constant comparison value c_j (by definition 21), the sign of $\hat{a}_j^0(\hat{t}) - a_j^0(t)$ will always be known if the comparison value is included as a landmark in the quantity space of a_j. The sign of the partial derivatives $\boldsymbol{d_q}$, $\boldsymbol{d_a}$ can be determined by inspection of the QDEs of the systems to be compared.

The generality of theorem 7 can be estimated by seeing how it is equally applicable to intra-model and inter-model CA problems. In the former case, the qualitative models of the two systems, and thus all corresponding functional relations, are identical. This makes the functional relations comparable from the outset. The vector $\boldsymbol{a}$ will always be the null vector $\mathbf{0}$, which simplifies equation (4.2) to:

$$\hat{p}(\hat{t}) - p(t) = \boldsymbol{d_q}(\hat{\boldsymbol{q}}(\hat{t}) - \boldsymbol{q}(t)).$$

This will not usually be the case, however, when we are dealing with inter-model CA problems.

4.3.3 State variables between pairs of comparison

Behavior fragments

The question one would like to answer is how the RVs of variables at a pair of comparison pc_1 relate to those at a predecessor pair of comparison pc_0. For this, we have to take into account how variables of the systems change during the respective behavior fragments between the pairs of comparison. First a framework for comparison across intervals is advanced which will then be elaborated by means of linear system theory.

We will assume that the behavior fragments start simultaneously, i.e. $t_0 = \hat{t}_0$. If this is not the case, then the behavior fragments should be synchronized first, by shifting the behavior fragment of the second system in such a way that both fragments start at t_0.

Consider the primitive behavior fragments of figure 4.2(a), for which we have $pc_0 = \langle t_0, \hat{t}_0 \rangle$ and $pc_1 = \langle t_1, \hat{t}_1 \rangle$. $T = t_1 - t_0$ and $\hat{T} = \hat{t}_1 - \hat{t}_0$ represent the *durations* of the two behavior fragments. For the comparison of composite behavior fragments the definition of durations is analogous. The composite behavior fragment of the second system in figure 4.3, with $pc_0 = \langle t_0, \hat{t}_0 \rangle$ and $pc_1 = \langle t_1, \hat{t}_2 \rangle$, has a duration $\hat{T} = \hat{t}_2 - \hat{t}_0$ made up of the durations of its two composing sub-intervals, $\hat{t}_2 - \hat{t}_1$ and $\hat{t}_1 - \hat{t}_0$.

We distinguish several cases in the comparison of two behavior fragments by analyzing how their durations relate. Figure 4.5(a) shows three different situations: $\hat{T} < T$, $\hat{T} = T$, or $\hat{T} > T$. For the comparison of a primitive behavior fragment with a behavior fragment composed of two sub-intervals, we obtain the five cases in figure 4.5(b): $\hat{T} < T$, $\hat{T} = T$,

$\hat{T} > T$ and $\hat{t}_1 < t_1$, $\hat{T} > T$ and $\hat{t}_1 = t_1$, or $\hat{T} > T$ and $\hat{t}_1 > t_1$. The comparison of other pairs of behavior fragments will lead to classifications that are different from those displayed in figure 4.5, but obtainable through the same principle of systematically enumerating the possible relations between the durations of (sub-)intervals.

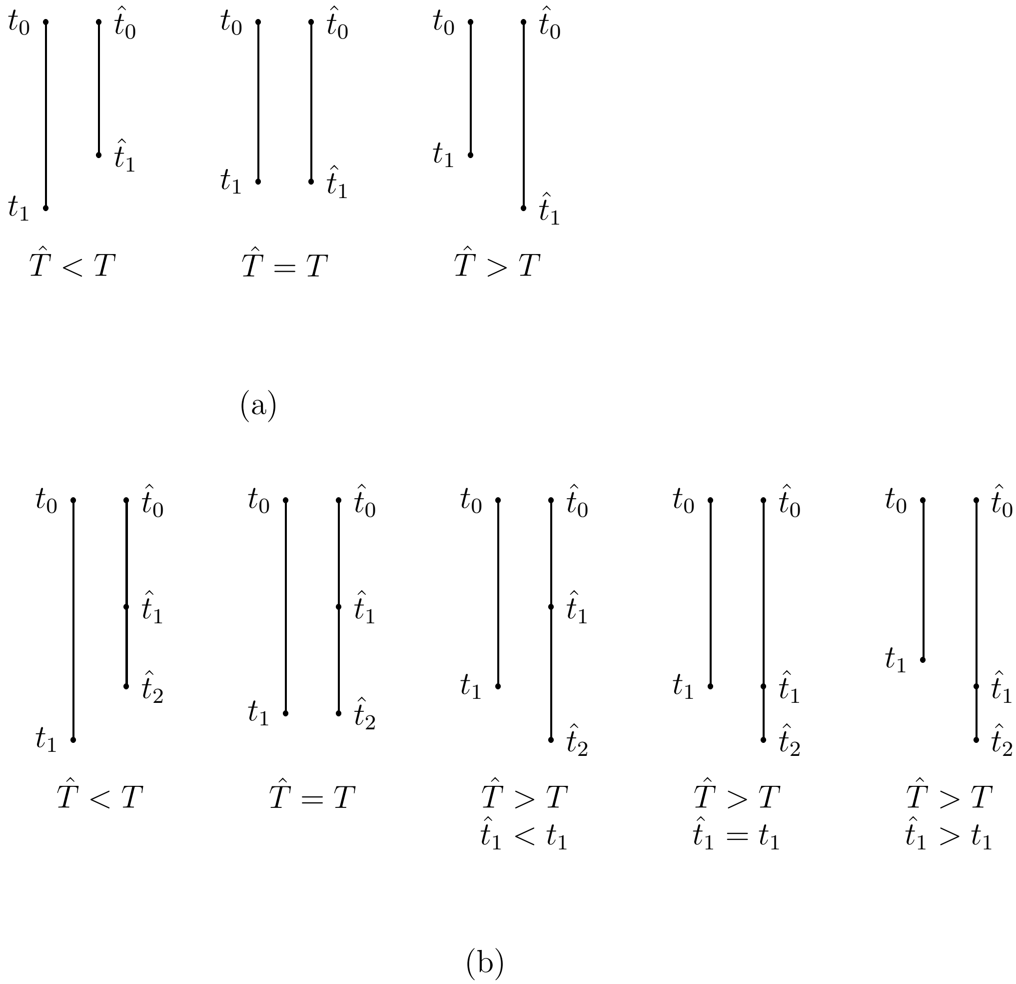


Figure 4.5: *Different cases in the comparison of two behavior fragments. In (a) two primitive behavior fragments and in (b) a primitive and a composite behavior fragment are compared.*

Figure 4.5 suggests a way to rewrite $\hat{x}(\hat{t}_1) - x(t_1)$ when comparing two primitive behavior fragments:

$$\hat{x}(\hat{t}_1) - x(t_1) = \begin{cases} x(\hat{t}_1) - x(t_1) + \hat{x}(\hat{t}_1) - x(\hat{t}_1) & , \hat{T} < T, \\ \hat{x}(\hat{t}_1) - x(\hat{t}_1) & , \hat{T} = T, \\ \hat{x}(\hat{t}_1) - \hat{x}(t_1) + \hat{x}(t_1) - x(t_1) & , \hat{T} > T. \end{cases} \tag{4.4}$$

In an analogous fashion, $\hat{x}(\hat{t}_2) - x(t_1)$ can be rewritten when comparing a primitive behavior fragment with a composite one.

$$\hat{x}(\hat{t}_2) - x(t_1) = \begin{cases} x(\hat{t}_2) - x(t_1) + \hat{x}(\hat{t}_2) - x(\hat{t}_2) & , \hat{T} < T, \\ \hat{x}(\hat{t}_2) - x(\hat{t}_2) & , \hat{T} = T, \\ \hat{x}(\hat{t}_2) - \hat{x}(t_1) + \hat{x}(t_1) - x(t_1) & , \hat{T} > T \text{ and } \hat{t}_1 < t_1, \\ \hat{x}(\hat{t}_2) - \hat{x}(t_1) + \hat{x}(t_1) - x(t_1) & , \hat{T} > T \text{ and } \hat{t}_1 = t_1, \\ \hat{x}(\hat{t}_2) - \hat{x}(\hat{t}_1) + \hat{x}(\hat{t}_1) - \hat{x}(t_1) + \hat{x}(t_1) - x(t_1) & , \hat{T} > T \text{ and } \hat{t}_1 > t_1. \end{cases} \tag{4.5}$$

Consider the case $\hat{T} < T$ in (4.4). The difference $\hat{x}(\hat{t}_1) - x(t_1)$ is split up into two components: $x(\hat{t}_1) - x(t_1)$ and $\hat{x}(\hat{t}_1) - x(\hat{t}_1)$. The sign of the former term directly follows from the qualitative behavior of the system, whereas the sign of the latter term must be determined. The behavior of $\hat{x} - x$ on $[\hat{t}_0, \hat{t}_1]$ (or $[t_0, \hat{t}_1]$ since $t_0 = \hat{t}_0$ by synchronicity of the behavior fragments) is determined by the QDEs of the two systems.

Analogously, the case $\hat{T} < T$ in (4.5) can be analyzed. Again $\hat{x}(\hat{t}_2) - x(t_1)$ is split up into two components $x(\hat{t}_2) - x(t_1)$ and $\hat{x}(\hat{t}_2) - x(\hat{t}_2)$, and the sign of the former term is known from the qualitative behavior of the system. This time however, we cannot consider the behavior of $\hat{x} - x$ on $[\hat{t}_0, \hat{t}_2]$ in one stroke. We must distinguish between $\hat{x} - x$ on $[\hat{t}_0, \hat{t}_1]$ and $[\hat{t}_1, \hat{t}_2]$, for the QDE of the second system may change at $\hat{t}_1$. This somewhat complicates the analysis, as will be seen below.

Comparison systems

For the general problem of establishing the difference $\hat{x}(t) - x(t)$ in terms of the difference of variables at an initial time-point t_a, we have to study the QDEs which define the dynamical behavior of the two systems. First, we introduce three classes of variables: shared state variables $\boldsymbol{x}$, shared input variables $\boldsymbol{u}$, and auxiliary variables $\boldsymbol{a}$. A variable v is (1) a *shared state* variable if it is shared and a state variable in at least one of the systems,[4] (2) a *shared input* variable if it is shared and an input variable in at least one of the systems, or (3) an *auxiliary* variable if it is not shared. As usual, a hat accent refers to variables of the second system. Thus, $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{u}}$, and $\hat{\boldsymbol{a}}$ represent the shared state, shared input, and auxiliary variables for the second system.

From the QDEs of the two systems one can derive the state equations:

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{f_1}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{a}(t)) \quad \text{and} \quad \frac{d\hat{\boldsymbol{x}}(t)}{dt} = \boldsymbol{f_2}(\hat{\boldsymbol{x}}(t), \hat{\boldsymbol{u}}(t), \hat{\boldsymbol{a}}(t)). \tag{4.6}$$

These transformations bring the models of the two systems into a form in which their dynamics can be more easily compared. This comparison is effected in a so-called *comparison system*:

Definition 22 (Comparison system) The comparison system for two systems whose state equations have been brought into the form (4.6) is a system whose dynamics is defined by:

[4] For our purpose, state variables are variables whose derivatives are specified in the QDE.

$$\frac{d\hat{\boldsymbol{x}}(t)}{dt} - \frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{f_2}(\hat{\boldsymbol{x}}(t), \hat{\boldsymbol{u}}(t), \hat{\boldsymbol{a}}(t)) - \boldsymbol{f_1}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{a}(t)). \tag{4.7}$$

What do we gain by introducing this concept? The following theorem shows how a comparison system can be transformed into a more attractive form: a linear comparison system.

Theorem 8 (Linear comparison system) Consider the comparison system of definition 22. If the functions $\boldsymbol{f_1}$ and $\boldsymbol{f_2}$ are comparable through a function $\boldsymbol{h}$, such that $\boldsymbol{f_1}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{a}) = \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{a^0})$ and $\boldsymbol{f_2}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{a}}) = \boldsymbol{h}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{a}}^{\boldsymbol{0}})$, then the state equation of the comparison system can be brought into the following linear form:

$$\frac{d}{dt}(\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t)) = \boldsymbol{A}(t)(\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t)) + \boldsymbol{B}(t)(\hat{\boldsymbol{u}}(t) - \boldsymbol{u}(t)) + \boldsymbol{E}(t)(\hat{\boldsymbol{a}}^{\boldsymbol{0}}(t) - \boldsymbol{a^0}(t)). \tag{4.8}$$

$\boldsymbol{A}(t)$, the *state matrix*, $\boldsymbol{B}(t)$, the *input matrix*, and $\boldsymbol{E}(t)$, the *auxiliary matrix*, are time-varying matrices of continuous functions:

$$a_{ij}(t) = \frac{\partial}{\partial x_j} h_i(\bar{\boldsymbol{x}}(t), \bar{\boldsymbol{u}}(t), \bar{\boldsymbol{a}}^{\boldsymbol{0}}(t)), \quad b_{ik}(t) = \frac{\partial}{\partial u_k} h_i(\bar{\boldsymbol{x}}(t), \bar{\boldsymbol{u}}(t), \bar{\boldsymbol{a}}^{\boldsymbol{0}}(t)),$$
$$e_{il}(t) = \frac{\partial}{\partial a_l} h_i(\bar{\boldsymbol{x}}(t), \bar{\boldsymbol{u}}(t), \bar{\boldsymbol{a}}^{\boldsymbol{0}}(t)) \tag{4.9}$$

$\bar{x}_j(t)$ lies between $\hat{x}_j(t)$ and $x_j(t)$, $\bar{u}_k(t)$ between $\hat{u}_k(t)$, and $u_k(t)$, and $\bar{a}_l^0(t)$ between $\hat{a}_l^0(t)$ and $a_l^0(t)$. The number of rows in $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$, and $\boldsymbol{E}(t)$ equals the number of shared state variables; the number of columns equals the number of shared state, shared input, and auxiliary variables, respectively.

The validity of the theorem follows directly from the definition of comparable functions and the application of theorem 7 to each h_i.

We will study linear comparison systems on finite closed intervals $[t_a, t_b]$ in which neither of the original systems has a distinguished time-point; consequently, qualitative values of variables do not change in $]t_a, t_b[$. From the qualitative behavior of the systems to be compared we know the sign of $\boldsymbol{x}$, $\hat{\boldsymbol{x}}$ and $\boldsymbol{u}$, $\hat{\boldsymbol{u}}$ and $\boldsymbol{a^0}$, $\hat{\boldsymbol{a}}^{\boldsymbol{0}}$ on $[t_a, t_b]$. This information may allow one to derive the sign of $\bar{\boldsymbol{x}}$, $\bar{\boldsymbol{u}}$, and $\bar{\boldsymbol{a}}^{\boldsymbol{0}}$ on $[t_a, t_b]$. Together with the QDEs that determine h, the sign of the matrix elements $\partial h_i/\partial x_j$, $\partial h_i/\partial u_k$, and $\partial h_i/\partial a_l$ on $[t_a, t_b]$ can now be established. Notice that the sign of the elements of $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$, $\boldsymbol{E}(t)$ is *not* necessarily known over the interval. For one reason, if $v > 0$ and $\hat{v} < 0$ on $[t_a, t_b]$, then the sign of $\bar{v}$, lying between v and $\hat{v}$, is indeterminate.

As an example, consider the comparison of watertight and leaky cascaded tanks, with the leak located at the bottom. The relevant vectors of variables in this case are $\boldsymbol{x} = [a_u \ a_l]'$, $\boldsymbol{u} = [i \ r_u \ r_l]'$, and $\boldsymbol{a} = o_h$. By means of the QDEs in figures 3.1(b) and 4.1(b), one can determine the following matrices in the state equation of linear comparison system (4.8):

$$\boldsymbol{A}(t) = \begin{bmatrix} -\frac{\partial}{\partial a_u} f(\bar{a}_u(t), \bar{r}_u(t)) & 0 \\ \frac{\partial}{\partial a_u} f(\bar{a}_u(t), \bar{r}_u(t)) & -\frac{\partial}{\partial a_l} g(\bar{a}_l(t), \bar{r}_l(t)) \end{bmatrix}, \ \boldsymbol{E}(t) = \begin{bmatrix} -1 \\ 0 \end{bmatrix},$$

$$\boldsymbol{B}(t) = \begin{bmatrix} 1 & -\frac{\partial}{\partial r_u} f(\bar{a}_u(t), \bar{r}_u(t)) & 0 \\ 0 & \frac{\partial}{\partial r_u} f(\bar{a}_u(t), \bar{r}_u(t)) & -\frac{\partial}{\partial r_l} g(\bar{a}_l(t), \bar{r}_l(t)) \end{bmatrix}, \quad (4.10)$$

The signs of the matrix elements are directly determinable in this case, independently from the interval on which we study the comparison system, since the partial derivatives of f and g are known to be positive (figures 3.1(b) and 4.1(b)).

Again, it is useful to point out how intra-model CA is treated as a special case of inter-model CA. The comparison of two identical cascaded-tanks systems with a watertight upper tank leads to $\boldsymbol{a} = \boldsymbol{0}$, since all variables are shared in a structurally identical system. This implies that the term $\boldsymbol{E}(t)(\hat{\boldsymbol{a}}^{\boldsymbol{0}}(t) - \boldsymbol{a}^{\boldsymbol{0}}(t))$ drops out of (4.8).

Solving comparison systems

The formulation of the comparison system as a linear system proves to be the pivotal move in deriving RV constraints between pairs of comparison, since an explicit solution of state equation (4.8) is a basic result in linear system theory (e.g., Rugh [1996]; Chen [1970]). Because the form of (4.8) deviates somewhat from that usually found in textbooks, the form of the solution is also slightly different. The solution given here is easily proven sound by reducing (4.8) to the conventional form.

Suppose that the difference in shared input variables $\hat{\boldsymbol{u}} - \boldsymbol{u}$ and the difference in auxiliary variables $\hat{\boldsymbol{a}}^{\boldsymbol{0}} - \boldsymbol{a}^{\boldsymbol{0}}$ is continuous on $[t_a, t]$, $t > t_a$, and the matrix elements are continuous on the same interval, then equation (4.8) has the unique, continuously differentiable solution

$$\begin{aligned} \hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t) &= \boldsymbol{\Phi}(t, t_a)(\hat{\boldsymbol{x}}(t_a) - \boldsymbol{x}(t_a)) \\ &\quad + \int_{t_a}^{t} \boldsymbol{\Phi}(t, \tau) \left[\boldsymbol{B}(\tau)(\hat{\boldsymbol{u}}(\tau) - \boldsymbol{u}(\tau)) + \boldsymbol{E}(\tau)(\hat{\boldsymbol{a}}^{\boldsymbol{0}}(\tau) - \boldsymbol{a}^{\boldsymbol{0}}(\tau)) \right] d\tau \end{aligned} \quad (4.11)$$

If additionally shared input variables are assumed to be constant during $[t_a, t]$, then $\hat{\boldsymbol{u}}(\tau) - \boldsymbol{u}(\tau) = \hat{\boldsymbol{u}}(t_a) - \boldsymbol{u}(t_a)$, and (4.11) becomes:[5]

[5] This condition is often enforced by model-builders in qualitative reasoning, like QPE (Forbus [1989]) and QPC (Farquhar [1994]). In the case that an input variable is constant in only one of the physical systems to be compared, additional precautions have to be taken in the formulation of the comparison system.

$$\begin{aligned}\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t) &= \boldsymbol{\Phi}(t, t_a)(\hat{\boldsymbol{x}}(t_a) - \boldsymbol{x}(t_a)) \\ &\quad + \int_{t_a}^{t} \boldsymbol{\Phi}(t, \tau) \left[\boldsymbol{B}(\tau)(\hat{\boldsymbol{u}}(t_a) - \boldsymbol{u}(t_a)) + \boldsymbol{E}(\tau)(\hat{\boldsymbol{a}}^{\mathbf{0}}(\tau) - \boldsymbol{a}^{\mathbf{0}}(\tau))\right] d\tau \end{aligned} \tag{4.12}$$

The matrix function $\boldsymbol{\Phi}(t, \tau)$ is called the *transition matrix*. Transition matrices form the topic of the next section.

Now return to the comparison of two synchronous behavior fragments, one of the first system and one of the second. If the two behavior fragments are primitive, equation (4.4) provides an expression for the difference $\hat{\boldsymbol{x}}(\hat{t}_1) - \boldsymbol{x}(t_1)$. The general solution of a linear comparison system over $[t_a, t]$ helps us in elaborating this expression.

Theorem 9 (State variables I) Suppose that two primitive behavior fragments are compared. The difference between the shared state variables $\boldsymbol{x}$ at $pc_1 = \langle t_1, \hat{t}_1 \rangle$ can be expressed in terms of differences at $pc_0 = \langle t_0, \hat{t}_0 \rangle$ as:

$$\hat{\boldsymbol{x}}(\hat{t}_1) - \boldsymbol{x}(t_1) = \begin{cases} \boldsymbol{x}(\hat{t}_1) - \boldsymbol{x}(t_1) + \boldsymbol{F}(\hat{t}_1, \hat{t}_0) & , \hat{T} < T, \\ \boldsymbol{F}(t_1, t_0) & , \hat{T} = T, \\ \hat{\boldsymbol{x}}(\hat{t}_1) - \hat{\boldsymbol{x}}(t_1) + \boldsymbol{F}(t_1, t_0) & , \hat{T} > T, \end{cases} \tag{4.13}$$

where the function $\boldsymbol{F}$ is defined as follows:

$$\begin{aligned}\boldsymbol{F}(t_b, t_a) &= \boldsymbol{\Phi}(t_b, t_a)(\hat{\boldsymbol{x}}(t_a) - \boldsymbol{x}(t_a)) \\ &\quad + \int_{t_a}^{t_b} \boldsymbol{\Phi}(t_b, \tau) \left[\boldsymbol{B}(\tau)(\hat{\boldsymbol{u}}(t_a) - \boldsymbol{u}(t_a)) + \boldsymbol{E}(\tau)(\hat{\boldsymbol{a}}^{\mathbf{0}}(\tau) - \boldsymbol{a}^{\mathbf{0}}(\tau))\right] d\tau. \end{aligned} \tag{4.14}$$

Proof. Consider the case $\hat{T} < T$, for which we have by (4.4): $\boldsymbol{x}(\hat{t}_1) - \boldsymbol{x}(t_1) + \hat{\boldsymbol{x}}(\hat{t}_1) - \boldsymbol{x}(\hat{t}_1)$. The term $\hat{\boldsymbol{x}}(\hat{t}_1) - \boldsymbol{x}(\hat{t}_1)$ is elaborated by means of first constructing a linear comparison system for $\hat{\boldsymbol{x}} - \boldsymbol{x}$, $\hat{\boldsymbol{u}} - \boldsymbol{u}$, and $\hat{\boldsymbol{a}}^{\mathbf{0}} - \boldsymbol{a}^{\mathbf{0}}$, and then finding a solution for this system on $[\hat{t}_0, \hat{t}_1]$ (or, equivalently, on $[t_0, \hat{t}_1]$). By assumption, $\hat{\boldsymbol{u}} - \boldsymbol{u}$ and $\hat{\boldsymbol{a}}^{\mathbf{0}} - \boldsymbol{a}^{\mathbf{0}}$ are continuous on $[\hat{t}_0, \hat{t}_1]$, and $\hat{\boldsymbol{u}}$ and $\boldsymbol{u}$ constant, so that we can apply (4.12) with $t_a = \hat{t}_0$ and $t = \hat{t}_1$, which gives $\boldsymbol{F}(\hat{t}_1, \hat{t}_0)$. □

Comparing a primitive behavior fragment of the first system with a composite behavior fragment of the second one, as shown in figure 4.5(b), turns out to be an obvious generalization of the comparison of two primitive behavior fragments. The main difference is that in some cases we now need to construct *two* comparison systems instead of one.

Consider the situation $\hat{T} < T$ in the figure. Apart from a comparison system for $[\hat{t}_0, \hat{t}_1]$, a comparison system for $[\hat{t}_1, \hat{t}_2]$ is needed. This second comparison system possibly deviates from the first, because $\hat{t}_1$ is a distinguished time-point at which a region transition can occur. For instance, the water in the upper tank of the leaky cascaded-tanks system may reach the level of the leak and initiate an extra outflow. In order to simplify the analysis, restrictions are put upon the change in model occurring at $\hat{t}_1$. In particular, it is assumed that the shared state variables of the comparison system for $[\hat{t}_0, \hat{t}_1]$ equal those for $[\hat{t}_1, \hat{t}_2]$, that is, $\boldsymbol{x} = \boldsymbol{x_1} = \boldsymbol{x_2}$.

To find the difference in shared state variables at $\hat{t}_2$, the difference between the shared state variables obtained from solving the comparison system for the first sub-interval $[\hat{t}_0, \hat{t}_1]$ is used as initial value when solving the comparison system defined for the second sub-interval $[\hat{t}_1, \hat{t}_2]$. In other words, we consecutively solve the state equation of the two comparison systems. More formally stated, this leads to the following theorem.

Theorem 10 (State variables II) Suppose that a primitive behavior fragment is compared with a composite behavior fragment. The difference between the shared state variables $\boldsymbol{x}$ at $pc_1 = \langle t_1, \hat{t}_2 \rangle$ can be expressed in terms of differences at $pc_0 = \langle t_0, \hat{t}_0 \rangle$ as:

$$\hat{\boldsymbol{x}}(\hat{t}_2) - \boldsymbol{x}(t_1) = \begin{cases} \boldsymbol{x}(\hat{t}_2) - \boldsymbol{x}(t_1) + \boldsymbol{G}(\hat{t}_2, \hat{t}_1, \hat{t}_0) & , \hat{T} < T, \\ \boldsymbol{G}(t_1, \hat{t}_1, \hat{t}_0) & , \hat{T} = T, \\ \hat{\boldsymbol{x}}(\hat{t}_2) - \hat{\boldsymbol{x}}(t_1) + \boldsymbol{G}(t_1, \hat{t}_1, \hat{t}_0) & , \hat{T} > T \text{ and } \hat{t}_1 < t_1, \\ \hat{\boldsymbol{x}}(\hat{t}_2) - \hat{\boldsymbol{x}}(t_1) + \boldsymbol{F}(t_1, t_0) & , \hat{T} > T \text{ and } \hat{t}_1 = t_1, \\ \hat{\boldsymbol{x}}(\hat{t}_2) - \hat{\boldsymbol{x}}(\hat{t}_1) + \hat{\boldsymbol{x}}(\hat{t}_1) - \hat{\boldsymbol{x}}(t_1) + \boldsymbol{F}(t_1, t_0) & , \hat{T} > T \text{ and } \hat{t}_1 > t_1, \end{cases} \tag{4.15}$$

where the functions $\boldsymbol{F}$ and $\boldsymbol{G}$ are defined as follows:

$$\begin{aligned} \boldsymbol{F}(t_b, t_a) = {} & \boldsymbol{\Phi_1}(t_b, t_a)(\hat{\boldsymbol{x}}(t_a) - \boldsymbol{x}(t_a)) \\ & + \int_{t_a}^{t_b} \boldsymbol{\Phi_1}(t_b, \tau) \left[\boldsymbol{B_1}(\tau)(\hat{\boldsymbol{u}}_{\boldsymbol{1}}(t_a) - \boldsymbol{u_1}(t_a)) + \boldsymbol{E_1}(\tau)(\hat{\boldsymbol{a}}^{\boldsymbol{0}}{}_{\boldsymbol{1}}(\tau) - \boldsymbol{a}^{\boldsymbol{0}}_{\boldsymbol{1}}(\tau)) \right] d\tau, \end{aligned} \tag{4.16}$$

$$\begin{aligned} \boldsymbol{G}(t_c, t_b, t_a) = {} & \boldsymbol{\Phi_2}(t_c, t_b)\boldsymbol{F}(t_b, t_a) \\ & + \int_{t_b}^{t_c} \boldsymbol{\Phi_2}(t_c, \tau) \left[\boldsymbol{B_2}(\tau)(\hat{\boldsymbol{u}}_{\boldsymbol{2}}(t_b) - \boldsymbol{u_2}(t_b)) + \boldsymbol{E_2}(\tau)(\hat{\boldsymbol{a}}^{\boldsymbol{0}}{}_{\boldsymbol{2}}(\tau) - \boldsymbol{a}^{\boldsymbol{0}}_{\boldsymbol{2}}(\tau)) \right] d\tau, \end{aligned} \tag{4.17}$$

with $\boldsymbol{u_1}$, $\boldsymbol{a}^{\boldsymbol{0}}_{\boldsymbol{1}}$, $\boldsymbol{B_1}$, $\boldsymbol{E_1}$ defined by the comparison system over the first sub-interval, and $\boldsymbol{u_2}$, $\boldsymbol{a}^{\boldsymbol{0}}_{\boldsymbol{2}}$, $\boldsymbol{B_2}$, $\boldsymbol{E_2}$ by the comparison system over the second sub-interval. $\boldsymbol{\Phi_1}$ and $\boldsymbol{\Phi_2}$ are the transition matrices for these comparison systems.

Proof. Consider the case $\hat{T} < T$ for which we have by (4.5): $\boldsymbol{x}(\hat{t}_2) - \boldsymbol{x}(t_1) + \hat{\boldsymbol{x}}(\hat{t}_2) - \boldsymbol{x}(\hat{t}_2)$. For $[\hat{t}_0, \hat{t}_2]$ we can construct two linear comparison systems, one on $[\hat{t}_0, \hat{t}_1]$ and one on $[\hat{t}_1, \hat{t}_2]$. $\hat{\boldsymbol{x}}(\hat{t}_1) - \boldsymbol{x}(\hat{t}_1)$ is given by the solution of the comparison system on $[\hat{t}_0, \hat{t}_1]$, and is used as the initial value for the solution of the comparison system on $[\hat{t}_1, \hat{t}_2]$. This yields $\boldsymbol{G}(\hat{t}_2, \hat{t}_1, \hat{t}_0)$ after recursively applying (4.12). $\hat{\boldsymbol{u}}_{\boldsymbol{1}} - \boldsymbol{u_1}$ and $\hat{\boldsymbol{a}}^{\boldsymbol{0}}{}_{\boldsymbol{1}} - \boldsymbol{a}^{\boldsymbol{0}}_{\boldsymbol{1}}$ are continuous on $[\hat{t}_0, \hat{t}_1]$ and $\hat{\boldsymbol{u}}_{\boldsymbol{2}} - \boldsymbol{u_2}$ and $\hat{\boldsymbol{a}}^{\boldsymbol{0}}{}_{\boldsymbol{2}} - \boldsymbol{a}^{\boldsymbol{0}}_{\boldsymbol{2}}$ on $[\hat{t}_1, \hat{t}_2]$, while $\hat{\boldsymbol{u}}_{\boldsymbol{1}} - \boldsymbol{u_1}$ and $\hat{\boldsymbol{u}}_{\boldsymbol{2}} - \boldsymbol{u_2}$ are constant. □

Theorem 10 is a proper generalization of theorem 9, since it reduces to the latter when setting $\hat{t}_2 = \hat{t}_1$. The generalization to other pairs of behavior fragments follows the same principle of repeatedly constructing and solving a linear comparison system. Although not difficult to achieve, the resulting expressions may become cumbersome for composite behavior fragments encompassing several distinguished time-points. (They were not necessary for dealing with the examples presented here and in de Jong [1996]).

Notice that by theorems 9 and 10 the difference in shared state variables is propagated from pc_0 to pc_1. In order to derive the difference in the other shared variables at pc_1, theorems 6 and 7 need to be used.

Determination of transition matrix

The usefulness of theorems 9 and 10 hinges on the possibility to determine the transition matrix $\mathbf{\Phi}(t,\tau)$ for all τ in $[t_a, t]$. From linear system theory (Chen [1970]) we know that this matrix is the unique solution of

$$\frac{\partial}{\partial t}\mathbf{\Phi}(t,\tau) = \boldsymbol{A}(t)\mathbf{\Phi}(t,\tau), \quad \mathbf{\Phi}(\tau,\tau) = \boldsymbol{I}.$$

Explicitly solving for $\mathbf{\Phi}(t,\tau)$ is generally a difficult task, but under certain conditions, imposed on $\boldsymbol{A}(t)$, a closed-form expression for $\mathbf{\Phi}(t,\tau)$ can be found. Appendix B reviews some relevant results from linear system theory and shows that when comparing a watertight and a leaky cascaded-tanks system, with $\boldsymbol{A}(t)$ given by (4.10), we have

$$\mathbf{\Phi}(t,\tau) = \begin{bmatrix} \phi_{11}(t,\tau) & 0 \\ \int_\tau^t \phi_{22}(t,\sigma)\frac{\partial}{\partial a_u} f(\bar{a}_u(\sigma), \bar{r}_u(\sigma))\phi_{11}(\sigma,\tau)d\sigma & \phi_{22}(t,\tau) \end{bmatrix}, \tag{4.18}$$

with

$$\begin{aligned} \phi_{11}(t,\tau) &= \exp\left(\int_\tau^t -\frac{\partial}{\partial a_u} f(\bar{a}_u(\sigma), \bar{r}_u(\sigma))d\sigma\right), \\ \phi_{22}(t,\tau) &= \exp\left(\int_\tau^t -\frac{\partial}{\partial a_l} g(\bar{a}_l(\sigma), \bar{r}_l(\sigma))d\sigma\right). \end{aligned} \tag{4.19}$$

It is immediately seen that $\phi_{11}(t,\tau), \phi_{22}(t,\tau), \phi_{21}(t,\tau) > 0$.

Once comparison systems and transition matrices have been determined, the equations in theorems 9 and 10 can be evaluated to obtain RV constraints between successive pairs of comparison. Suppose we compare a behavior of the cascaded-tanks system with a watertight upper tank and a behavior of a system with an upper tank having a leak at the bottom (figure 4.2). For propagation from pc_0 to pc_1 theorem 9 applies, since the two systems are compared over primitive behavior fragments $[t_0, t_1]$ and $[\hat{t}_0, \hat{t}_1]$. Evaluating (4.13) for $\hat{T} < T$ by using the vectors $\boldsymbol{x}$, $\boldsymbol{u}$, $\boldsymbol{a^0}$ and the matrices $\boldsymbol{A}$, $\boldsymbol{E}$, $\mathbf{\Phi}$ determined above yields:

$$\hat{a}_u(\hat{t}_1) - a_u(t_1) = a_u(\hat{t}_1) - a_u(t_1) + c_1(\hat{a}_u(\hat{t}_0) - a_u(t_0)) + c_2(\hat{i}(\hat{t}_0) - i(t_0)) - c_3(\hat{r}_u(\hat{t}_0) - r_u(t_0)) - c_4 \int_{\hat{t}_0}^{\hat{t}_1} \hat{o}_h(\tau) d\tau \quad (4.20)$$

$$\hat{a}_l(\hat{t}_1) - a_l(t_1) = a_l(\hat{t}_1) - a_l(t_1) + c_5(\hat{a}_u(\hat{t}_0) - a_u(t_0)) + c_6(\hat{a}_l(\hat{t}_0) - a_l(t_0)) + c_7(\hat{i}(\hat{t}_0) - i(t_0)) + c_8(\hat{r}_u(\hat{t}_0) - r_u(t_0)) - c_9(\hat{r}_l(\hat{t}_0) - r_l(t_0)) - c_{10} \int_{\hat{t}_0}^{\hat{t}_1} \hat{o}_h(\tau) d\tau \quad (4.21)$$

with all $c_i > 0$, except for c_8 whose sign is ambiguous. The constant factors are transition matrix elements (e.g., $c_1 = \phi_{11}(\hat{t}_1, \hat{t}_0)$) or integrals involving transition matrix elements and input matrix elements (e.g., $c_3 = -\int_{\hat{t}_0}^{\hat{t}_1} \phi_{11}(\hat{t}_1, \tau) b_{12}(\tau) d\tau$). From the qualitative behavior in fig. 4.1(c) we see that the amounts of water in the leaky upper tank and watertight lower tank of the first system are increasing between t_0 and t_1, so $a_u(\hat{t}_1) - a_u(t_1) < 0$ and $a_l(\hat{t}_1) - a_l(t_1) < 0$ (since $\hat{t}_1 < t_1$). Furthermore, $\hat{o}_h$ is positive in $]\hat{t}_0, \hat{t}_1]$ and the integrals $\int_{\hat{t}_0}^{\hat{t}_1} \hat{o}_h(\tau) d\tau$ are consequently positive as well.

The above equations can be directly translated into RV constraints:

$$\text{if } RV(T, pc_0, pc_1) = \Downarrow, \text{ then } RV(a_u, pc_1) < RV(a_u, pc_0) + RV(i, pc_0) - RV(r_u, pc_0)$$
$$\text{if } RV(T, pc_0, pc_1) = \Downarrow \text{ and } RV(r_u, pc_0) = \|,$$
$$\text{then } RV(a_l, pc_1) < RV(a_u, pc_0) + RV(a_l, pc_0) + RV(i, pc_0) - RV(r_l, pc_0)$$

Since the case $\hat{T} < T$ has been considered, the behavior fragment of the second system has a shorter duration and we have the condition $RV(T, pc_0, pc_1) = \Downarrow$. The additional condition $RV(r_u, pc_0) = \|$ in the second constraint serves to compensate for the ambiguity of c_8 in (4.21).

Knowing the RVs of some variables allows one to derive the RVs of others. For instance, from the first constraint it follows that if $T \Downarrow_{pc_0 \to pc_1}$ and $a_u \|_{pc_0}$, $i \|_{pc_0}$, $r_u \|_{pc_0}$, then $a_u \Downarrow_{pc_1}$. The second RV constraint implies, by contraposition, that from $a_l \|_{pc_0}$, $a_l \|_{pc_1}$, $a_u \Downarrow_{pc_0}$, $i \|_{pc_0}$, $r_l \|_{pc_0}$, $r_u \|_{pc_0}$ we must conclude that the relative duration from pc_0 to pc_1 cannot be shorter.

For $\hat{T} > T$ or $\hat{T} = T$, the other two cases in theorem 9, RV constraints can be derived in a similar fashion. For $\hat{T} > T$ one arrives at:

$$\text{if } RV(T, pc_0, pc_1) = \Downarrow, \text{ then } RV(a_u, pc_1) \gtreqless RV(a_u, pc_0) + RV(i, pc_0) - RV(r_u, pc_0)$$
$$\text{if } RV(T, pc_0, pc_1) = \Downarrow \text{ and } RV(r_u, pc_0) = \|,$$
$$\text{then } RV(a_l, pc_1) \gtreqless RV(a_u, pc_0) + RV(a_l, pc_0) + RV(i, pc_0) - RV(r_l, pc_0)$$

The constraints are vacuous, because they do not put any restrictions on the RVs of the variables. As a consequence, they can be omitted.

As a second example, consider the comparison of a behavior of cascaded tanks with a watertight upper tank and with an upper tank having a leak at the side (figure 4.3). The two systems are compared between pc_0 and pc_1, so that theorem 10 applies and equation (4.15) should be evaluated. For the case $\hat{T} < T$ we need two comparison systems, since at $\hat{t}_1$ a region transition occurs, as explained above. On $[\hat{t}_0, \hat{t}_1]$ we must compare two structurally equal systems (with watertight upper tanks) and on $[\hat{t}_1, \hat{t}_2]$ two structurally different systems (with a watertight and a leaky upper tank).

The two comparison systems are straightforward to determine. We have $\boldsymbol{x_1} = \boldsymbol{x_2} = [a_u \; a_l]'$, $\boldsymbol{u_1} = \boldsymbol{u_2} = [i \; r_u \; r_l]'$, and $\boldsymbol{a_1} = \mathbf{0}$, $\boldsymbol{a_2} = o_h$. Furthermore,

$$\boldsymbol{A_1}(t) = \boldsymbol{A_2}(t) = \begin{bmatrix} -\frac{\partial}{\partial a_u} f(\bar{a}_u(t), \bar{r}_u(t)) & 0 \\ \frac{\partial}{\partial a_u} f(\bar{a}_u(t), \bar{r}_u(t)) & -\frac{\partial}{\partial a_l} g(\bar{a}_l(t), \bar{r}_l(t)) \end{bmatrix},$$

$$\boldsymbol{E_1}(t) = \mathbf{0}, \;\; \boldsymbol{E_2}(t) = \begin{bmatrix} -1 \\ 0 \end{bmatrix},$$

$$\boldsymbol{B_1}(t) = \boldsymbol{B_2}(t) = \begin{bmatrix} 1 & -\frac{\partial}{\partial r_u} f(\bar{a}_u(t), \bar{r}_u(t)) & 0 \\ 0 & \frac{\partial}{\partial r_u} f(\bar{a}_u(t), \bar{r}_u(t)) & -\frac{\partial}{\partial r_l} g(\bar{a}_l(t), \bar{r}_l(t)) \end{bmatrix}. \tag{4.22}$$

The equality of $\boldsymbol{A_1}(t)$ and $\boldsymbol{A_2}(t)$ leads to equal transition matrices $\boldsymbol{\Phi_1}(t, \tau)$ and $\boldsymbol{\Phi_2}(t, \tau)$ which have already been determined and are given by (4.18).

Substituting the matrices and vectors in (4.17), elaborating and simplifying the expressions, and transforming them into RV constraints gives rise to

$$\begin{aligned} &\text{if } RV(T, pc_0, pc_1) = \Downarrow, \text{ then } RV(a_u, pc_1) < RV(a_u, pc_0) + RV(i, pc_0) - RV(r_u, pc_0) \\ &\text{if } RV(T, pc_0, pc_1) = \Downarrow \text{ and } RV(r_u, pc_0) = \|, \\ &\qquad \text{then } RV(a_l, pc_1) < RV(a_u, pc_0) + RV(a_l, pc_0) + RV(i, pc_0) - RV(r_l, pc_0) \end{aligned}$$

Notice that these RV constraints are the same as those for the comparison of a watertight upper tank and an upper tank with a leak at the bottom. This is in conformity with the intuition that from a qualitative perspective it does not really matter at which height the leak is located, as long as it causes an additional water outflow to occur in comparison with the watertight tank. When the systems with a leak at the bottom and with a leak at a certain height are compared, however, a difference in outcome is expected.

4.3.4 Constants between pairs of comparison

Besides state variables, constants also have their difference propagated between pairs of comparison.

Theorem 11 (Constants) Suppose two behavior fragments defined by $pc_0 = \langle t_0, \hat{t}_0 \rangle$ and $pc_1 = \langle t_1, \hat{t}_1 \rangle$ are compared. If the variable c has a constant value during the behavior fragments, then

$$\hat{c}(\hat{t}_1) - c(t_1) = \hat{c}(\hat{t}_0) - c(t_0).$$

The result follows immediately with the observation that constancy of c in both systems implies $d\hat{c}/dt = 0$ and $dc/dt = 0$.

The equality of differences between two successive pairs of comparison enables one to formulate RV constraints. For instance, the constancy of the inflow in the cascaded-tanks behaviors compared in figure 4.2 leads to the constraint $RV(i, pc_1) = RV(i, pc_0)$.

4.3.5 Variables at a region transition

The behavior of the leaky cascaded tanks in figure 3.7 exhibits a *region transition*, a change of model. At $\hat{t}_1$ the QDE of the system changes in order to account for the flow out of the leak. As we saw, the comparison of this behavior with the behavior of a watertight cascaded-tanks system can be handled by theorem 10 in a straightforward way. More generally spoken, the theorems 9 and 10 remain applicable when region transitions occur in the behavior fragments. Even the cases in figure 4.6, where two successive pairs of comparison span a region transition, are covered. For example, in (a) we have two primitive behavior fragments of *a single time-point*, $[t_0, t_0]$ and $[\hat{t}_0, \hat{t}_0]$, in which the QDEs of the first region are assumed to hold.[6] By assumption the behavior fragments are synchronous, so $\hat{t}_0 = t_0$. Theorem 4 can now be applied for every state variable x, leading to $\hat{x}(\hat{t}_0) - x(t_0) = \boldsymbol{F}(t_0, t_0) = \hat{x}(\hat{t}_0) - x(t_0)$. The comparison of two cascaded tanks with a leak at the side and a behavior given by figure 3.7 leads to the situation shown in (a).

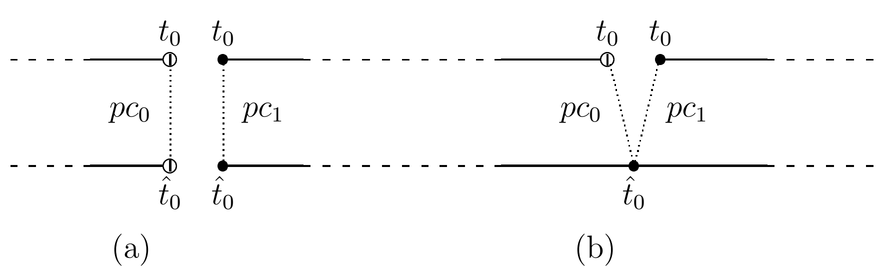


Figure 4.6: *Two examples of region transitions that coincide with successive pairs of comparison.*

The theorems 9 and 10 apply to state variables only. In the case of region transitions of the type displayed in figure 4.6, a stronger statement can be made for variables that are continuous across the region transition.

[6]The choice of the QDEs of the first region is conventional. We could equally well have chosen the QDEs of the second region.

Theorem 12 (Region transitions) Suppose two instantaneous, primitive behavior fragments are compared, with both time-points being transition time-points or one time-point being a transition time-point and the other one a regular time-point (figure 4.6). Let v be a variable shared by the two systems that is continuous across the region transition(s). The difference $\hat{v} - v$ at $pc_1 = \langle t_0, \hat{t}_0 \rangle$ equals that at $pc_0 = \langle t_0, \hat{t}_0 \rangle$ and for the durations of the behavior fragments we have $\hat{T} = T = 0$.

This theorem implies that for every shared variable v we have $RV(v, pc_1)$=$RV(v, pc_0)$, and $RV(t_1, pc_0, pc_1) =\|$. The theorem is an obvious consequence of the continuity of variables across region transitions.

4.4 The CEC* algorithm

The RV constraints are employed in the *CEC* algorithm*, which transforms a triple input of qualitative behaviors, qualitative models (QDEs), and relative values into a comparative envisionment, an output structure representing the RVs of shared variables at the pairs of comparison. The propagation process achieving this transformation will be discussed in this section by introducing the CEC* algorithm. An application of predictive and explanatory CA in the cascaded-tanks examples illustrates the working of the algorithm.

4.4.1 Description of the algorithm

A comparative analysis is performed by propagating RVs from the top to the bottom, or from the bottom to the top, of the OPC structure. In the remainder of this section the direction of propagation is assumed to be from top to bottom, that is, from pc_0 to pc_n. Recall that this direction corresponds with performing a predictive CA (section 4.2.7).

The propagation of RVs starts at the first pair of comparison. The initial RVs are completed to one or more comparative states at pc_0. There may be several comparative states at a pair of comparison, because the RV of a variable can be ambiguous. The algorithm continues by repeatedly selecting a current pair of comparison for which the possible comparative states at *all* direct predecessor pairs of comparison have been determined. A pair of comparison may have one or several direct predecessor pairs of comparison, depending on the OPC structure. The comparative states at the current pair of comparison that are consistent with both the RV constraints and comparative states at the predecessor pairs of comparison are then generated. Since there may be several comparative states at each predecessor pair of comparison, we have to repeat this for every (combination of) comparative state(s) at the predecessor pairs of comparison.

Each of the resulting comparative states is linked to the comparative states at the predecessor pairs of comparison that have been used for its generation. The activities of determining comparative states at a current pair of comparison and linking them to comparative states at predecessor pairs of comparison continue until the bottom pair of comparison pc_n has been processed. The propagation of RVs from the first to the last pair of comparison is then completed. The output is an a-cyclic, directed hypergraph whose nodes represent comparative states and whose edges represent transitions from one or more comparative states at predecessor pairs of comparison to a comparative state at the current pair of comparison. The edges are labeled with the relative durations $RV(T)$ of the behavior fragments between the pairs of comparison. This graph will be called a *comparative envisionment* (de Jong, Mars & van der Vet [1996]).

That the comparative envisionment is a *hyper*graph is a natural consequence of the fact that the current pair of comparison may have multiple predecessors, each of which contributes a comparative state when computing a comparative state at the current pair of comparison. In fact, the edges have more than two end-points, if and only if a pair of comparison has multiple predecessors in the OPC structure.

A comparative envisionment differs from a qualitative behavior tree produced by QSIM in two respects. First, each state in a qualitative behavior has only a single predecessor state. Second, a particular qualitative state may occur several times in the behavior tree, in different branches of the tree, whereas a comparative state occurs only once in the envisionment. Examples of comparative envisionments are shown in figures 4.7 and 4.8.

The possible comparative behaviors of the systems are represented by connected subgraphs of the comparative envisionment. A comparative behavior contains exactly one comparative state at each pair of comparison. Obviously, a comparative behavior needs to be a connected subgraph, otherwise there would be comparative states not reachable through transitions. If every pair of comparison in the OPC structure has a *single* predecessor, the resulting comparative behaviors are paths in the envisionment (figure 4.8 shows three paths). In case there are pairs of comparison with several predecessors, a comparative behavior is a more complicated kind of subgraph.

This informal description of the algorithm can be summarized in the following procedure for *comparative envisionment construction*:

Algorithm 2 (CEC*) Given two qualitative behaviors, one for each system, which define a set of ordered pairs of comparison (OPC structure). The RVs at the first comparative state have been (partially) specified as $CS(init)$. The function $v(pc_i)$ returns the status of pair of comparison pc_i, i.e., 'done' or 'not done'. A comparative envisionment can now be constructed as follows:

Step 1 Assign the status 'not done' to each pair of comparison in the OPC structure, so $v(pc_i) =$'not done' for all pc_i. Initialize the comparative envisionment as the null graph.

Step 2 Determine a pair of comparison pc_i and its set of predecessor pairs of comparison $\mathrm{pre}(pc_i)$, such that $v(pc_i) =$'not done' and either for every $pc_j \in \mathrm{pre}(pc_i)$ it holds that $v(pc_j) =$'done' or $\mathrm{pre}(pc_i)$ is empty. If no such pc_i can be found, then finish.

Step 3 Retrieve the RV constraints appropriate for propagating from the pairs of comparison in $\mathrm{pre}(pc_i)$ to pc_i.

Step 4 For each $pc_j \in \mathrm{pre}(pc_i)$, retrieve the comparative states $CS_{pc_j,r}$, $r \geq 0$, from the comparative envisionment. If for some pc_j no comparative states can be found, then finish. Find all consistent comparative states at pc_i for every combination of comparative states at the predecessor pairs of comparison. Call these comparative states $CS(pc_i)_0, \ldots, CS(pc_i)_m$, $m \geq 0$.

Step 5 Add the new comparative states $CS(pc_i)_0, \ldots, CS(pc_i)_m$ to the comparative envisionment and link them to the comparative states of the predecessor pairs of comparison that were used to obtain them. Label the edges with $RV(T, pc_j, pc_i)$ for each $pc_j \in \mathrm{pre}(pc_i)$. Assign the status 'done' to pc_i and return to step 2.

When the algorithm starts at pc_0, no predecessor pairs of comparison exist, so $\mathrm{pre}(pc_0)$ is an empty set. As a consequence, in step 3 only RV constraints valid for pc_0 are retrieved. They are used in step 4 to turn the possibly partially specified initial comparative state $CS(init)$ into one or more completely specified comparative states $CS(pc_0)_m$. By analogy with established QR terminology this pass of the algorithm is called *state completion*.

Thus far, the algorithm has been described as designed for predictive CA, reasoning from differences in initial conditions to differences in response of the two systems. Notice, however, that the algorithm can be used for explanatory CA as well, after some small adaptations. We then propagate from the last pair of comparison in the OPC structure, the bottom element in the $\preceq$-ordering, to the first pair of comparison, the top element.

4.4.2 Implementation of the algorithm

The CEC* algorithm has been implemented in Common Lisp, as a program interacting with the QSIM implementation (appendix D; van Raalte & de Jong [1997]). It counts approximately 11,000 lines of code, and includes modules for computing the OPC structure from a pair of qualitative behaviors, retrieving appropriate RV constraints, generating comparative states consistent with the RV constraints, and constructing a comparative envisionment. A teletype interface allows the user to specify the input of a comparative analysis (qualitative models, qualitative behaviors, initial relative values), the processing mode (predictive or explanatory CA), and a trace level. The output is a formal representation of the comparative envisionment which can be transformed into a pictorial representation like the ones shown in the next section.

A basic step in the algorithm is the repeated determination of possible comparative states at a pair of comparison. This step has been formalized as the generation and solution of a constraint satisfaction problem (CSP) consisting of the variables shared by the two systems, domains $\{\Downarrow, \Uparrow, \|\}$ of their possible relative values, and constraints relating the RVs of variables at the current and at predecessor pairs of comparison. The CSP algorithm Cfilter (Kuipers [1994], ch. 4), embedded in the QSIM implementation, is employed by CEC* as well to find solutions for the CSP in a correct and efficient way.

The RV constraints in the CSP are instantiated from a database with templates derived from the theorems 6, 7, 9, 10, 11, and 12. The templates link features of the qualitative models and qualitative behaviors to RV constraints. Especially for RV constraints directed at the propagation of state variables between pairs of comparison (theorems 9 and 10), the computations to arrive at suitable templates, though following a strict procedure, are often tedious to perform by hand. In later versions of the implementation we hope to employ mathematical packages like Maple and Mathematica to derive RV constraints (semi-)automatically from the behaviors and models, thus obviating the shortcut of a user-provided database with templates.

The CEC* algorithm and its implementation have been applied to a range of intra-model and inter-model CA questions concerning simple and more advanced QR systems, such as bathtubs, watertanks filled by a pump, masses on a spring, heat exchangers, sliding blocks, and materials submitted to external stress. These examples generalize upon the cascaded-tanks example elaborated in section 4.4.3 by requiring CEC* to deal with topologically different qualitative behaviors, region transitions, and composite behavior fragments between pairs of comparison.

4.4.3 Examples of comparative envisionments

Predictive CA of cascaded tanks

An example of a comparative envisionment produced by CEC* is shown in figure 4.7. The envisionment answers the predictive CA question: How does the equilibrium state of a cascaded-tanks system with a leak at the bottom of the upper tank differ from that with a watertight upper tank when the systems are considered under the same conditions? In other words, we start the comparative analysis from pc_0 with the initial state information

$$CS(init) = \langle a_u \|_{pc_0}, a_l \|_{pc_0}, r_u \|_{pc_0}, r_l \|_{pc_0}, i \|_{pc_0} \rangle.$$

Each path from the first to the last comparative state in the envisionment forms a comparative behavior. Although ambiguities arise with respect to a_l,n_l at pc_1, the five possible comparative behaviors predict a single outcome: the leak tends to lower the amount of water in the tanks at equilibrium ($CS(pc_2)_0$).[7] The comparative envisionment shows how a structural difference between two systems can cause a differential response, even though the systems are made to evolve under the same conditions. The RV constraints do not allow one to derive an unambiguous conclusion about the relative durations to reach the equilibrium state, given the ambiguities in the relative durations from pc_0 to pc_1 and from pc_1 to pc_2.

As the example illustrates, the output of the CEC* algorithm, the comparative envisionment, lends itself for the generation of user explanations. The sequence of comparative states in a comparative behavior traces the differential development of shared system variables over time. It should be kept in mind, however, that although a suitable starting-point, there is more to explanation than following the links between comparative states in a comparative behavior (see Weld [1990], ch. 4, for issues in the context of CA).

[7] Actually, there are more than five comparative behaviors. The figure abstracts from ambiguities with respect to relative durations by subsuming the three alternative $RV(T)$ values by a single edge labeled $T?$ in the graph.

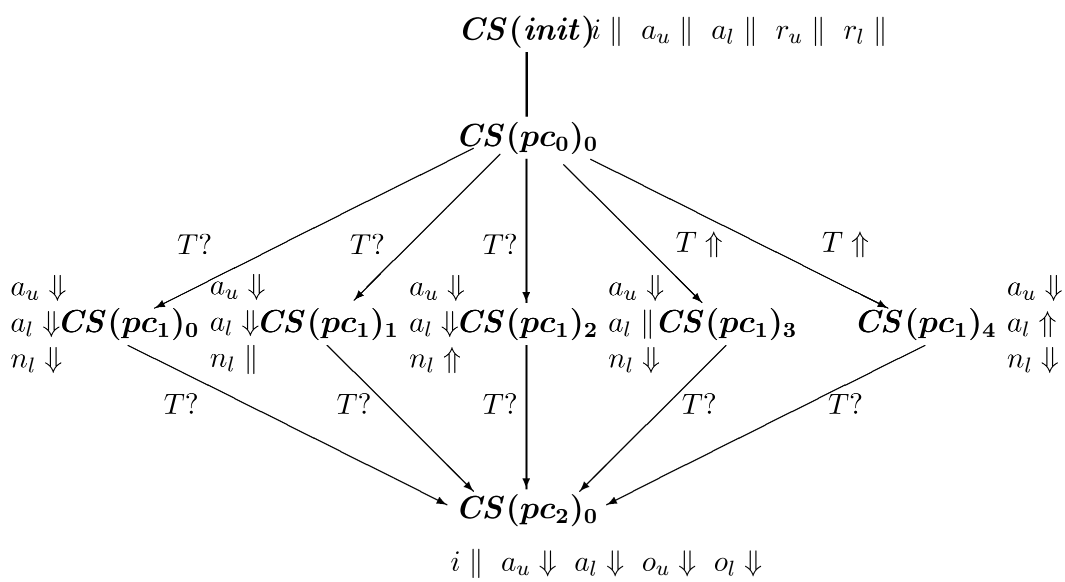


Figure 4.7: *Comparative envisionment arising from the comparison of a watertight and a leaky cascaded-tanks system (figure 4.2). The comparative envisionment is produced in response to the predictive CA question: How does the equilibrium state of the system with a leak at the bottom differ from that with a watertight upper tank when the systems are considered under the same conditions? A few distinctive RVs are indicated at the comparative states (CS).*

Explanatory CA of cascaded tanks

The same cascaded-tanks example can be used to illustrate explanatory reasoning as well. Consider the following CA question: When the relative amounts of water at pc_1 and pc_2 have been measured for watertight and leaky cascaded tanks filled from empty, and the orifices have the same size, which differences could account for the lower equilibrium water amounts in the leaky system? The initial comparative state information is

$$CS(init) = \langle a_u \Downarrow_{pc_2}, a_l \Downarrow_{pc_2}, r_u \|_{pc_2}, r_l \|_{pc_2}, i \|_{pc_2} \rangle,$$

and, in addition, it is known by observation that $a_u \Downarrow_{pc_1}$ and $a_l \|_{pc_1}$.

CEC* concludes that the lower equilibrium water amounts in the second system must be attributed to one of the following three causes: (1) a leaky upper tank and lower inflow ($CS(pc_0)_0$), (2) a leaky upper tank and higher inflow ($CS(pc_0)_2$), and (3) a leaky upper tank only ($CS(pc_0)_1$). The leak, making the second system structurally different from the first, tends to lower the equilibrium amounts at pc_2. An additional difference in initial conditions may strengthen ($i \Downarrow_{pc_0}$) or attenuate ($i \Uparrow_{pc_0}$) this effect. In the latter case the structural difference and the differential antecedent conditions work in opposite directions, but from the outcome, the observations to be explained, we know that the contribution of the leak will always dominate the higher inflow.

The number of explanations generated by CEC* could be reduced by seeking new information which can rule out alternative branches in the comparative envisionment, possibly already *during* the CA process. For instance, a measurement of the difference in inflow at a pair of comparison would allow one to prune two of the three comparative behaviors.

Another way to bring some order in the set of explanations proposed by CEC* is prioritization of alternatives. A very simple scheme is to apply Occam's Razor, that is, to focus on the explanations that hypothesize the least number of differences. In the example of figure 4.8, $CS(pc_0)_1$ would receive the highest priority, because it does not assume anything about a difference in the inflow rate. There is nothing, however, that forbids an observed difference to be explained by a combination of differences in the structure of the systems and the initial conditions. The strategy is clearly heuristic, a means to single out the most likely candidate assuming that the hypothesized differences occur independently.

Dead-ends in a comparative envisionment

Suppose that a watertight and a leaky cascaded-tanks system are compared, and an explanation is sought for fact that the two systems produce the same response at pc_2:

$$CS(init) = \langle i \|_{pc_2}, a_u \|_{pc_2}, a_l \|_{pc_2}, r_u \|_{pc_2}, r_l \|_{pc_2} \rangle.$$

In the first pass of the algorithm, one finds that the initial comparative state information cannot be completed into one or more comparative states at pc_2, because there

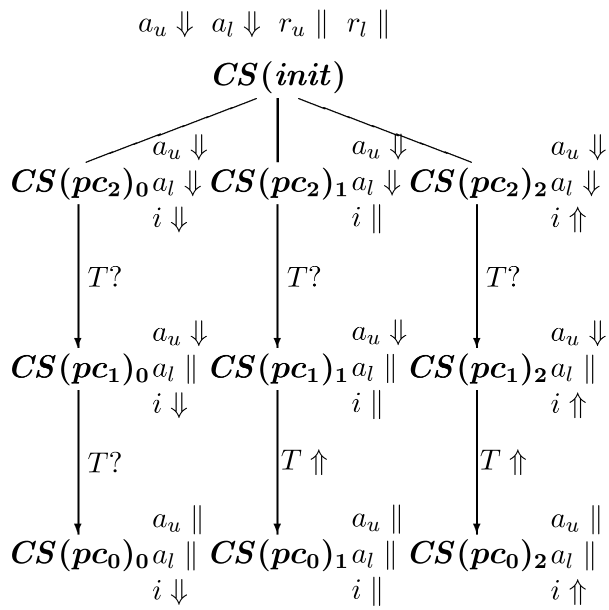


Figure 4.8: *Comparative envisionment arising from the comparison of a watertight and a leaky cascaded-tanks system when both reach equilibrium in the upper tank first (figure 4.2). The comparative envisionment is produced in response to the explanatory CA question: When the relative amounts of water at* pc_1 *and* pc_2 *have been measured for watertight and leaky cascaded tanks filled from empty, and the orifices have the same size, which differences could account for the lower equilibrium water amounts in the leaky system?*

are no states consistent with the RV constraints retrieved in step 3. The RV constraints require the outflow of the upper tank, o_u, to be both $\|$ and $\Downarrow$.

More generally, it may turn out that no transition to a direct successor pair of comparison is possible from a particular comparative state. Such a comparative state will be called a *dead-end* in the comparative envisionment. The algorithm continues the analysis when a dead-end is detected, unless *all* comparative states at a pair of comparison are dead-ends. No comparative states are then generated at a successor pair of comparison and the algorithm is aborted at step 4 in the next pass.

Dead-ends represent situations that cannot possibly occur given the constraints determined by the structure and the behavior of the systems, so it would be desirable to remove them from the comparative envisionment afterwards. Since such post-processing may introduce new dead-ends, the process should be recursive. In the special case that all comparative states at a pair of comparison are dead-ends, recursive pruning of dead-ends leads to a comparative envisionment that is the null graph. An empty comparative envisionment signalizes an *inconsistency* in the input of the comparative analysis. The inconsistency implies that at least one of the three input elements – qualitative models, qualitative behaviors, and initial relative values – must be revised.[8]

In the case of the cascaded watertanks, no comparative state is found at pc_2 and the analysis yields an empty comparative envisionment. On reflection, it is clear that a leaky system with the same size of the orifices and the same inflow as a watertight system cannot reach the same equilibrium water amounts when both systems are filled from empty. The leak tends to lower the equilibrium water amounts and there are no further differences to counter this tendency.

4.5 Properties of the CEC* algorithm

In order to prepare the evaluation in section 4.7, I will investigate in some detail the correctness (i.e., soundness and completeness), the ability to deal with inconsistent input, and the computational complexity of the CEC* algorithm. The notions of soundness, completeness, and inconsistency in comparative analysis are defined analogously to their counterparts in qualitative simulation (section 3.5).

4.5.1 Comparative behavior abstraction

In the previous chapter, I have reviewed how abstraction relations between QDEs and qualitative behaviors on the one hand, and ODEs and solutions to ODEs on the other, are defined in QSIM (theorems 1 and 2). These abstraction relations can be extended with a *comparative behavior abstraction*.

[8] In Weld's DQ analysis (Weld [1990]) the qualitative behavior of the second system is chosen by default as the culprit of an inconsistency. The inconsistency is attributed to a violation of the assumption of topological equality of the behaviors of the original system and the disturbed system.

Theorem 13 (Comparative behavior abstraction) Let $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ be vectors of reasonable functions, which represents the time behavior of the variables of two dynamical systems on the bounded intervals $[a, b]$ and $[\hat{a}, \hat{b}]$. Further, $\boldsymbol{q}$ is a vector of shared variables. Then a comparative behavior CB can be defined, consisting of the tuple of comparative states and relative durations

$$CB(\boldsymbol{q}) = \langle CS(\boldsymbol{q}, pc_0), \ldots, CS(\boldsymbol{q}, pc_n), RV(T, pc_{i_1}, pc_{j_1}), \ldots, RV(T, pc_{i_k}, pc_{j_k}) \rangle.$$

There are n pairs of comparison, with $pc_0 = \langle a, \hat{a} \rangle$, $pc_n = \langle b, \hat{b} \rangle$, and $pc_{i_1} \preceq pc_{j_1}, \ldots, pc_{i_k} \preceq pc_{j_k}$ the k direct successor relations in the OPC graph. CB is the qualitative behavior abstraction of $\boldsymbol{q}$.

The proof is obvious and follows by the definitions of relative values, comparative states, and comparative behaviors in section 4.2. Notice that a given comparative behavior is satisfied, in general, by a whole class of reasonable functions $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$. Extending the QSIM terminology, two solutions $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ are said to *satisfy* or to be *consistent with* a comparative behavior CB. On the other hand, CB *abstracts from* or *describes* $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$. The formal connection between concepts from comparative analysis and the theory of differential equations plays an important role in defining and establishing properties of CEC*.

4.5.2 Defining soundness, completeness, and inconsistency

Recall that productive way to look at QSIM is to view the program as a theorem prover deriving theorems of the following form:

$$QSIM \vdash QDE \wedge QS(init) \rightarrow QB_1 \vee \ldots \vee QB_m.$$

QSIM proves from a QDE and initial qualitative state information $QS(init)$ a disjunction of qualitative behaviors $QB_1, \ldots, QB_m$. In an analogous way, we can interpret CEC* as a theorem prover:

$$CEC^* \vdash [QDE \wedge QS(init) \wedge QB] \wedge [\hat{QDE} \wedge \hat{QS}(init) \wedge \hat{QB}] \wedge CS(init) \rightarrow CB_1 \vee \ldots \vee CB_n.$$

CEC* proves from (1) two qualitative behaviors QB and $\hat{QB}$ with their corresponding QDEs and initial qualitative state information,[9] and (2) a set of RVs at the first pair of comparison $CS(init)$, a disjunction of comparative behaviors: $CB_1 \vee \ldots \vee CB_n$.

The behavior tree generated by QSIM may contain spurious qualitative behavior, that is, behaviors QB_i which describe no solution to any initial value problem $ODE \wedge \boldsymbol{v}(t_0) = \boldsymbol{v_0}$ satisfying $QDE \wedge QS(init)$ (definition 9). The analogy between the CEC* inference and the QSIM inference can be pushed one step further by introducing the notion of *spurious comparative behavior*:

[9]In order to simplify the analysis, it is here assumed that the input behaviors do not exhibit region transitions. Generalization to the case that region transitions do occur is straightforward.

Definition 23 (Spurious comparative behavior) A comparative behavior CB_i in a CEC* prediction

$$[QDE \wedge QS(init) \wedge QB] \wedge [\hat{QDE} \wedge \hat{QS}(init) \wedge \hat{QB}] \wedge CS(init) \rightarrow CB_1 \vee \ldots \vee CB_n$$

is a spurious comparative behavior, iff it does not describe the difference between any two solutions $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ to initial value problems $ODE \wedge \boldsymbol{v}(t_0) = \boldsymbol{v_0}$ and $\hat{ODE} \wedge \hat{\boldsymbol{w}}(\hat{t}_0) = \hat{\boldsymbol{w}}_\mathbf{0}$, such that

1. $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ satisfy QB and $\hat{QB}$, and
2. $ODE \wedge \boldsymbol{v}(t_0) = \boldsymbol{v_0}$ and $\hat{ODE} \wedge \hat{\boldsymbol{w}}(\hat{t}_0) = \hat{\boldsymbol{w}}_\mathbf{0}$ satisfy $QDE \wedge QS(init)$ and $\hat{QDE} \wedge \hat{QS}(init)$.

In contrast, CB_i is a *genuine comparative behavior* if it can actually occur, that is, if it does describe the difference between two solutions $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ to $ODE \wedge \boldsymbol{v}(t_0) = \boldsymbol{v_0}$ and $\hat{ODE} \wedge \hat{\boldsymbol{w}}(\hat{t}_0) = \hat{\boldsymbol{w}}_\mathbf{0}$, and the differential equations, initial conditions, and solutions satisfy their qualitative counterparts used in the prediction.

The definition of a spurious comparative behavior has the following direct consequence:

Corollary 1 Given a CEC* prediction, if QB or $\hat{QB}$ is a spurious *qualitative* behavior, then every CB_i in $CB_1 \vee \ldots \vee CB_n$ is a spurious *comparative* behavior.

The corollary is obvious, since (1) and (2) in definition 23 cannot hold at the same time when QB or $\hat{QB}$ is spurious. If one of the two qualitative behaviors to be compared cannot possibly occur, then the resulting comparative behavior also cannot occur.

Where does this formal apparatus bring us? Its main use lies in precisely specifying what should be understood by soundness and completeness:

Definition 24 (Soundness) CEC* is sound, iff all genuine comparative behaviors are generated by the algorithm.

Definition 25 (Completeness) CEC* is complete, iff every comparative behavior generated by the algorithm is genuine.

The definitions can be interpreted as follows. Consider two solutions $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ describing actual behaviors of the two systems to be compared. Soundness means that the comparative behavior abstracting from the difference between $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ is included in the disjunction generated by CEC*. (Since the disjunction covers *all* genuine comparative behaviors, one of them must describe the difference between $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$.) The algorithm is called sound, because a disjunction is always true when one of its disjuncts is true.

Soundness of CEC* does not guarantee that no comparative behaviors are generated that are spurious, i.e., that do not correspond to any pair of real behaviors of the systems.

If CEC* generates a disjunction that includes spurious comparative behaviors, it fails to prove the stronger theorem without the spurious behaviors, so it is properly called incomplete.

Incompleteness of CEC* arises from two causes : (1) spurious qualitative behaviors QB and $\hat{QB}$ in the input, or (2) the failure of the RV propagation process to weed out impossible comparative states. I will assume that the first cause can always be taken away by selecting only genuine qualitative behaviors as input for CEC*. This requires that additional information on the qualitative behaviors be added to the QSIM output, since the QSIM behavior tree may include spurious qualitative behaviors. In many situations this requirement can be met, in particular if we know (by observation or otherwise) the actual behavior of a system (section 4.7).

In section 4.4 a situation was encountered in which the input of CEC* turned out to be inconsistent. By means of the concepts introduced above, inconsistency can be more rigorously defined.

Definition 26 (Inconsistent input) The input

$$[QDE \wedge QS(init) \wedge QB] \wedge [\hat{QDE} \wedge \hat{QS}(init) \wedge \hat{QB}] \wedge CS(init)$$

of the CEC* algorithm is inconsistent, iff no conceivable comparative behavior for this input is a genuine comparative behavior.

An inconsistency could be caused by an incompatibility between QB and $\hat{QB}$ and the RVs at the initial pair of comparison, as in the cascaded-tanks example in section 4.4.3. As in QSIM, there is an interesting relation between inconsistencies in the input and the completeness of the algorithm:

Theorem 14 If CEC* is complete, then it will detect every inconsistency in the input not arising from spurious qualitative behaviors.

Proof. Completeness means that no spurious comparative behaviors are generated (definition 25). Since an inconsistent input cannot lead to genuine comparative behaviors (definition 26), no comparative behaviors are generated. The absence of comparative behaviors (i.e., the comparative envisionment being a null graph) consequently signals the inconsistency of the input. □

4.5.3 CEC* is sound and incomplete

Now that soundness and completeness have been given a clear definition by establishing a mapping to the domain of differential equations, we can evaluate these properties for the CEC* algorithm of the previous section. As will be seen, CEC* is sound and incomplete.

Before we can conclude this, we first have to establish a relationship between genuine comparative behaviors and the RV constraints discussed in section 4.3.

Lemma 1 Suppose RV constraints have been derived from the qualitative differential equations QDE and $\hat{QDE}$ of two systems and corresponding qualitative behaviors QB and $\hat{QB}$. If $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ are solutions of ordinary differential equations satisfying QDE and $\hat{QDE}$, and $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ satisfy QB and $\hat{QB}$, then the comparative behavior $CB(\boldsymbol{q})$ abstracted from $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ is consistent with the RV constraints.

Proof. The RV constraints have been deduced from QDE, $\hat{QDE}$ and QB, $\hat{QB}$ by means of the theorems 6, 7, and 9 to 12. The RV constraints put restrictions on relative values at pairs of comparison and relative durations between pairs of comparison. By theorem 2, the solutions $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ satisfy (the qualitative constraints of) QDE and $\hat{QDE}$. Further, they satisfy QB and $\hat{QB}$. As a consequence, $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ are consistent with the RV constraints derived by means of theorems 6, 7, and 9 to 12. More specifically, the qualitative abstraction of (1) the difference in value of the shared variables $\boldsymbol{q}$ at the pairs of comparison (i.e., $RV(q, pc_i)$), and (2) the difference in duration of the intervals between pairs of comparison (i.e., $RV(T, pc_i, pc_j)$), are consistent with the RV constraints. $CB(\boldsymbol{q})$ is this qualitative abstraction, so $CB(\boldsymbol{q})$ is consistent with the RV constraints. □

The lemma helps in proving the following result:

Theorem 15 CEC* is sound.

Proof. Analogously to the proof of theorem 4 in Kuipers [1994], the basic idea behind the proof is that every genuine comparative behavior is generated and only spurious ones are filtered out (though maybe not all of them, see below). The proof follows the mathematical induction schema.

Suppose $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ are solutions of ODEs satisfying QDE and $\hat{QDE}$, and $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ satisfy QB and $\hat{QB}$. By theorem 13 the behaviors $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ define a comparative behavior

$$CB(\boldsymbol{q}) = \langle CS(\boldsymbol{q}, pc_0), \ldots, CS(\boldsymbol{q}, pc_n), RV(T, pc_{i_1}, pc_{j_1}), \ldots, RV(T, pc_{i_k}, pc_{j_k})\rangle.$$

Of course, under these conditions $CB(\boldsymbol{q})$ is not spurious.

At pc_0 we have the qualitative models QDE and $\hat{QDE}$, and the qualitative states $QS(t_0)$ and $\hat{QS}(\hat{t}_0)$. Together they generate a number of RV constraints, as explained in section 4.3. Further, we have domain restrictions $CS(init)$ that are assumed to be consistent with the comparative state $CS(\boldsymbol{q}, pc_0)$. A correct CSP algorithm (e.g., Cfilter in Kuipers [1994], ch. 4) generates all assignments of RVs to variables at pc_0 that are consistent with the RV constraints and the domain restrictions. In other words, it generates all consistent comparative states. The comparative state $CS(\boldsymbol{q}, pc_0)$ is consistent with those constraints and domain restrictions, for $CB(\boldsymbol{q})$ is an abstraction of the difference between the solutions $\boldsymbol{v}$ and $\hat{\boldsymbol{w}}$ (lemma 1). Therefore, $CS(\boldsymbol{q}, pc_0)$ is generated by the algorithm.

Now suppose that the elements of the set of predecessor pairs of comparison of pc_i, pre(pc_i), have the status 'done' and for each $pc_j \in$ pre(pc_i) the corresponding comparative state $CS(\boldsymbol{q}, pc_j)$ of $CB(\boldsymbol{q})$ has been generated. Consider the RV constraints and domain restrictions that are created for propagation of relative values from a certain pc_j to pc_i. During the relevant behavior fragments of QB and $\hat{QB}$, QDE and $\hat{QDE}$ are valid, together with the qualitative states at the time-points and intervals of the behavior fragments. By means of the theorems in section 4.3, constraints on the RVs of the variables can be generated. $CS(\boldsymbol{q}, pc_j)$ provides domain restrictions.

Thus a CSP is created by repeating the above procedure for all $pc_j \in \mathrm{pre}(pc_i)$. Cfilter generates all assignments of RVs to variables on pc_i, and to the relative duration of the two behavior fragments between each pc_j and pc_i, that are consistent with the RV constraints and domain restrictions. Since $CS(\boldsymbol{q}, pc_j)$ and the $RV(T, pc_j, pc_i)$s in $CB(\boldsymbol{q})$ are consistent with those constraints and domain restrictions (lemma 1), they are generated by the algorithm.

Taking it all together, $CS(pc_0)$ is generated and each successor comparative state $CS(\boldsymbol{q}, pc_j)$ in the comparative behavior $CB(\boldsymbol{q})$ is generated from its predecessors together with the corresponding $RV(T, pc_j, pc_i)$s. Consequently, $CB(\boldsymbol{q})$is generated step by step. Because the number of pairs of comparison is finite, the algorithm will terminate in a finite number of steps. □

All comparative behaviors that can actually occur will be generated by CEC*, as a consequence of its soundness. However, not all comparative behaviors generated can actually occur:

Theorem 16 CEC* is incomplete, even if it starts from genuine (non-spurious) qualitative behaviors.

Proof. The incompleteness of CEC* can easily be shown by means of a counterexample, that is, an example in which spurious comparative behaviors are generated (definition 23).

Consider again the ideal mass-spring system of the previous chapter (figure 3.8). Figure 4.9 shows two qualitative behaviors of the undamped mass-spring system and the OPC structure derived from them. Only one period of the behaviors is shown. Notice that we are dealing with an instance of intra-model CA with topologically equal behaviors. The pairs of comparison define primitive behavior fragments.

In order to determine the RV constraints linking $pc_0 = \langle t_0, \hat{t}_0 \rangle$ and $pc_1 = \langle t_1, \hat{t}_1 \rangle$, we must apply theorem 9. As before, a comparison system will be constructed. The shared state variables $\boldsymbol{x}$ and the shared input variables $\boldsymbol{u}$ are found to be $[x\ v]'$ and $[k\ m]'$, respectively. Since two systems with the same structure are compared, there are no auxiliary variables, $\boldsymbol{a} = \boldsymbol{0}$. The state, input, and auxiliary matrices are determined as

$$\boldsymbol{A}(t) = \begin{bmatrix} 0 & 1 \\ -\dfrac{\bar{k}(t)}{\bar{m}(t)} & 0 \end{bmatrix}, \qquad \boldsymbol{E}(t) = \boldsymbol{0}, \qquad \boldsymbol{B}(t) = \begin{bmatrix} 0 & 0 \\ -\dfrac{\bar{x}(t)}{\bar{m}(t)} & \dfrac{\bar{k}(t)\bar{x}(t)}{\bar{m}^2(t)} \end{bmatrix}, \quad (4.23)$$

with $\bar{x}$ lying between $\hat{x}$ and x, $\bar{k}$ between $\hat{k}$ and k, and $\bar{m}$ between $\hat{m}$ and m. Although $k, \hat{k}$ and $m, \hat{m}$ are constants, $\bar{k}$ and $\bar{m}$ are not necessarily so. In appendix A.2 it is proven

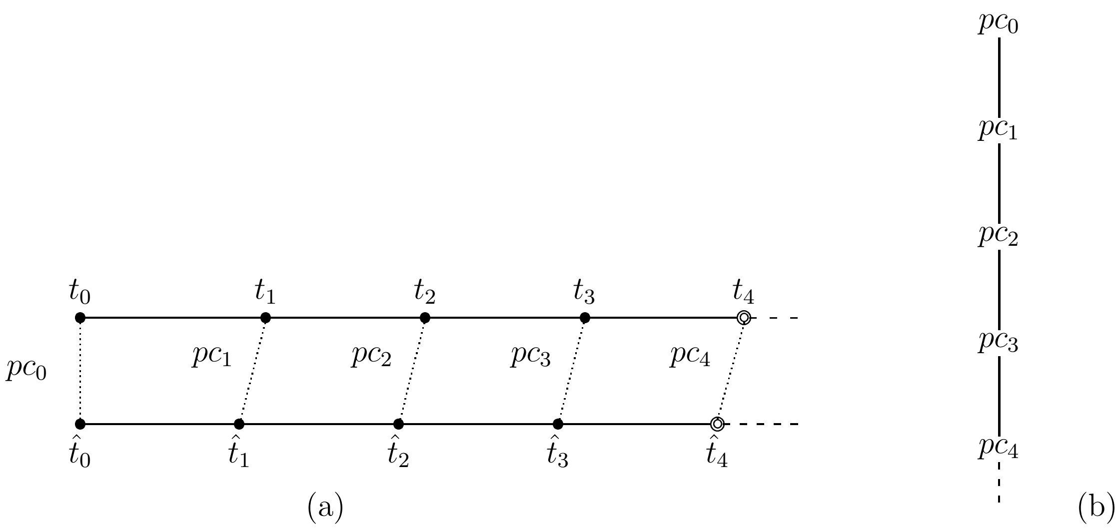


Figure 4.9: *(a) The comparison of two ideal mass-spring behaviors (figure 3.8(c)). (b) The set of pairs of comparison ordered by the $\preceq$-relation.*

separately that the term $-\bar{k}(t)/\bar{m}(t)$ appearing in the state matrix is time-invariant, so that we arrive at $\boldsymbol{A}(t) = \boldsymbol{A}$. By means of appendix B the state matrix is then found to be

$$\boldsymbol{\Phi}(t,\tau) = \begin{bmatrix} \cos\sqrt{\dfrac{\bar{k}}{\bar{m}}}(t-\tau) & \sqrt{\dfrac{\bar{m}}{\bar{k}}}\sin\sqrt{\dfrac{\bar{k}}{\bar{m}}}(t-\tau) \\ -\sqrt{\dfrac{\bar{k}}{\bar{m}}}\sin\sqrt{\dfrac{\bar{k}}{\bar{m}}}(t-\tau) & \cos\sqrt{\dfrac{\bar{k}}{\bar{m}}}(t-\tau) \end{bmatrix}. \tag{4.24}$$

As can be seen, the transition matrix is periodic with period $\bar{P} = 2\pi\sqrt{\dfrac{\bar{m}}{\bar{k}}}$.

Now substitute the vectors and matrices into equation (4.13) of theorem 9 for the case that $\hat{T} > T$. In order to determine the sign of the elements of $\boldsymbol{F}(t_1, t_0)$, we need to know the sign of the functions $\cos\sqrt{\dfrac{\bar{k}}{\bar{m}}}(t_1-\tau)$ and $\sin\sqrt{\dfrac{\bar{k}}{\bar{m}}}(t_1-\tau)$ for all τ on $[t_0, t_1]$, since the sign of these functions determines the sign of the elements of $\boldsymbol{\Phi}(t,\tau)$. Phrased in an alternative way, we need to know how the period $\bar{P}$ of the transition matrix relates to $t_1 - t_0 = T$. This information is not available from the qualitative behaviors, so theorem 9 yields very weak RV constraints like

$$\text{if } RV(T)_{pc_0\to pc_1} = \Uparrow, \text{ then } RV(x)_{pc_1} > ?\,RV(x)_{pc_0}\,?\,RV(v)_{pc_0}\,?\,RV(k)_{pc_0}\,?\,RV(m)_{pc_0}, \tag{4.25}$$

where the question mark means that the RV term following it is prefixed by the + or $-$ operator, or omitted altogether. If $x\,\|_{pc_0}$, $v\,\|_{pc_0}$, $k\,\|_{pc_0}$, and $m\,\|_{pc_0}$, one can conclude that from $RV(T, pc_0, pc_1) = \Uparrow$ it follows that $x \Uparrow_{pc_1}$, but stronger inferences are not supported by the constraint.

Suppose we perform a CA with initial RVs $CS(init) = \langle x \Downarrow_{pc_0}, v \parallel_{pc_0}, k \parallel_{pc_0}, m \parallel_{pc_0} \rangle$ at pc_0, that is, the spring of the second system is compressed more at the beginning. (Notice that x and $\hat{x}$ are negative under compression, see figure 3.8(c).) When applying CEC* to this example, using RV constraints derived from theorems 6, 7, 9, and 11, we obtain a comparative envisionment of which a fragment is shown in figure 4.10.

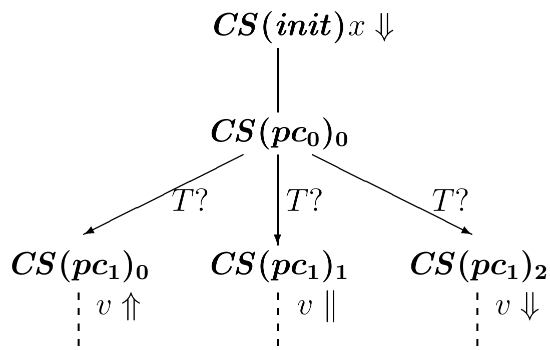


Figure 4.10: *Fragment of the comparative envisionment resulting from a predictive comparative analysis of two ideal mass-spring systems. The spring of the second system is initially compressed more.*

This result is too weak, as confirmed by the analytic solution of ideal mass-spring systems starting from rest:

$$v(t) = -\sqrt{\frac{k}{m}}x(t_0)\sin\sqrt{\frac{k}{m}}t \quad \text{and} \quad \hat{v}(\hat{t}) = -\sqrt{\frac{\hat{k}}{\hat{m}}}\hat{x}(\hat{t}_0)\sin\sqrt{\frac{\hat{k}}{\hat{m}}}\hat{t}.$$

The periods of the solutions are given by $P = 2\pi\sqrt{\frac{m}{k}}$ and $\hat{P} = 2\pi\sqrt{\frac{\hat{m}}{\hat{k}}}$. At pc_1 both systems have completed one quarter of their period, so

$$\hat{T} - T = \frac{\pi}{2}\left(\sqrt{\frac{\hat{m}}{\hat{k}}} - \sqrt{\frac{m}{k}}\right) \quad \text{and} \quad \hat{v}(\hat{t}_1) - v(t_1) = -\left(\sqrt{\frac{\hat{k}}{\hat{m}}}\hat{x}(\hat{t}_0) - \sqrt{\frac{k}{m}}x(t_0)\right),$$

where $T = t_1 - t_0$ and $\hat{T} = \hat{t}_1 - \hat{t}_0$ represent, as usual, the durations of the two behavior fragments between pc_0 and pc_1. Given that $k \parallel$, $m \parallel$, and $x \Downarrow$ at pc_0, one can immediately verify that $T \parallel_{pc_0 \to pc_1}$ and $v \Uparrow_{pc_1}$.

Although this answer is covered by the comparative envisionment, eight more (fragments of) comparative behaviors are generated. These behaviors are spurious, so CEC* is incomplete. □

By theorem 3, the incompleteness of CEC* has an immediate consequence:

Corollary 2 CEC* does not detect every inconsistency in the input.

However, CEC*'s soundness does allow one to draw a few useful conclusions (the proofs are simple and will be omitted):

Corollary 3 If no comparative behavior is generated by CEC*, then the input is inconsistent.

Corollary 4 If the input is consistent and CEC* generates a single comparative behavior, then this is the only genuine comparative behavior.

4.5.4 Global constraints against CEC*'s incompleteness

The observed incompleteness of comparative analysis is caused by the local nature of the propagation process in a comparative analysis. The constraints relating the RVs of variables at pair of comparison pc_i to those at a direct predecessor pair of comparison pc_j do not take into account global information about the behaviors of the two systems to be compared. The results of a comparative analysis can be improved by refining and extending the set of RV constraints on the basis of such information, analogously to strategies for countering the incompleteness of qualitative simulation (section 3.5.3) and the qualitative analysis of dynamical systems (e.g., Sacks [1990b]). Three different directions will be briefly explored here: (1) exploiting analytical solutions of the systems to be compared, (2) employing energy constraints, and (3) involving non-direct predecessor pairs of comparison in the propagation process.

Return to the counter example in the proof of CEC*'s incompleteness. Appendix A.3 shows that when comparing two mass-spring systems starting from rest one can prove that

$$\min\{T,\hat{T}\} = \min\{\frac{1}{4}P,\frac{1}{4}\hat{P}\} \leq \frac{1}{4}\bar{P} \leq \max\{\frac{1}{4}P,\frac{1}{4}\hat{P}\} = \max\{T,\hat{T}\}. \tag{4.26}$$

Evaluation of equation (4.13) in theorem 9 for $\hat{T} > T$ is now straightforward, since the sine and cosine functions are found to be positive on the whole interval $]t_0, t_1]$. Transformation of the resulting expressions to RV constraints yields:

if $RV(T, pc_0, pc_1) = \Uparrow$, then $RV(x, pc_1) > RV(x, pc_0) + RV(v, pc_0) + RV(k, pc_0) - RV(m, pc_0)$
if $RV(T, pc_0, pc_1) = \Uparrow$, then $RV(v, pc_1) > -RV(x, pc_0) + RV(v, pc_0) + RV(k, pc_0) - RV(m, pc_0)$

Similar constraints are obtained for the cases $\hat{T} = T$ and $\hat{T} < T$. Together with the constraint

$$RV(T, pc_0, pc_1) = RV(m, pc_0) - RV(k, pc_0)$$

from the periodicity of the linear mass-spring systems, CEC* now produces the single correct comparative behavior for the initial comparative state $CS(init) = \langle x \Downarrow_{pc_0}, v \parallel_{pc_0}, k \parallel_{pc_0}, m \parallel_{pc_0} \rangle$.

A similar use of analytical solutions helps in obtaining tighter RV constraints when comparing an ideal mass-spring system with a frictional mass-spring system. In the case that two non-linear mass-spring systems are considered, in which accelerations are defined as $a = -f(x)$ and $\hat{a} = -f(\hat{x})$, with f being an unspecified monotonically increasing function, RV constraint (4.25) cannot be further restricted. Actually, it is the tightest possible constraint given the weak information contained in the QDEs. The availability of mathematical packages with symbolic reasoning facilities makes the opportunity to integrate knowledge about analytical solutions of the systems to be compared more attractive.[10]

When two ideal mass-spring systems are compared with initial comparative state $CS(init) = \langle x \parallel_{pc_0}, v \parallel_{pc_0}, k \Uparrow_{pc_0}, m \parallel_{pc_0} \rangle$, an ambiguity arises at pc_1 with respect to the RV of v, even when the RV constraints refined above have been added. Viewed from an energetic perspective, the RV constraints do not capture the insight that in conservative systems like the ideal mass-spring systems, the sum of the potential and kinetic energy stored in the systems does not change during the behavior. By making energy functions (or more generally Lyapunov functions) explicit, one can introduce new RV constraints. These energy constraints may have been specified already, in order to allow QSIM to refine its predictions by pruning spurious qualitative behaviors (section 3.5.3). By means of explicit energy functions, CEC* succeeds in resolving the ambiguity by establishing that $v \Uparrow_{pc_1}$, that is, a higher stiffness of the spring results in a higher velocity when it passes through the rest position.

A further way to improve upon the incompleteness of CEC*, can be found in involving non-direct predecessor pairs of comparison in the propagation process. At pc_4 the mass-spring systems have returned to their initial state and the variables have attained exactly the same qualitative value. As a consequence, the RVs at pc_4 must equal those at pc_0, which can be captured in a simple RV constraint. More generally, exploiting information about the change in qualitative value of variables occurring between the current pair of comparison and those directly or indirectly preceding it enables one to formulate additional RV constraints. This approach to non-local RV propagation amounts to providing a generalization of theorem 11 and is reminiscent of the mechanism of corresponding values in qualitative simulation (section 3.3.1).

The exploitation of global information of various sorts can thus be pursued to prevent CEC* from generating spurious comparative behaviors. The clear mathematical foundation of CEC* facilitates the formulation and integration of such global RV constraints. It is not yet clear how this strategy affects the incompleteness of comparative analysis in general, that is, whether interesting classes of CA problems exist for which completeness

[10] This is especially so in the case of linear systems, like the mass-spring systems compared here. Other CA techniques also use information about the linearity of dynamical systems, but do so in an either ad-hoc or roundabout way. DQ analysis formulates a dedicated inference rule for the case of ideal harmonic oscillators (Weld [1990], prop. 10) and RSIM introduces the rather intricate notion of P-values enabling one to be more specific in the case of linearity constraints (Neitzke [1997]).

can be guaranteed. This seems an important topic to explore, especially as it is closely connected with a similar open problem encountered in the discussion of qualitative simulation. In conclusion, it should be emphasized that the results of incompleteness are not bound to be disastrous. The CEC* algorithm will proceed in the ordinary way, but may encounter a higher number of ambiguities and consequently generate a larger comparative envisionment.

4.5.5 Computational complexity of CEC*

A good metric of the computational complexity is the number of comparative states generated by CEC* in step 4 of the algorithm. In the worst case, we have to generate at each pair of comparison, for each predecessor pair of comparison, the maximum number of comparative states. That is, we may have to generate 3^{rq} comparative states, where q is the (maximum) number of variables shared by the two systems and r the maximum number of predecessor pairs of comparison. Notice that base 3 is a consequence of the three possible relative values for a variable ($\Downarrow$,$\|$,$\Uparrow$). The worst-case situation occurs when the RVs of the variables are not constrained in any way. The lack of constraints may arise from vacuous RV constraints and particular combinations of RVs for variables. Since there are n pairs of comparison, we have a computational complexity of the order $\mathcal{O}(n3^{rq})$.[11]

Although the computational complexity is exponential in theory, the average-case behavior is much more agreeable. The upper bound of 3^{rq} comparative states at a pair of comparison is very conservative, for one thing since it assumes that the variables are not constrained in any way, a rare event. More practically, for the examples that have been studied, CEC* generates the comparative envisionments within a few seconds on a SUN SPARCstation 5, while the code has not been optimized for run-time performance.

The number of comparative behaviors in an envisionment strongly depends on the nature of the differential initial conditions or differential responses in combination with structural differences. In particular, several differences having opposite tendencies, such as a higher inflow combined with greater orifices in the cascaded tanks, are sometimes found to lead to unwieldy comparative envisionments as a consequence of the ambiguities they introduce. I expect this problem to become more pertinent when larger systems are considered. It is important to emphasize that this is *not* a problem caused by incompleteness of the algorithm, but rather a consequence of the restriction to qualitative information about the structure and behavior of the physical systems to be compared.

As in qualitative simulation, additional information could be employed during the construction of the comparative envisionment to reduce the number of comparative behaviors. Two strategies for countering this problem, integrating new observations and prioritizing alternative comparative behaviors, were briefly mentioned in section 4.4.3. Further inspiration may be found in efforts to avoid intractable branching in qualitative simulation.

[11] It is here assumed that the first pair of comparison is also unconstrained; in other words, $CS(init)$ is effectively empty.

An especially promising strategy would be to extend CEC* to a form of semi-quantitative comparative analysis interacting with a semi-quantitative simulation algorithm. Using information about numerical bounds on parameter values and monotonic functions, this algorithm should be able to determine more precise intervals for the difference $\hat{q} - q$ (instead of the intervals $[-\infty, 0[$ and $]0, \infty]$ denoted by $\Downarrow$ and $\Uparrow$). In addition, the algorithm should guarantee that progressive refinement of the numerical bounds causes comparative analysis to converge to the trivial case of comparing two numerical solutions of the systems to be compared. Extending CEC* to a technique for performing *semi-quantitative* comparative analysis would enable one to put constraints on the net effect of opposing tendencies.

4.6 Related work on comparative analysis

Weld's *DQ analysis* (Weld [1990]) is one of the first attempts in QR to perform a comparative analysis of dynamical systems. CEC* resembles DQ analysis in a number of respects, notably in the formalization of comparative analysis and in the general approach of employing explicit constraints to propagate relative values across pairs of comparison. However, when going down to details the two techniques show considerable differences. DQ analysis has been designed for intra-model, predictive CA with topologically equal behaviors,[12] and its basic concepts are consequently less general than their counterparts in CEC*. In particular, DQ analysis uses specific types of relative values (relative change values), pairs of comparison (transition points), and theorems defining propagation constraints (DQ inference rules).

A major difference is DQ's ability to compare the RVs of variables *during* corresponding intervals between pairs of comparison, using the notion of *perspectives*. CEC* restricts its relative values to the time-points bounding the intervals. Though mathematically sophisticated, the comparison of variables during corresponding intervals has the disadvantage that RVs may be undefined during corresponding intervals, that is, be neither $\Uparrow$, $\Downarrow$, or $\|$ (proposition 21 in Weld [1990]). No answer is produced when such an undefined RV is encountered; the algorithm simply terminates. In addition, DQ analysis halts when an RV is ambiguous. As a consequence, in terms of section 4.5, DQ analysis is neither sound nor complete. However, *when* DQ analysis produces an answer, which is by definition an unambiguous answer (a single comparative behavior), it is guaranteed to be sound and complete. This matches with the guarantees for CEC* in such a situation (theorem 15 and corrolary 4). Of course, when ambiguities arise, CEC* also produces a sound answer in the form of alternative comparative behaviors. When ambiguities arise branching in the face of ambiguities is possible, because at a pair of comparison an RV is either $\Uparrow$, $\Downarrow$, or $\|$. This allows CEC* to produce a sound answer in the form of alternative comparative behaviors.

[12] Although DQ analysis is sometimes able to detect violations of the topological equality assumption (i.e., to detect inconsistencies in its input), it cannot proceed the analysis with a topologically different behavior.

The restriction of DQ analysis to intra-model CA has been relaxed by embedding the method into a framework for model sensitivity analysis. *Model sensitivity analysis (MSA)* (Weld [1992]) addresses the question how a change in model will affect the resulting behavior. In the MSA approach, a qualitative behavior arising from a simple model is compared with a behavior from a more complex model. The simple model must be a *fitting approximation* of the more complex model, that is, the complex model must have an independent variable, a fitting parameter, such that, when this parameter is taken to a certain approximation limit, the RVs of the shared variables become $\|$. An example of a fitting parameter is the coefficient of friction in the model of a block sliding down a plane. As this coefficient tends to 0, the sliding behavior with friction approaches the sliding behavior without friction. The advantage of introducing fitting approximations is that the *inter*-model CA problem can now under certain conditions be reduced to an *intra*-model CA problem in the complex model (Weld [1992], proposition 2.3). If we know how a perturbation of the fitting parameter in the complex model affects the RVs of other system variables (intra-model CA), we can derive RVs for the variables when the complex and the simple model are compared (inter-model CA).

Since the MSA approach reduces inter-model CA questions to intra-model CA questions that are solved by DQ analysis, it inherits the weaknesses of the latter technique. In addition, the limitation of having to formulate a fitting approximation makes the MSA approach not generally applicable to inter-model CA problems. In the formulation of the cascaded-tanks problem of figure4.1, for example, the leaky system does not differ from the watertight one by an independent variable. More basically, the reformulation of inter-model CA problems to intra-model CA problems by means of fitting approximations appears less straightforward than the relation between these two classes of problems achieved in CEC*. CEC* simply treats intra-model CA as a special case of inter-model CA within the same general framework.

DQ analysis and CEC* are *analytical* approaches, that is, approaches in which constraints on relative values at successive pairs of comparison are inferred from the characteristics of the underlying qualitative models and behaviors. The CA technique proposed by Chiu & Kuipers [1992] carries the analytical approach even further. It presents rules for algebraic manipulation and simplification by which the solution for a restricted class of intra-model CA problems can be directly obtained from the QDEs.

As an alternative to the analytical approaches towards comparative analysis, *simulation* approaches have been developed. Instead of inferring constraints that relate RVs at successive pairs of comparison from the qualitative models, model-independent transition rules are derived from the continuity properties of the variables in the qualitative behaviors. Weld's exaggeration technique EXAG is an early example of such an approach (Weld [1990]). EXAG is restricted to intra-model CA problems and predictive reasoning. Unlike DQ analysis, the technique is neither sound nor complete; even if it does produce an answer, this answer is not guaranteed to be correct.

Simulation is also used in the CA technique RSIM (Neitzke [1997]; Neitzke & Neumann [1994]). RSIM integrates comparative analysis with qualitative simulation into a *relative simulator* by adding relative descriptions to the qualitative descriptions of a variable. The relative descriptions record deviations of a value from an implicit reference value. In addition to the transition rules of QSIM, which provide successors of qualitative values, special transition rules are defined by which possible successors of relative values can be found. A change in the qualitative *or* relative description of a variable gives rise to a new state in the simulation. Successive states are connected in an extended behavior tree, which forms the output of RSIM.

The operation of RSIM can be interpreted as generating the Cartesian product of (1) all possible qualitative behaviors and (2) all possible comparative behaviors arising from the comparison of each qualitative behavior with an implicit reference behavior. In contrast, CEC* separates the simulation and comparative analysis activities and takes two explicit qualitative behaviors as input. The latter approach has the appeal of conceptual clarity.[13]

Advantages of RSIM are that it handles changes in behavioral topology and that the transition rules are not coupled to qualitative models. The algorithm seems to be sound and incomplete, in the terminology of section 4.5. A disadvantage of the method is the potentially large number of branches in the extended behavior tree, which may lead to time-consuming computations and to extensive and sometimes confusing output (Neitzke [1997]). Further, RSIM is only suitable for predictive reasoning and, with the exception of differences in the monotonicity of functional relations, requires the two systems to be compared to have identical models. Thus, unlike CEC*, RSIM is restricted to intra-model comparative analysis.

4.7 Evaluation

CEC* is a general technique for predictive and explanatory comparative analysis which is able to handle the comparison of structurally different and structurally identical systems within a single formalism. With the exception of Weld's MSA, inter-model CA problems have not been addressed in the literature. Since the comparative analysis of experimental systems is concerned with general problems of this type, CEC* seems to be an excellent tool for elaborating model-based conflict resolution and error identification. Experimental systems may be structurally different as a consequence of differences in the performance of an experiment. The next chapter will explain in more detail how qualitative simulation and comparative analysis can be used to formalize model-based measurement analysis.

The mathematical foundation of the technique permits one to relate its concepts to the theory of differential equations, similar as in QSIM. Thus I have been able to define what should be understood by the correctness of comparative analysis and prove that CEC* is sound but incomplete. CEC* guarantees a sound answer to every CA question, but the usefulness of the answer can occasionally be undermined by the occurrence of spurious comparative behaviors. A few examples of the integration of global information to eliminate spurious behaviors were given.

[13] As an extension to the basic algorithm, Neitzke describes how an explicit reference behavior can be taken into account (Neitzke [1997], ch. 7).

Even if the comparative envisionment contains only genuine comparative behaviors, their sheer number may sometimes complicate an interpretation of the effects of experimental conditions deviating from the conditions expected in an ideal experiment or the causes of a conflict between two measurements. Practical experience with the techniques in the case-study of chapter 7 will give an indication of how serious this problem is in practice. It should be kept in mind, however, that it is not properly considered a failure of the *technique*. Rather, it is a consequence of the *lack of information* about the experimental systems being compared. In section 4.5.5, supplementary techniques were mentioned which reduce the number of comparative behaviors through 'graceful extension' (Forbus [1984]) of the available information about the experimental systems.

Chapter 5

Measurement Analysis System

Chapter 2 gave an outline of methods to accomplish the tasks of conflict detection, conflict resolution, and error identification in a systematic way. This chapter will formally define the methods and describe how they have been implemented in a measurement analysis system. The techniques for the qualitative simulation and comparative analysis of experimental systems will be seen to play a key role in the formalization of model-based conflict resolution and error identification (de Jong et al. [1998]; de Jong, Mars & van der Vet [1998]).

To a varying degree, all of the methods employ knowledge about experimental systems, the experimental conditions imposed upon these systems, and the properties that were measured. A formal specification of this knowledge in section 5.1 will precede the elaboration of the methods in the next three sections. Section 5.5 describes the KIMA system for measurement analysis. The chapter closes with a brief evaluation of the methods and a look ahead at the application of the KIMA system in the case-study (section 5.6).

5.1 Measurements of properties

5.1.1 Experimental systems

In an experiment scientists create and sustain an experimental system on which they perform measurements to determine properties of physical entities. These controlled physical systems are conceived of as dynamical systems that can be described in (qualitative) mathematical terms, as explained in the previous two chapters. More specifically, experimental systems are modeled by qualitative differential equations.

The possible models that can be used to describe an experimental system investigated in a particular type of experiment are contained in a *model space*. The models in a model space are mutually exclusive in the sense that no system can be adequately described by two models from the model space at the same time; the models represent experimental systems with a different structure.

Definition 27 (Model space) For every experiment of type et a finite set $SPACE_{et}$ of

QDEs is defined. $SPACE_{et}$ is called the model space of the experimental systems created in the type of experiment.

The definition asserts a model space to be finite and tied to an experiment of a certain type. The first characteristic is basically a relevance decision. Implicitly, the model space determines which models are relevant for the conflict resolution or error identification problem at hand. A criterion for the dividing line between relevant and irrelevant could be borrowed from the experimenters themselves; only those aspects of the systems mentioned in critical discussions of the experiments need to be taken into account. This approach has been followed in the materials science case-study in part III. The second characteristic is straightforward; a physical system created in one type of experiment cannot usually be described in an adequate way by a model of a system created in another type of experiment.

The structure of a particular experimental system will be described by a subset of an applicable model space. This subset, the set of *candidate models*, comprises alternative models of the system. Although at most one of the candidate models forms an adequate description of the system, there is not enough information to discard the others. For instance, if no information is provided about the kind of sample holder used in a melting temperature experiment, we cannot decide whether we have to describe the system by a QDE accounting for reactions between the sample and holder or not. Each candidate model is accompanied by *candidate experimental conditions*.

Definition 28 (Experimental system) An experimental system investigated in an experiment of type et is described by a set CM of tuples $\langle QDE, QS(init)\rangle$ of candidate models and candidate conditions. The candidate models are drawn from the model space for et, i.e. $QDE \in SPACE_{et}$.

Notice that this description is a short-cut, representing the outcome of a complex model-building process. For instance, in selecting a subset of the model space as candidate models of the experimental system, we are taking for granted unarticulated knowledge about the execution of the experiment. In chapter 8 I will return to this topic and discuss how the present approach could be extended by explicating background knowledge about the experiments and experimental systems.

5.1.2 Measurements and measured states

The description of the quantities of an experimental system follows the QSIM convention.

Definition 29 (Quantities) The quantities of an experimental system are described by a vector $\boldsymbol{q}$ of variables which are reasonable functions of time.

There is an obvious relation between the QDEs in the model space of an experimental system and the vector $\boldsymbol{q}$ used to describe the quantities of the system: the variables occuring in the QDEs are elements of $\boldsymbol{q}$.

Measurements are empirically determined values of quantities of an experimental system at a certain time-point. A measurement provides a numerical magnitude relative to some unit of measure.

Definition 30 (Measurement) Given an experimental system with quantities $\boldsymbol{q}$ which is investigated on a time-interval $[a, b]$. A measurement of a quantity q in $\boldsymbol{q}$ at time-point u, $a \leq u \leq b$, is described by a tuple

$$MV(q, u) = \langle nval, um \rangle.$$

nval is a numerical value tuple and *um* is drawn from a set *UM* of constants referring to units of measure.[1] A numerical value tuple is a tuple $\langle m, \sigma_m \rangle$ of two real numbers, standing for the estimated mean (m) of a population of values for q and the standard error of the mean (σ_m).

Usually, several quantities are measured at a time-point, which gives rise to the concept of *measured state*. In many cases only a subset of the quantities of an experimental system will have been measured, for the reason that measurements of some of the quantities are deemed irrelevant, too difficult, or even impossible to obtain.

Definition 31 (Measured state) Assume that $\boldsymbol{q}$ is the vector of quantities of an experimental system considered on the interval $[a, b]$. Let $\boldsymbol{p}$ be a vector of quantities contained in $\boldsymbol{q}$ and let $\boldsymbol{p}$ have k elements. A measured state is a tuple of m measurements at a particular time-point $u \in [a, b]$:

$$MS(\boldsymbol{p}, u) = \langle MV(p_1, u), \ldots, MV(p_k, u) \rangle.$$

When the quantities of the system are measured at successive time-points $u_0, u_1, \ldots, u_m$ on $[a, b]$ we have a sequence of measured states:

Definition 32 (Measured state sequence) Assume that an experimental system considered on an interval $[a, b]$. A measured state sequence is a tuple of measured states of the experimental system. The measured states have been determined at successive time-points $u_0, u_1, \ldots, u_m$ on $[a, b]$:

$$MSS(\boldsymbol{p}, a, b) = \langle MS(\boldsymbol{p}, u_0), MS(\boldsymbol{p}, u_1), \ldots, MS(\boldsymbol{p}, u_m) \rangle.$$

[1] In line with the EngMath ontology for engineering mathematics, a unit of measure is conceived of as a constant quantity (Gruber & Olsen [1994]).

For simplicity, I will assume that $a = u_0$ and $b = u_m$. That is, the interval on which we will consider the system is bounded by the first and the last measurement.

For our purposes it will be necessary to map the time-points $u_0, \ldots, u_m$ of the measured states of an experimental system to the distinguished time-points $t_0, \ldots, t_n$ of a qualitative behavior (section 5.3.1). This is not straightforward, since the former time-points do not need to coincide with the latter. The mapping will be effected in an indirect way, by adding a qualitative state description to a measured state. The qualitative state description specifies the qualitative value of some of the variables in $\boldsymbol{q}$ at the instance of measurement.

Definition 33 (Extended measured state) An extended measured state at a time-point $u \in [a, b]$ is a measured state as in definition 31, supplemented by a qualitative state description $QS_{meas}(u)$:

$$MS(\boldsymbol{p}, u) = \langle MV(p_1, u), \ldots, MV(p_k, u), QS_{meas}(u) \rangle.$$

$QS_{meas}(u)$ is a tuple of qualitative values.

For instance, when we start with loading a material in a fracture test the applied stress σ_a will be 0 and increasing. As a consequence, the qualitative value $\langle o, inc \rangle$ for σ_a is included in QS_{meas}.

5.1.3 Properties and property measurements

The performance of a measurement in an experiment is directed at the determination of a property of the physical system. Measuring a property amounts to determining the value of a certain quantity when the experimental system has been brought into a certain state. For instance, the melting temperature property of a material is defined as the value of the temperature quantity when the material starts to melt. The quantity to be measured when determining a property will be called the *definition quantity* and the state in which the experimental system is required to be the *definition state*.[2]

Definition 34 (Property) A property is described by a tuple consisting of a type of experiment et, a definition quantity q_{def}, and a definition state QS_{def}, that is, a tuple

$$\langle et, q_{def}, QS_{def} \rangle.$$

The definition state QS_{def} equals a tuple of qualitative values for selected quantities or a symbol *FIN* referring to the final state of the experimental system.

[2]To further characterize a property, one could extend the definition state of a property with a *definition behavior fragment* of states that should precede the definition state. I will omit this refinement here.

Following this definition, the *value of a property* will be interpreted as the value of the definition quantity when the experimental system is in the definition state. For simplicity, I assume that there is only a single definition state for a property. Since different kinds of physical processes may occur in experimental systems investigated in different types of experiment (section 2.1.3), the definition of a property should be tied to an experiment of a certain type.

Notice that the definition is based on a *qualitative* characterization of the definition state. The state at which the definition quantity is to be measured will usually be initiated by some distinct and recognizable event, like a material sample beginning to melt. If landmarks referring to this event are included in the quantity spaces of the relevant variables, a qualitative description of the definition state will be sufficient to specify when the definition quantity is to be measured.

A definition state is usually partially specified in that it does not contain a qualitative value for every quantity of an experimental system on which measurements are conducted. In the extreme case no qualitative values are specified and the definition state is directly specified as the final state *FIN* of the experimental system being investigated. I will restrict the analysis to this case while observing that extension to the general case is not difficult to achieve.

The definition brings out that a property can be measured under different experimental conditions, as long as we ensure that the experimental system has been brought into the required definition state. Since the definition state covers only a subset of the quantities, the property might be measured for different values of the remaining quantities. In addition, only qualitative restrictions are put on the state in which the experimental system has to be brought, so that in general a variety of numerical values for the quantities in this state will be compatible with the restrictions. Suppose the fracture strength is defined as the external stress applied to the material when the maximum stress at the crack-tip reaches the theoretical strength. One can then measure the fracture strength at different temperatures, different grain sizes and porosities, different geometries of the specimen, etc.

A *property measurement* is the measurement of a particular property in an experiment. The measurement bases analyzed in this thesis are conceived of as sets of property measurements.

Definition 35 (Property measurement) A property measurement is described by a tuple

$$\langle p, es, MSS \rangle,$$

with p a property, es an experimental system, and MSS a sequence of measured states of the experimental system.

5.2 Conflict detection

In section 2.2.1 the detection of conflicts in a measurement set was described as consisting of the pairwise comparison of property measurements to see whether they are incompatible and the subsequent storage of conflicts in a conflict matrix (figure 5.1). Here I will elaborate this method by first discussing criteria for deciding whether two property measurements are in conflict and then giving a simple algorithm for conflict detection.

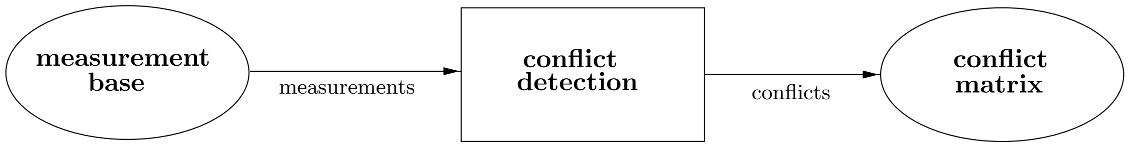


Figure 5.1: *Schematic overview of conflict detection. In this figure, and in the other figures of this chapter, boxes represent processes, ovals knowledge sources, and arrows information flows.*

5.2.1 Conflict criteria

A property measurement contains a sequence of measured states, one of which provides a measured value of the property. The property value takes the form of an estimation of the mean of a population of values of the definition quantity q_{def} when the system has been brought in the definition state QS_{def} under certain experimental conditions. A conflict between two property measurements exists when it is unlikely that the two underlying populations have the same mean value, and hence unlikely that the measurements *both* represent the value of the property at the given conditions. The difference between the two reported mean values is too large to be solely attributed to random errors.

In order to be able to say that two (individual or aggregated) measurements are in conflict, we need a *conflict criterion*. I will use a basic statistical criterion (Barry [1978]). Suppose that the measurements have been specified as an estimation m of the population mean supplemented by a standard error σ_m or standard deviation σ_s. (The standard error can be immediately calculated from the standard deviation of a sample of k measurements: $\sigma_m = \sigma_s/\sqrt{k}$.) If it is additionally assumed that the populations of measured quantity values have a normal distribution, a reasonable assumption for measurements in science and engineering, then we have the following definition of a conflict:

Definition 36 (Conflict) Two property measurements pm_1 and pm_2 determine the value of a property p in terms of estimated population means m_1, m_2 and standard errors $\sigma_{m,1}$, $\sigma_{m,2}$. They are in conflict with probability $Pr(d)$ if the following relation obtains:

$$\mid m_1 - m_2 \mid = d\sqrt{\sigma_{m,1}^2 + \sigma_{m,2}^2}, \qquad d \geq d_{limit},$$

for some value of d_{limit}.

The probability $Pr(d)$ is determined from the normal distribution function. For $d = 2$, $Pr(d)$ is approximately 95%, and for $d = 3$ approximately 99% (Barry [1978]). The choice of d_{limit} fixes the threshold for speaking about a conflict. For $d_{limit} = 2$, we need to be at least 95% sure before we can conclude that the difference between m_1 and m_2 is significant, that is, not due to random deviations from a common population mean.

The statistical conflict criterion is applicable under the assumption that the standard errors or standard deviations are known and the quantity has a normal distribution (or another distribution for which a similar criterion can be formulated). When these conditions are not satisfied in practice, we will have to take recourse to heuristic criteria, such as testing whether the confidence intervals of pm_1 and pm_2 arrived at by informal judgment overlap. A heuristic evaluation of conflicts in data sets can be recognized in the reviews of the melting temperatures of refractory oxides (Hlaváč [1982]; Coutures & Rand [1989]).

5.2.2 Algorithm for conflict detection

The task of conflict detection is directed at finding all conflicts between property measurements in a measurement set. The simple algorithm introduced below accomplishes this task by retrieving pairs of property measurements from the measurement set, applying a suitable conflict criterion, and storing the results in a conflict matrix.

In order to find out whether one property measurement is in conflict with another, the measured values of the property have to be retrieved from the information contained in the measured states of the experimental systems (definition 35). More particularly, the measured value of a property equals the value of the definition quantity in a measured state consistent with the definition state of the property. From the sequence $MSS(\boldsymbol{p})$ of measured states we have to select a measured state $MS(\boldsymbol{p}, t)$ which (1) contains the measured value $MV(q_{def}, t)$ of the definition quantity q_{def}, and which (2) matches the requirements imposed on it by the definition state QS_{def}. I will assume that such a measured state can always be found.

Although the other measured states in $MS(\boldsymbol{p}, t)$ do represent measurements performed on the experimental system, they are *not* measurements of the property in question. For example, in the case of a fracture strength measurement we are interested in the value of the applied stress when a pre-existing crack starts to propagate and *not* when the process of elongating the sample starts, at the very beginning of the experiment.

The conflict criterion of definition 36 is used to decide whether two measured values of a property are in conflict. The result of this comparison of property values is represented by a cell in the *conflict matrix*.

Definition 37 (Conflict matrix) Given a measurement base PM. The conflicts in PM can be represented in a conflict matrix C_{PM}, where each element $C_{PM}(i, j)$ denotes whether property measurements pm_i and pm_j are in conflict. The values for $C_{PM}(i, j)$ are tuples of the form $\langle c, d_{limit} \rangle$, where c is a symbol drawn from the set $\{conf, noconf, undet\}$ and d_{limit} is the threshold for detecting a conflict by means of the statistical criterion. Of course, $C_{PM}(i, j)$ equals $C_{PM}(j, i)$.

The following simple algorithm for conflict detection finds all conflicts between property measurements in a measurement set.

Algorithm 3 (Conflict detection) Given a measurement base PM of n property measurements and an initial conflict matrix C_{PM} in which all cells contain the value *undet*. The conflicts in the measurement set can now be determined as follows:

Step 1 Take two property measurements pm_i and pm_j from PM which have not yet been tested for conflicts. If all pairs of property measurements have been tested, then stop.

Step 2 Retrieve the measured values MV_i and MV_j of the property from the measured states specified in pm_i and pm_j.

Step 3 Determine whether MV_i and MV_j are in conflict according to the statistical criterion. Store the result in $C_{PM}(i,j)$ and $C_{PM}(j,i)$, and return to step 1.

5.3 Conflict resolution

The task of conflict resolution is directed at finding possible explanations of a conflict between two property measurements. A model-based method for performing this task has been sketched in section 2.2.2 and will be elaborated below by means of the QR techniques of the previous two chapters.

Figure 5.2 summarizes the basic steps in the method. The property measurements state which candidate models and candidate conditions describe the two physical systems created and controlled in the experiments. The candidate models are simulated to obtain the candidate qualitative behaviors of the systems. The method proceeds by comparing every combination of a candidate model and candidate behavior of the first experimental system with every combination of a candidate model and candidate behavior of the second system. This comparison takes the form of an explanatory comparative analysis in which (initial) RVs are supplied by the measured states of the experimental system and by the user. The alternative comparative behaviors in the comparative envisionments thus produced represent explanations of the conflict which are consistent with the information provided by the property measurements. I will prove that all explanations of a conflict are found, but that some of the explanations may be spurious.

5.3.1 Candidate behaviors of an experimental system

An experimental system investigated in a particular experiment is described by a set CM of tuples of a *candidate model* and *candidate experimental conditions* (definition 28). Each tuple represents a different possible structure of the system and a different possible set of experimental conditions. For every pair $\langle QDE, QS(init)\rangle \in CM$, one can predict a set of *candidate behaviors* of the experimental system by performing a qualitative simulation and then eliminating the qualitative behaviors which do not agree with

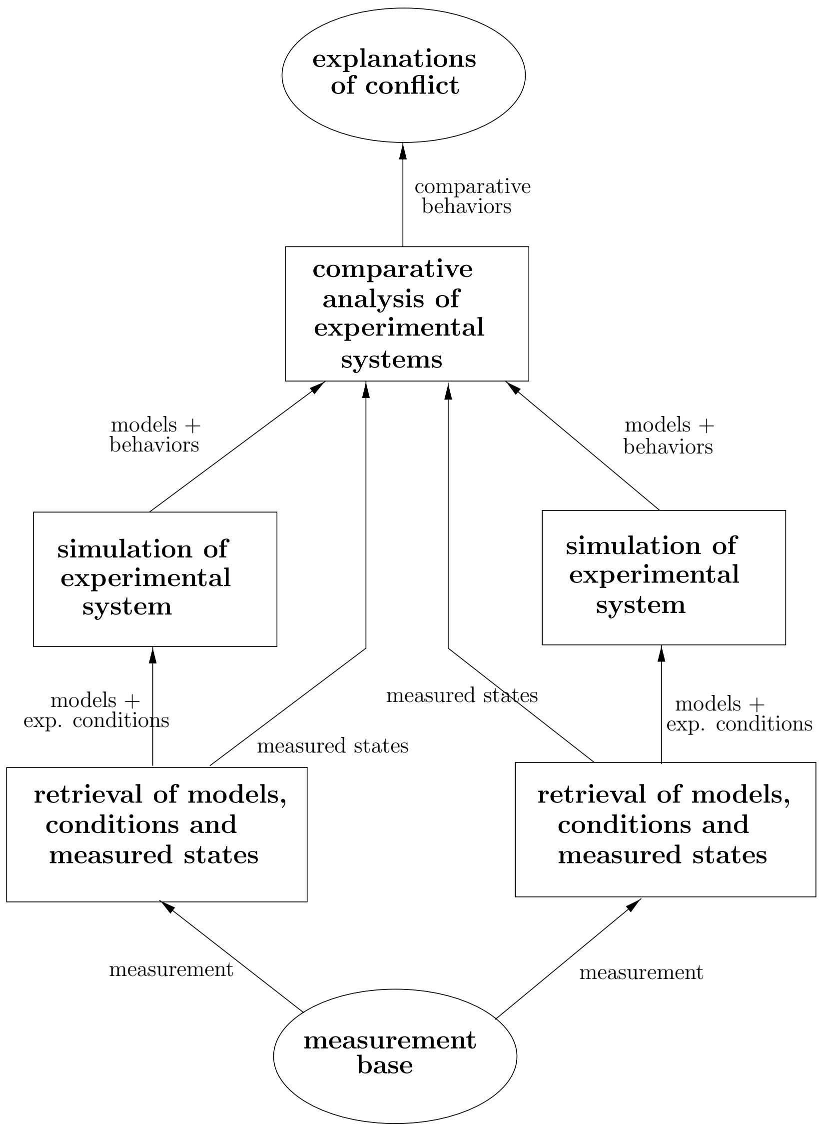


Figure 5.2: *Schematic overview of model-based conflict resolution.*

the measured states of the system. The candidate behaviors are alternative qualitative descriptions of the behavior of the experimental system. I will use QSIM to derive the behaviors from QDE and $QS(init)$.

In order to count as a candidate behavior of the experimental system, a behavior resulting from simulation must be consistent with the measured states of the system. A sequence MSS of (extended) measured states $MS(u_0), \ldots, MS(u_m)$ at consecutive time-points $u_0, \ldots, u_m$ invalidates a potential candidate behavior QB when it conflicts with the sequence of qualitative states $QS(t_0), \ldots, QS(t_n)$ in QB. In particular, we need to check whether the sequence $QS_{meas}(u_0), \ldots, QS_{meas}(u_m)$ of qualitative state descriptions agrees with $QS(t_0), \ldots, QS(t_n)$. This problem, which lies at the heart of work on measurement interpretation (Forbus [1987]; DeCoste [1991]; Dvorak & Kuipers [1991]), has been solved by the simple algorithm described in appendix C.

The information about the experimental system provided by the measured states may occasionally rule out *all* qualitative behaviors for a particular candidate model QDE, that is, it may render the set of candidate behaviors empty. Assuming that the measurements are correct, this implies that the experimental system is not adequately represented by QDE and $QS(init)$, and we must remove the tuple $\langle QDE, QS(init)\rangle$ from the set CM of candidate models.

5.3.2 Explanations of a conflict

Suppose we have taken two conflicting property measurements from a measurement set PM. From the information contained in the property measurements we can derive sets CMB and $\hat{CMB}$ of tuples of a candidate model, candidate conditions, and a candidate behavior. These tuples $\langle QDE, QS(init), QB\rangle$ and $\langle \hat{QDE}, \hat{QS}(init), \hat{QB}\rangle$ are called *candidate model-behavior triples*.

By definition 34, a property value is the empirically determined value of a definition quantity q_{def} when the experimental system is in a certain state, the definition state QS_{def}. To simplify matters, attention is restricted to properties whose definition state coincides with the final time-point of the interval on which the experimental system is considered. Differently put, the definition state corresponds with the final state in the qualitative behavior of the experimental system.

A conflict between two measured property values can consequently be viewed as a differential final response of the physical systems. And conflict resolution amounts to performing an explanatory comparative analysis to find possible causes of the differential response. To be more precise, the resolution of a conflict requires one to perform an explanatory comparative analysis for *every combination* of candidate model-behavior triples from CMB and $\hat{CMB}$. Different candidate model-behavior triples imply different assumptions about the structure and behavior of the experimental system and might lead to different explanations of the conflict.

An explanatory CA starts with a set $CS(init)$ of RVs at the last pair of comparison (section 4.4). In the case of conflict resolution, the initial RVs can be derived from the measured states of the experimental systems. The final time-points of the intervals $[a, b]$

and $[\hat{a}, \hat{b}]$ on which the experimental systems are investigated correspond with the last pair of comparison pc_n. A comparison of the measured values of quantities at this pair of comparison yields the RVs in $CS(init)$. If the values of a quantity q are significantly different according to the criterion in definition 36, we obtain $q \Downarrow_{pc_n}$ or $q \Uparrow_{pc_n}$, depending on the sign of the difference. If the values are not significantly different, we have $q \|_{pc_n}$. Additional RVs in $CS(init)$ may be supplied by the user.

The conflict between two measured property values will appear as a relative value $\Uparrow$ or $\Downarrow$ for q_{def} in the initial set of RVs $CS(init)$. For each combination of a candidate model and a candidate behavior of the first system and a candidate model and a candidate behavior of the second system, CEC* will return a comparative envisionment. The comparative behaviors in the envisionment explain how differences at the first pair of comparison pc_0 can account for the observed differences at the last pair of comparison pc_n. This suggests how the notion of *explanation* of a conflict can be defined in a precise way.

Definition 38 (Explanation of conflict) Given two conflicting property measure ments with sets of candidate model-behavior triples CMB and $\hat{CMB}$ describing the experimental systems investigated. Suppose an explanatory CA is carried out involving a model-behavior triple $\langle QDE, QS(init), QB \rangle \in CMB$ and a triple $\langle \hat{QDE}, \hat{QS}(init), \hat{QB} \rangle \in \hat{CMB}$. In addition, we have a set $CS(init)$ of initial RVs at the last pair of comparison pc_n. The comparative envisionment constructed in a comparative analysis has comparative behaviors $CB_1, \ldots, CB_m$.

Each comparative behavior CB_k, $1 \leq k \leq m$, in conjunction with the candidate models QDE, $\hat{QDE}$ and candidate behaviors QB, $\hat{QB}$, will be called a possible explanation of the conflict.

This definition matches the informal account of explanations of a conflict in chapter 2. The cause of the conflict between two property measurements, it was said, must be sought in differences in the structure of the experimental systems, in differences in the experimental conditions imposed upon the systems, or in a combination of these. Differences in the experimental conditions appear as different qualitative behaviors or as relative values $\Uparrow$ or $\Downarrow$ for quantities in the comparative state at pc_0. A structural difference occurs when the candidate models of the two systems being compared are unequal. The comparative behavior explains the conflict by showing how differences in models and behaviors and differences in initial conditions lead to the observed discrepancy between the values of q_{def} at pc_n.

The explanations of a conflict must be consistent with the measured states of the experimental systems. That is, we must verify for each comparative behavior CB whether it is consistent with RVs derived from the sequences MSS and $\hat{MSS}$ of measured states. In appendix C an algorithm for performing this consistency check is given. Comparative behaviors which are not consistent with the measured states are eliminated.

5.3.3 Algorithm for conflict resolution

The qualitative simulation of candidate models of experimental systems, and the comparative analysis of the models and the resulting candidate behaviors, form the main elements of the model-based method for conflict resolution. How exactly these elements mesh together is summarized in the following algorithm: algorithm

Algorithm 4 (Conflict resolution) Given two property measurements which are in conflict according to the conflict matrix. The property measurements specify sets CM and $\hat{CM}$ of tuples of candidate models and candidate conditions of the experimental systems and sequences MSS and $\hat{MSS}$ of measured states. The possible explanations of the conflict are produced as follows:

Step 1 Perform a qualitative simulation for every tuple $\langle QDE, QS(init)\rangle \in CM$ and every tuple $\langle \hat{QDE}, \hat{QS}(init)\rangle \in \hat{CM}$. Call the resulting sets of candidate model-behavior triples CMB and $\hat{CMB}$.

Step 2 Use the measured states MSS and $\hat{MSS}$ to reduce the sets of candidate model-behavior triples CMB and $\hat{CMB}$. Candidate behaviors and models that are not consistent with the measurements are eliminated.

Step 3 Perform an explanatory comparative analysis for every combination of a candidate model-behavior triple $\langle QDE, QS(init), QB\rangle \in CMB$ and a candidate model-behavior triple $\langle \hat{QDE}, \hat{QS}(init), \hat{QB}\rangle \in \hat{CMB}$. The initial comparative state information $CS(init)$ is obtained from the measured states MSS and $\hat{MSS}$, and from the user.

Step 4 Eliminate comparative behaviors in the envisionments that are not consistent with the measured states MSS and $\hat{MSS}$.

Step 5 Return the consistent comparative envisionments generated in the previous steps. Each comparative behavior in a comparative envisionment represents a possible explanation of the conflict.

A comparative envisionment returned by the algorithm may occasionally be empty, that is, reveal an inconsistency in the input (corollary 3). The inconsistency implies that the particular combination of candidate models QDE, $\hat{QDE}$, candidate behaviors QB, $\hat{QB}$, and measured states MSS, $\hat{MSS}$ does not adequately describe the experimental systems, and hence cannot explain the conflict. If it is assumed that the measurements are correct, we must conclude that at least one of the candidate models QDE, $\hat{QDE}$ or candidate behaviors QB, $\hat{QB}$ does not form an adequate description of the structure or behavior of the experimental systems. However, without additional information it is not possible to tell which models and/or behaviors are to blame.

An extreme case of inconsistency in the input of CEC* remains to be discussed, namely when *all* comparative envisionments produced by the algorithm for conflict resolution are empty. That is, there are *no* explanations for the conflict between the property measurements. The only way to account for this inconsistency, given that the measurements are correct, is to conclude that *none* of the candidate models and candidate conditions of the first experimental system, the second experimental system, or both are acceptable. In what follows I will leave this type of inconsistency out of consideration (but see chapter 8).

5.3.4 Properties of the algorithm

Which guarantees can be given on the outcome of the conflict resolution algorithm? Intuitively, one would demand from a possible explanation of the conflict that it does not rest upon a spurious comparative behavior of the experimental systems, that is, a comparative behavior that does not adequately describe the differential dynamics of the systems. The reason is obvious: it must be possible for the hypothesized cause to actually produce the conflict. This consideration leads to the definition of *genuine* and *spurious explanations* of a conflict.

Definition 39 (Spurious explanation) Given a possible explanation of a conflict consisting of the qualitative models QDE and $\hat{QDE}$, qualitative behaviors QB and $\hat{QB}$, and a comparative behavior CB. The explanation is spurious, iff CB is a spurious comparative behavior. A genuine explanation is an explanation that is not spurious.

Recall that a spurious comparative behavior may arise as a consequence of the spuriousness of one the qualitative behaviors being compared (definition 23).

Using the theorems on the soundness and incompleteness of both QSIM and CEC* (theorems 4, 5, 15, and 16), a clear statement on the correctness of the algorithm for conflict resolution can be made.

Theorem 17 The algorithm for conflict resolution generates all genuine explanations of a conflict.

Theorem 18 The algorithm for conflict resolution may occasionally generate spurious explanations of a conflict.

The proofs are obvious consequences of the fact that QSIM and CEC* generate all genuine qualitative and comparative behaviors of dynamical systems, but occasionally spurious behaviors as well.

The computational complexity of the conflict resolution algorithm is determined by the computational complexity of CEC*. The number of comparative states generated by the algorithm, a good metric of its complexity, is the product of the number of comparative envisionments generated and the number of comparative states generated to construct each of them. The first number is a product of the sizes of the sets CMB and $\hat{CMB}$, whereas the second number is of the order $\mathcal{O}(n3^{rq})$, with n representing the number of pairs of comparison, q the (maximum) number of shared quantities of the systems, and r the maximum number of predecessors in the OPC structure (section 4.5.5). Clearly, the computational complexity of conflict resolution is exponential. The average-case behavior of the algorithm will be less severe for the reasons given in section 4.5.5.

5.4 Error identification

Although conflict resolution suggests possible explanations of a conflict between two measurements of a property, it does not pronounce a judgment on the accuracy of the individual measurements. At least one of them is inaccurate, that is all one can conclude from a conflict. In section 2.2.3, a model-based method for analyzing the accuracy of a property measurement has been proposed, a method by which the presence and direction of a systematic error in the measured value of a property can be assessed. As in the case of conflict resolution, the qualitative simulation and comparative analysis techniques will be used to elaborate the method.

Figure 5.3 gives an outline of the basic steps in model-based error identification. The property measurement whose accuracy is to be evaluated specifies a set of candidate models of the experimental system and candidate experimental conditions. In the manner discussed above, each candidate model can be simulated to obtain a set of candidate behaviors. This property measurement is compared with an ideal property measurement, specifying an ideal experimental system and ideal experimental conditions. The model and behavior of the ideal experimental system are compared with every combination of a candidate model and candidate behavior of the experimental system actually investigated. The comparative analysis results in alternative comparative behaviors which predict an RV for the definition quantity at the last pair of comparison. An RV equal to $\Downarrow$ or $\Uparrow$ represents a possible systematic error in the measured property value. It is not difficult to prove that any systematic error in the measured value is predicted, although the method is not guaranteed to exclude spurious systematic errors.

5.4.1 Ideal experimental systems

In order to evaluate the quality of a measurement one needs a reference, a standard to which it can be compared. In chapter 2, I have introduced such a standard in the guise of a measurement obtained in an *ideal experiment*. In an ideal experiment an *ideal experimental system* is realized, that is, an experimental system with an ideal structure and an ideal behavior exhibited when the system evolves under ideal conditions. What counts as 'ideal' is dependent upon the aim of the measurement, e.g., the approximation of an idealized theoretical situation. Recall that it may be impossible to carry out an ideal experiment, so that an ideal experimental system is best conceived of as a hypothetical system and the ideal value of the property a hypothetical value.

In terms of the definitions in section 5.1.1, the property measurement obtained in an ideal experiment refers to an ideal experimental system with a structure QDE investigated under ideal experimental conditions $QS(init)$.

5.4.2 Predictions of systematic errors

Systematic errors in the measured value of a property can be detected by comparing the property measurement to be evaluated with the ideal property measurement. Let $\hat{CMB}$ be the set of candidate model-behavior triples of the experimental system actually

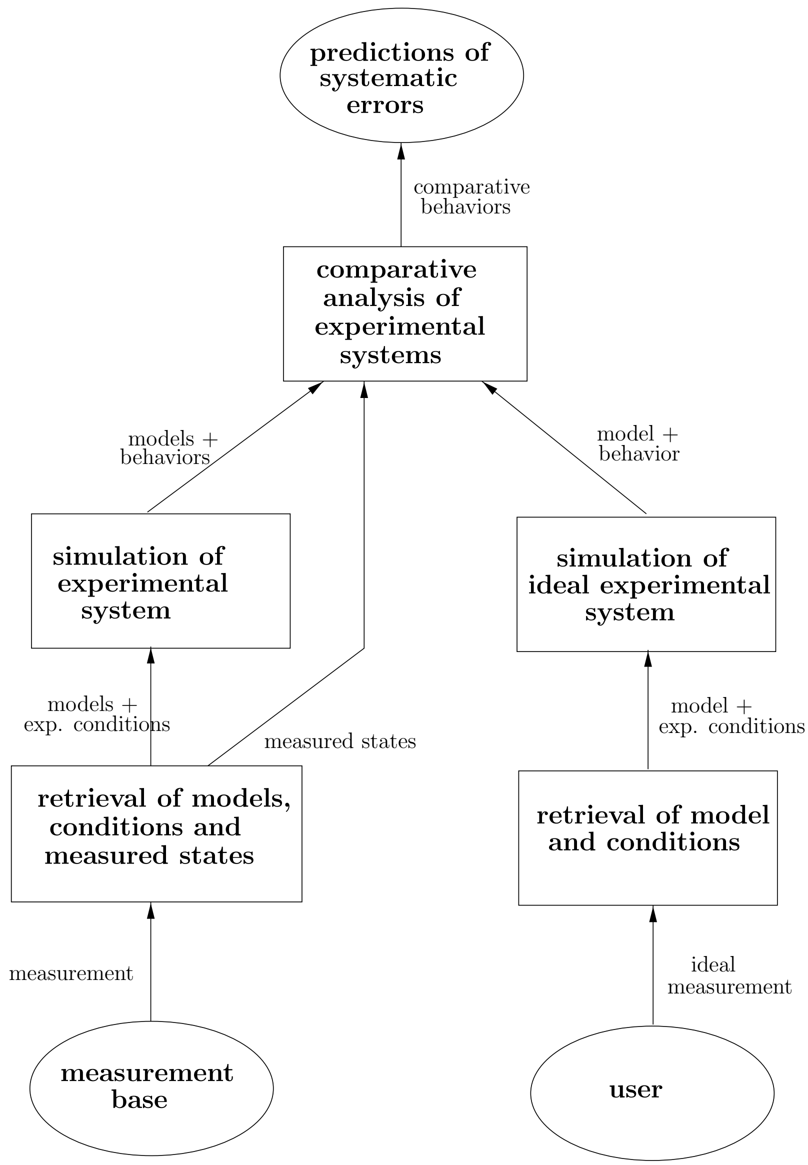


Figure 5.3: *Schematic overview of model-based error identification.*

investigated. $\hat{CMB}$ can be obtained through qualitative simulation of the candidate models of the system under given experimental conditions, as described in section 5.3.1. Further, let $\langle QDE, QS(init), QB\rangle$ represent the structure, initial state, and behavior of the hypothetical ideal experimental system. The behavior QB of the ideal system is determined by qualitative simulation of QDE under the ideal experimental conditions $QS(init)$ and, in the case of ambiguities, selection of one of the resulting behaviors.

The property to be measured has been defined as the value of the definition quantity q_{def} in the definition state QS_{def} (definition 34). Since the definition state equals the final state of the experimental system, a systematic error can be seen as a differential final response of the two systems, more specifically a difference in the value of q_{def}. The task of detecting a systematic error can then be reformulated as performing a predictive comparative analysis. The goal of this analysis is to predict whether the response of the experimental system actually investigated would differ from the response of the ideal experimental system if the latter were realized in an experiment. The model QDE and behavior QB of the ideal experimental system have to be compared with *every* candidate model-behavior triple in $\hat{CMB}$, since different assumptions about the model and behavior of the actual physical system may yield different predictions of a systematic error.

A predictive CA by means of CEC* starts with a set $CS(init)$ of RVs at the first pair of comparison pc_0 which represents differences in the experimental conditions of the systems (section 4.4). In the case of error identification, initial RVs can be derived by comparing measurements of the initial state of the actual system with a quantitative specification of the ideal experimental conditions. If the values of a quantity q are significant different according to the statistical conflict criterion, we obtain $q \Downarrow_{pc_0}$ or $q \Uparrow_{pc_0}$, depending on the sign of the difference. Otherwise we have $q \,\|_{pc_0}$. Additional RVs in $CS(init)$ may be directly supplied by the user.

A comparative analysis involving the model QDE and behavior QB of the ideal experimental system, as well as a candidate model $\hat{QDE}$ and candidate behavior $\hat{QB}$ of the experimental system actually investigated, yields a comparative envisionment. A possible systematic error in the measured property value appears as a relative value $\Uparrow$ or $\Downarrow$ for q_{def} at pc_n in a comparative behavior of the envisionment.[3] More precisely, the RV of the definition quantity is a qualitative abstraction of the systematic error. The comparative behavior explains how specific deviations from the ideal experimental conditions at pc_0, added to deviations from the structure and behavior of the ideal experimental system, result in a systematic error at pc_n. The notion of *systematic error* is formally captured as follows.

Definition 40 (Systematic error) Given an ideal property measurement and an actual property measurement. The experimental systems of the property measurements are described by a model-behavior triple $\langle QDE, QS(init), QB\rangle$ and a set of candidate model-behavior triples $\hat{CMB}$, respectively. Suppose a predictive CA is carried out involving $\langle QDE, QS(init), QB\rangle$ and $\langle \hat{QDE}, \hat{QS}(init), \hat{QB}\rangle \in \hat{CMB}$. In addition, we have a set

[3] If the RV is $\|$, the systematic error in the measured property value equals 0. Adhering to common usage of the term, I will say that there is no systematic error in this case.

$CS(init)$ of initial RVs at the first pair of comparison pc_0. The comparative envisionment constructed has comparative behaviors $CB_1, \ldots, CB_m$.

Let q_{def} be the definition quantity of the property. Each comparative behavior CB_k, $1 \leq k \leq m$, in which $q_{def} \Uparrow_{pc_n}$ or $q_{def} \Downarrow_{pc_n}$ indicates a possible systematic error in the measured value of the property. $q_{def} \Uparrow_{pc_n}$ or $q_{def} \Downarrow_{pc_n}$ is the qualitative abstraction of this systematic error.

It is important to emphasize that *different* comparative behaviors in a comparative envisionment may point at *different* systematic errors, even when they have the same RV. The RV of q_{def} at pc_n is a qualitative abstraction of a numerical systematic error which is not guaranteed to be the same in different comparative behaviors. Consider once again the example of the melting-temperature experiments. If in one experiment the sample reacts with its container (a source of systematic error) and in another the sample reacts with its container and has a low purity in addition (another source of systematic error), there will be a systematic error in both observed melting temperatures. However, these systematic errors do not have to be the same.

5.4.3 Algorithm for error identification

The algorithm for error identification given below summarizes how a measurement can be evaluated by detecting possible systematic errors in the reported value of the property. Notice the analogy with the algorithm for conflict resolution. Whereas conflict resolution performs an explanatory CA involving two experiments that have been actually carried out, error identification performs a predictive CA involving an actual experiment and an ideal experiment. No measured states of the experimental system are available in the case of an ideal experiment.

Algorithm 5 (Error identification) Given an actual property measurement and an ideal property measurement. The ideal property measurement refers to a hypothetical experimental system described by the qualitative model QDE and considered under experimental conditions $QS(init)$. The actual property measurement specifies the set $\hat{CM}$ of tuples of candidate models and candidate conditions, and the sequence $\hat{MSS}$ of measured states. The following algorithm detects possible systematic errors in the measured value of the property:

Step 1 Perform a qualitative simulation for every tuple $\langle \hat{QDE}, \hat{QS}(init) \rangle \in \hat{CM}$. Call the resulting sets of candidate model-behavior triples $\hat{CMB}$. Similarly, determine a behavior QB for the ideal experimental system from QDE and $QS(init)$.

Step 2 Use the measured states $\hat{MSS}$ to reduce the sets of candidate model-behavior triples $\hat{CMB}$. Candidate behaviors and models that are not consistent with the measurements are eliminated.

Step 3 Perform a predictive comparative analysis for the model-behavior triple $\langle QDE, QS(init), QB \rangle$ and every candidate model-behavior triple $\langle \hat{QDE}, \hat{QS}(init),$

$\hat{QB}\rangle \in \hat{CMB}$. The initial comparative state information $CS(init)$ is obtained from the ideal experimental conditions and the measured states $\hat{MSS}$ supplemented by user-specified RVs.

Step 4 Return the comparative envisionments generated in the previous step. Each comparative behavior in which $q_{def} \Uparrow_{pc_n}$ or $q_{def} \Downarrow_{pc_n}$ indicates a possible systematic error.

As in conflict resolution, the algorithm may occasionally produce empty comparative envisionments. Under the assumptions that the measurements are correct and the specification of the ideal experimental system is consistent, we must conclude that the combination of a candidate model $\hat{QDE}$ and candidate behavior $\hat{QB}$ does not adequately describe the structure and behavior of the experimental system realized in the experiment.

5.4.4 Properties of the algorithm

The reliability of the algorithm for error identification can be assessed by inquiring into its ability to find all and only possible systematic errors in a measured property value. Obviously, a systematic error originating from a spurious comparative behavior is not possible, since a spurious comparative behavior does not adequately describe differences in the behaviors of the actual and ideal experimental system. This insight underlies the definition of *spurious* and *genuine systematic errors*.

Definition 41 (Spurious systematic error) Given a comparative behavior CB indicating a possible systematic error by predicting that $q_{def} \Uparrow_{pc_n}$ or $q_{def} \Downarrow_{pc_n}$. The systematic error is spurious, iff CB is a spurious comparative behavior. A genuine systematic error is a systematic error that is not spurious.

This definition allows one to prove the following guarantees on the outcome of error identification.

Theorem 19 The algorithm for error identification generates all genuine systematic errors in a measured property value.

Theorem 20 The algorithm for error identification may occasionally generate spurious systematic errors in a measured property value.

The theorems are obvious consequences of the theorems 4, 5, 15, and 16 in chapter 3 and 4.

As for conflict resolution, the computational complexity of the error identification algorithm is determined by the computational complexity of CEC*. Although error identification is exponential in the worst case, the average-case behavior of the algorithm is expected to be much more agreeable. The arguments are similar to those given in the discussion of conflict resolution.

5.5 KIMA system

The methods to accomplish the MA tasks studied in this thesis – conflict detection, conflict resolution, and error identification – are embedded in an implemented system for *Knowledge-Intensive Measurement Analysis (KIMA)*. As can be seen in figure 5.4, the three methods are centered around the measurement base which contains the property measurements to be analyzed.

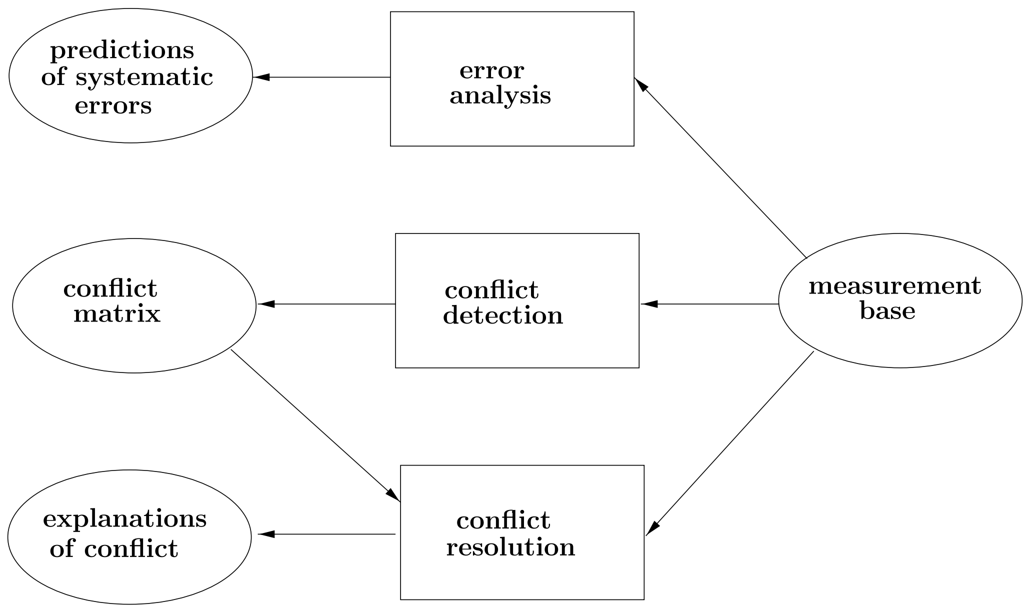


Figure 5.4: *Outline of the KIMA system. The tasks of conflict detection, conflict resolution, and error identification operate upon the measurement base.*

Several interactions exist between the KIMA methods. In the first place, conflict resolution builds upon the results of conflict detection stored in the conflict matrix. In the second place, error identification may indirectly influence conflict detection and conflict resolution by necessitating corrections of the property measurements in order to compensate for systematic errors. Of course, there may be interactions between KIMA and other tools used by scientists. In the concluding chapter I will return to this topic.

The algorithms of the methods making up the KIMA system have been implemented in Common Lisp (appendix D; de Wit & de Jong [1998]). The KIMA implementation comprises about 4000 lines of code and is built on top of the implementations of QSIM and CEC* whose main functions are repeatedly called. It consists of a module in which property measurements, experimental systems, measured states, etc. are defined as Lisp structures, alongside modules for conflict detection, conflict resolution, and error identification with functions operating upon these structures. A teletype interface allows the user to load a measurement base, select an MA method, and provide additional information required by the methods, such as initial RVs in conflict resolution and error identification. The output of the KIMA system is either a conflict matrix or comparative envisionments representing explanations of conflicts or predictions of systematic errors.

5.6 Evaluation

In summary, this chapter has connected the conceptualization of knowledge-based measurement analysis of chapter 2 with the qualitative reasoning techniques of chapters 3 and 4. The methods for conflict resolution and error identification have been formalized in terms of algorithms for the qualitative simulation and comparative analysis of experimental systems. This turns the description of measurement analysis in chapter 2 into a set of neatly defined algorithms with proven guarantees. The methods have been implemented in the KIMA system for Knowledge-Intensive Measurement Analysis which will be used in the case-study considered in the next two chapters.

An interesting point to notice is the interaction between qualitative and quantitative knowledge achieved in the methods. The qualitative models and behaviors of the experimental systems are compared with available measurements, and rejected when an inconsistency is observed. This interaction could be intensified when semi-quantitative models of the experimental systems can be formulated and used in combination with algorithms for semi-quantitative simulation and comparative analysis. The predictions of systematic errors would become more precise and it might be possible to rule out some of the explanations of a conflict. In chapter 8 I will return to this topic.

In chapter 2 a distinction between individual property measurements and aggregated property measurements was introduced. An aggregated measurement is a summary of a set of individual measurements obtained by repeating the experiment a number of times. Strictly speaking, the conflict resolution and error identification methods of this chapter apply to individual measurements. The sequence of measured states of an experimental system is determined in a single execution of an experiment. Repeating the experiment will generally lead to a different sequence of measured states. Also, the physical systems investigated in two executions of the same experiment may have a different structure, and may therefore be described by different models. Notwithstanding these complications, it is possible to derive conclusions about aggregated measurements on the basis of an analysis of individual measurements. This will be illustrated in the case-study.

Part III

Case Study

Chapter 6

Fracture Strength of Ceramic Materials

The measurement analysis methods implemented in the KIMA system will be applied in a case-study in materials science. I will illustrate the ability of the system to identify systematic errors in measurements of the fracture strength of ceramic materials as well as resolve conflicts between these measurements. This chapter will introduce the domain of the case-study, while the next chapter discusses the results of applying KIMA to sets of fracture strength measurements.

In the first section, the knowledge extraction project Plinius is described which forms the context of the case-study. Plinius is concerned with the computer analysis of natural-language texts on mechanical properties of ceramic materials. Section 6.2 gives a brief introduction to this domain, with a focus on the fracture strength property. Section 6.3 and section 6.4 review theories of brittle fracture and two well-known fracture tests. These theories underlie the models of the experimental systems required for conflict resolution and error identification (section 6.5).

6.1 The Plinius project

The aim of the Plinius project presently carried out at the University of Twente is to develop computer programs for semi-automatically extracting domain-specific knowledge from the title and abstract of scientific publications in the field of ceramic materials. A computer analysis of natural-language texts could be helpful in overcoming the problems experienced in building large knowledge bases by hand, in particular the large time and money investments required. The increasing availability of on-line text resources, such as bibliographic databases, electronic journals, and machine-readable lexica, adds to the attractiveness of semi-automatic knowledge extraction. This section gives a brief outline of the Plinius project; for more details and the state of the art, one may consult Mars et al. [1994] and van der Vet et al. [1994].

Figure 6.1 shows the flow of the documents through the Plinius processes and the knowledge sources employed in these processes. At the start, document descriptions consisting of an identifier, title, and abstract are prepared for natural-language processing. These (mainly syntactic) manipulations together constitute the *preprocess*.

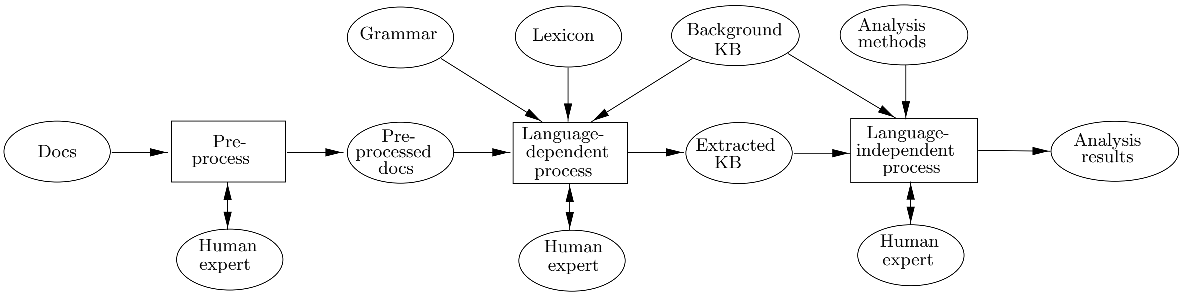


Figure 6.1: *Schematic overview of the Plinius processes and knowledge sources, represented by boxes and ovals, respectively. The horizontal arrows describe the flow of the documents through the knowledge extraction processes, while the vertically-oriented arrows illustrate the use of different knowledge sources during these processes.*

Then, in the *language-dependent process*, a syntactic, semantic, and discourse analysis is carried out on the preprocessed abstracts, leading to the representation in a formal language of the knowledge contained in the documents. These formal assertions are stored in the *extracted knowledge base*. The language-dependent process makes use of a number of knowledge sources: a sublanguage grammar for the syntactic analysis of sentences in the abstracts, a lexicon containing both syntactic features and semantic translations of the words contained in the sentences, and a background knowledge base to assist in semantic and discourse analysis. For an overview of the language-dependent process, with a focus on the interpretation of nominal compounds, see ter Stal [1996].

Much work has been done on the specification of the knowledge sources employed in the language-dependent process and the knowledge bases generated by it. A domain ontology has been developed that provides concepts for expressing knowledge on ceramic materials and their properties (van der Vet & Mars [1993]; van der Vet & Mars [1998]). The ontology specifies the semantic translations of words in the lexicon and defines the structure of the assertions stored in the background knowledge base and the extracted knowledge base. How the knowledge sources can be formally expressed and implemented in a knowledge representation system has been addressed in a separate study (Speel [1995]).

The extracted knowledge base obtained from the language-dependent process forms the input for the *language-independent process*. The language-independent process aims at the analysis of the knowledge in the extracted knowledge base. The work described in this thesis can be considered a (partial) realization of the language-independent process with a focus on the analysis of measurements. It borrows from ideas for knowledge integration proposed earlier in the Plinius project (de Jong [1992]; de Jong, Mars & van der Vet [1993]). Although the MA methods are thus used in the context of a project in which the measurement base is semi-automatically extracted from natural-language sources, they are equally well applicable to measurement bases obtained by manual data entry or by other means.

The knowledge extraction system to be developed in the Plinius project is not fully automated, but allows for interaction with human experts and users in all three processes. In the language-dependent process a human operator will deal with such problems as natural-language ambiguities. During the language-independent process, the user will be able to direct the application of the methods by intervening at particular decision points.

6.2 Mechanical properties of ceramic materials

The domain chosen for the Plinius project, and for the case-study in this thesis, is a subfield of materials science: *mechanical properties of ceramic materials*. Ceramic materials are those materials that "have as their essential component, and are composed in large part of, inorganic nonmetallic materials" (Kingery, Bowen & Uhlmann [1976], p. 3). They include pottery, porcelain, refractories, abrasives, cements, glass, and ferro-electrics. The mechanical properties of a material can be understood as describing the material's behavior when subjected to (external) forces: strength, elasticity, hardness, wear resistance, etc.

Of these properties we will be especially concerned with the *fracture strength* of a material, the stress that needs to be applied to initiate failure. Factors influencing the strength of ceramic materials are summarized in figure 6.2. As can be seen, the fracture strength is highly dependent on the structure of the material, on the atomic as well as on the micro- and macro-structural scale. Important structural features of a material include its chemical composition, the phase distribution and grain size, pores and cracks, and the condition of the surface. The raw materials, the way they are processed and fired, and the way the resulting specimens are finished can all affect the structure and hence the fracture strength of a material. In addition, the conditions under which a material is tested, including the temperature and the environment, are important determinants of its strength. In experimental determinations of the fracture strength, scientists will make provisions to obtain specimens with desired micro- and macro-structural properties and impose appropriate test conditions.

We will consider the fracture strength of ceramic materials at low temperatures. This means that we can restrict attention to brittle fracture, since at low temperatures most ceramic materials fail in a brittle manner. *Brittle fracture* means that failure of the material is not preceded by macroscopic plastic deformation, unlike *ductile fracture* which does involve an appreciable amount of plastic deformation (Courtney [1990], ch. 9). More specifically, the kind of fracture that will be taken into consideration is brittle fracture occurring by the propagation of preexisting cracks or flaws in the material. In section 6.3 I discuss a simple criterion specifying the conditions under which crack propagation will be initiated.

It should be borne in mind that brittle fracture and ductile fracture are extremes on a continuum of fracture modes. Thus, many materials that would be classified as failing in a brittle manner exhibit some macroscopic or microscopic plastic flow. For example, brittle fracture may be accompanied by microscopic plasticity at the vicinity

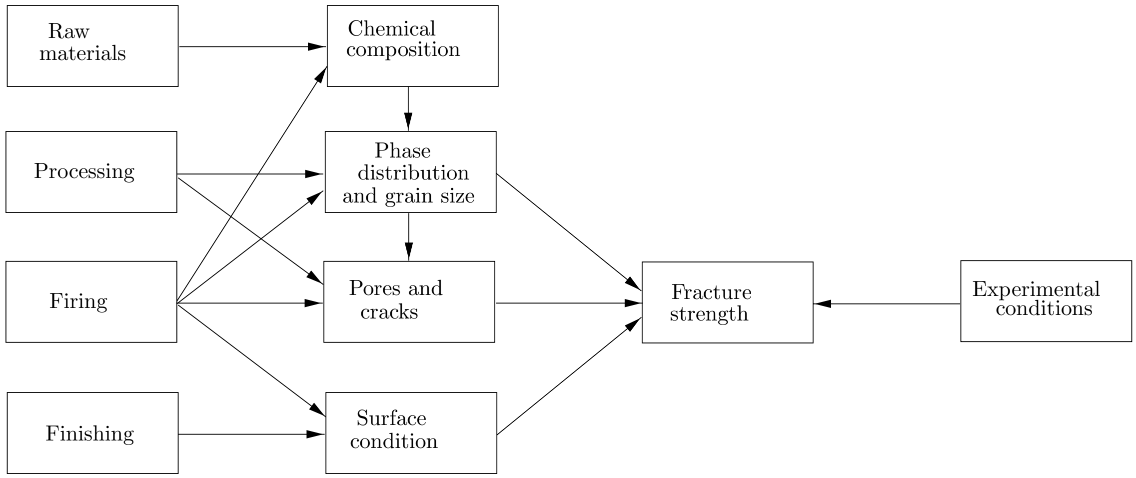


Figure 6.2: *Factors influencing the mechanical properties of ceramic materials (adapted from Davidge [1979]).*

of the crack tip due to the high stress concentrations developing there when the material is being loaded.

A variety of tests to determine the fracture strength of a material have been developed. Marschall & Rudnick [1974] and Quinn [1991] discuss a few of these, including tension tests, three-point and four-point bend tests, pressurized ring tests, and diametral compression tests. I will focus on tension tests and four-point bend tests, because they are widely used, well-documented, and conceptually simple. The tests cause the material to fracture by application of a uniaxial or *tensile* stress (figures 6.5 and 6.7). Section 6.4 discusses tension tests and four-point bend tests in more detail.

A reported value for the strength of a material is usually an aggregated measurement, that is, an average value obtained from measuring a batch of specimens in the same experiment. In the case of brittle materials, fracture strength measurements often substantially vary from specimen to specimen, even if they have been fabricated and finished in the same way. As a consequence, the reported aggregated value may not be very precise. The cause of this imprecision must be sought in the large sensitivity of the fracture strength to uncontrollable variations in the distribution, size, shape, and orientation of flaws in the material. In contrast with ductile materials, such variations cannot be alleviated by plastic deformation prior to fracture.

Besides variations *within* an experiment, variations *between* experiments are liable to occur. More often than not, different fabrication and surface finishing methods are used, which may lead to significant deviations of the fracture strength measured in one experiment from that found in another. Differences in loading conditions and differences in specimen dimensions add to these deviations.

In order to counter the difficulties in controlling conventional fracture tests, researchers in ceramics often prefer to measure the *fracture toughness* instead of the fracture strength of a material. The fracture toughness is a material property parameterized for the known size, shape, and location of a crack which has been deliberately introduced in the sample to initiate fracture (Courtney [1990], ch. 9).

6.3 Brittle fracture

Theoretical analyses of brittle fracture relie on the basic insight that cracks in a material act as stress concentrators and that fracture occurs through extension of a *critical crack*, that is, the crack effecting the greatest stress concentration. As cracks move with a velocity quickly approaching the velocity of sound, the material fails immediately after the initiation of crack propagation. The main thing to establish, therefore, are the circumstances under which a crack starts to propagate. A simple criterion for fracture initiation compares the maximum stress at the crack tip with the theoretical strength of a material (Davidge [1979], ch. 3; Lawn & Wilshaw [1975], ch. 7).

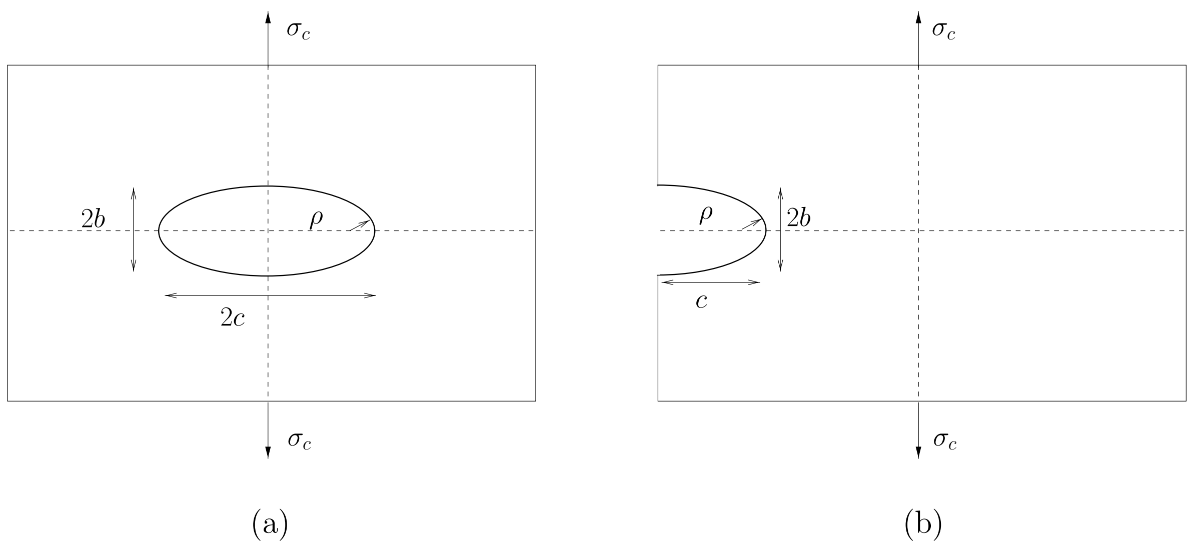


Figure 6.3: *Elliptical hole (a) inside and (b) at the edge of a plate subjected to a uniform applied stress σ_c. For $b \ll c$ the holes represent an interior and a surface crack, respectively.*

Consider an elliptical hole of semi-axes b and c in a uniformly stressed plate (figure 6.3(a)). When $b \ll c$ this is a good representation of an interior crack of length $2c$. One can now examine the modifying effect of the hole on the stress distribution in the plate. Under the assumption that linear elasticity holds everywhere in the plate, that the boundary of the hole is stress-free, and that the axes b and c are small in comparison with the plate dimensions, the stress distribution can be calculated. The *maximum stress* σ_m occurs at the crack tip, and is given by

$$\sigma_m = \sigma_c(1 + 2\sqrt{\frac{c}{\rho}}), \tag{6.1}$$

with ρ the *crack tip radius* and σ_c the applied stress. Bearing in mind that ρ may be equated to b^2/c, we can for $b \ll c$ simplify equation (6.1) to

$$\sigma_m = \sigma_c 2\sqrt{\frac{c}{\rho}}. \tag{6.2}$$

The ratio $2\sqrt{c/\rho}$ is called the *stress concentration factor* which can be much larger than unity when $b \ll c$. Figure 6.4 displays the stress concentration near one side of the crack. Notice that major perturbations to the applied stress field occur only within a distance of about c from the crack tip.

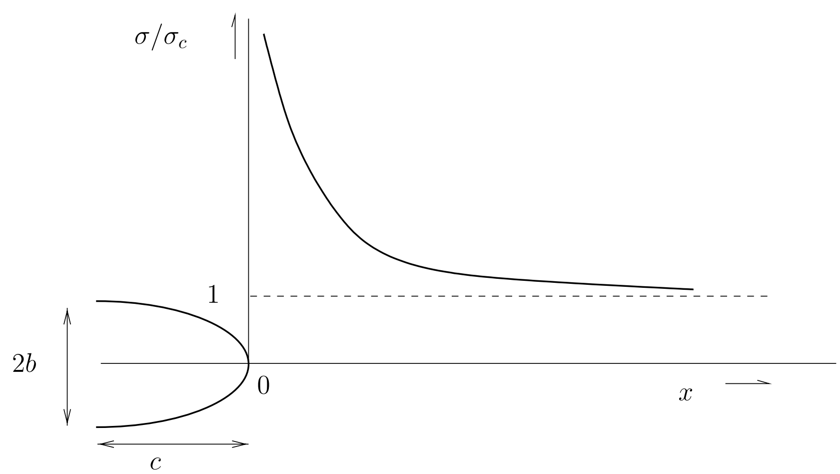


Figure 6.4: *Stress concentration at a distance x from the crack tip.*

Equation (6.2) directly applies to a surface crack of length c (figure 6.3(b)). For a surface crack of length $2c$ we then have

$$\sigma_m = \sigma_c 2\sqrt{2}\sqrt{\frac{c}{\rho}}, \tag{6.3}$$

so that a surface crack is immediately seen to be a stronger stress concentrator than an interior crack.

Strictly speaking, the maximum stress formulas hold only for a uniformly stressed plate. In the examples of the case-study, however, we will frequently encounter situations characterized by a non-uniform stress distribution in the material, for example when an error in the experimental procedure causes a local stress concentration. If the stress gradient is low in the region of the crack, (6.2) and (6.3) are good approximations of the maximum stress near the crack tip.

One can imagine the plate to fracture when the maximum stress σ_m at the crack tip equals the *theoretical strength* σ_{th} of the material. The theoretical strength of a material is the stress required to rupture the atomic bonds across the fracture plane. Under the assumptions of uniform stressing and linear elasticity, it can be expressed in terms of Young's modulus E, the interatomic spacing a_0, and the surface energy per unit area γ:

$$\sigma_{th} = \sqrt{\frac{E\gamma}{a_0}}. \tag{6.4}$$

Thus a high theoretical strength is favored by a large Young's modulus and surface energy, and by close packing of atomic planes.

The equality of the maximum stress at the crack tip and the theoretical strength as a requirement for the initiation of fracture, i.e.,

$$\sigma_m = \sigma_{th}, \tag{6.5}$$

is called the *Orowan criterion* for brittle fracture. It is used in the tension test and bend test systems considered in the next chapter.

An alternative, more sophisticated criterion for brittle fracture is based on the principle that the extension of a preexisting crack should be accompanied by a decrease in system energy. This *Griffith criterion* lies at the heart of modern fracture mechanics (Lawn & Wilshaw [1975]). It is considerably more involved than the Orowan criterion, and I will not discuss it here. It should be noted, though, that both approaches arrive at more or less the same conclusions as to the parameters influencing the fracture strength of a material.

6.4 Fracture strength experiments

Two kinds of experiment for determining the fracture strength of a material are analyzed below: tension tests and four-point bend tests. The MA methods will be applied to strength measurements obtained by means of these experiments.

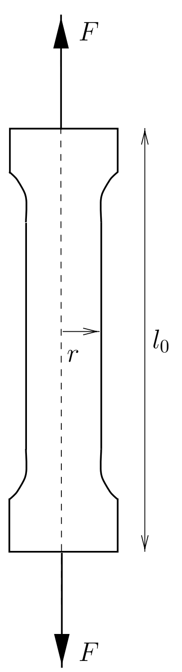
Marschall & Rudnick [1974] state in their review of strength testing methods that the tension test, while it is "the simplest … to analyze, it is among the most difficult to conduct satisfactorily" (p. 77). With this in mind, the analytical simplicity of the tension test is often traded for the experimental simplicity of a bend test, but "in many cases the desired experimental simplicity is not obtained" (p. 85). Tension tests and four-point bend tests thus seem particularly well-suited as examples: on the one hand they are simple to understand (making them convenient illustrations of the methods) and on the other hand difficult to perform (making the occurrence of conflicts and systematic errors a real possibility).

The discussion of the tests is structured as follows. First, I describe the basic theory of the experiments; then, I proceed with modifications of the basic theory which become necessary when the assumptions underlying the theory are violated in an actual experiment. This approach is motivated by the fact that deviations from the conditions under which the theory holds are usually seen as disturbances of the ideal fracture strength experiment. This is nicely summarized in Marschall & Rudnick [1974]:

> "Stress analyses are based on certain assumptions and idealizations relating to material characteristics, loading conditions, and specimen geometry. The accuracy in evaluating mechanical properties is related directly to the extent that the assumed conditions are realized; deviations from conditions assumed in the analysis of raw data are sources of error" (p. 70).[1]

6.4.1 Tension test

The *tension test* is schematically displayed in figure 6.5. In a tension test a (cylindrical) specimen of the material is slowly elongated at a specified rate by application of a tensile force along the neutral axis. The force and elongation are measured by means of a load cell and an extensometer, respectively. The maximum stress which the sample can withstand is called its (tensile) fracture strength.



(a) (b)

Figure 6.5: *(a) Schematic representation of the tension test and (b) impression of the experimental set-up (reproduced from Courtney [1990]). The interpretation of the variables is as follows: F applied force, l_0 initial length of specimen, r radius of gage section of specimen.*

[1] See Newnham [1975] and Baratta [1984] for similar remarks.

The two fundamental parameters in a tension test are the stress σ and the strain ϵ, defined as the force per unit area and the percentage elongation. The stress and strain may vary across the material due to certain material and loading characteristics. The stress and strain distributions in the material are uniform, however, when the material is perfectly homogeneous and isotropic, that it, when its properties do not vary with position and direction. Under the additional assumption of linear elasticity, the stress and strain are related by a form of Hooke's law:

$$\epsilon = \frac{1}{E}\sigma. \tag{6.6}$$

A measure of the average stress and strain in a material are the *nominal* stress σ_a and the *nominal* strain ϵ_a. They are defined as the ratio of applied force F and sample cross-section area A and the ratio of elongation $l - l_0$ and initial sample length l_0, respectively:

$$\sigma_a = \frac{F}{A}, \qquad \epsilon_a = \frac{l - l_0}{l_0}. \tag{6.7}$$

In the case of uniform stress and strain distributions, the stress and strain equal the nominal stress and strain everywhere in the material. The strength value usually reported in a tension test is the nominal applied stress at the moment of fracture.

The ideal tension test is performed under circumstances such that uniform stress and strain distributions satisfying (6.7) are realized. Marschall & Rudnick [1974] identify a number of deviations from the ideal circumstances which are considered to be important sources of systematic errors. The disturbances are sometimes hard to account for in a fully quantitative way, so that we may have to be satisfied with qualitative specifications of their effects. It should be kept in mind that, although the disturbances are discussed separately, they can occur in combination.

Eccentric loading

Eccentric loading is the most widely recognized problem in conducting tension tests (figure 6.6(b)). In eccentric loading the force is not applied along the neutral axis of the specimen, thereby causing bending of the specimen. In the case of linear elasticity, we then obtain a non-uniform, linear stress distribution defined by

$$\sigma(x) = \sigma_a(1 + \frac{4ex}{r^2}) \tag{6.8}$$

at a position x relative to the neutral axis, with e the eccentricity (the distance from the neutral axis to the axis of load application). Notice that the stress reaches its maximum at the surface of the specimen, in the direction of eccentricity, when $x = r$. Given a homogeneous distribution of cracks, fracture is likely to occur through a crack located at

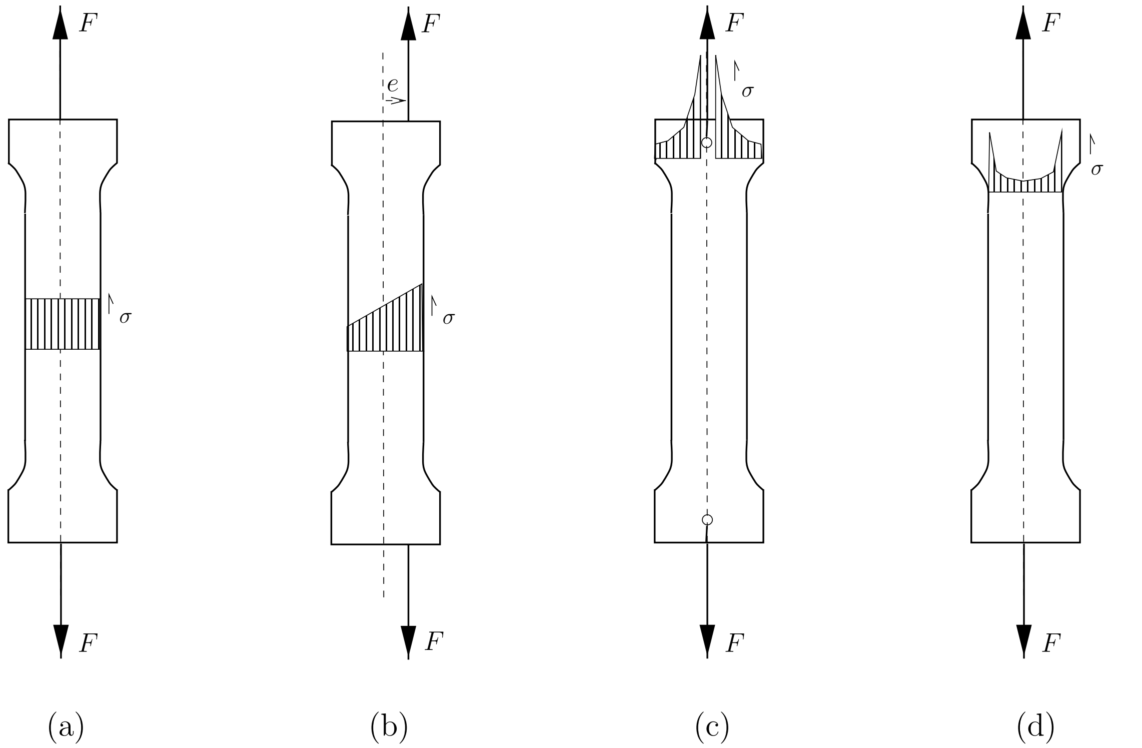


Figure 6.6: *(a) Stress distribution achieved in the ideal tension test, and deviations from this distribution in (b) an eccentrically loaded specimen, (c) a specimen loaded through stress-concentrating grips, and (d) a specimen with a sudden change in cross-section at its shoulders (adapted from Marschall & Rudnick [1974]).*

or near the surface. Eccentricity can lead to local stresses 80% higher than the reported nominal stress.

Eccentric loading might be caused by faults in the loading system and specimen geometry and by non-uniform specimen properties. Marschall & Rudnick [1974] list a number of measures that could be taken to avoid or reduce eccentricity, such as loading systems with grips specifically designed so as to align the applied force with the neutral axis of the specimen.

Load application through stress-concentrating grips

Another common disturbance in a tension test is the occurrence of large stress concentrations near the grips through which the force is applied to the material. Specimens may for example be gripped around the end surfaces by pins through the head of the specimen (figure 6.6(c)). As a consequence of these stress concentrations, fracture of the sample is liable to be initiated by a crack near the grips instead of a crack in the gage section. The publications on the tension test that I consulted do not further specify the stress concentration, so that it will be accounted for by the following functional relation:

$$\sigma_c = f(\sigma_a), \text{ with } f \in M_0^+, \text{ and } f(\sigma_a) > \sigma_a \text{ when } \sigma_a > 0.$$

Ohji [1988] describes possible measures to induce stress relaxation, such as the application of a buffer material between test piece and grips.

Sudden change in cross-section at sample shoulders

Parasitic stress concentrations, as Marschall & Rudnick [1974] call them, may also arise at the shoulders of the specimen, at the transition from the gripped heads to the reduced gage section (figure 6.6(d)). For cylindrical specimens, the magnitude of this stress concentration is a function of the ratio of the head diameter d_h to the gage-section diameter d and of the ratio of the transition radius r_t to the gage-section diameter. More specifically, we have

$$\sigma_c = g(\frac{d_h}{d}, \frac{r_t}{d})\sigma_a, \quad g \in M^{+-}.$$

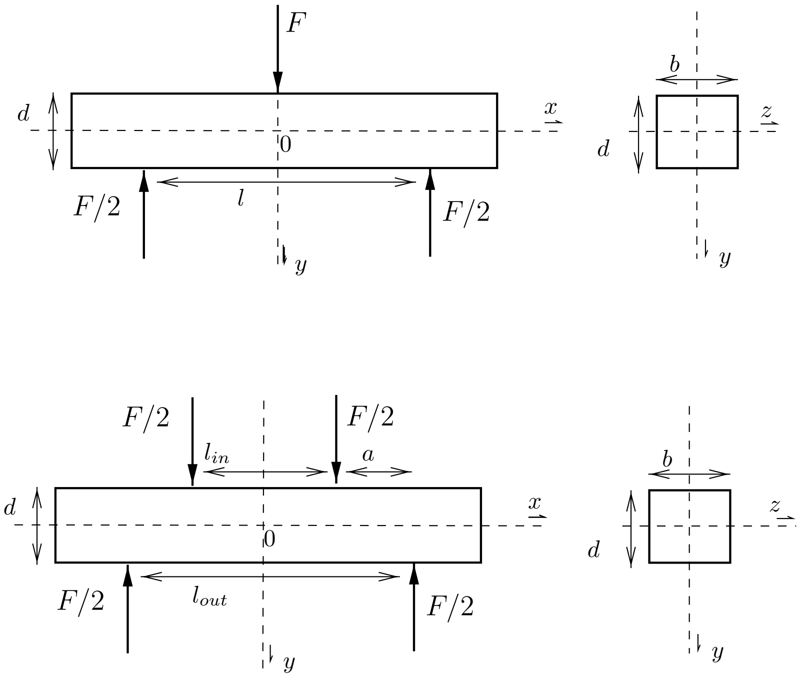
By carefully tapering the specimen, that is, by reducing d_h/d or enlarging r_t/d, one can avoid failure through a stress concentration at the shoulders (see Quinn [1991] for examples of standard specimen geometries).

6.4.2 Four-point bend test

In a *bend test* or *flexure test* the tensile stress is not directly, but indirectly applied to the material by means of bending. A beam is fixated in a test machine and slowly deflected until it breaks (figure 6.7). The maximum stress developed at the surface of the specimen at the instant of fracture is defined to be the fracture strength of the material, also referred to as the *flexure strength* or *modulus of rupture*. The number of load points gives rise to a distinction between *three-point* and *four-point* bend tests. In the former case we have two support points and one load point, and in the latter case two support points and two load points. Since four-point bend tests are more often used and simpler to analyze, I will focus on them.

Bending of the specimen gives rise to a non-uniform stress distribution in the material. The distribution can be derived from the theory of elasticity when making a few assumptions, including that the beam is homogeneous and isotropic, exhibits linear stress-strain behavior, and has a maximum deflection which is small compared to the beam depth (see Baratta [1984] for a detailed discussion of the assumptions). At a distance x from the mid-length line and y from the mid-depth line, the stress is given by:

$$\sigma(y) = \begin{cases} \frac{F}{2I}ay & \text{, if } 0 < \mid x \mid < (\frac{l_{out}}{2} - a), \\ \frac{F}{2I}(\frac{l_{out}}{2} - x)y & \text{, if } (\frac{l_{out}}{2} - a) < \mid x \mid < \frac{l_{out}}{2}, \\ 0 & \text{, if } \mid x \mid > \frac{l_{out}}{2}, \end{cases} \tag{6.9}$$



(a)

(b)

Figure 6.7: *(a) Schematic representation of the three-point and four-point bend tests (adapted from van der Laan & Bach [1987]) and (b) impression of the experimental set-up of the latter with A the specimen and B the left-hand pair of rollers for loading and supporting the specimen (reproduced from Hoagland, Marschall & Duckworth [1976]). The interpretation of the variables in (a) is as follows: F applied force, l span length in three-point bend test, l_{in}, l_{out} inner and outer span length in four-point bend test, a distance between support point and load point, d depth of the beam, b width of the beam, x distance from the mid-length line, y distance from the mid-depth line.*

with F the applied force, I the moment of inertia of the beam, l_{out} the distance between the support points (the *outer span*), and a the distance between a support point and a load point.

A few qualitative features of (6.9) stand out. First, for a given y, the applied stress is uniform between the load points; second, the stress is 0 on the neutral axis of the beam ($y = 0$); and third, the maximum and minimum stress are reached between the load points at the surface of the beam ($y = \pm d/2$). The stress above the neutral axis ($y > 0$) is a compressive stress, whereas the stress below this axis ($y < 0$) is a tensile stress. Given that a brittle material is more liable to fracture under tension than under compression, the critical crack will be located in the lower half of the beam. What is more, the stress gradient in the y direction will favor fracture through a crack at or near the surface of the specimen.

For a beam of depth d and width b the moment of inertia about the neutral axis is given by:

$$I = \frac{bd^3}{12}.$$

If we take $a = l_{out}/4$, as is common practice, and then substitute for I and a in (6.9), we obtain the stress:

$$\sigma(y) = \sigma_a \frac{2y}{d},$$

between the load points (i.e., $(l_{out}/2 - a) <| x |< l_{out}/2$), with

$$\sigma_a = \frac{3Fl_{out}}{4bd^2}. \tag{6.10}$$

The nominal applied stress σ_a represents the maximum tensile stress at the lower surface of the beam ($y = d/2$). The reported fracture strength in a four-point bend test is the value of σ_a when the specimen breaks. Figure 6.8 shows the stress in the beam as a function of x and y.

As in the tension test, the stress and strain in the beam are related to each other by means of the simple formula:

$$\epsilon = \frac{1}{E}\sigma,$$

where the modulus of elasticity in tension is supposed to equal the modulus of elasticity in compression. The strain ϵ is imposed by deflecting the specimen over a distance d. Baratta [1984] finds the following expression for the nominal strain at the tensile surface

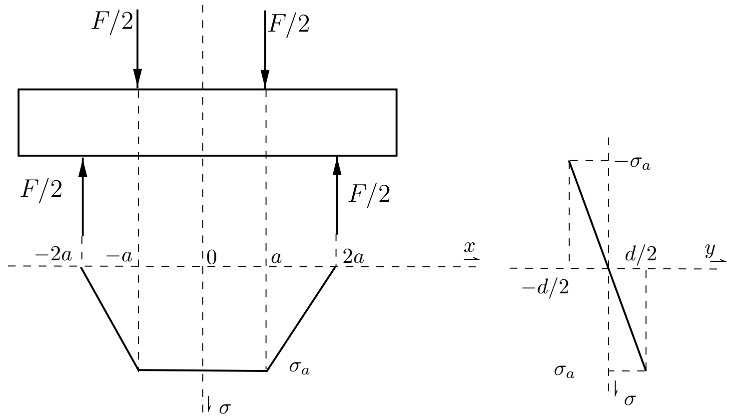


Figure 6.8: *Stress distribution achieved in the ideal four-point bend test (adapted from van der Laan & Bach [1987]). The variation with x of the tensile stress at the lower surface is shown, as well as the variation with y of the (tensile and compressive) stress between the load points. As above, x refers to the distance from the mid-length line and y to the distance from the mid-depth line.*

$$\epsilon_a = \frac{3y_d}{\frac{a}{d}(3l_{out} - 4a)}.$$

With $a = l_{out}/4$ this simplifies to

$$\epsilon_a = \frac{6y_d d}{l_{out}^2}. \tag{6.11}$$

The assumptions compiled into the calculation of the stress distribution (6.9), and contained in the loading conditions shown in figure 6.7, implicitly specify the ideal four-point bend test. That is, the accuracy of a fracture strength measurement depends on the extent to which the execution of the experiment succeeds in approximating the conditions implied by the assumptions. Using Hoagland, Marschall & Duckworth [1976], supplemented by Newnham [1975], Baratta [1984], van der Laan & Bach [1987], and Marschall & Rudnick [1974], I will review a number of important deviations from the ideal circumstances. The stress analysis is considerably more complicated than in the case of tension tests, and most authors constrain themselves to specifying the maximum deviation from the ideal stress distribution. This practice will be followed here as well.

Unequal loading

Figure 6.9(a) shows an unequal distribution of the load over the two load points. This gives rise to an asymmetrical stress distribution with a maximum stress achieved near the left load point, at the tensile surface of the beam. This stress σ_c is given by:

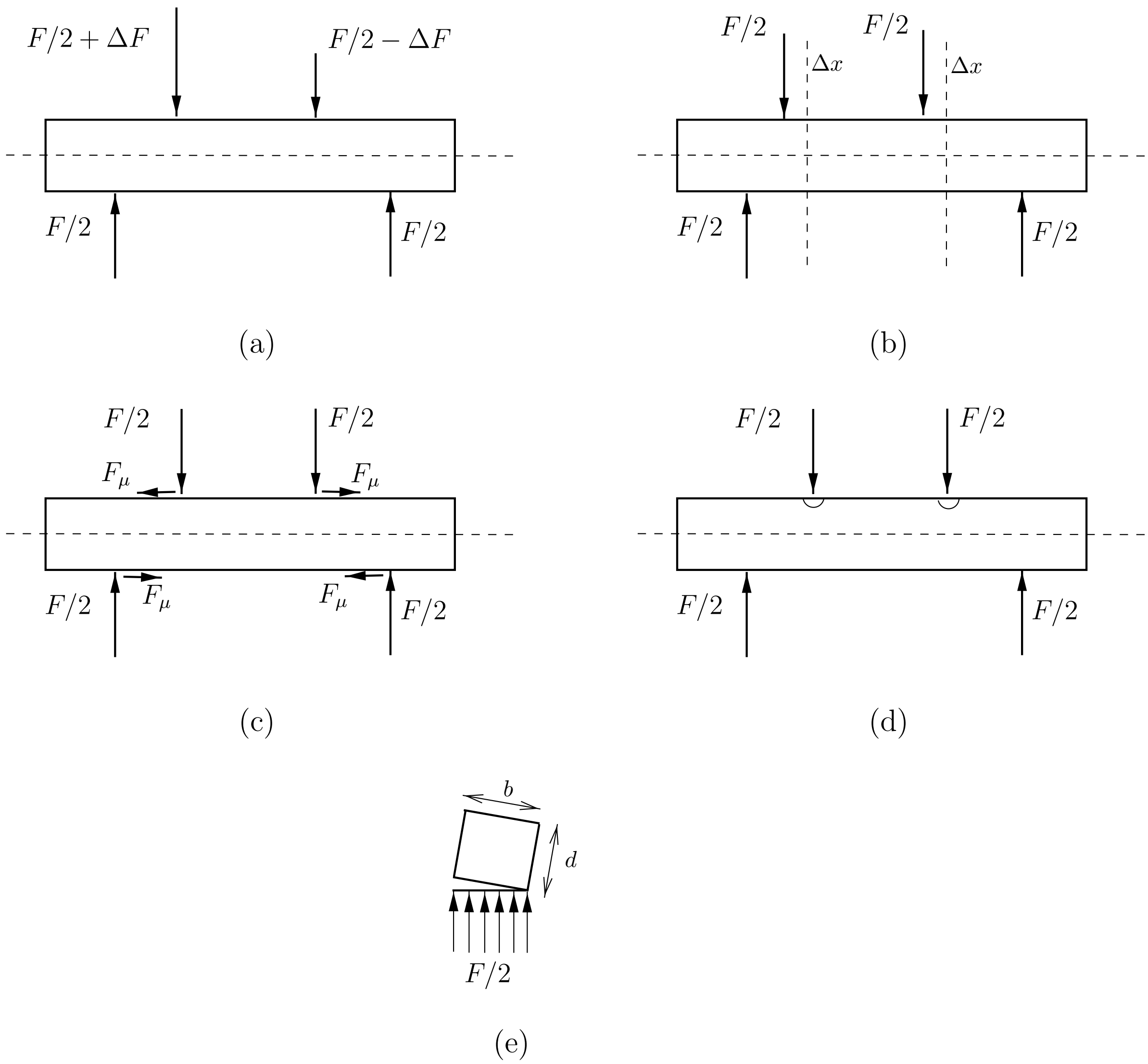


Figure 6.9: *Deviations from the ideal loading conditions of the four-point bend test: (a) unequal loading, (b) shift of load points, (c) friction, (d) wedging, and (e) twisting (adapted from van der Laan & Bach [1987]).*

$$\sigma_c = (1 + 2\frac{\Delta F}{F})\sigma_a,$$

where ΔF represents the load difference. Assuming a homogeneous crack distribution, the specimen will tend to fracture near the load point at which the maximum stress is attained.

Load mislocation

A related source of systematic error is a shift of the load points along the length of the beam (figure 6.9(b)). Since bend fixtures are often designed so as to have a fixed inner and outer span, the left and right load point will both shift over the same distance Δx. Again, we have an asymmetrical stress distribution, with the maximum stress now occurring near the right loading point, at the tensile surface of the beam. This stress is estimated by

$$\sigma_c = (1 + 2\frac{\Delta x}{l_{out}})\sigma_a.$$

Wedging

Loads on bend test specimens are usually applied through rollers contacting the surface on lines perpendicular to the specimen axis. In a two-dimensional view the load lines reduce to load points, as shown in figure 6.7(b). The resulting high stresses under the rollers, analogous to the conditions generated by forcing a sharp wedge into the material, superpose an additional tensile stress on the stress induced by bending (figure 6.9(d)). In particular, at a distance $\pm d/4$ from the load points, the maximum stress at the tensile surface of the beam is given by

$$\sigma_c = (1 + 0.1157\frac{d}{l_{out}})\sigma_a. \tag{6.12}$$

Notice that the effect of *wedging* is local and small for reasonable values of d/l_{out}.

Twisting

A potentially significant source of error in the bend test is torsion due to a nonuniform load distribution along the rollers, as shown schematically in figure 6.9(e). This *twisting* effect might easily arise if the beam is skewed over its length or if the pair of rollers at one end of the specimen is not parallel to the pair at the other end. In the extreme, the rollers make contact at one edge of the specimen only, and the maximum stress at the tensile surface is estimated by:

$$\sigma_c = \left(\frac{1}{2} + \sqrt{(\frac{b}{l_{out}})^2 \frac{1}{36k_2^2} + \frac{1}{4}}\right) \sigma_a, \tag{6.13}$$

with k_2 being a function of b/d. Notice that σ_c/σ_a is larger than 1, so that twisting tends to decrease the measured strength of a material.

A possible way to minimize twisting is the use of rollers that can rotate about an axis parallel to the specimen z axis.

Friction

Another possible disturbance in the execution of a four-point bend test is the development of *friction* forces due to the contact between the cylindrical rollers and the specimen. Friction forces resist surface straining. An estimate of their effect at the tensile (lower) surface of the beam is

$$\sigma_c = (1 - 2\mu\frac{d}{l_{out}})\sigma_a,$$

where μ represents the coefficient of static friction. The net effect of friction is a reduction of the stress and hence an increase in the measured fracture strength. The magnitude of the frictional effect can be significant.

A simple arrangement to remove friction is to design a bend fixture with rollers that are allowed to roll freely along the specimen surface.

6.5 Evaluation

This chapter reviewed theories describing the stress distribution in a specimen and the conditions under which fracture occurs. Different ways of performing an experiment may lead to different stress distributions, and thus to different outcomes of the experiments. In the next chapter the theories will be used to build models of the physical systems created and controlled in the experiments.

An interesting aspect of the discussion is the manner in which the theory of elasticity underlying the analysis of bend tests and tension tests is closely related with what counts as the ideal execution of the experiments. This is an instance of the intricate relations between theory and experiment that have been studied by philosophers of science (e.g., Kuhn [1977], Hacking [1983], and Galison [1987]). Deviations from the assumptions made in the theoretical analysis are considered to be error sources, maybe because the fracture strengths reported in tension tests and four-point bend tests (equations (6.7) and (6.10)) presuppose the applicability of the theory. Besides violations of the theoretical assumptions, other kinds of error sources should be mentioned, such as deviations from desired micro-structural properties (e.g., grain size and porosity) and surface conditions of the material. In particular, “the material for the test specimen should be prepared in the same way as that for the actual component” and “the surface and edge finish used in the test should model that which a component will possess in practice” (Creyke, Sainsbury & Morrell [1982], p. 79-80).

Chapter 7

Analysis of Fracture Strength Measurements

"Systematic property variations lend themselves to detection."

C.W. Marschall and A. Rudnick, 'Conventional strength testing', 1973

This chapter presents the results of applying KIMA in a case-study on a realistic though simplified problem: the analysis of measurements of the fracture strength of ceramic materials. After some preliminary remarks of a general and methodological kind (section 7.1), I will discuss how KIMA successfully explains conflicts and predicts systematic errors in a number of examples involving measurements obtained in tension tests and four-point bend tests (sections 7.2 to 7.3). The chapter closes with a summary of the results (section 7.4).

7.1 Application of the MA methods in the case-study

The aim of the case-study is to illustrate the MA methods by means of examples of conflicts and systematic errors in a realistic domain, the fracture strength of ceramic materials. I will show how the experimental systems investigated in fracture tests are modeled by means of QDEs, how experimental results are represented as measured state sequences, and how the models are employed to resolve conflicts and identify systematic errors in the property measurements. The examples have been chosen in such a way that they illustrate salient features of conflicts and systematic errors, such as a potentially large number of explanations of a conflict, the elimination of explanations by adding further information, the interaction of different disturbances in bringing about a systematic error, and the failure to explain a conflict with given sets of candidate models and candidate conditions. The use of the methods in a realistic domain is in conformity with the application requirement in section 1.2.

The ceramic material selected for the case-study is *aluminium oxide* (Al_2O_3), also called *alumina* (Dörre & Hübner [1984]). It is one of the most extensively investigated ceramic materials and is regarded as a typical representative of a whole family of materials. It exhibits brittle fracture until about 1200 ^{0}C and is characterized by a low fracture strength as compared to its theoretical strength. The strength of alumina is, as with other brittle materials, subject to large variations within a single experiment and between different experiments for the reasons cited in section 6.2. Experimental results summarized in Wachtman [1996], ch. 23, show that flexure strengths of polycrystals with grain size ± 10 μm, measured at room temperature, range from 200 to 600 MPa with standard deviations sometimes as large as 100 MPa.

In the ideal case strength measurements, and information about the strength tests in which they were obtained, would be taken from the materials science literature, together with assessments of conflicts and errors in the measurements. Unfortunately, no such critical reviews of empirical determinations of the fracture strength of alumina are available as, for example, compiled for the melting temperature of refractory oxides (Hlaváč [1982]; Coutures & Rand [1989]). Although property handbooks like Dörre & Hübner [1984] do list experimental results, it is only rarely that the measurements are annotated with details about the experiments and compared with other measurements.

As a consequence, I have decided to use hypothetical measurement sets in the case-study. The strength values have not been chosen at random, though; it has been assured that they fall within the bounds mentioned above and that they agree with a number of empirically established relations between the measured strength of a material and characteristics of the experiments. An example of such a relation, reported in the literature, is the assertion that 'a surface finishing procedure which does not remove damage caused by grinding operations tends to lead to a lower measured fracture strength'. I could have chosen measurements and descriptions of experiments at random, and the MA methods would have resolved conflicts and identified systematic errors with the guarantees given in chapter 5, but the answers would then not have been of much use for illustrative purposes.

The models required for the analysis of the measurements were obtained from a study of the literature on brittle fracture and strength tests, summarized in the previous chapter, and additional consultation of a domain expert. Both for the tension test and the bend test, a space of possible qualitative models and associated initial conditions has been constructed. The experimental systems described by the models are most naturally, and in conformity with the usual approach in the literature, equated with the specimen to which the force is applied. The test machine, including the grips and rollers to fasten the specimen, is considered only in so far as it influences the stress distribution in the specimen.

The models are QDEs describing the shape and location of the critical crack and the stress field in its vicinity. The stress field is created by loading the specimen in a specific manner. Deviations from the ideal loading conditions, like eccentric loading in a tension test or load misloaction in a bend test, lead to stress fields deviating from the one predicted by the basic theory of elasticity (section 6.4). Both the ideal and deviating loading conditions are accounted for in models of the experimental systems. With an occasional exception, I will not consider models for situations in which *combinations* of deviating loading conditions occur.

Unfortunately, it is not possible to predict the shape and location of the critical crack, and thus determine the appropriate model for the experimental system, from the loading conditions and surface finish of the specimen. There is no way that the distribution of cracks in a piece of material, and their length and sharpness, can be controlled by experimenters (Creyke, Sainsbury & Morrell [1982], ch. 3). The specimen might fracture through a crack in the region where the highest stress is achieved, but also through a rather large crack in a region with a lower stress. However, if it is assumed that cracks are sufficiently numerous and randomly distributed throughout the material, a special case of the homogeneity assumption in chapter 6, then the critical crack is most likely to be located in the region with the highest stress concentration.

In order to simplify the analysis of tension test and bend test results, I will only consider the *most likely* fracture mode of a specimen for given loading conditions and surface finishing procedures, and hence the model that is *most likely* to be adequate. Accordingly, the explanations of a conflict and the predictions of a systematic error generated by the MA methods are the *most likely* explanations and predictions given the information about the experiments. For instance, in a bend test the critical crack is most likely to be a surface crack located in the area between the two load points where the applied tensile stress is at a maximum (equation (6.9)). It should be noted that a similar form of probabilistic reasoning is often employed by material scientists themselves to account for the results of their experiments.

All analyses reported in this chapter have been actually carried out by means of KIMA. In appendix E a trace of the second tension test example is included.

7.2 Conflict resolution and error identification in tension tests

7.2.1 Model space

The model space for the tension test, $SPACE_{tt}$, is composed of eight QDEs:

$$SPACE_{tt} = \{QDE_{ttsurf}, QDE_{tteccsurf}, QDE_{ttgripsurf}, QDE_{ttshsurf}, QDE_{surf}, QDE_{ttecc}, QDE_{ttgrip}, QDE_{ttsh}\}.$$

The models QDE_{ttsurf}, $QDE_{tteccsurf}$, $QDE_{ttgripsurf}$, and $QDE_{ttshsurf}$, with their associated initial conditions, are shown in figure 7.1 and 7.2. They have been derived from the analysis of tension tests in section 6.4 and describe the following systems:

QDE_{ttsurf} A specimen which is concentrically loaded and fractures through a surface crack;

$$\begin{aligned}
\epsilon_a &= \frac{1}{E}\sigma_a \\
\epsilon_a &= \frac{l - l_0}{l}, \qquad \sigma_a = \frac{F}{A} \\
\sigma_c &= \sigma_a \\
\frac{dl}{dt} &= rl \\
\sigma_{th} &= h(\frac{\gamma E}{a_0}), \qquad h \in M_0^+ \\
\sigma_m &= 2\sqrt{2}\sigma_c f(\frac{c}{\rho}), \qquad f \in M_0^+
\end{aligned}$$

(QDE_{ttsurf})

$$\begin{aligned}
\epsilon_a &= \frac{1}{E}\sigma_a \\
\epsilon_a &= \frac{l - l_0}{l}, \qquad \sigma_a = \frac{F}{A} \\
\sigma_c &= \boldsymbol{f_{ec}}\sigma_a \\
\boldsymbol{f_{ec}} &\boldsymbol{= 1 + \frac{4e}{r} x_{rel}} \\
x_{rel} &= \frac{x}{r} \\
\frac{dl}{dt} &= rl \\
\sigma_{th} &= h(\frac{\gamma E}{a_0}), \qquad h \in M_0^+ \\
\sigma_m &= 2\sqrt{2}\sigma_c f(\frac{c}{\rho}), \qquad f \in M_0^+
\end{aligned}$$

$(QDE_{tteccsurf})$

$$\begin{aligned}
\epsilon_a &= \frac{1}{E}\sigma_a \\
\epsilon_a &= \frac{l - l_0}{l}, \qquad \sigma_a = \frac{F}{A} \\
\sigma_c &= \boldsymbol{g}(\sigma_a), \qquad \boldsymbol{g \in M_0^+} \\
\frac{dl}{dt} &= rl \\
\sigma_{th} &= h(\frac{\gamma E}{a_0}), \qquad h \in M_0^+ \\
\sigma_m &= 2\sqrt{2}\sigma_c f(\frac{c}{\rho}), \qquad f \in M_0^+
\end{aligned}$$

$(QDE_{ttgripsurf})$

$$\begin{aligned}
\epsilon_a &= \frac{1}{E}\sigma_a \\
\epsilon_a &= \frac{l - l_0}{l}, \qquad \sigma_a = \frac{F}{A} \\
\sigma_c &= \boldsymbol{g(\frac{d_h}{d}, \frac{r_t}{d})}\sigma_a, \qquad \boldsymbol{g \in M^{+-}} \\
\frac{dl}{dt} &= rl \\
\sigma_{th} &= h(\frac{\gamma E}{a_0}), \qquad h \in M_0^+ \\
\sigma_m &= 2\sqrt{2}\sigma_c f(\frac{c}{\rho}), \qquad f \in M_0^+
\end{aligned}$$

$(QDE_{ttshsurf})$

Figure 7.1: *Possible models in* $SPACE_{tt}$ *of the experimental system investigated in a tension test. The differences between* $QDE_{tteccsurf}$*,* $QDE_{ttgripsurf}$*, and* $QDE_{ttshsurf}$ *on the one hand, and* QDE_{ttsurf} *on the other, are shown in* **bold**. *The variables have the following interpretation:* σ_a *nominal applied stress,* σ_c *applied stress near crack,* F *applied force,* A *cross-sectional area of specimen,* l *and* l_0 *instantaneous and initial length,* r *radius,* d *diameter,* d_h *diameter at heads,* ϵ_a *nominal strain,* c *crack half-length,* γ *surface energy per unit area,* E *Young's modulus,* σ_{th} *theoretical strength,* σ_m *maximum stress at crack tip,* a_0 *interatomic distance,* ρ *crack tip radius,* e *eccentricity,* x *distance between crack and neutral axis,* f_{ec} *stress concentration factor,* rl *elongation rate,* r_t *transition radius of shoulders. The parameters* A, l_0, r, d, c, γ, E, a_0, ρ, e, x, x_{rel}, rl, r_t, d_h *are constants.*

$$\begin{aligned}
QS(init)_{ttsurf} = {} & \langle QV(l,t_0) = \langle\langle 0,\infty\rangle, _\rangle, QV(l_0,t_0) = \langle\langle 0,\infty\rangle, std\rangle, \\
& QV(E,t_0) = \langle\langle 0,\infty\rangle, std\rangle, QV(A,t_0) = \langle\langle 0,\infty\rangle, std\rangle, \\
& QV(rl,t_0) = \langle\langle 0,\infty\rangle, std\rangle, QV(\gamma,t_0) = \langle\langle 0,\infty\rangle, std\rangle, \\
& QV(a_0,t_0) = \langle\langle 0,\infty\rangle, std\rangle, QV(c,t_0) = \langle\langle 0,\infty\rangle, std\rangle, \\
& QV(\rho,t_0) = \langle\langle 0,\infty\rangle, std\rangle\rangle \\
QS(init)_{tteccsurf} = {} & QS(init)_{ttsurf} \oplus \\
& \langle QV(r,t_0) = \langle\langle 0,\infty\rangle, std\rangle, QV(x_{rel},t_0) = \langle 1, std\rangle, QV(e,t_0) = \langle 0, std\rangle\rangle \\
QS(init)_{ttgripsurf} = {} & QS(init)_{ttsurf} \\
QS(init)_{ttshsurf} = {} & QS(init)_{ttsurf} \oplus \\
& \langle QV(d,t_0) = \langle\langle 0,\infty\rangle, std\rangle, QV(d_h,t_0) = \langle\langle 0,\infty\rangle, std\rangle, \\
& QV(r_t,t_0) = \langle\langle 0,\infty\rangle, std\rangle\rangle
\end{aligned}$$

Figure 7.2: *Initial conditions for the models in figure 7.1. The initial conditions for $QDE_{tteccsurf}$, $QDE_{ttgripsurf}$, and $QDE_{ttshsurf}$ are expressed in terms of those for QDE_{ttsurf}. The $\oplus$ operation merges the argument tuples into a new tuple.*

$QDE_{tteccsurf}$ A specimen which is eccentrically loaded and fractures through a surface crack;

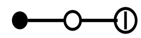
$QDE_{ttgripsurf}$ A specimen which is concentrically loaded by means of stress-concentrating grips and fractures through a surface crack at the hole;

$QDE_{ttshsurf}$ A specimen which is concentrically loaded and fractures through a surface crack at the shoulders.

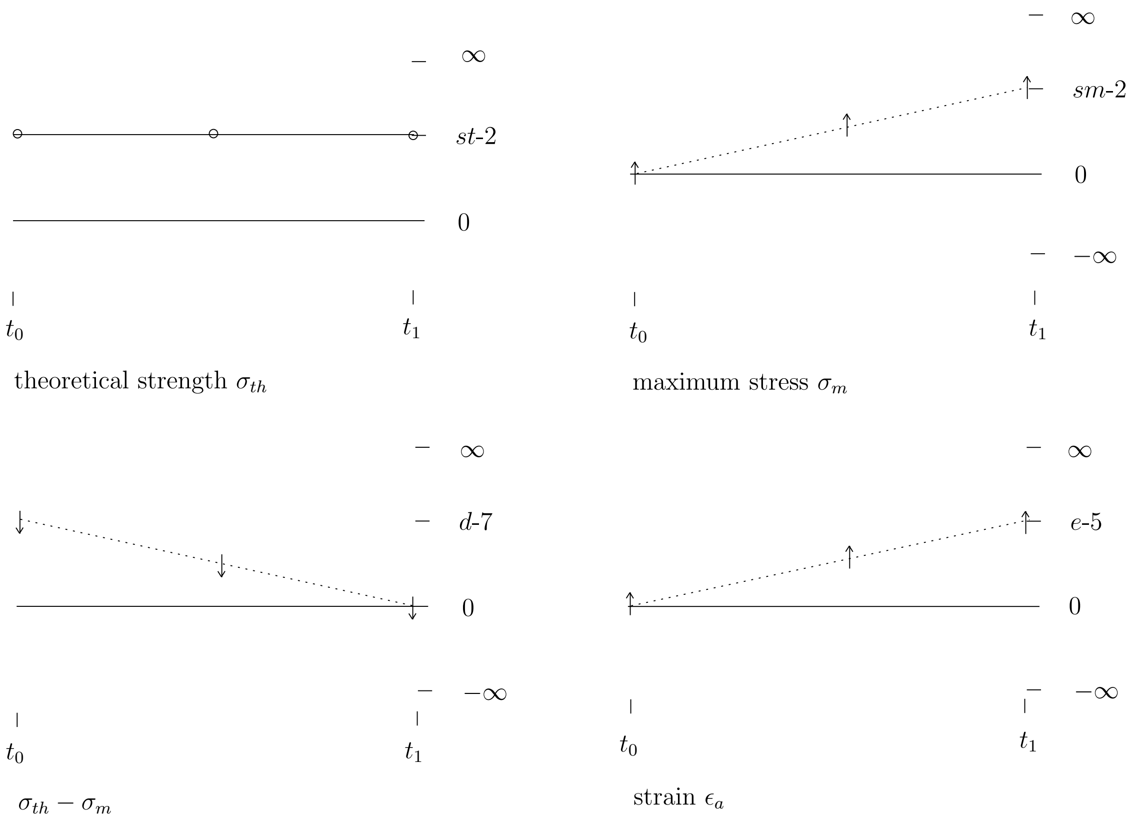
Notice that in $QDE_{tteccsurf}$ the critical crack is located on the side where the maximum stress due to bending is attained (see figure 6.6(b)). Fracture through a crack on the other side is unlikely to occur and hence not considered. The models QDE_{tt}, QDE_{ttecc}, QDE_{ttgrip}, and QDE_{ttsh} refer to specimens which fracture through an interior crack. They are obtained by a straightforward modification of the above QDEs, in particular by changing the numerical factor in the maximum stress formula from $2\sqrt{2}$ to 2 (compare equation (6.2) with equation (6.3)).

After simulation with the appropriate experimental conditions, each of the models of the experimental system in figure 7.1 gives rise to a single (non-spurious) qualitative behavior. The behaviors predicted from the models agree with each other as to the main features of the dynamics of the system (see figure 7.3(b) for a few key quantities). The behaviors are quite simple. As can be seen, the stress and strain increase until the stress at the crack tip reaches the theoretical strength ($\sigma_m = \sigma_{th}$), after which the material fails instantaneously. The fracture stress is the value of σ_a at t_1.

The ideal experimental system consists of a specimen which under loading has a stress distribution determined by simple tension theory (equation (6.7)). In addition, the



(a)



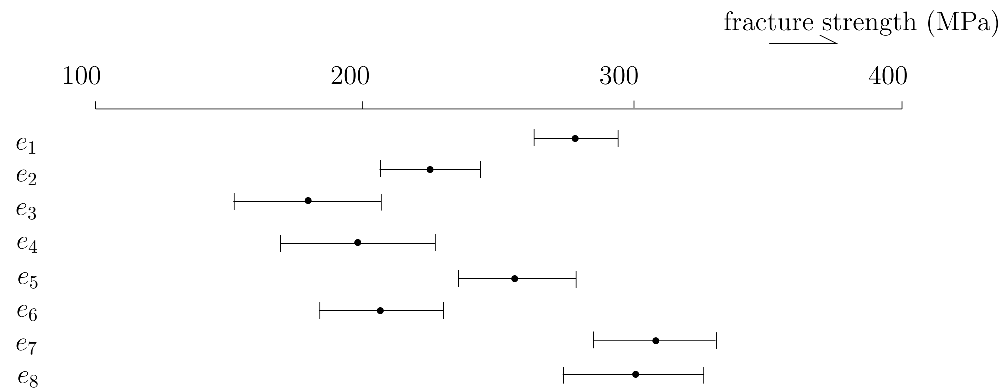
(b)

Figure 7.3: *(a) Behavior tree and (b) qualitative behavior obtained by simulation of the tension test models in figure 7.1. (For the notation, see figure 3.6.)*

specimen has the desired micro- and macro-structural properties created by appropriate fabrication and finishing processes. The model of the ideal system will therefore be equated with QDE_{ttsurf} or QDE_{tt}, depending on the way in which the surface of the specimen has been treated. Given a homogeneous crack distribution, the critical crack is most likely to be a surface crack, but extensive lapping and polishing may smoothen the surface of the specimen to the point that the specimen is more likely to fracture through an interior crack (Ohji [1988]).

7.2.2 Measurement sets and conflicts

Figure 7.4(a) shows the means and standard deviations of eight hypothetical sets of fracture strength measurements stored in a measurement base. Each measurement set is assumed to consist of 25 individual measurements performed on alumina specimens in a tension test at room temperature.



(a)

C_{tt}	pm_{e_1}	pm_{e_2}	pm_{e_3}	pm_{e_4}	pm_{e_5}	pm_{e_6}	pm_{e_7}	pm_{e_8}
pm_{e_1}	*noconf*	*conf*	*conf*	*conf*	*noconf*	*conf*	*conf*	*noconf*
pm_{e_2}		*noconf*	*conf*	*noconf*	*conf*	*noconf*	*conf*	*conf*
pm_{e_3}			*noconf*	*noconf*	*conf*	*noconf*	*conf*	*conf*
pm_{e_4}				*noconf*	*conf*	*noconf*	*conf*	*conf*
pm_{e_5}					*noconf*	*conf*	*conf*	*conf*
pm_{e_6}						*noconf*	*conf*	*conf*
pm_{e_7}							*noconf*	*noconf*
pm_{e_8}								*noconf*

(b)

Figure 7.4: *(a) Aggregated measurements of the fracture strength of alumina obtained in experiments* $e_1, \ldots, e_8$*. The experiments are tension tests performed at room temperature. (b) Conflict matrix* C_{tt} *for the aggregated measurements* $pm_{e_1}, \ldots, pm_{e_8}$*.*

In figure 7.5 an example of an individual measurement obtained in experiment e_1 is depictured. The two quantities measured during a tension test are the nominal stress σ_a and the nominal strain ϵ_a. Actually, these are indirect measurements calculated from measurements of the applied force F and the elongation $l - l_0$, additional measurements of the dimensions of the specimen, and the formulas in (6.7). Further, the elongation rate fixed by the test machine equals the reported strain rate $d\epsilon_a/dt$. The successive measurements of the stress and strain are summarized in the form of a *stress-strain diagram* in which the measurement points are tagged with a time stamp. The stress-strain diagram is represented as a measured state sequence (definition 32).

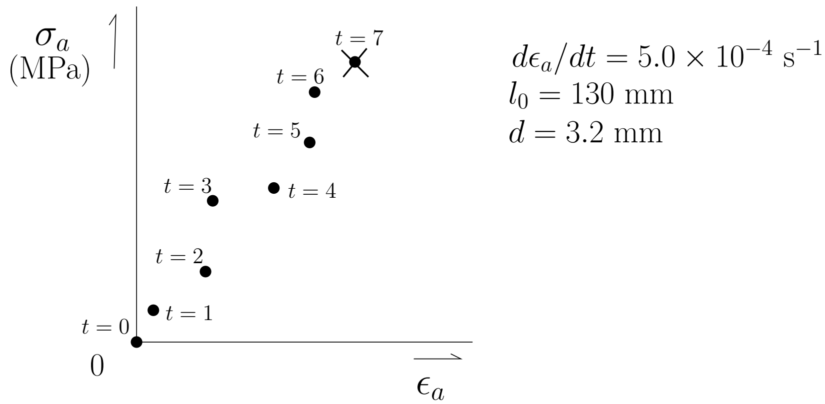


Figure 7.5: *Example of a fracture strength measurement obtained in experiment* e_1. *A cross indicates that the specimen fractures at that point.*

The error in the force and elongation measurements is set to 1%. Given the spread in fracture strength measurements, readily verified in figure 7.4(a), a large proportion of the measurements performed on individual specimens will be mutually conflicting. In addition, many pairs of aggregated measurements are in conflict, as shown in the conflict matrix of figure 7.4(b). The conflicts between the measurements have been calculated by means of the statistical criterion in definition 36.

7.2.3 Example 1: Grips

Suppose we compare a property measurement from e_5 with a property measurement from e_6. The first measurement claims the fracture strength of alumina to be 270 ± 3 MPa, whereas the second says it is 225 ± 2 MPa. Obviously, these measurements are in conflict. What might be the cause?

In experiment e_5 precautions have been taken to minimize eccentric loading and counteract undesired stress concentrations at the shoulders and grips of the specimens. In experiment e_6, the same precautions have been taken except for the prevention of high stress concentrations near the grips. As a consequence, the experimental systems in the two tests are described by the following candidate models:

$$CM = \{\langle QDE_{ttsurf}, QS(init)_{ttsurf}\rangle\},$$
$$\hat{CM} = \{\langle Q\hat{D}E_{ttgripsurf}, \hat{Q}S(init)_{ttgripsurf}\rangle\}.$$

That is, the specimen of the first measurement is most likely to have broken through a surface crack in the gage section, whereas the specimen of the second measurement is most likely to have broken through a crack on the surface of the hole through which the specimen is gripped. From equations (6.2) and (6.3) it is clear that surface cracks are greater stress concentrators than interior cracks and, as Creyke, Sainsbury & Morrell [1982] note:

> "[g]ripping a brittle test piece in jaws, or using a screw thread or other clamping device leads to such stress concentrations that failure is most likely at the grip, possibly before the test is proper under way" (p. 76).

The candidate models each give rise to a single candidate behavior (figure 7.3(b)) which is consistent with the measurements of σ_a and ϵ_a during the execution of the tests. In both experiments the stress and strain are positive and increasing after $t = 0$ s.

The conflict between the strength measurements is explained by the pairwise comparison of the candidate models and behaviors of the first system with the candidate models and behaviors of the second system. A single comparative analysis needs to be performed. When comparing the behavior derived from QDE_{ttsurf} with the behavior derived from $QDE_{ttgripsurf}$, two pairs of comparison are found: $pc_0 = \langle t_0, \hat{t}_0\rangle$ and $pc_1 = \langle t_1, \hat{t}_1\rangle$. The conflict between the fracture strengths is represented by $\sigma_a \Downarrow_{pc_1}$ and a comparison of the measured states at pc_1 further yields $\epsilon_a \Downarrow_{pc_1}$. When asserting that the specimens have the same microstructure, surface conditions, and dimensions, and that the elongation rate is the same in both experiments, the following initial comparative state information is obtained:

$$CS(init) = \langle \sigma_a \Downarrow_{pc_1}, \epsilon_a \Downarrow_{pc_1}, E \parallel_{pc_1}, l_0 \parallel_{pc_1}, \gamma \parallel_{pc_1}, rl \parallel_{pc_1}, a_0 \parallel_{pc_1}\rangle.$$

The RV constraints employed by CEC* are inferred from the QDEs and qualitative behaviors in the usual fashion (chapter 4). Actually, they are quite simple to obtain in this case since the experimental systems are modeled as first-order systems. More specifically, theorems 6, 7, 9, and 11 have been used to derive RV constraints at and between pc_0 and pc_1.

The comparative envisionment derived from $CS(init)$, when the first system is described by QDE_{ttsurf} and the second by $Q\hat{D}E_{ttgripsurf}$, is shown in figure 7.6. In order not to clutter the envisionment with comparative behaviors expressing distinctions that do not add much to the overall picture, the single crack shape parameter $s_c = c/\rho$ has been used. There are three (genuine) comparative behaviors and hence three explanations of the conflict. The difference in the structure of the systems (interior crack versus surface crack and stress concentration at the grips) is already sufficient for explaining the difference in fracture strength (see the comparative behavior ending in $CS(pc_0)_1$). However,

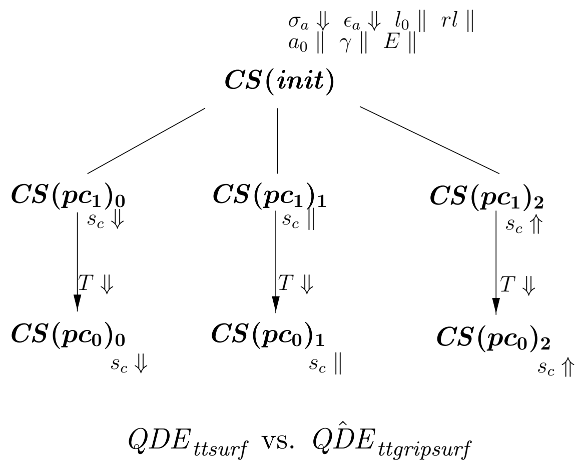


Figure 7.6: *The comparative envisionment arising from the comparison of an experimental system free from the disturbances modeled in $SPACE_{tt}$ and an experimental system with a stress concentration at the grips. The envisionment represents possible explanations of the conflict $\sigma_a \Downarrow_{pc_1}$. A few distinctive RVs are indicated at the comparative states.*

a difference in crack shape may strengthen ($s_c \Uparrow_{pc_0}$ in $CS(pc_0)_0$) or weaken ($s_c \Downarrow_{pc_0}$ in $CS(pc_0)_2$) the structural difference. In the former case, we have a longer and/or sharper crack and in the latter a shorter and/or blunter crack. Notice that all three comparative behaviors hypothesize the relative duration $RV(T)$ of the intervals between pc_0 and pc_1 to be $\Downarrow$. In other words, it takes longer in the first experiment to break the sample.

Figure 7.4(b) shows that there is also a conflict between the aggregated measurements obtained in experiments e_5 and e_6. This conflict can be explained by extending the above analysis to the populations of individual measurements sampled in e_5 and e_6. The aggregated measurements are estimations of the mean of these populations. Given a normal distribution of the individual fracture strength measurements, the conflict between the aggregated measurements implies that if we take an arbitrary individual measurement from the first population and an arbitrary individual measurement from the second population, we are most likely to find $\sigma_a \Downarrow_{pc_1}$. The fracture modes of the rods on which the measurements are performed are most likely described by QDE_{ttsurf} and $Q\hat{D}E_{ttgripsurf}$. The comparative envisionment in figure 7.6 reveals that at the given differences in experimental conditions this may lead to $\sigma_a \Downarrow_{pc_1}$ in the case of two arbitrary individual measurements, and hence a lower fracture strength on average.

7.2.4 Example 2: Eccentric loading and surface finishing

As a second example, suppose an individual strength measurement obtained from e_1 is compared with an individual strength measurement from e_2. The second measurement (210 ± 2 MPa) is substantially lower than the first (285 ± 3 MPa), so we have a conflict that needs to be explained.

Whereas the first measurement has been made in an experiment in which the specimens were carefully polished after machining so as to remove surface damage induced by inappropriate grinding, the second measurement originates from an experiment in which no such precautions were taken. In both experiments the specimens are tapered and a buffer material is used to reduce the stress concentrations at the grips. Further, in both experiments concentric loading of the specimen cannot be assured. This results in the following candidate models for the experimental systems investigated in the tests:

$$CM = \{\langle QDE_{tt}, QS(init)_{tt}\rangle, \langle QDE_{ttecc}, QS(init)_{ttecc}\rangle\},$$
$$\hat{CM} = \{\langle Q\hat{D}E_{ttsurf}, \hat{Q}S(init)_{ttsurf}\rangle, \langle Q\hat{D}E_{tteccsurf}, \hat{Q}S(init)_{tteccsurf}\rangle\}.$$

That is, the specimen of the first measurement is most likely to have fractured through an interior crack, possibly under conditions of eccentric loading. Smoothening the surface of the specimen tends to increase the chance of internal fracture to the point that it is more likely to occur than surface fracture. (For instance, Ohji [1988] reports an 84% chance of internal fracture in the case of silicon nitride and 50 μm surface removal.) The second specimen is most likely to have fractured through a surface crack, even more so because cracks caused by damaging machining operations tend to be larger than intrinsic cracks (Davidge [1979]).

Again, we simulate the models with their associated experimental conditions and perform a comparative analysis with the models and resulting qualitative behaviors. Since QSIM finds a single behavior for every model (figure 7.3), a total of four comparative analyses needs to be performed. Consider the comparison of QDE_{tt} and $Q\hat{D}E_{ttcecsurf}$, and assume that the measured states and additional information about the experiments yield

$$CS(init) = \langle \sigma_a \Downarrow_{pc_1}, \epsilon_a \Downarrow_{pc_1}, E \,\|_{pc_1}, l_0 \,\|_{pc_1}, \gamma \,\|_{pc_1}, rl \,\|_{pc_1}, a_0 \,\|_{pc_1}, s_c \Uparrow_{pc_1}\rangle,$$

where $s_c \Uparrow_{pc_1}$ accounts for the fact that cracks due to surface damage are usually larger than intrinsic cracks.

CEC* finds a single genuine comparative behavior explaining the conflict between the two measurements. This comparative behavior is depictured in figure 7.7. The comparative behavior explains the lower fracture strength of the second measurement by the larger stress-concentrating effect of a surface crack as compared to an interior crack, the greater length of a crack due to surface damage, and the non-uniform stress-distribution caused by eccentric loading. Notice that these differences in the structure of the physical systems and the experimental conditions not only work in the same direction, but enhance each other's effect. From equation (6.8) it follows that the maximum effect of eccentric loading is reached at the surface.

The results confirm the remarks by several authors that surface finish is an important determinant of the fracture strength and that avoiding machining damage is critical, "lest the test merely become a measure of machining damage" (Quinn [1991], p. 589).

Similar analyses have been carried out for the other combinations of candidate models and candidate behaviors. In all, the conflict resolution method produces eight different

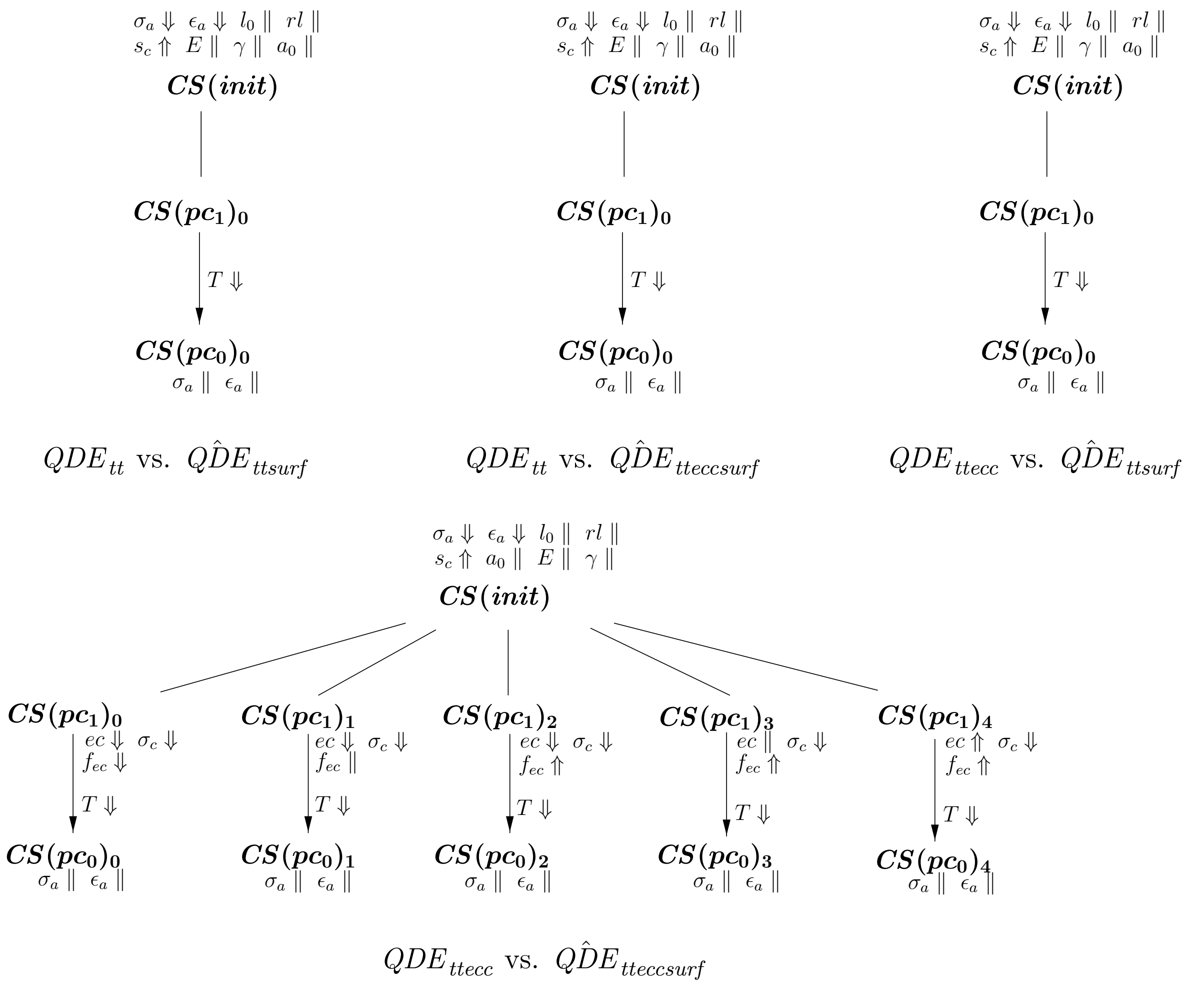


Figure 7.7: *The comparative envisionments arising from the comparison of two experimental systems with different surface finishes. The first specimen has been polished after machining, whereas the second specimen is tested 'as-machined'. Both experimental systems could be eccentrically loaded. The envisionments represent possible explanations of the conflict $\sigma_a \Downarrow_{pc_1}$. A few distinctive RVs are indicated at the comparative states.*

explanations of the conflict using the models. When comparing the models QDE_{ttecc} with $Q\hat{D}E_{tteccsurf}$, which both assume eccentric loading, five comparative behaviors are obtained. The cause of the ambiguities lies in the fact that we do not know in which of the two experimental systems the eccentricity e is larger.

Figure 7.4(b) shows that the aggregated measurements from experiments e_1 and e_2 are in conflict too. The above explanations of the conflict between two individual measurements provide a clue for the resolution of the conflict on the aggregated level. Again assume that we compare an arbitrary individual measurement from the population of e_1 with an arbitrary individual measurement from the population of e_2. If the first system is eccentrically loaded and the second is not, we have QDE_{ttecc} and $Q\hat{D}E_{ttsurf}$ as the fracture modes most likely to occur. Figure 7.7 shows that this may lead to $\sigma_a \Downarrow_{pc_1}$ at the given differential experimental conditions. Hence, when a large number of measurements from the two populations are compared, the fracture strength could be lower on average. For other assumptions about the loading conditions, a similar explanation of the conflict on the aggregated level can be given.

7.2.5 Example 3: Eccentricity and surface damage

In the above two examples conflicts between measurements were analyzed, while questions about the accuracy of the measurements were not posed. Here I will consider the accuracy of a measurement 195 ± 2 MPa obtained in experiment e_4.

On the basis of the description of the experiment, it is assumed that the specimens used in the experiments were probably damaged because grinding was performed by a motion perpendicular to, rather than parallel to, the specimen length axis (which causes flaws that tend to be critical, see Baratta [1984] and Ohji [1988]). Further, eccentric loading cannot be excluded as a possible error source. We arrive at the following candidate models for the actual experimental systems:

$$C\hat{M} = \{\langle Q\hat{D}E_{ttsurf}, \hat{Q}S(init)_{ttsurf}\rangle, \langle Q\hat{D}E_{tteccsurf}, \hat{Q}S(init)_{tteccsurf}\rangle\},$$

while the model of the ideal experimental system is described by QDE_{ttsurf} with initial conditions $QS(init)_{ttsurf}$ (section 7.2.1).

It is likely that the specimen on which the measurement was performed failed through a machining crack on the surface, in a region with a local applied stress $\hat{\sigma}_c$ higher than the nominal applied stress $\hat{\sigma}_a$ due to eccentric loading.

Systematic errors in the property measurement are identified by systematically comparing every possible qualitative model and behavior of the actual experimental system with the model and behavior of the ideal experimental system. Hence two comparative analyses must be performed: QDE_{ttsurf} versus $Q\hat{D}E_{ttsurf}$ and QDE_{ttsurf} versus $Q\hat{D}E_{tteccsurf}$.

Consider the comparison of QDE_{ttsurf} and $Q\hat{D}E_{tteccsurf}$. If the specimens used in the actual experiment have the desired microstructure and are loaded at the same rate, we have

$$CS(init) = \langle E \parallel_{pc_0}, l_0 \parallel_{pc_0}, \gamma \parallel_{pc_0}, rl \parallel_{pc_0}, a_0 \parallel_{pc_0}, s_c \Uparrow_{pc_0} \rangle,$$

with again $s_c \Uparrow_{pc_0}$ accounting for the fact that cracks due to surface damage are usually larger than intrinsic cracks. Obviously, $\sigma_a \parallel_{pc_0}$ and $\epsilon_a \parallel_{pc_0}$, because loading starts from an unstressed state.

The resulting comparative envisionment, consisting of a single comparative behavior, is shown in figure 7.8(a). The measured fracture strength is predicted to be too low in comparison with the ideal case. Eccentric loading and surface damage due to inappropriate machining contribute to the systematic error in the measurement; both error sources working in the same direction and even enhancing each other's effect. As revealed by the figure, the comparison of QDE_{ttsurf} versus $\hat{QDE}_{ttsurf}$ also leads to $\sigma_a \Downarrow_{pc_1}$. The conclusion to be drawn is that the measured fracture strength 195 ± 2 MPa is likely to be too low. The strength-decreasing effects of eccentric loading and surface damage are widely reported in the fracture test theory.

Now suppose instead that the specimens used in the actual experiment are stiffer than desired, i.e. $E \Uparrow_{pc_0}$. This might be caused by a lower volume fraction porosity of the specimens (Dörre & Hübner [1984]). On repeating the comparative analyses with the new $CS(init)$, the envisionments in figure 7.8(b) are obtained. Whereas a higher elasticity modulus increases σ_{th}, and thus tends to make the measured fracture strength too high, surface damage and eccentric loading tend to make it too low. The effect of the higher elasticity modulus counteracts the effects of surface damage and eccentric loading, which causes ambiguities. As shown in the figure, there may be a positive systematic error in the measured value ($\sigma_a \Uparrow_{pc_1}$), a negative systematic error ($\sigma_a \Downarrow_{pc_1}$), or no systematic error at all due to the masking of one error by another ($\sigma_a \parallel_{pc_1}$). In a similar way, ambiguities about the relative durations of the ideal and actual experiments arise.

The above assessments of the possible systematic errors in an individual measurement from experiment e_4 help in judging the accuracy of the aggregated measurement (197 ± 30 MPa). First consider the case that $E \parallel$. When taking an arbitrary individual measurement from the population sampled in e_4, it is likely to be too low, whether the experimental system has been eccentrically loaded or not. So the average of a large number of measurements from the population is likely to be too low as well, and the true value of the fracture strength will be higher than 197 ± 30 MPa. No such conclusion can be drawn when $E \Uparrow$, due to ambiguities in the predicted systematic error.

7.2.6 Example 4: Non-brittle fracture

As a final tension test example, consider the conflicting property measurements 175 ± 2 MPa and 326 ± 3 MPa, performed in experiments e_3 and e_7, respectively. The measured states of the experimental systems on which the measurements were performed are shown in figure 7.9.

The experiments e_3 and e_7 are assumed to have been carried out in such a way that the disturbances accounted for by the models in $SPACE_{tt}$ cannot have occurred. As a

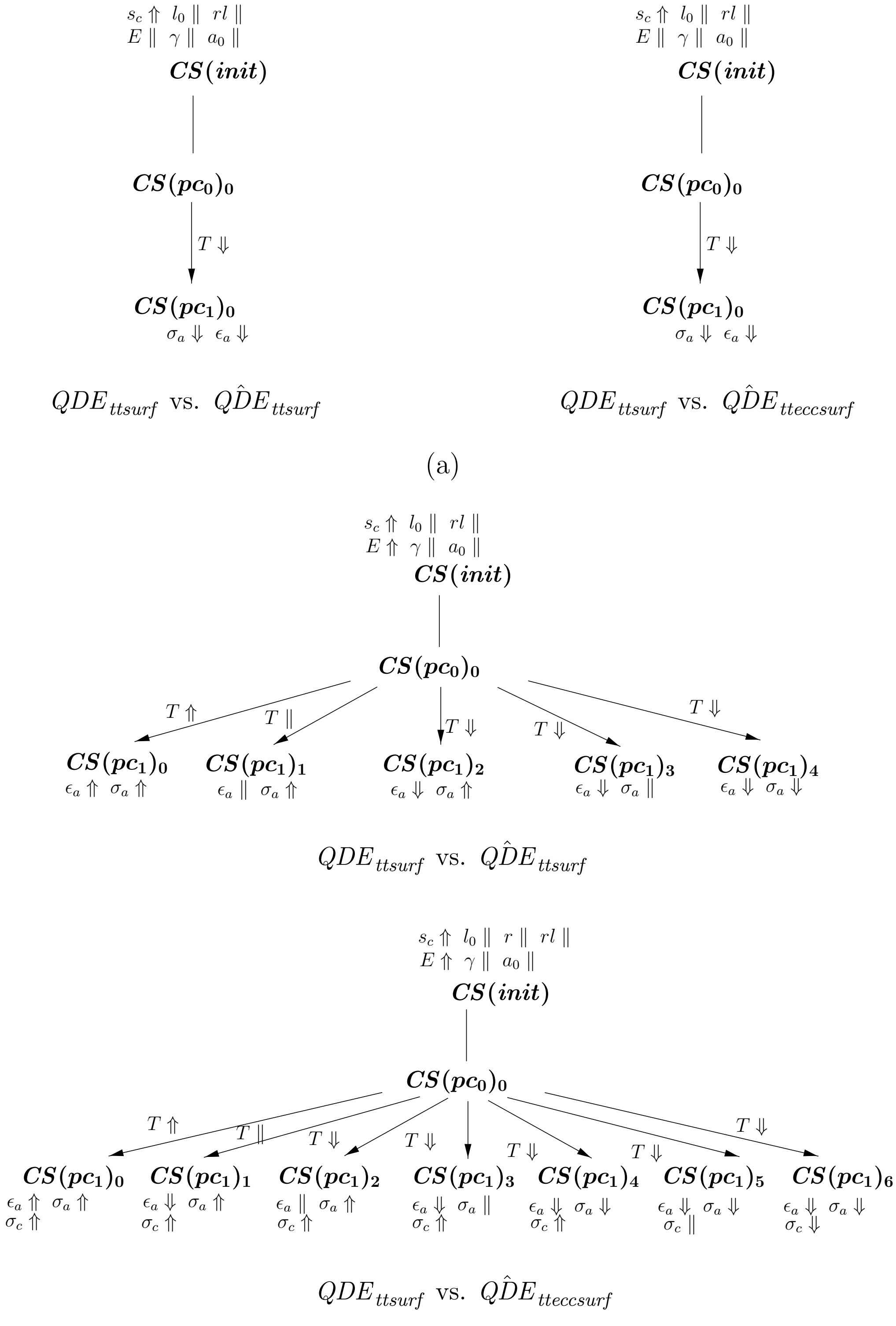


Figure 7.8: *(a) The comparative envisionments arising from the comparison of the ideal tension test system and an experimental system being damaged at the surface and possibly eccentrically loaded as well. In (b) the same comparisons are carried out with the additional deviation $E \Uparrow$. The comparative behaviors predict $RV(\sigma_a, pc_1)$.*

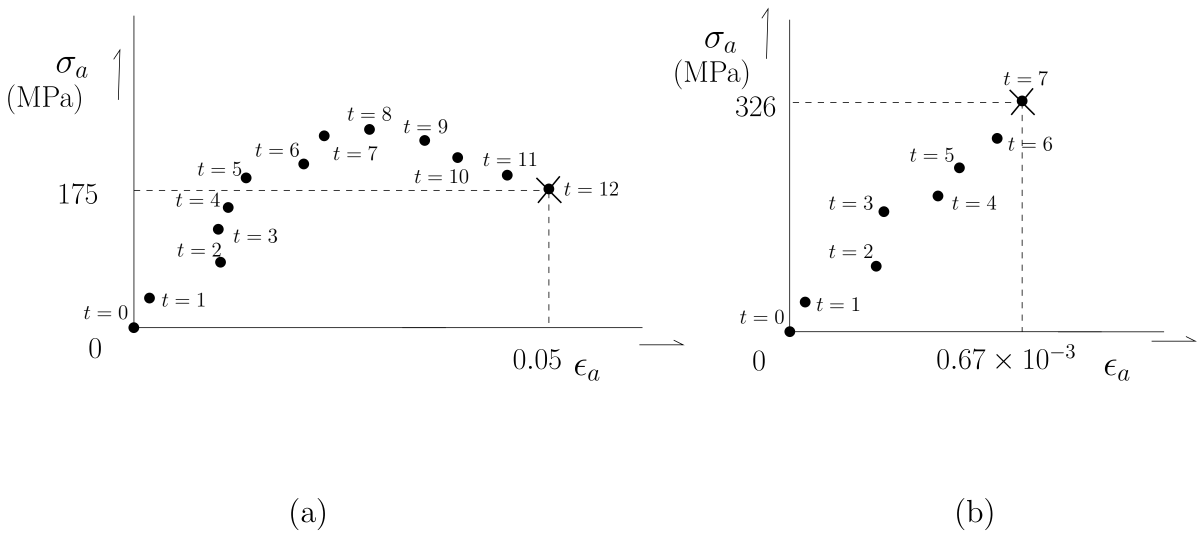


Figure 7.9: *Fracture strength measurements performed on the experimental systems investigated in (a) experiment* e_3 *and (b) experiment* e_7*.*

consequence, the sets of candidate models consist of a single (ideal) model:

$$CM = \{\langle QDE_{ttsurf}, QS(init)_{ttsurf}\rangle\},$$
$$\hat{CM} = \{\langle \hat{QDE}_{ttsurf}, \hat{QS}(init)_{ttsurf}\rangle\}.$$

The conflict resolution algorithm first performs a qualitative simulation and then proceeds with an attempt to eliminate qualitative behaviors in the behavior tree that are not consistent with the measured states. Up to now, the latter step was taken for granted and assumed not to rule out any of the behaviors in the behavior tree. When we look at the measured states in figure 7.9, however, we see that the only qualitative behavior found for QDE_{ttsurf} does not agree with the stress measurements. Whereas a monotonically increasing applied stress is predicted to accompany an increasing strain (figure 7.3), the measured states show that σ_a reaches a maximum value and subsequently decreases. Eliminating the only qualitative behavior for QDE_{ttsurf} implies that the model itself needs to be ruled out as well, and CM becomes empty. In other words, it is not possible to give an explanation of the conflict.

Assuming that the measurements of the applied stress are correct, we must conclude that the model space $SPACE_{tt}$ is incomplete. Figure 7.9 suggests that the specimen does not fail in a brittle way. Judging from the shape of the stress-strain curve, a substantial amount of plastic deformation must have occurred. We might hypothesize that experiment e_3 is a high-temperature determination of the fracture strength of alumina, instead of a determination carried out at room temperature (see Davidge [1979], ch. 1, for exemplary stress-strain diagrams for brittle and ductile fracture).

7.3 Conflict resolution and error identification in four-point bend tests

7.3.1 Model space

For the analysis of four-point bend tests the following model space will be used:

$$SPACE_{bt} = \{QDE_{btsurf}, QDE_{btfricsurf}, QDE_{btuneqsurf}, QDE_{btmissurf}, QDE_{btwgsurf}, QDE_{bttwsurf}, QDE_{btmisfricsurf}\}. \quad (7.1)$$

The models in $SPACE_{bt}$, with their associated initial conditions, are shown in figure 7.10 and 7.11. They describe the following experimental systems:

QDE_{btsurf} A specimen which is loaded according to simple bend theory and fractures through a crack at the tensile surface;

$QDE_{btfricsurf}$ A specimen which experiences friction forces during loading and fractures through a crack at the tensile surface;

$QDE_{btuneqsurf}$ A specimen which is unequally loaded and fractures through a crack at the tensile surface;

$QDE_{btmissurf}$ A specimen which is loaded through shifted load points and fractures through a crack at the tensile surface;

$QDE_{btwgsurf}$ A specimen which experiences wedging forces at the load points and fractures through a crack at the tensile surface;

$QDE_{bttwsurf}$ A twisted specimen which fractures through a crack at the tensile surface;

$QDE_{btmisfricsurf}$ A specimen which is loaded through shifted load points and experiences friction forces in addition.

In all of the above models the critical crack is located between the load points ($x < |a|$), at the tensile surface of the specimen. We could have included models in $SPACE_{bt}$ accounting for (1) fracture between the load points and support points, (2) fracture in the interior of the specimen, and (3) fracture at the compressive surface. However, the existence of large stress gradients between the tensile surface ($y = d/2$) and the mid-depth line ($y = 0$), and between the load points ($x = |a|$) and support points ($x = |2a|$), added to the stronger stress-concentrating effect of surface cracks, makes it unlikely that the fracture modes described by such models will occur (Davidge [1979]). As a consequence, they are omitted from consideration.

As a further simplification, it is implicitly assumed that in the case of unequal loading, load mislocation, wedging, and twisting, fracture will be initiated near the place where the *maximum* deviation from the uniform stress distribution is reached. The formulas in

$$
\begin{aligned}
\epsilon_a &= \frac{1}{E}\sigma_a \\
\epsilon_a &= \frac{6y_d d}{l_{out}^2}, \qquad \sigma_a = \frac{3Fl_{out}}{4bd^2} \\
\sigma_c &= \sigma_a \\
\frac{dy_d}{dt} &= ry_d \\
\sigma_{th} &= h(\frac{\gamma E}{a_0}), \qquad h \in M_0^+ \\
\sigma_m &= 2\sqrt{2}\sigma_c f(\frac{c}{\rho}), \qquad f \in M_0^+
\end{aligned}
$$

(QDE$_{btsurf}$)

$$
\begin{aligned}
\epsilon_a &= \frac{1}{E}\sigma_a \\
\epsilon_a &= \frac{6y_d d}{l_{out}^2}, \qquad \sigma_a = \frac{3Fl_{out}}{4bd^2} \\
\sigma_c &= \boldsymbol{f_{fric}}\sigma_a \\
\boldsymbol{f_{fric}} &\boldsymbol{= 1 - 2\mu\frac{d}{l_{out}}} \\
\frac{dy_d}{dt} &= ry_d \\
\sigma_{th} &= h(\frac{\gamma E}{a_0}), \qquad h \in M_0^+ \\
\sigma_m &= 2\sqrt{2}\sigma_c f(\frac{c}{\rho}), \qquad f \in M_0^+
\end{aligned}
$$

(QDE$_{btfricsurf}$)

$$
\begin{aligned}
\epsilon_a &= \frac{1}{E}\sigma_a \\
\epsilon_a &= \frac{6y_d d}{l_{out}^2}, \qquad \sigma_a = \frac{3Fl_{out}}{4bd^2} \\
\sigma_c &= \boldsymbol{f_{uneq}}\sigma_a \\
\boldsymbol{f_{uneq}} &\boldsymbol{= 1 + 2\frac{\Delta F}{F}} \\
\frac{dy_d}{dt} &= ry_d \\
\sigma_{th} &= h(\frac{\gamma E}{a_0}), \qquad h \in M_0^+ \\
\sigma_m &= 2\sqrt{2}\sigma_c f(\frac{c}{\rho}), \qquad f \in M_0^+
\end{aligned}
$$

(QDE$_{btuneqsurf}$)

$$
\begin{aligned}
\epsilon_a &= \frac{1}{E}\sigma_a \\
\epsilon_a &= \frac{6y_d d}{l_{out}^2}, \qquad \sigma_a = \frac{3Fl_{out}}{4bd^2} \\
\sigma_c &= \boldsymbol{f_{mis}}\sigma_a \\
\boldsymbol{f_{mis}} &\boldsymbol{= 1 + 2\frac{\Delta x}{l_{out}}} \\
\frac{dy_d}{dt} &= ry_d \\
\sigma_{th} &= h(\frac{\gamma E}{a_0}), \qquad h \in M_0^+ \\
\sigma_m &= 2\sqrt{2}\sigma_c f(\frac{c}{\rho}), \qquad f \in M_0^+
\end{aligned}
$$

(QDE$_{btmissurf}$)

Figure 7.10: *Possible models in SPACE$_{bt}$ of the experimental system investigated in a four-point bend test. The differences between QDE$_{btfricsurf}$, QDE$_{btuneqsurf}$, QDE$_{btmissurf}$, QDE$_{btwgsurf}$, QDE$_{bttwsurf}$, and QDE$_{btmisfricsurf}$ on the one hand, and QDE$_{btsurf}$ on the other, are shown in* **bold**. *The variables have the following interpretation: σ_a nominal applied stress, σ_c applied stress near crack, F applied force, d and b depth and width of specimen, l_{out} outer span length, ϵ_a nominal strain, c crack half-length, γ surface energy per unit area, E Young's modulus, σ_{th} theoretical strength, σ_m maximum stress at crack tip, a_0 interatomic distance, ρ crack tip radius, y_d deflection, ry_d deflection rate, μ coefficient of static friction, ΔF load difference, Δx shift of load points, k_2 twisting coefficient. The variables f_{fric}, f_{uneq}, f_{mis}, f_{wg}, and f_{tw} are stress concentration factors. The parameters b, d, l_{out}, c, γ, E, a_0, ρ, μ, Δx, ΔF, ry_d, k_2 are constants.*

$$\epsilon_a = \frac{1}{E}\sigma_a$$
$$\epsilon_a = \frac{6y_d d}{l_{out}^2}, \qquad \sigma_a = \frac{3Fl_{out}}{4bd^2}$$
$$\sigma_c = \boldsymbol{f_{wg}}\sigma_a$$
$$\boldsymbol{f_{wg} = 1 + 0.1157\frac{d}{l_{out}}}$$
$$\frac{dy_d}{dt} = ry_d$$
$$\sigma_{th} = h(\frac{\gamma E}{a_0}), \qquad h \in M_0^+$$
$$\sigma_m = 2\sqrt{2}\sigma_c f(\frac{c}{\rho}), \qquad f \in M_0^+$$

($QDE_{btwgsurf}$)

$$\epsilon_a = \frac{1}{E}\sigma_a$$
$$\epsilon_a = \frac{6y_d d}{l_{out}^2}, \qquad \sigma_a = \frac{3Fl_{out}}{4bd^2}$$
$$\sigma_c = \boldsymbol{f_{tw}}\sigma_a$$
$$\boldsymbol{f_{tw} = \frac{1}{2} + g(\frac{b}{l}, k_2), \qquad g \in M^{+-}}$$
$$\frac{dy_d}{dt} = ry_d$$
$$\sigma_{th} = h(\frac{\gamma E}{a_0}), \qquad h \in M_0^+$$
$$\sigma_m = 2\sqrt{2}\sigma_c f(\frac{c}{\rho}), \qquad f \in M_0^+$$

($QDE_{bttwsurf}$)

$$\epsilon_a = \frac{1}{E}\sigma_a$$
$$\epsilon_a = \frac{6y_d d}{l_{out}^2}, \qquad \sigma_a = \frac{3Fl_{out}}{4bd^2}$$
$$\sigma_c = \boldsymbol{f_{fric} f_{mis}}\sigma_a$$
$$\boldsymbol{f_{mis} = 1 + 2\frac{\Delta x}{l_{out}}}$$
$$\boldsymbol{f_{fric} = 1 - 2\mu\frac{d}{l_{out}}}$$
$$\frac{dy_d}{dt} = ry_d$$
$$\sigma_{th} = h(\frac{\gamma E}{a_0}), \qquad h \in M_0^+$$
$$\sigma_m = 2\sqrt{2}\sigma_c f(\frac{c}{\rho}), \qquad f \in M_0^+$$

($QDE_{btmisfricsurf}$)

Figure 7.10: *(Continued)*

$$
\begin{aligned}
QS(init)_{btsurf} =\ & \langle QV(l_{out}, t_0) = \langle\langle 0, \infty\rangle, std\rangle, QV(y_d, t_0) = \langle 0, \_\rangle, \\
& QV(E, t_0) = \langle\langle 0, \infty\rangle, std\rangle, QV(b, t_0) = \langle\langle 0, \infty\rangle, std\rangle, \\
& QV(d, t_0) = \langle\langle 0, \infty\rangle, std\rangle, QV(ry_d, t_0) = \langle\langle 0, \infty\rangle, std\rangle, \\
& QV(\gamma, t_0) = \langle\langle 0, \infty\rangle, std\rangle, QV(a_0, t_0) = \langle\langle 0, \infty\rangle, std\rangle, \\
& QV(c, t_0) = \langle\langle 0, \infty\rangle, std\rangle, QV(\rho, t_0) = \langle\langle 0, \infty\rangle, std\rangle\rangle \\
QS(init)_{btfricsurf} =\ & QS(init)_{btsurf} \oplus \langle QV(\mu, t_0) = \langle\langle 0, \infty\rangle, std\rangle\rangle \\
QS(init)_{btuneqsurf} =\ & QS(init)_{btsurf} \oplus \langle QV(\Delta F, t_0) = \langle\langle 0, \infty\rangle, std\rangle\rangle \\
QS(init)_{btmissurf} =\ & QS(init)_{btsurf} \oplus \langle QV(\Delta x, t_0) = \langle\langle 0, \infty\rangle, std\rangle\rangle \\
QS(init)_{btwgsurf} =\ & QS(init)_{btsurf} \\
QS(init)_{bttwsurf} =\ & QS(init)_{btsurf} \oplus \langle QV(k_2, t_0) = \langle\langle 0, \infty\rangle, std\rangle\rangle \\
QS(init)_{btmisfricsurf} =\ & QS(init)_{btsurf} \oplus \langle QV(\mu, t_0) = \langle\langle 0, \infty\rangle, std\rangle, QV(\Delta x, t_0) = \langle\langle 0, \infty\rangle, std\rangle\rangle
\end{aligned}
$$

Figure 7.11: *Initial conditions of the models in figure 7.10. As in figure 7.2, the $\oplus$ symbol represents a merge operation between tuples.*

section 6.4.2, on which the QDEs are based, describe this maximum deviation. Given a homogeneous crack distribution, one can reasonably make this assumption.

Except for $QDE_{btmisfricsurf}$, every model in $SPACE_{bt}$ gives rise to a single (non-spurious) qualitative behavior when a simulation is performed with the experimental conditions in figure 7.10. The behaviors are straightforward and show how the stress σ_a and the strain ϵ_a increase until the Orowan criterion is satisfied (i.e., $\sigma_m = \sigma_{th}$), after which failure occurs immediately (figure 7.12(c)). Again, the fracture stress is the value of σ_a at t_1. $QDE_{btmisfricsurf}$ leads to three qualitative behaviors which only differ in the qualitative value of $f_{fric} f_{mis}$. Since the qualitative magnitude of f_{fric} is $]0, 1[$ and that of f_{mis} $]1, \infty[$, their product can be larger than, equal to, or smaller than 1. In the first case the effect of the shift in load points outweighs the friction effect, in the second the effects cancel, and in the third the latter effect outweighs the former.

Ideally, the fracture strength should be measured on a beam failing through an intrinsic surface crack in the uniformly stressed region between the load points. The stress distribution in the beam should be as predicted by simple beam theory (equation (6.9)), which makes the stress in the vicinity of the critical crack (σ_c) equal to the reported nominal applied stress (σ_a). On the basis of these considerations, the model $QDE_{btsurf} \in SPACE_{bt}$ is an adequate representation of the structure of the ideal four-point bend test system.

7.3.2 Measurement sets and conflicts

In figure 7.13(a) the means and standard deviations of nine sets of measurements of the fracture strength of alumina are shown. Each measurement set consists of 25 individual measurements performed in a four-point bend test at room temperature.

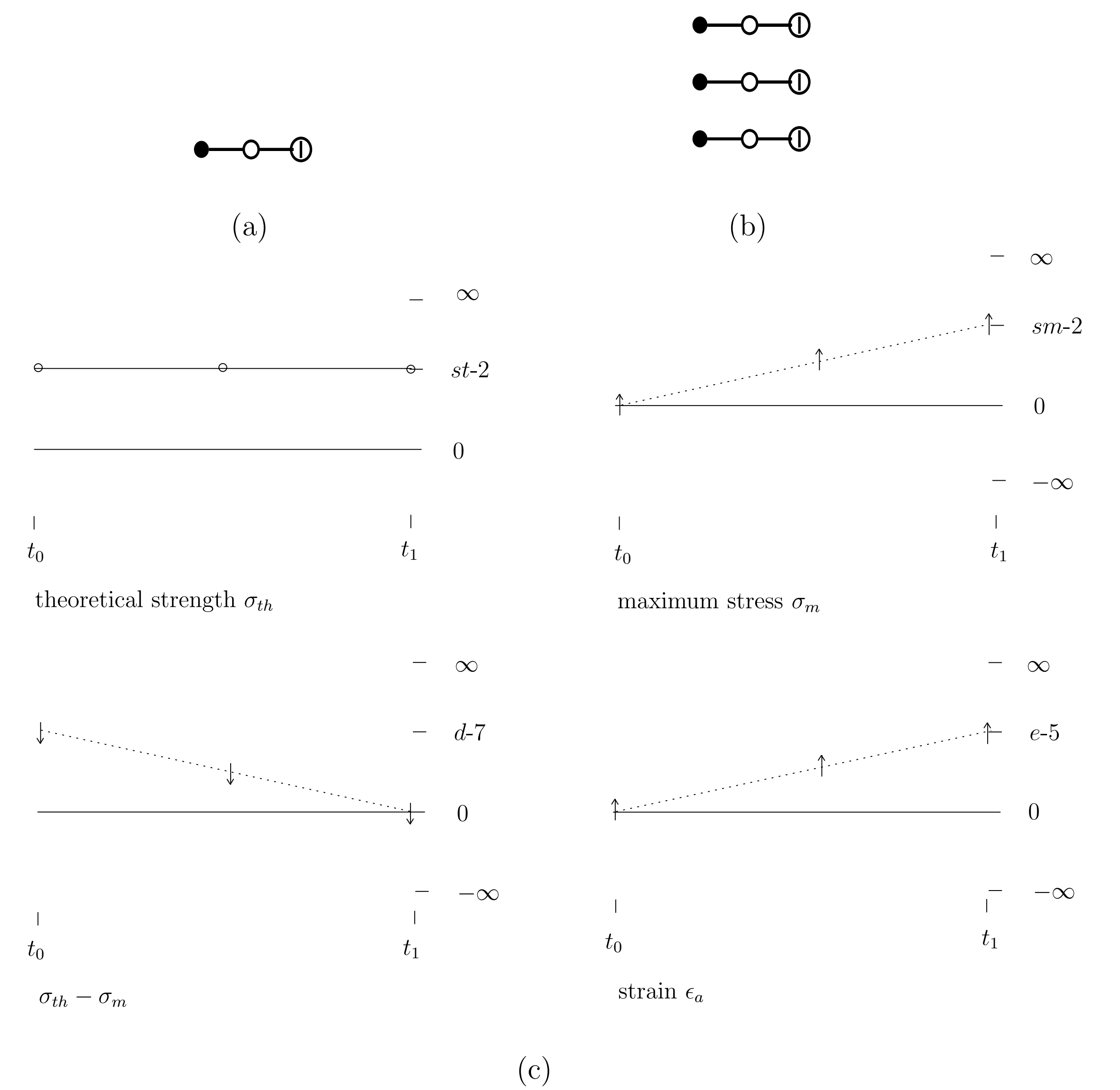
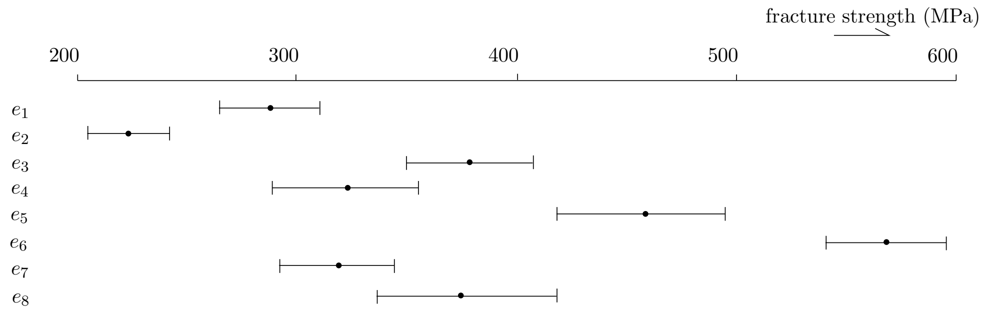


Figure 7.12: *(a)-(b) Behavior trees and (c) qualitative behavior obtained by simulation of the four-point bend test models in figure 7.10.*



(a)

C_{bt}	pm_{e_1}	pm_{e_2}	pm_{e_3}	pm_{e_4}	pm_{e_5}	pm_{e_6}	pm_{e_7}	pm_{e_8}
pm_{e_1}	*noconf*	*conf*	*conf*	*noconf*	*conf*	*conf*	*noconf*	*conf*
pm_{e_2}		*noconf*	*conf*	*conf*	*conf*	*conf*	*conf*	*conf*
pm_{e_3}			*noconf*	*conf*	*conf*	*conf*	*conf*	*noconf*
pm_{e_4}				*noconf*	*conf*	*conf*	*noconf*	*noconf*
pm_{e_5}					*noconf*	*conf*	*conf*	*conf*
pm_{e_6}						*noconf*	*conf*	*conf*
pm_{e_7}							*noconf*	*conf*
pm_{e_8}								*noconf*

(b)

Figure 7.13: *(a) Aggregated measurements of the fracture strength of alumina obtained in experiments* $e_1, \ldots, e_8$. *The experiments are four-point bend tests performed at room temperature. (b) Conflict matrix* C_{bt} *for the aggregated measurements* $pm_{e_1}, \ldots, pm_{e_8}$.

The nominal stress σ_a and nominal strain ϵ_a are the quantities measured in a four-point bend test. The quantities are indirectly measured, in the sense that they are calculated from equations (6.10) and (6.11) by substitution of direct measurements of the applied force F and the deflection y_d and additional measurements of the beam dimensions. The experiment is assumed to be carried out at a fixed deflection rate $dy_d/dt = ry_d$ which determines the reported strain rate $d\epsilon_a/dt$ according to (6.11). Stress-strain diagrams like the one shown in figure 7.14 provide an overview of the sequence of measured states of the specimen in the experiment.

The error in the force and length measurements underlying the reported stress and strain values is assumed to be 1%, as in the tension test examples.

7.3.3 Example 5: Twisting and specimen dimensions

Consider two individual property measurements 353 ± 4 MPa and 408 ± 4 MPa obtained in the experiments e_4 and e_5. Clearly, these measurements are in conflict according to the criterion of definition 36.

Now suppose that in both experiments provisions have been made to rule out all of

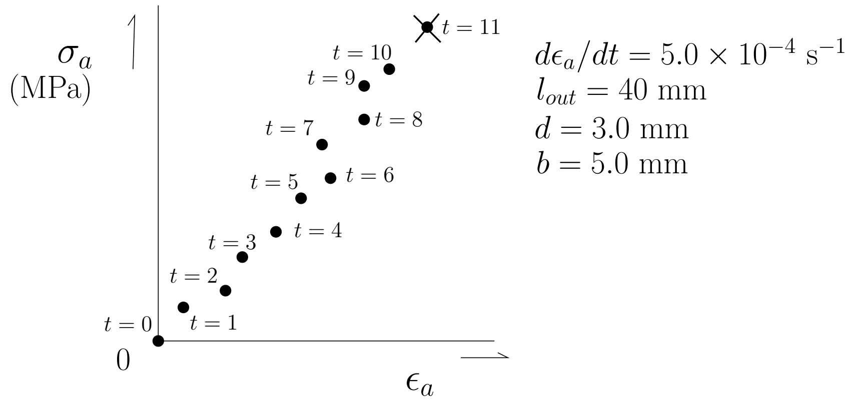


Figure 7.14: *Example of a fracture strength measurement obtained in experiment* e_1.

the disturbances accounted for in the model space except for twisting of the specimens. We know that the specimens were probably warped and that the bending fixtures did not have rotatable rollers to compensate for this deformation. This leads to the following sets of candidate models for the experimental systems:

$$CM = \{\langle QDE_{bttwsurf}, QS(init)_{bttwsurf}\rangle\},$$
$$\hat{CM} = \{\langle \hat{QDE}_{bttwsurf}, \hat{QS}(init)_{bttwsurf}\rangle\}.$$

Both specimens are most likely to have fractured through a surface crack located in a region with a stress concentration due to twisting.

At the given experimental conditions, QSIM finds a single candidate behavior for each candidate model (figure 7.12(c)). This behavior is assumed to be consistent with the measurements of σ_a and ϵ_a that were made during the execution of the tests.

In order to resolve the conflict, CEC* compares the candidate model and behavior of the first experimental system with the candidate model and behavior of the second experimental system. It finds the pairs of comparison $pc_0 = \langle t_0, \hat{t}_0\rangle$ and $pc_1 = \langle t_1, \hat{t}_1\rangle$. From the measured states of the specimens during the execution of the tests the initial RVs $\sigma_a \Uparrow_{pc_1}$ and $\epsilon_a \Uparrow_{pc_1}$ are derived. Assuming that the specimens have been fabricated and finished in the same way, and are deflected at the same speed, we arrive at the additional RVs $E \parallel_{pc_1}$, $\gamma \parallel_{pc_1}$, $a_0 \parallel_{pc_1}$, and $ry_d \parallel_{pc_1}$. The length of the outer span, width, and depth of the specimens in experiment e_4 are given by $40 \times 5 \times 3$ mm. The width and depth of the the specimens used in e_5 are specified as 5×3 mm, but the length of the outer span is unknown. This information about the experiment results in the following partial description of the initial comparative state:

$$CS(init) = \langle \sigma_a \Uparrow_{pc_1}, \epsilon_a \Uparrow_{pc_1}, E \parallel_{pc_1}, b \parallel_{pc_1}, d \parallel_{pc_1}, \gamma \parallel_{pc_1}, ry_d \parallel_{pc_1}, a_0 \parallel_{pc_1}, k_2 \parallel_{pc_1}\rangle.$$

The RV constraints needed by the CEC* algorithm are deduced from the QDEs and qualitative behaviors by means of the theorems in chapter 4. The bend test models in

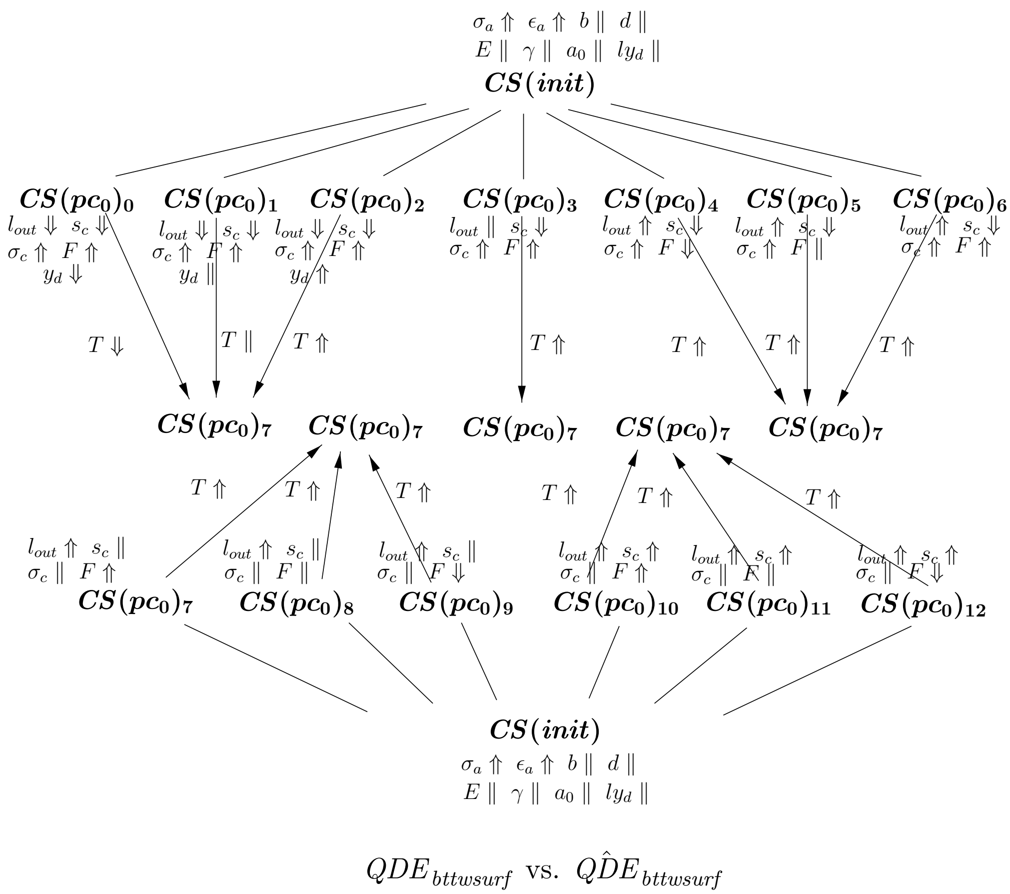


Figure 7.15: *The comparative envisionment arising from the comparison of two twisted specimens. The envisionment represents possible explanations of the conflict* $\sigma_a \Uparrow_{pc_1}$. *A few distinctive RVs are indicated at the comparative states.*

$SPACE_{bt}$ are first-order systems, so that the derivation of RV constraints between pc_0 and pc_1 is straightforward. Theorem 9 has been used for this purpose.

Starting with the behaviors of the models $QDE_{bttwsurf}$, $\hat{QDE}_{bttwsurf}$ and with $CS(init)$, CE finds the thirteen comparative behaviors displayed in figure 7.15. All of the behaviors are genuine. The conflict can be explained by five combinations of RVs for l_{out} and s_c. $l_{out} \Uparrow$ is in itself enough to account for $\sigma_a \Uparrow_{pc_1}$, but a difference in crack shape may add to ($s_c \Downarrow$) or detract from ($s_c \Uparrow$) the consequences of a greater outer span length. If l_{out} is larger, then the effect of twisting is reduced through the ratio b/l_{out} (equation (6.13)). As a consequence, at the same theoretical strength σ_{th}, a larger nominal stress σ_a needs to be applied to satisfy the crack propagation condition $\sigma_m = \sigma_{th}$. In the case of $l_{out} \parallel$ and $l_{out} \Downarrow$ the effect of twisting is the same or strengthened, so that the critical crack must necessarily be shorter and/or blunter in order to arrive at $\sigma_a \Uparrow_{pc_1}$.

Apart from influencing the size of the twisting effect, a difference in outer span length also influences the force that needs to be applied in order to effect a certain stress and strain at the tensile surface. From the measurements we know that $\sigma_a \Uparrow_{pc_1}$ and $\epsilon_a \Uparrow_{pc_1}$. Equations (6.10) and (6.11) determine the combinations of $RV(F, pc_1)$ and $RV(y_d, pc_1)$ that are consistent with $RV(\epsilon_a, pc_1)$, $RV(\sigma_a, pc_1)$, and $RV(l_{out}, pc_1)$. The majority of the ambiguities arise because $l_{out} \Uparrow$ tends to increase the bending moment and hence σ_a, while it is not sure whether this is enough to account for $\sigma_a \Uparrow_{pc_1}$. Notice that access to the direct measurements of F and y_d, from which σ_a and ϵ_a have been computed, allows one to resolve ambiguities at pc_1 and thus reduce the number of comparative behaviors to five.

The aggregated property measurements obtained in the experiments e_4 and e_5 are also in conflict, as indicated by the conflict matrix (figure 7.13): a higher average fracture strength is found in e_5. As a consequence, if two arbitrary measurements are taken from the populations sampled in e_4 and e_5, we are likely to find $\sigma_a \Uparrow_{pc_1}$. We know that the beams on which these measurements are performed are likely to fracture through a critical crack in a region with a higher stress concentration due to twisting ($QDE_{bttwsurf}$ and $\hat{QDE}_{bttwsurf}$). Now assume that $l_{out} \Uparrow$. The specimens in e_4 and e_5 have been manufactured and finished in the same way, so that except for $RV(s_c)$ the other experimental conditions are the same. The comparative behaviors in figure 7.15 show that for any relative value of s_c, the fracture strength could be higher when comparing two arbitrary measurements, so that it could be higher on average.

Interestingly, this argument leads to an inconsistency when it is assumed that $l_{out} \parallel$ or $l_{out} \Downarrow$. As can be verified in the comparative envisionment, under these conditions the measured fracture strength is only expected to be higher when the critical crack is likely to be shorter and/or blunter in experiment e_5 (i.e., $s_c \Downarrow$). Since the specimens in e_4 and e_5 have been manufactured and finished in the same way, there is no reason to suppose this to be the case. Consequently, the assumptions $l_{out} \parallel$ and $l_{out} \Downarrow$ do not seem warranted and can be dismissed. This reduces the number of comparative behaviors in figure 7.15 from thirteen to nine. Notice how the failure to explain a conflict between *aggregated* measurements thus helps in ruling out explanations of a conflict between *individual* measurements.

7.3.4 Example 6: Friction and load mislocation

As an example of error identification, I will consider the measurement 577 ± 6 MPa selected from the measurement set obtained in experiment e_6. Suppose that the measurement has been performed under circumstances in which twisting, wedging, and unequal loading can be excluded, so that we have the following set of candidate models:

$$\hat{CM} = \{\langle Q\hat{D}E_{btsurf}, \hat{Q}S(init)_{btsurf}\rangle, \langle Q\hat{D}E_{btmissurf}, \hat{Q}S(init)_{btmissurf}\rangle,$$
$$\langle Q\hat{D}E_{btfricsurf}, \hat{Q}S(init)_{btfricsurf}\rangle, \langle Q\hat{D}E_{btmisfricsurf}, \hat{Q}S(init)_{btmisfricsurf}\rangle\}.$$

The model of the ideal experimental system is given by QDE_{btsurf}, as explained in section 7.3.1. The candidate models of the actual experimental system state that the specimen fractured through a crack at the tensile surface, either at the nominal applied stress ($Q\hat{D}E_{btsurf}$), at a stress lower than the nominal applied stress due to friction ($Q\hat{D}E_{btfricsurf}$), at a stress higher than the nominal stress due to mislocation of the load points ($Q\hat{D}E_{btmissurf}$), or at an indeterminate stress due the opposing effects of friction and load mislocation ($Q\hat{D}E_{btmisfricsurf}$). In the latter two cases, the critical crack is likely to be found in the region where the maximum effect of the shift in load points is attained.

Assuming that the specimen in question has been fabricated and finished in the desired way, and it is deflected at the desired speed, we have the following initial comparative state information:

$$CS(init) = \langle E\,\|_{pc_0}, b\,\|_{pc_0}, d\,\|_{pc_0}, l_{out}\,\|_{pc_0}, s_c\,\|_{pc_0}, \gamma\,\|_{pc_0}, ry_d\,\|_{pc_0}, a_0\,\|_{pc_0}\rangle.$$

The error identification method requires six comparative analyses to be performed for $CS(init)$, where QDE_{btsurf} and $Q\hat{D}E_{btmisfricsurf}$ are to be compared for all three behaviors inferred by QSIM for the latter. The results of the analysis are shown in figure 7.16. The question whether the measurement is accurate or inaccurate cannot be straightforwardly answered. Some comparative behaviors indicate that $\sigma_a\,\|_{pc_1}$, whereas others suggest $\sigma_a \Uparrow_{pc_1}$ or $\sigma_a \Downarrow_{pc_1}$. Additional information about the experiment is necessary to distinguish between these cases, in particular information about the test machine used and the manner in which the specimen has been fixed.

Notice that even if we can make sure that both friction and load mislocation have occurred, the outcome is still indeterminate. The effect of friction may outweigh the effect of a shift of load points, the effect of a shift of load points may outweigh the effect of friction, or the effects may cancel out. This suggests an interesting strategy to avoid systematic errors in an experiment. As Galison [1987] points out, when certain contrary deviations cannot be completely eliminated but are sufficiently well controlled, they can be deliberately introduced to compensate for each other's effects. This seems not possible in the case of friction and load mislocation. For one reason, of the errors discussed in section 6.4.2, "frictional effects … are most difficult to predict quantitatively" (Hoagland, Marschall & Duckworth [1976], p. 192).

Just as the accuracy of an individual measurement performed in experiment e_6 cannot be ascertained, the accuracy of the aggregated measurement is indeterminate.

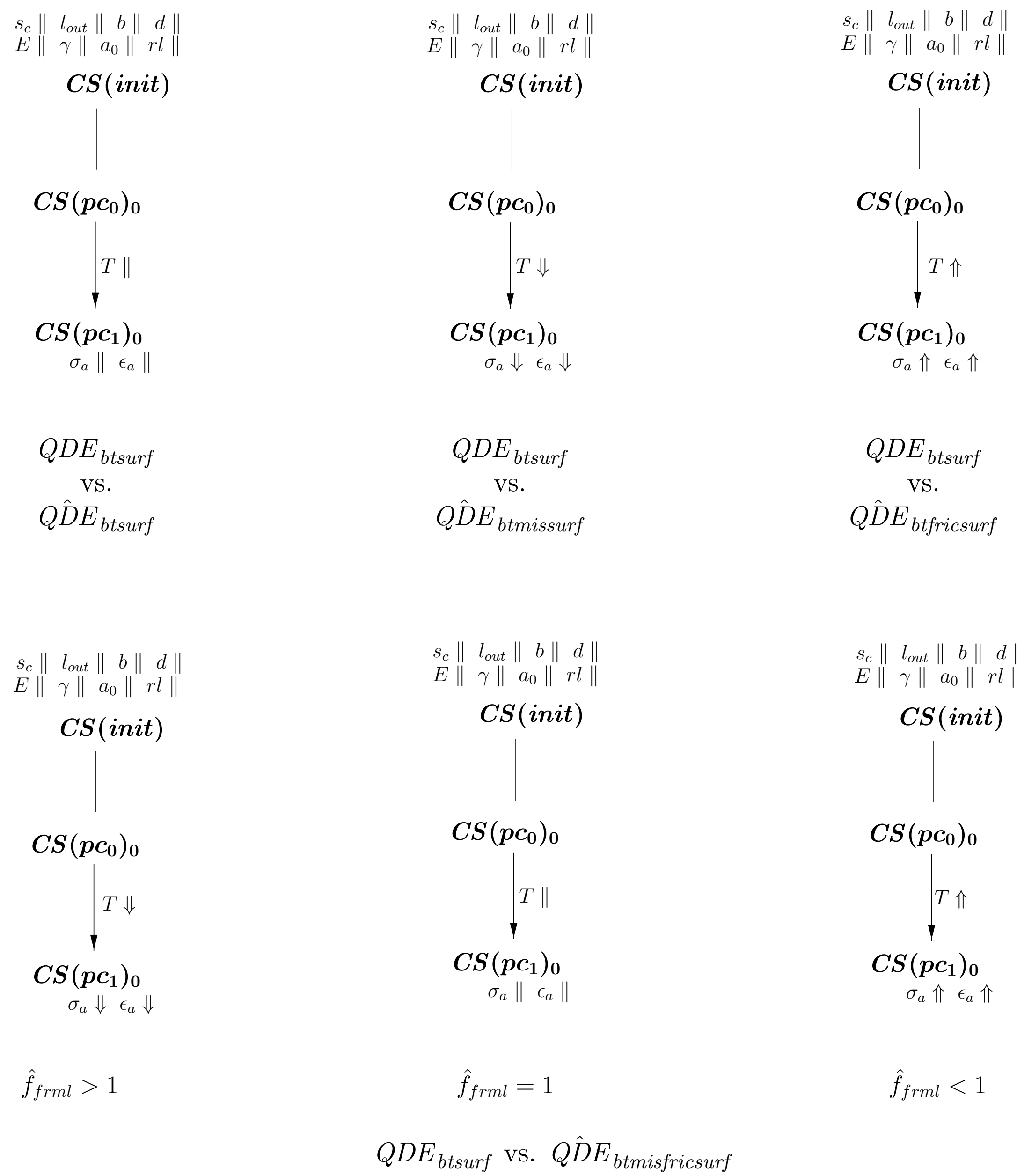


Figure 7.16: *The comparative envisionments arising from the comparison of the ideal four-point bend test system and an experimental system with possibly deviating loading conditions. The comparative behaviors predict* $RV(\sigma_a, pc_1)$*.*

7.3.5 Example 7: Unequal loading

In the previous examples, the two measurements to be compared all came from *different* measurement sets. Now suppose we compare the fracture strength measurements 380 ± 4 Mpa and 357 ± 4 Mpa from the *same* measurement set obtained in experiment e_3. Which circumstances could explain the conflict between these two measurements?

Experiment e_3 has been carried out in such a way that all perturbations of the ideal loading conditions can be excluded, except for an unequal distribution of the load over the two load points. The candidate models of the two experimental systems are thus found to be

$$CM = \{\langle QDE_{btsurf}, QS(init)_{btsurf}\rangle, \langle QDE_{btuneqsurf}, QS(init)_{btuneqsurf}\rangle\},$$
$$\hat{CM} = \{\langle Q\hat{D}E_{btsurf}, \hat{Q}S(init)_{btsurf}\rangle, \langle Q\hat{D}E_{btuneqsurf}, \hat{Q}S(init)_{btuneqsurf}\rangle\}.$$

That is, fracture of the specimens is likely to have initiated at the tensile surface, in the case of unequal loading at the height of the load point where the force $F + \Delta F$ is applied.

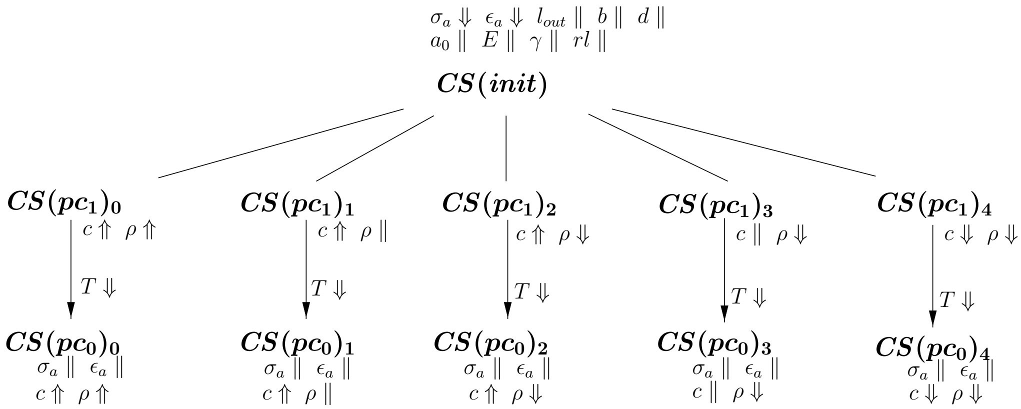
Since the measurements have been obtained in the same experiment, under the same conditions, we need to compare QDE_{btsurf} and $Q\hat{D}E_{btsurf}$, and $QDE_{btuneqsurf}$ and $Q\hat{D}E_{btuneqsurf}$. It makes no sense to compare QDE_{btsurf} and $Q\hat{D}E_{btuneqsurf}$, for when unequal loading affects the second measurement it must also have affected the first one. The other experimental conditions must also have been the same, except for the critical crack length c and tip radius ρ which cannot be controlled by the experimenters. This yields the following initial state information:

$$CS(init) = \langle \sigma_a \Downarrow_{pc_1}, \epsilon_a \Downarrow_{pc_1}, E \,\|_{pc_1}, b \,\|_{pc_1}, d \,\|_{pc_1}, d \,\|_{pc_1}, \gamma \,\|_{pc_1}, ry_d \,\|_{pc_1}, a_0 \,\|_{pc_1}, \Delta F \,\|_{pc_1}\rangle.$$

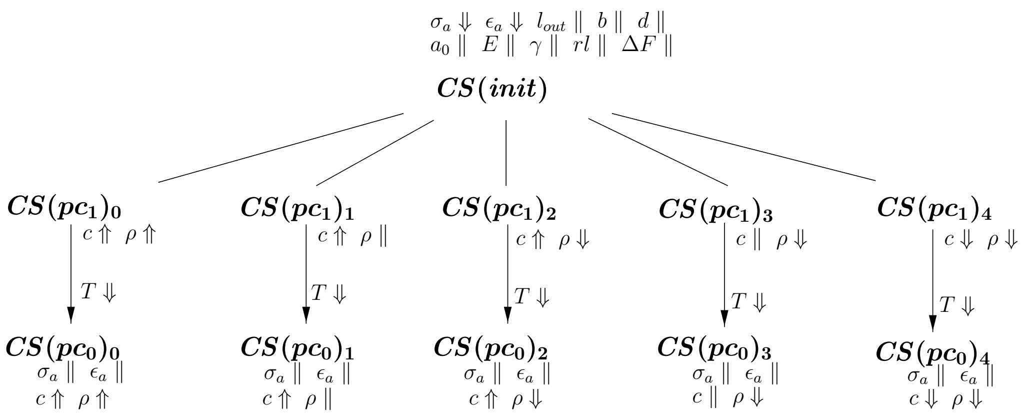
The results of the comparative analyses are summarized in figure 7.17. As might be expected, the conflict is most likely explained by differences in the shape of the critical crack (here elaborated as differences in c and ρ). The specimens belong to the same batch, so they have been fabricated and finished in the same way. In addition, they have been loaded in the same way. A difference in fracture strength must therefore be attributed to parameters varying from one specimen to the next, that is, to a difference in crack length and crack tip radius. This is in conformity with the literature on fracture tests (see, e.g., Creyke, Sainsbury & Morrell [1982], ch. 3).

7.4 Summary of results

The examples in the case-study illustrate how KIMA deals with conflict resolution and error identification problems involving measurements of the fracture strength of a brittle material. The case-study is intended to be an illustration of the methods, but in fulfilling this role it also says something about the capabilities of the MA methods. Various aspects of the results will be used in the next chapter to assess the strengths and limitations of



QDE_{btsurf} vs. $Q\hat{D}E_{btsurf}$



$QDE_{btuneqsurf}$ vs. $Q\hat{D}E_{btuneqsurf}$

Figure 7.17: *The comparative envisionments arising from the comparison of two specimens tested in experiment* e_5. *Both specimens may be either loaded in the ideal way or loaded in such a way that the load at the first load point is higher than the load at the second load point. The envisionments represent possible explanations of the conflict* $\sigma_a \Downarrow_{pc_1}$. *A few distinctive RVs are indicated at the comparative states.*

the methods. As a preliminary to this evaluation, I will briefly summarize what the case-study has shown.

First, the case-study reveals that the experimental systems investigated in a realistic domain can be conveniently modeled by means of qualitative differential equations. The QDE formalism corresponds with the way in which materials scientists conceptualize the physical systems they investigate. Second, it shows how successive measurements of the state of an experimental system can be used to rule out possible behaviors and models of the systems, and explanations of a conflict. Third, it shows that the MA methods can successfully reproduce a number of interesting empirically established relations, pointed out by domain scientists, between measured property values and features of the experiments. For instance, the methods indicate the strength decreasing effects of surface damage and inappropriate gripping, the strength increasing effects of frictional contact points and polishing, and the occurrence of a large spread in strength measurements in a single batch caused by uncontrollable differences in critical crack parameters.

What these findings imply for the practical use of the methods for model-based measurement analysis will be discussed in section 8.7.

Part IV

Evaluation

Chapter 8

Discussion and Related Work

This chapter will discuss the strengths and limitations of model-based measurement analysis in the context of related work in AI and neighboring fields. To that end, the methods described in part II, and their application in the case-study of part III, are reconsidered with seven main topics in mind: model-based and other forms of knowledge-based measurement analysis (section 8.1), the use of qualitative knowledge (section 8.2), the relationship of model-based measurement analysis with the model-based diagnosis of dynamical systems (section 8.3), the construction and revision of models of experimental systems (section 8.4 and 8.5), the generality of the methods (section 8.6), and the practical use of the methods (section 8.7).

8.1 Varieties of knowledge-based measurement analysis

Conflict resolution and error identification have been described as activities requiring background knowledge for their completion (chapter 1). In order to explain a conflict or find a systematic error in a measurement, one needs knowledge about the physical systems on which the measurements were performed and the experimental circumstances under which the systems were investigated. In the case-study this knowledge was supplied in the form of QDEs and initial qualitative values describing the structure of the strength test systems and the experimental conditions, respectively.

Taking a model-based approach towards measurement analysis has the advantage that answers explicitly relate to the physical principles that are believed to govern the systems being investigated. The models of the systems are based upon theories generally subscribed to in a field, like the theories of brittle fracture and strength tests in materials science, which favors the understanding, acceptance, and use of the answers by scientists working with the system. Moreover, a model-based approach facilitates further investigation, for instance by suggesting additional measurements or new experiments to exclude certain explanations in the case of conflict resolution.

As a disadvantage of model-based MA one should mention the substantial efforts required to obtain the models employed by the methods. Although these tasks can in turn be alleviated by computer support, as section 8.4 will argue, in some situations users may not be prepared to make the initial investments. Also, for their purposes, they may be satisfied with more approximate explanations of conflicts or predictions of systematic errors. I will therefore briefly discuss four other approaches towards knowledge-based MA that are less knowledge-intensive than model-based MA.

First of all, one can think of the identification of *outliers* in a measurement base, either by robust statistical criteria or by visual inspection (Rousseeuw & Leroy [1987]; Tukey [1977]; Tufte [1983]). An outlying measurement deviates from the bulk of the measurements and is consequently involved in many conflicts in the conflict matrix. These conflicts could be resolved by simply assuming that the outlying measurement has been obtained in an experiment substantially differing from those in which the other measurements have been obtained, and that these experimental differences are to blame for the conflict. Moreover, one could declare the outlying measurement to be inaccurate, because the relative proximity of the other measurements makes it unlikely that it represents the true value of the property.

Of course, this method cannot give any guarantees: the outlying measurement may well be accurate and the other measurements inaccurate. Still heuristic, but of a more knowledge-intensive kind, is a method based on the use of experiential rules like 'If the experiment giving rise to one of the measurements involved in the conflict is incompletely specified, then blame a difference in the unspecified part of the experiment for the conflict' or 'If a measurement was obtained by a primitive method, then doubt its accuracy'. Although the rules may be founded on a (theoretical) understanding of the experiment, this understanding is not made explicit but instead compiled into the antecedents and consequents of the rules. The use of such experiential heuristics can be recognized in the assessment of melting temperature measurements in Hlaváč [1982], where it is for instance argued that "more weight must be given to more recent values obtained under better defined experimental conditions" (p. 683).

A more systematic and domain-specific way of taking into account information about the experiments is the use of check lists with features of experiments that are deemed to have a modifying influence on the measured value of a property. Conflicts between measurements can be resolved by running down the check list and attributing the observed difference in the property value to differences in mentioned aspects of the design and execution of experiments. Also, by regarding the items on the list as methodological criteria to be satisfied, the check list can function as an informal specification of the ideal experiment and hence be useful in error identification as well. Check lists with methodological criteria have been developed in the field of meta-analysis (Hunter & Schmidt [1990]; Bouter [1994]), for the purpose of weighing measurements from different experimental studies that need to be integrated into a single property value.

A check list identifies modifying influences on the measured value of a property, but does not say in which direction these influences work. In order to add this information, one could extend the check list with a set of rules. In the fracture strength example, such rules might state that eccentric loading in a tension test leads to a lower measured strength, or that friction at the load points in a four-point bend test leads to a higher measured strength. However, no information is provided about the mechanism through which the deviation from the ideal experiment effects the measured value. Also, a large number of rules are needed to cover all possible combinations of deviations, which raises questions about the consistency and completeness of the set of rules, a problem familiar from experiences with rule-based expert systems.

Going from outlier identification techniques to rules specifying the effect of modifying influences, an increasing amount of domain knowledge is brought to bear on the resolution of conflicts and the identification of systematic errors. However, even the use of rules of the latter kind is considerably less knowledge-intensive than model-based measurement analysis, and can only give a hint at an explanation of why and how the error or conflict occurred.

8.2 Use of qualitative knowledge

The domain knowledge used by the methods for conflict resolution and error identification is of a qualitative nature. In chapter 1 two general arguments for using qualitative models of experimental systems were given, both of which are confirmed by experiences with the case-study. Often it is difficult to give an adequate quantitative model, or quantitative initial conditions, because this information is simply not available. The most detailed description of the tension test that I have been able to find specifies a stress concentration factor $g(d_h/d, r_t/d)$ due to a sudden change in cross-section (equation (6.4.1)), without giving any details about the function g beyond it being a member of the class M^{+-}. This qualitative knowledge is enough, however, to understand how specimen geometry influences the occurrence of large stress concentrations at the shoulders and also how these influences can be avoided in new and improved experiments.

The use of qualitative knowledge implies that causes of a conflict and systematic errors are specified in terms of relative values $\Downarrow$, $\|$, or $\Uparrow$ for certain quantities. A more precise statement on the difference in value cannot be made. Also, as a further implication of the use of qualitative knowledge, a large number of explanations of a conflict and predictions of systematic errors may result from the application of the MA methods, due to ambiguities in the qualitative and relative values of quantities.

In the case-study, ambiguities did not present a real problem. Although about twenty quantities are involved, the experimental systems can be described by first-order differential equations which (except in the case of $QDE_{btmisfric}$) give a single qualitative behavior after simulation. This limits the number of comparative analyses that need to be performed in conflict resolution and error identification, and hence the number of explanations of conflicts and predictions of systematic errors (even more so because the comparative envisionments are usually small for the given $CS(init)$s).

It should be added immediately, however, that the results of the case-study are far from trivial. As the examples in chapter 7 show, the MA questions require the methods to deal with combinations of several structural differences and differences in the initial conditions. Manually determining the answers is time-consuming and error-prone. Also, even though the examples seem fairly straightforward, they may give rise to surprising answers. For example, I was initially puzzled by the ambiguities in figure 7.7, arising from the comparison of QDE_{ttecc} and $QDE_{tteccsurf}$, until I realized that the difference in eccentricity $RV(e)$ is unknown and leads to alternative explanations of the conflict.

One can think of situations in which the number of explanations and predictions will not be small (section 8.7). In chapters 3 and 4 a number of extensions of the QR techniques were mentioned that can help in eliminating behaviors from the behavior tree or comparative envisionment, such as requests for new measurements, order-of-magnitude reasoning, and semi-quantitative reasoning.

Especially the use of semi-quantitative techniques for simulation and comparative analysis in the MA methods would be interesting. Not only would it allow one to reduce the number of alternative explanations and predictions, by exploiting knowledge about numerical bounds on landmark values and numerical envelopes around monotonic functions, but it would also be a means to improve upon another limitation of qualitative knowledge: its lack of precision. A prediction of a systematic error $q \Uparrow$ means that $(\hat{q} - q) \in]0, \infty]$. Semi-quantitative comparative analysis could lead to a refinement of this interval, and thereby enable the users to assess whether for their purposes the disturbance of the ideal experimental system is serious enough to warrant further attention.

The decomposition of model-based conflict resolution and error identification into a sequence of qualitative simulation and comparative analysis tasks has the advantage that the extensions mentioned above can be readily incorporated into the MA methods. What is more, if guarantees on the correctness of the techniques are not affected by the extensions, guarantees on the correctness of the methods will remain valid as well. Semi-quantitative simulation with Q3 is sound but incomplete (Berleant & Kuipers [1997]), so that its inclusion in the conflict resolution algorithm of chapter 5 will not change the conclusions expressed by the theorems 17 and 18: all genuine explanations of the conflict will be generated, but occasionally spurious ones as well.

8.3 Conflict resolution, error identification, and model-based diagnosis

Although a model-based approach towards conflict resolution and error identification seems unexplored thus far, there are interesting parallels with *model-based diagnosis (MBD)* of technical devices (Hamscher, Console & de Kleer [1992]). An interpretation of model-based measurement analysis within an MBD framework helps to clarify both the problem and the methods. Generally speaking, model-based diagnosis is concerned with finding explanations for deviations of the behavior of an observed system from that of a known reference system. The MBD approach proceeds by hypothezising fault models and fault conditions for the observed system, making predictions from these models and conditions, and matching the predictions with the measured states of the system. If the predictions derived from a fault hypothesis conflict with the measured states, then the fault hypothesis can be ruled out.

There are many ways to realize this general approach when diagnosing dynamical systems, using a variety of modeling formalisms and reasoning techniques, such as semi-quantitative models and semi-quantitative simulation (Dvorak & Kuipers [1991]), qualitative models in combination with qualitative simulation and comparative analysis (Neitzke [1997]), or qualitative models and state consistency checking without simulation (Malik & Struss [1996]). Although both the general idea and the techniques are similar, there are a number of differences between model-based diagnosis and model-based measurement analysis.

In the problems usually addressed by MBD we compare an observed system with a reference system whose structure and behavior are known. Measurements performed on the observed system help in determining why its behavior deviates from that of the reference system. In the case of model-based conflict resolution, however, there is no such reference system. We compare *two* experimental systems of which we generally do not know the precise structure and behavior, and measurements performed on *both* systems are used to exclude potential explanations of the conflict.

Model-based error identification also differs from MBD. In the first place, the reference system used for error analysis is a hypothetical system. Since the ideal experiment has not been carried out, the true value of the property is not known and a discrepancy between the true value and measured value cannot be directly detected. At best, a broad interval in which the true value is believed to lie can be derived by means of accepted theories and background knowledge. The unavailability of the true value leads to a second difference, a difference in the kind of solution that is expected. Whereas model-based diagnosis attempts to *explain* observed deviations from the behavior of the reference system, model-based error identification is directed at the *prediction* of deviations that would be observed if the ideal experiment were carried out.

The relationship between model-based measurement analysis and model-based diagnosis calls attention to a number of issues that have been raised in the context of the latter, but that are equally relevant for the former. For instance, the remark in Davis & Hamscher [1988] that "all model-based reasoning is only as good as the model" points at an aspect of model-based measurement analysis that has been left implicit thus far. Although the theorems in chapter 5 guarantee that *all* possible explanations of a conflict and *all* possible systematic errors will be found, they do so relative to the available models of the experimental systems. If the models are too coarse, for instance, the conflict resolution method will not be able to generate a subtle explanation of a conflict. This raises the questions how an adequate model is obtained from a description of the experiment, and how it is revised when it turns out to be inadequate. The issues of model building and model revision will be discussed in the next two sections.

8.4 Constructing models of experimental systems

In chapter 5 the notion of model space was introduced. A model space contains the possible models by which one can describe an experimental system investigated in a particular type of experiment. The candidate models of an experimental system are manually selected from the model space on the basis of a description of the experiment.

The construction of a model space and the selection of candidate models from the model space have been taken for granted thus far, because the development of methods to reason with the models has been at the center of attention. However, since model building touches upon fundamental issues in the analysis of conflicts and systematic errors (chapter 2), it warrants further discussion. This will be done by comparing the simple approach adopted here with work on computer-supported modeling of physical systems (see Xia & Smith [1996] and Schut & Bredeweg [1996] for overviews).

Model building is often viewed as *model composition*, that is, the composition of a model of a physical system from a knowledge base with *model fragments*, given a scenario description (e.g., Forbus [1984]; Low & Iwasaki [1992]; Farquhar [1994]). The model fragments describe distinguishable aspects of the system, such as the physical parts of which it is composed or the processes occurring in it. Each model fragment is valid under certain conditions and contributes one or several equations, or parts of equations, to the QDE. For instance, in the case study one can define a model fragment consisting of the frictional factor $f_{fric} = 1 - 2\mu(d/l_{out})$, which is valid in a four-point bend test when the rollers are not free to rotate. Model composition proceeds by determining the valid model fragments from a description of the experiment and then combining and integrating the model fragments into a QDE.

For a given description of the experiment, the set of valid model fragments may be quite large and certain model fragments mutually contradictory. In deciding which model fragments to omit, one can introduce explicit *modeling assumptions* about the domain and label the model fragments in the knowledge base with the modeling assumptions under which they are applicable (e.g., Addanki, Cremonini & Penberthy [1989]; Falkenhainer & Forbus [1991]; Nayak [1995]; Iwasaki & Levy [1994]). Which modeling assumptions are made in a particular situation, and thus which model fragments are considered to be applicable, depends on the question to be answered. A number of modeling assumptions have been made in the case-study, and underlie the models in the figure 7.1 and 7.10. For instance, I have focussed on brittle fracture in a homogeneous and isotropic material sample which constitutes the experimental system. Even more basically, the models are assumed to be used for conflict resolution and error identification, which motivates why deviations from the ideal experiment have been taken into account (instead of neglected).

In order to build a knowledge base with model fragments and formulate descriptions of experiments, at least to do so in a structured and principled way, one would like to have a formal ontology of scientific experiments. An *ontology* is a structured system of concepts and relations between concepts which can be used to formally specify knowledge about a particular domain (van der Vet & Mars [1993]). An ontology of scientific experiments would provide concepts and relations to express knowledge about the physical systems being investigated, and the materials, methods, and instruments used to create and sustain these systems. General ontologies of physical systems have been developed in recent years, like the CML ontology (Falkenhainer et al. [1994]), YMIR ontology (Alberts [1993]), and PhysSys ontology (Borst, Akkermans & Top [1997]). I do not know, however, of any ontologies for scientific experiments. Some inspiration for the development of such an ontology might be found in philosophy of science, especially with authors who try to give a systematic description of the elements making up experimental workplaces (e.g., Hacking [1992]).

When viewed from the perspective of compositional modeling, the approach adopted in chapter 5 is a shortcut which abstracts from a complex model-building process. The model spaces of chapter 5 could be induced by the composition of all possible models from a knowledge base of model fragments given certain modeling assumptions. Also, the candidate models of an experimental system could be composed from a knowledge base with model fragments given certain modeling assumptions and a description of the experiment. Compositional modeling thus suggests a direction in which model-based measurement analysis could be extended (section 9.2).

8.5 Revising models of experimental systems

Two fundamental aspects of the role of models in measurement analysis, mentioned in section 2.1.6, can be rephrased within a model composition framework. In the first place, which explanations of a conflict or predictions of a systematic error are obtained depends on the modeling assumptions chosen, for instance the assumption to omit the test apparatus from consideration. Different modeling assumptions give rise to different models, and thus to different explanations and predictions. In the second place, when no acceptable explanations of a conflict are found, or systematic errors suspected on other grounds are not predicted, one may start to doubt the adequacy of the models of the experimental systems. Within the compositional modeling framework this means that the appropriateness of the modeling assumptions or the validity of the model fragments are put up for discussion. This gives rise to the problem of *model revision*.

Model revision changes the role of the models in measurement analysis. Whereas they are resources to explain conflicts or identify systematic errors in the methods discussed in chapter 5, in model revision they have become objects of investigation themselves. The two handles for model revision in the compositional approach are the modeling assumptions and the model fragments. In example 4 of the case-study, in which a conflict could not be explained because the single qualitative behavior of one of the experimental systems was in conflict with the sequence of measured states, the suggested solution was to drop the assumption that we are dealing with brittle fracture. Alternatively, one can revise the model by modifying the model fragments giving rise to it. One could for instance change the equations that a model fragment contributes to the model or change the conditions under which the model fragment is applicable.

Changes in model fragments amount to revisions of the underlying theory or theories of the domain. Although in domains with a firmly established theoretical framework this will not be the option considered first, it does occasionally present a way to resolve otherwise unresolvable conflicts or identify systematic errors thus far not taken into account. In the case-study, a comparison of the discussion of bend tests in Marschall & Rudnick [1974] and Baratta [1984] learns that new disturbing processes have been identified, and knowledge about the effects of already known disturbing processes refined, in the ten years intervening between the articles. Theory development is obviously an open-ended process.

Model revision is a difficult problem which has been investigated in different guises in a variety of fields within and outside AI. Belief revision (e.g., Gärdenfors [1988]), computer-supported modeling (e.g., Pos, Akkermans & Top [1998]), validation and verification of knowledge-based systems (e.g., Gupta [1991]), computer-supported discovery (e.g., Shrager & Langley [1991]), and (computational) philosophy of science (e.g., Thagard [1988]; Darden [1991]) are just a few examples of work addressing issues relevant for model revision in the context of model-based measurement analysis. Nevertheless, an approach to model revision which is at the same time general, formal, and practical seems to be an open research problem.

8.6 Generality of the measurement analysis methods

What about the generality of the methods developed in this thesis? It is readily seen that neither the techniques of chapters 3 and 4, nor the methods of chapter 5, make any assumptions about the particular domain in which measurement analysis problems will be studied. They only require the models of the experimental systems investigated in these domains to be expressed as (qualitative) differential equations. So, in principle at least, the methods for conflict resolution and error identification are directly applicable to any problem formalized in terms of differential equations. This is a pleasant consequence, since differential equations are a common modeling formalism in science and engineering. They are used in domains as diverse as biochemistry, population dynamics, and fluid mechanics.

The applicability of the methods could be pushed even further when it is realized that the basic ideas underlying model-based conflict resolution and error identification do not strictly depend on the particular modeling formalism and reasoning techniques used in this thesis. Figure 5.2 and figure 5.3 suggest that we could choose another modeling formalism, e.g. causal models or signed directed graphs, with associated techniques for inferring behaviors of the experimental systems and comparing the behaviors. Of course, the algorithms filling in the details of the methods would then need to be redefined, and the guarantees of the methods redetermined, but conflict resolution and error identification would still proceed by a simulation phase and a comparison phase, in which measured states of the experimental systems limit the search space. In this way, the MA methods might be generalizable to domains where modeling formalisms other than differential equations are more commonly used, for instance the medical and social sciences.

8.7 Upscaling and practical use

The MA methods proved successful in the case-study in the sense that KIMA has been able to reproduce a number of interesting observed phenomena in a realistic domain.

The results indicate that the methods form a fruitful basis for approaching measurement analysis problems. However, in order to be applicable in a real-life context, and transform the KIMA system into a practical aid for working scientists, a number of *upscaling* issues need to be addressed.

The QR techniques underlying the methods did not run into trouble in the case-study: no spurious (qualitative or comparative) behaviors were generated and the number of ambiguities remained within acceptable bounds (section 8.2). This will not usually be the case when experimental systems are described by higher-order differential equations, the model spaces are larger, and the descriptions of the experiments are more incomplete. In order to function well in these situations, the techniques need to be improved and extended along the lines sketched in earlier chapters. In particular, semi-quantitative simulation and comparative analysis in combination with the possibility to integrate further measurements are desirable extensions.

Upscaling of the methods to life-size MA problems also requires one to make substantial efforts in obtaining models of the physical systems. The modeling approaches discussed in sections 8.4 and 8.5 could be helpful in this, but they entail their own upscaling problems. For instance, one will need to construct large knowledge bases with model fragments, annotated with appropriate modeling assumptions. The collaborative construction of such knowledge bases has received attention recently (e.g., Iwasaki et al. [1997]), in the context of a growing interest in knowledge sharing and reuse (Neches et al. [1991]).

Somewhat broader conceived, upscaling also refers to the increasing number of people who are becoming involved in the development of an MA tool and who start to use it, the increasing number of research questions at which the system is directed, and its increasing entrenchment in a network of tools in a scientific practice. These broader upscaling issues will be further explored in the next chapter.

Chapter 9

Conclusions

9.1 Achievements

Model-based methods have been developed for the resolution of conflicts between measurements contained in large scientific measurement bases and for the identification of systematic errors in these measurements. The methods have been implemented in the measurement analysis system KIMA.

In order to put these achievements into perspective, I will briefly reconsider the functionality requirement, the application requirement, and the generality requirement that were mentioned in section 1.2. In accordance with the functionality requirement, formally specified methods with proven correctness and efficiency guarantees have been developed, which produce comparative envisionments that can be used to provide explanations. The implementation of the methods has been successfully used in a case-study in a realistic domain (conform the application requirement). The qualitative simulation and comparative analysis techniques, which lie at the heart of the MA methods, make no assumptions about the particular domain under study. This ensures that the methods are directly applicable in any domain in which the experimental systems can be modeled by (qualitative) differential equations (thus satisfying the generality requirement).

The contributions of the thesis are twofold. In the first place, the methods for model-based conflict resolution and error identification supplement conventional statistical analyses of conflict and errors. They allow one to exploit knowledge about the domain, in particular knowledge about the controlled physical systems investigated in scientific experiments, to give possible explanations of conflicts and predictions of systematic errors. In the second place, a general and mathematically well-founded technique for comparative analysis has been developed which improves upon existing approaches in that it is able to compare structurally different dynamical systems. The CA technique CEC^* has potential applications in the fields of diagnosis, design, discovery, and model-building.

9.2 Further work

Just as the contributions of the thesis have a technical and application dimension, opportunities for further research can be thus classified. I will identify a few interesting directions for continuing the work presented here.

As indicated on several occasions already one could generalize CEC* to a *semi-quantitative technique for comparative analysis*. In combination with existing algorithms for semi-quantitative simulation, this would permit the development of semi-quantitative methods for conflict resolution and error identification. The possible explanations of conflicts and predictions of systematic errors could then be refined and their number reduced (section 8.2). Initial explorations suggest that the proposed generalization of CEC* is a non-trivial but quite promising direction to pursue.[1] The basic idea is to use the theorems 6 to 12, in combination with numerical bounds on landmark values and monotonic functions, to obtain constraints on the numerical interval containing the difference $\hat{\boldsymbol{q}} - \boldsymbol{q}$ of the shared variables at the pairs of comparison.

The availability of a method for semi-quantitative comparative analysis would enable the extension of the MA methods considered in this thesis with a method for *error correction*. Instead of deriving possible systematic errors $\Uparrow$ or $\Downarrow$ for a measured property value, semi-quantitative error identification would result in numerical bounds on the systematic errors. At present error correction is restricted to pronouncing that the true value will be lower or higher than the measured value. The numerical bounds obtained in semi-quantitative error identification would permit one to estimate by how much the true value is expected to be lower or higher.

Another interesting direction for further research would be to connect the methods for conflict resolution and error identification with a *model-building application*. In order to achieve this, an ontology of scientific experiments needs to be developed which supplements existing ontologies of physical systems. The informal conceptualization of scientific experiments in chapter 2, which centers around the idea of an (ideal or disturbed) experimental system, seems to provide a clear starting-point for this. Although suitable model composition algorithms can be found in the literature on computer-supported modeling, substantial efforts may be required to develop a knowledge base with model fragments structured by appropriate modeling assumptions.

Augmenting the KIMA system with the capability to reason about semi-quantitative information and a facility to construct models from experimental scenarios forms an important requirement for *upscaling the prototype* to a real-life application. Some care should be exercised in selecting an appropriate domain for this application. Obviously, the scientists working in the domain would have to deal with large amounts of measurements, but in addition assessing the quality of these measurement should present a real problem. The experimental methods should be sufficiently complicated or non-standardized to give rise to systematic deviations between property measurements, and these deviations should be in need of explanation. Also, the domain would have to be computerized to some degree, in the sense that measurement bases containing infor-

[1] I am grateful to Ivayla Vatcheva, who did most of the work referred to here.

mation about experiments are available and that scientists are familiar with the use of advanced computer tools in their research practices. Above all, cooperation should be sought with the intended users of a measurement analysis application.

9.3 Some speculations

The actual use of KIMA in scientific practices can be seen as an exponent of the movement towards *computer-supported discovery environments* currently taking place in a number of scientific disciplines, molecular biology being one of the prime examples (de Jong & Rip [1997]; de Jong & Rip [1995]). By the term discovery environments I refer to the integrated systems of material and intellectual tools available to scientists in their search practices. Computer-supported discovery environments arise from conventional discovery environments by a process in which existing discovery tasks are increasingly taken over and coordinated by computer tools. Moreover, this gives rise to the emergence of new scientific problems predicated upon the capabilities of the computer tools.

A system for model-based measurement analysis might become part of a computer-supported discovery environment composed of measurement bases and knowledge bases with model fragments, and a variety of systems interacting with the measurement analysis tool and the measurement bases and knowledge bases. An intriguing conflict between two measurements of a property might lead scientists to construct models of the experimental systems by means of a model-building application, and then use these models to analyze the conflicts. The possible explanations produced by the measurement analysis tool could be fed into a computer program for generating designs of new experiments that attempt to discriminate between the alternatives. When after performing the proposed experiments in the laboratory it turns out that none of the explanations can be maintained, and hence the models should be considered inadequate, scientists may decide to engage in computer-supported model revision. It is only in the context of such a discovery environment that I expect systems for computer-supported measurement analysis to make worthwhile contributions to scientific inquiry (de Jong, Mars & van der Vet [1998]).

The above impression already suggests that the movement towards computer-supported discovery environments should not be viewed as a straightforward computerization of existing systems of tools. Perhaps the most interesting, and in the long range the most consequential aspect of the emergence of computer-supported discovery environments are the new ways of doing research co-evolving with the new computer tools. The availability of large-scale databases and knowledge bases, together with computer applications for analyzing, interpreting, and extending their contents, will have an influence on the kind of questions that are considered interesting and feasible, the kind of methods considered reliable and fruitful, and the kind of answers expected and considered acceptable. The development towards computer-supported discovery environments may lead to the establishment of *computer regimes* in which scientists adopt a way of working driven by the opportunities offered by the use of computer tools (de Jong & Rip [1997]; de Jong & Rip [1995]).

This fascinating development, which might lead to a *computer revolution in science*, has been one of my main motivations to work on the kind of applications considered in this thesis. In the next decade I expect it to provide computer science with a host of interesting problems of both practical and theoretical importance.

Part V

Appendices

Appendix A

Miscellaneous Proofs

A.1 Solution of linear comparison systems

In section 4.3.3 the solution of the state equation of the linear comparison system

$$\frac{d}{dt}(\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t)) = \boldsymbol{A}(t)(\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t)) + \boldsymbol{B}(t)(\hat{\boldsymbol{u}}(t) - \boldsymbol{u}(t)) + \boldsymbol{E}(t)(\hat{\boldsymbol{a}}^{\boldsymbol{0}}(t) - \boldsymbol{a}^{\boldsymbol{0}}(t))$$

over an interval $[t_a, t]$ was stated to be

$$\begin{aligned}\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t) &= \boldsymbol{\Phi}(t, t_a)(\hat{\boldsymbol{x}}(t_a) - \boldsymbol{x}(t_a)) \\ &\quad + \int_{t_a}^{t} \boldsymbol{\Phi}(t, \tau)\left[\boldsymbol{B}(\tau)(\hat{\boldsymbol{u}}(\tau) - \boldsymbol{u}(\tau)) + \boldsymbol{E}(\tau)(\hat{\boldsymbol{a}}^{\boldsymbol{0}}(\tau) - \boldsymbol{a}^{\boldsymbol{0}}(\tau))\right] d\tau\end{aligned}$$

In order to verify the validity of this statement, notice that the state equation can be rewritten as

$$\frac{d}{dt}(\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t)) = \boldsymbol{A}(t)(\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t)) + \boldsymbol{B}^{\boldsymbol{*}}(t)(\hat{\boldsymbol{u}}^{\boldsymbol{*}}(t) - \boldsymbol{u}^{\boldsymbol{*}}(t)) \tag{A.1}$$

where

$$\boldsymbol{B}^{\boldsymbol{*}} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{E} \end{bmatrix}, \qquad \boldsymbol{u}^{\boldsymbol{*}} = \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{a}^{\boldsymbol{0}} \end{bmatrix} \tag{A.2}$$

The solution of (A.1) over $[t_a, t]$ follows by theorem 4.3 in Chen [1970]:

$$\begin{aligned}\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t) &= \boldsymbol{\Phi}(t, t_a)(\hat{\boldsymbol{x}}(t_a) - \boldsymbol{x}(t_a)) \\ &\quad + \int_{t_a}^{t} \boldsymbol{\Phi}(t, \tau)\boldsymbol{B}^{\boldsymbol{*}}(\tau)(\hat{\boldsymbol{u}}^{\boldsymbol{*}}(\tau) - \boldsymbol{u}^{\boldsymbol{*}}(\tau))d\tau\end{aligned} \tag{A.3}$$

Substituting (A.2) into (A.3) gives the expression that was sought.

A.2 Time-invariant comparison systems for mass-spring systems

Given the state equation of the comparison system for two frictionless mass-spring systems

$$\frac{d}{dt}\begin{bmatrix}\hat{x}(t)-x(t)\\ \hat{v}(t)-v(t)\end{bmatrix}=\begin{bmatrix}\hat{v}(t)-v(t)\\ -\frac{\hat{k}}{\hat{m}}\hat{x}(t)-(-\frac{k}{m}x(t))\end{bmatrix}. \tag{A.4}$$

The spring constant k and the mass m of the block are constants.

Remember that a linear comparison system can be obtained by applying the generalized mean value theorem of the differential calculus (theorem 8):

$$\frac{d}{dt}\begin{bmatrix}\hat{x}(t)-x(t)\\ \hat{v}(t)-v(t)\end{bmatrix}=\begin{bmatrix}0 & 1\\ -\frac{\bar{k}(t)}{\bar{m}(t)} & 0\end{bmatrix}\begin{bmatrix}\hat{x}(t)-x(t)\\ \hat{v}(t)-v(t)\end{bmatrix}+\begin{bmatrix}0 & 0\\ -\frac{\bar{x}(t)}{\bar{m}(t)} & \frac{\bar{k}(t)\bar{x}(t)}{\bar{m}^2(t)}\end{bmatrix}\begin{bmatrix}\hat{k}-k\\ \hat{m}-m\end{bmatrix}. \tag{A.5}$$

We will proof that the matrix $\boldsymbol{A}(t)$ appearing in the state equation (A.5) is time-invariant, that is, $\boldsymbol{A}(t)=\boldsymbol{A}$.

Bearing in mind that $a=dv/dt$, equation (A.5) implies

$$\hat{a}(t)-a(t)=-\frac{\bar{k}(t)}{\bar{m}(t)}(\hat{x}(t)-x(t))-\frac{\bar{x}(t)}{\bar{m}(t)}(\hat{k}-k)+\frac{\bar{k}(t)\bar{x}(t)}{\bar{m}^2(t)}(\hat{m}-m). \tag{A.6}$$

Of course, $\bar{x}(t)$ lies between $\hat{x}(t)$ and $x(t)$, $\bar{m}(t)$ between $\hat{m}$ and m, and $\bar{k}(t)$ between $\hat{k}$ and k. The generalized mean value theorem further states that the point $(\bar{k}(t),\bar{m}(t),\bar{x}(t))$ lies on the line segment connecting $(k,m,x(t))$ and $(\hat{k},\hat{m},\hat{x}(t))$ (Courant [1959], vol. II, ch. 2). From this it is easy to show that in addition $\bar{k}(t)/\bar{m}(t)$ lies between k/m and $\hat{k}/\hat{m}$.

Although there are more general ways to proceed, linearity of the systems suggests that the point $(\bar{k}(t),\bar{m}(t),\bar{x}(t))$ lies exactly midway on this line segment, so that we have:

$$\begin{aligned}\bar{k}(t)&=\frac{1}{2}(k+\hat{k})\\ \bar{m}(t)&=\frac{1}{2}(m+\hat{m})\\ \bar{x}(t)&=\frac{1}{2}(x(t)+\hat{x}(t))\\ \frac{\bar{k}(t)}{\bar{m}(t)}&=\frac{1}{2}(\frac{k}{m}+\frac{\hat{k}}{\hat{m}})\end{aligned}$$

By substitution of these equations into (A.6) the suggestion can be verified. Indeed, we obtain

$$\hat{a}(t) - a(t) = -\frac{\hat{k}}{\hat{m}}\hat{x}(t) - (-\frac{k}{m}x(t)),$$

that is, the original expression for the difference in acceleration at t.

We have thus proven by construction that $\bar{k}(t)/\bar{m}(t)$ is time-invariant, so that $\boldsymbol{A}(t)$ is time-invariant.

A.3 Periodicity of mass-spring comparison systems

In section 4.5 two ideal mass-spring systems starting from rest were compared (figure 3.8). The linear comparison system derived from the QDEs has the state matrix:

$$\boldsymbol{A}(t) = \begin{bmatrix} 0 & 1 \\ -\frac{\bar{k}}{\bar{m}} & 0 \end{bmatrix}, \tag{A.7}$$

with $\bar{k}, \bar{m} > 0$. This state matrix gives rise to a periodic transition matrix (4.24) with period $\bar{P} = 2\pi\sqrt{\frac{\bar{m}}{\bar{k}}}$.

Suppose the mass-spring systems are compared over the first two pairs of comparison $pc_0 = \langle t_0, \hat{t}_0 \rangle$ and $pc_1 = \langle t_1, \hat{t}_1 \rangle$, see figure 4.9(a). These pairs of comparison define two behavior fragments with durations $T = t_1 - t_0$ and $\hat{T} = \hat{t}_1 - \hat{t}_0$. T and $\hat{T}$ represent one quarter of the periods P and $\hat{P}$ of the solutions of the ideal mass-spring systems, which by analytical solution are found to be $P = 2\pi\sqrt{\frac{m}{k}}$ and $\hat{P} = 2\pi\sqrt{\frac{\hat{m}}{\hat{k}}}$.

We will prove that

$$\min\{T, \hat{T}\} = \min\{\frac{1}{4}P, \frac{1}{4}\hat{P}\} \leq \frac{1}{4}\bar{P} \leq \max\{\frac{1}{4}P, \frac{1}{4}\hat{P}\} = \max\{T, \hat{T}\}. \tag{A.8}$$

First, it will be shown that $\min\{\frac{1}{4}P, \frac{1}{4}\hat{P}\} \leq \frac{1}{4}\bar{P}$. From the argument in appendix A.2 we know that the quotient $\bar{k}/\bar{m}$ is bounded by k/m,$\hat{k}/\hat{m}$, so that

$$\min\{\frac{m}{k}, \frac{\hat{m}}{\hat{k}}\} \leq \frac{\bar{m}}{\bar{k}} \leq \max\{\frac{m}{k}, \frac{\hat{m}}{\hat{k}}\}.$$

As a consequence, the period of the comparison system is bounded by the periods of the original mass-spring systems. We conclude that $\min\{\frac{1}{4}P, \frac{1}{4}\hat{P}\} \leq \frac{1}{4}\bar{P}$ and $\frac{1}{4}\bar{P} \leq \max\{\frac{1}{4}P, \frac{1}{4}\hat{P}\}$. The special case $\min\{\frac{1}{4}P, \frac{1}{4}\hat{P}\} = \frac{1}{4}\bar{P}$ occurs when $\frac{1}{4}P = \frac{1}{4}\hat{P} = \frac{1}{4}\bar{P}$, and consequently $\frac{m}{k} = \frac{\hat{m}}{\hat{k}} = \frac{\bar{m}}{\bar{k}}$.

The relations $\min\{T, \hat{T}\} = \min\{\frac{1}{4}P, \frac{1}{4}\hat{P}\}$ and $\max\{T, \hat{T}\} = \max\{\frac{1}{4}P, \frac{1}{4}\hat{P}\}$ are obvious when one considers that the durations T and $\hat{T}$ of the behavior fragments between the pairs of comparison correspond to a quarter period of the behavior of the systems (compare figure 4.9 with figure 3.8). If the systems do not start from rest, we have $\min\{T, \hat{T}\} < \min\{\frac{1}{4}P, \frac{1}{4}\hat{P}\}$.

Appendix B

Theorems for Determining $\Phi(t, \tau)$

An important step in deriving RV constraints between pairs of comparison is the calculation of the transition matrix appearing in the solution of a linear comparison system (theorems 9 and 10). Explicitly solving for $\boldsymbol{\Phi}(t,\tau)$ is a difficult task in general, but under certain conditions, imposed on $\boldsymbol{A}(t)$, a closed-form expression for $\boldsymbol{\Phi}(t,\tau)$ can be found.[1]

If for every value t it holds that $\boldsymbol{A}(t)$ and $\int_\tau^t \boldsymbol{A}(\sigma)d\sigma$ commute, that is,

$$\boldsymbol{A}(t)\left(\int_\tau^t \boldsymbol{A}(\sigma)d\sigma\right) = \left(\int_\tau^t \boldsymbol{A}(\sigma)d\sigma\right)\boldsymbol{A}(t), \tag{B.1}$$

then

$$\boldsymbol{\Phi}(t,\tau) = \exp\left(\int_\tau^t \boldsymbol{A}(\sigma)d\sigma\right). \tag{B.2}$$

This condition is for example satisfied by time-invariant and symmetrical matrices $\boldsymbol{A}(t)$. The exponential of a matrix can be computed in various ways (e.g., (Chen [1970]), ch. 2; (Rugh [1996]), ch. 4).

Conditions on the form of $\boldsymbol{A}(t)$ may decompose the problem of determining a transition matrix into a number of simpler problems. For instance (Rugh [1996]) shows that, if $\boldsymbol{A}(t)$ is partitioned as

$$\boldsymbol{A}(t) = \begin{bmatrix} \boldsymbol{A}_{11}(t) & \boldsymbol{0} \\ \boldsymbol{A}_{21}(t) & \boldsymbol{A}_{22}(t) \end{bmatrix}, \tag{B.3}$$

where $\boldsymbol{A}_{11}(t)$ and $\boldsymbol{A}_{22}(t)$ are square, then

[1] Actually, less strict conditions would be sufficient for our purposes, conditions which allow the *signs* of the elements of $\boldsymbol{\Phi}(t,\tau)$ to be determined. A direct way to find such a qualitative transition matrix is to trace the zero-input solutions of the linear comparison system (4.8) in the phase space for initial states consisting of the column vectors of the unit matrix $\boldsymbol{I}$ (see, e.g., Rugh [1996]). This method is applied in (de Jong [1996]).

$$\boldsymbol{\Phi}(t) = \begin{bmatrix} \boldsymbol{\Phi}_{11}(t,\tau) & \mathbf{0} \\ \boldsymbol{\Phi}_{21}(t,\tau) & \boldsymbol{\Phi}_{22}(t,\tau) \end{bmatrix}, \tag{B.4}$$

where

$$\frac{\partial}{\partial t}\boldsymbol{\Phi}_{jj}(t,\tau) = \boldsymbol{A}_{jj}(t)\boldsymbol{\Phi}_{jj}(t,\tau), \quad \boldsymbol{\Phi}(\tau,\tau) = \boldsymbol{I},\ j = 1,2,$$

and

$$\boldsymbol{\Phi}_{21}(t,\tau) = \int_{\tau}^{t} \boldsymbol{\Phi}_{22}(t,\sigma)\boldsymbol{A}_{21}(\sigma)\boldsymbol{\Phi}_{11}(\sigma,\tau)d\sigma.$$

As can be easily verified, the state matrix

$$\boldsymbol{A}(t) = \begin{bmatrix} -\frac{\partial}{\partial a_u} f(\bar{a}_u(t),\bar{r}_u(t)) & 0 \\ \frac{\partial}{\partial a_u} f(\bar{a}_u(t),\bar{r}_u(t)) & -\frac{\partial}{\partial a_l} g(\bar{a}_l(t),\bar{r}_l(t)) \end{bmatrix}$$

derived in section 4.3.3 for the comparison of systems of cascaded-tanks with a water-tight and a leaky upper tank has the form (B.3), so that the transition matrix $\boldsymbol{\Phi}(t,\tau)$ is given by (B.4). For $\boldsymbol{A}_{11}(t) = -\frac{\partial}{\partial a_u} f(\bar{a}_u(t),\bar{r}_u(t))$ and $\boldsymbol{A}_{22}(t) = -\frac{\partial}{\partial a_l} g(\bar{a}_l(t),\bar{r}_l(t))$ condition (B.1) applies and we have

$$\boldsymbol{\Phi}(t,\tau) = \begin{bmatrix} \phi_{11}(t,\tau) & 0 \\ \int_{\tau}^{t} \phi_{22}(t,\sigma)\frac{\partial}{\partial a_u} f(\bar{a}_u(\sigma),\bar{r}_u(\sigma))\phi_{11}(\sigma,\tau)d\sigma & \phi_{22}(t,\tau) \end{bmatrix}. \tag{B.5}$$

with

$$\begin{aligned} \phi_{11}(t,\tau) &= \exp\left(\int_{\tau}^{t} -\frac{\partial}{\partial a_u} f(\bar{a}_u(\sigma),\bar{r}_u(\sigma))d\sigma\right) \\ \phi_{22}(t,\tau) &= \exp\left(\int_{\tau}^{t} -\frac{\partial}{\partial a_l} g(\bar{a}_l(\sigma),\bar{r}_l(\sigma))d\sigma\right). \end{aligned} \tag{B.6}$$

The mass-spring example in section 4.5.3 has a linear comparison system with the time-invariant state matrix

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ -\frac{\bar{k}}{\bar{m}} & 0 \end{bmatrix}. \tag{B.7}$$

The commutation condition (B.1) applies and a transition matrix is found by computing the exponential of the integral of $\boldsymbol{A}$, as stated in (B.2):

$$\boldsymbol{\Phi}(t,\tau) = \begin{bmatrix} \cos\sqrt{\frac{\bar{k}}{\bar{m}}}(t-\tau) & \sqrt{\frac{\bar{m}}{\bar{k}}}\sin\sqrt{\frac{\bar{k}}{\bar{m}}}(t-\tau) \\ -\sqrt{\frac{\bar{k}}{\bar{m}}}\sin\sqrt{\frac{\bar{k}}{\bar{m}}}(t-\tau) & \cos\sqrt{\frac{\bar{k}}{\bar{m}}}(t-\tau) \end{bmatrix}, \tag{B.8}$$

As can be seen, the transition matrix is periodic with period $\bar{P} = 2\pi\sqrt{\frac{\bar{m}}{\bar{k}}}$.

Appendix C

Algorithms for Filtering Behaviors

This appendix discusses algorithms for filtering qualitative and comparative behaviors of experimental systems against the measured states of the systems. The algorithms are used for model-based conflict resolution and error identification in chapter 5.

The algorithm for *qualitative behavior filtering* checks whether the sequence of measured states, in particular the sequence of qualitative state descriptions, can be mapped onto the predicted sequence of qualitative states of the behavior.

Algorithm 6 (Qualitative behavior filter) Given a qualitative behavior QB consisting of a sequence of qualitative states $QS(t_0), \ldots, QS(t_n)$. The consistency of QB with a sequence of qualitative state descriptions $QS_{meas}(u_0), \ldots, QS_{meas}(u_m)$ from the measured state sequence MSS needs to be determined.

Step 1 Set $i = 0$ and $j = 0$.

Step 2 If $j = m + 1$, then QB is consistent and stop. If $i = n + 1$ and $j \leq m$, then QB is inconsistent and stop.

Step 3 Check the consistency of $QS(t_i)$ and $QS_{meas}(u_j)$ by comparing the qualitative values of the variables in the states.

Step 4 If $QS(t_i)$ and $QS_{meas}(u_j)$ are consistent, then $j = j + 1$ and go to step 2. If they are not consistent, then $i = i + 1$ and go to step 2.

The algorithm could be further refined by distinguishing between qualitative states at a time-point and qualitative states over an interval, and by additionally checking whether the first and last measured state agree with the first and last qualitative state in the behavior.

The algorithm for *comparative behavior filtering* checks whether the measured states can be mapped onto pairs of comparison in the OPC structure. If this is the case, RVs for shared variables are derived from the measured states and confronted with the predicted RVs in the comparative behavior.

Algorithm 7 (Comparative behavior filter) Given a comparative behavior CB consisting of comparative states $CS(pc_0), \ldots, CS(pc_r)$ at the pairs of comparison $pc_0, \ldots, pc_r$. CB has been derived from the qualitative models QDE, $\hat{QDE}$ and the qualitative behaviors QB, $\hat{QB}$ consisting of the qualitative states $QS(t_0), \ldots, QS(t_{n_1})$ and $QS(\hat{t}_0), \ldots, QS(\hat{t}_{n_2})$. The consistency of CB with the measured state sequences $MSS = \langle MS(u_0), \ldots, MS(u_{m_1})\rangle$ and $\hat{MSS} = \langle MS(\hat{u}_0), \ldots, MS(\hat{u}_{m_2})\rangle$ needs to be determined.

Step 1 Find the measured states in MSS and $\hat{MSS}$ corresponding with a single distinguished time-point in QB and $\hat{QB}$, respectively. That is, compare the qualitative state descriptions $QS_{meas}(u_0), \ldots, QS_{meas}(u_{m_1})$ with the qualitative states $QS(t_0), \ldots, QS(t_{n_1})$ to find all u_{i_1}, $0 \leq i_1 \leq m_1$, corresponding with a single t_{j_1}, $0 \leq j_1 \leq n_1$. Label $MS(u_{i_1})$ with t_{j_1}. In the same way, label the appropriate $MS(\hat{u}_{i_2})$ with $\hat{t}_{j_2}$ ($0 \leq i_2 \leq n_2$ and $0 \leq j_2 \leq n_2$).

Step 2 Find pairs of measured states labeled in the previous step which correspond with a pair of comparison. If $MS(u_{i_1})$ is labeled with t_{j_1}, and $MS(\hat{u}_{i_2})$ with $\hat{t}_{j_2}$, and $pc_h = \langle t_{j_1}, \hat{t}_{j_2}\rangle$ is a pair of comparison, then $MS(u_{i_1})$ and $MS(\hat{u}_{i_2})$ correspond with pc_h.

Step 3 Derive RVs from the labeled measured states corresponding to pairs of comparison pc_h, $0 \leq h \leq r$. If the RV derived for some variable p contradicts $RV(p, pc_h)$ in $CS(pc_h)$, then CB is inconsistent with MSS and $\hat{MSS}$. Otherwise, it is consistent with the measured states.

Appendix D

Implementation of CEC* and KIMA

This appendix discusses the implementations of the comparative analysis technique CEC* and the system for knowledge-based measurement analysis KIMA. CEC* has been implemented by Frank van Raalte (van Raalte & de Jong [1997]) and KIMA by Hans de Wit (de Wit & de Jong [1998]). Output traces of an example of comparative analysis by means of CEC* and an example of conflict resolution by means of KIMA are included in appendix E.

D.1 Implementation of CEC*

Version 1.0 of CEC* was developed on a Sun SparcStation5 with 64Mb of RAM, running Solaris 2.5. The code has been written in Lucid Common Lisp version 4.2.1 and counts approximately 11,000 lines. CEC* was built on top of the Common Lisp implementation of QSIM 3.0 (van Raalte & de Jong [1997]).

In figure D.1 an outline of the CEC* implementation is given in the form of a dataflow diagram. I will first briefly discuss the data flows between the data stores, processes, and inputs in the figure, and then continue with a more detailed look at three aspects of the system: the determination of pairs of comparison, the construction of the comparative envisionment, and the RV rules. For more details the reader is referred to the technical report mentioned above.

The user is invited to specify the identifiers of the qualitative models of the first and second experimental system which are stored in a model database. Also, he is asked to specify the identifiers of corresponding initial qualitative states. A simulation is then performed with the models and initial states which results in two behavior trees, one for each system. The user is subsequently asked to select one behavior for each system, after which the structure of ordered pairs of comparison is determined. A set of initial RVs is then specified and propagated through the OPC structure by means of appropriate RV rules stored in a database with RV rules. The *RV rules* are basically templates of RV constraints with activation conditions depending on specific features of the qualitative models and behaviors. The results of RV propagation are presented in the form of a comparative envisionment. A teletype interface allows the user to specify the input and to set global variables determining the processing mode (predictive or explanatory CA) and the trace level.

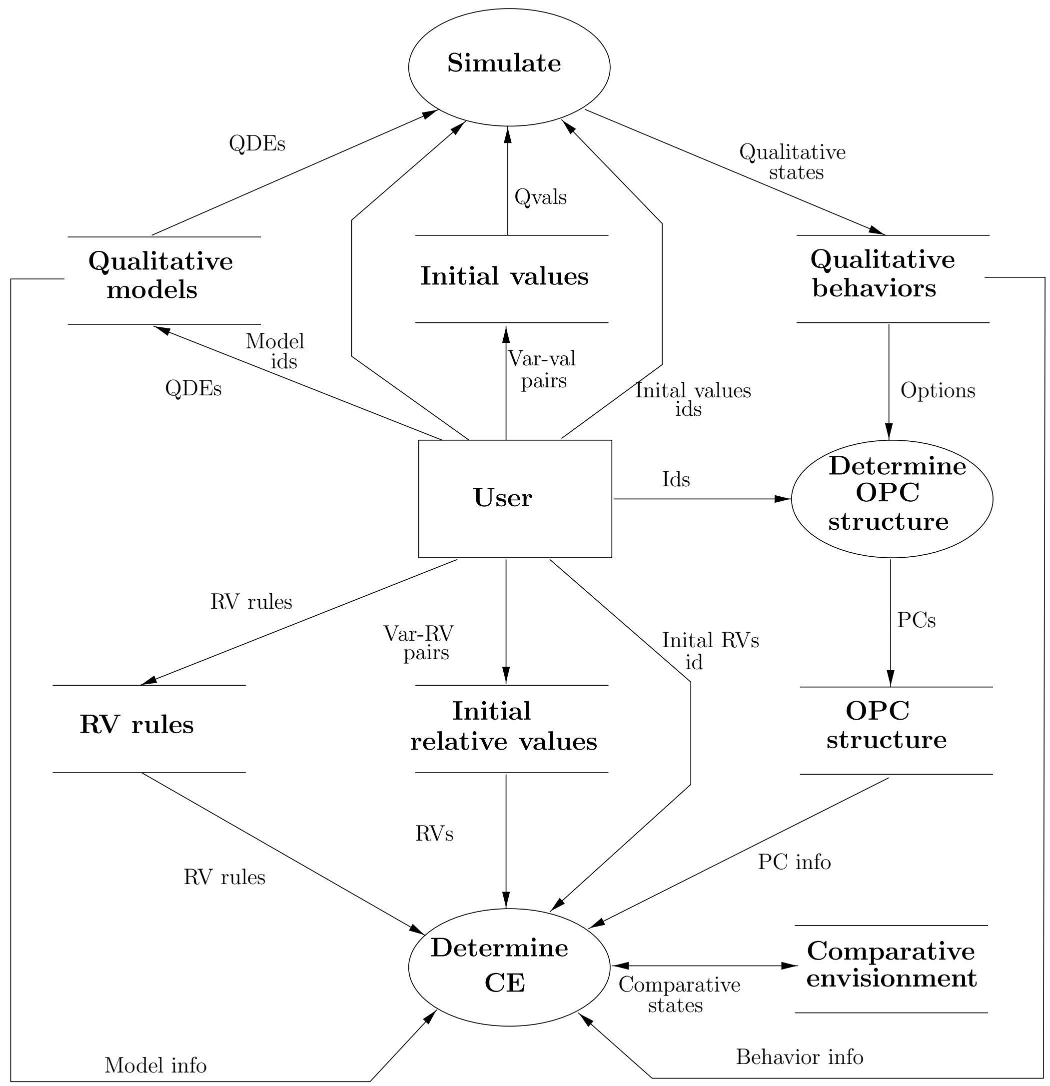


Figure D.1: Dataflow diagram of CEC*. Ovals represent processes, boxes represent inputs, and parallel bars represent data stores.

The implementation consists of the following modules:

`im-cec.lisp` Top-level functions of CEC*.

`opc.lisp` Functions for determining the OPC structure from two qualitative behaviors.

`qconstr.lisp` Definitions of RV constraints.

`rvconstr.lisp` Definitions of functions which can be used in the condition part of RV rules.

`rvrule.lisp` Functions achieving the selection of applicable RV rules from the RV rule database and the instantiation of the corresponding RV constraints included in the action part.

`rvruledb.lisp` RV rule database with domain-independent RV rules.

`cstate.lisp` Functions for determining a comparative envisionment from the ordered pairs of comparison and initial RVs.

`utils.lisp` General-purpose functions for constructing double-linked and hyperlink graphs. Comparative envisionments and OPC structures are implemented as hyperlink and double-linked graphs, respectively.

`interface.lisp` Interface functions.

`trace.lisp` Trace functions.

In order to perform a comparative analysis for a specific example, three additional input files need to be prepared. In particular, we need a file `example.lisp` specifying the QDEs of the systems in the domain, a file `example-ini.lisp` with possible sets of initial qualitative and relative values and settings of other CEC* parameters, and a file `example-rv.lisp` with domain-dependent RV rules. In appendix E the contents of these files will be shown.

The main function is named `im-cec` and shown in figure D.2. `im-cec` allows the user to initialize CEC* by changing the settings of the program, including the global variables `*SYS1-FORM*`, `*SYS2-FORM*`, `*SYS1-INITIAL-VID*`, and `*SYS2-INITIAL-VID*` which contain the identifiers of simulation functions and sets of initial qualitative values for the first and second experimental system. The function `obtain-selected-behavior` returns a qualitative behavior selected by the user after simulation has been performed by means of QSIM.

```
(defun im-cec ()
  (and
   (cec-initialize)
   (let*
       ((beh-sys1 (obtain-selected-behavior *SYS1-FORM* *SYS1-INITIAL-VID*))
        (beh-sys2 (obtain-selected-behavior *SYS2-FORM* *SYS2-INITIAL-VID*
                                            beh-sys1)))
     (setf *OPC-SET* (ask-opc-verification (opc-set beh-sys1 beh-sys2)))
     (comparative-states (opc-set-graph *OPC-SET*))
     (cs-graph-printer *CS-GRAPH*)))
)
```

Figure D.2: *The function* `im-cec` *in* `im-cec.lisp`.

The global variable `*OPC-SET*` contains a Lisp structure representing the ordered pairs of comparison. The OPC structure is created by the function `opc-set`, which takes as input two qualitative behaviors `beh-sys1` and `beh-sys1`. First, `opc-set` determines pairs of distinguished time-points at which the experimental systems can be meaningfully compared,[1] and then constructs a double-linked graph in which they have been ordered according to the $\preceq$-relation.

The function `comparative-states` determines the comparative envisionment from the OPC structure and the set of initial relative values referred to by `*INITIAL-RV-ID*` (figure D.3). The resulting comparative envisionment is a hypergraph structure stored in the global variable `*CS-GRAPH*`. The function `*CS-GRAPH-PRINTER*` prints the comparative envisionment on the screen.

```
(defun comparative-states (opc)
  (setf *CS-GRAPH*
        (cs-post-process
         opc
         (do ((pci (cs-valid-pc (dl-graph-root opc) opc)
                   (cs-valid-pc (dl-graph-root opc) opc)))
             ((null pci) *CS-GRAPH*)
             (comparative-states-pc (car pci) (cdr pci)))))
  (im-trace 'cs 'vars *CS-GRAPH* opc)
  *CS-GRAPH*
)

(defun comparative-states-pc (PC Preds)
  (im-trace 'cs 'pc PC Preds)
  (let ((qde-name (cs-generate-qde-name PC)))
    (cs-generate-qde qde-name PC
                     (cs-generate-qvars PC Preds)
                     (cs-deduce-constraints PC Preds))
    (cs-compute-comparative-states PC Preds qde-name (cs-value-sets Preds)
                                   (cs-generate-rv-time-qvar-names PC Preds)))
  (setf (pc-status (symbol-value PC)) 'done)
)
```

Figure D.3: *The functions* `comparative-states` *and* `comparative-states-pc` *in* `cstate.lisp`*. They realize the algorithm for constructing a comparative envisionment.*

The determination of the comparative envisionment by propagating initial RVs through the OPC structure is the most complicated part of the CEC* implementation. Basically, it follows the algorithm in chapter 4. At every pair of comparison a constraint satisfaction problem (CSP) is generated consisting of variables with domains $\{\Downarrow, \Uparrow, \|\}$ and RV

[1] In the present version 1.0 of CEC* the condition that a pair of comparison is not covered by a predecessor pair of comparison (see definition 15) has not yet been implemented.

```
(MSI
 MSI-1
 :system1
             ((model             ((ADD ?y ?z ?x)) ))
 :system2
             ((model             ((ADD ?y ?z ?x)) ))
 :result
             ((    (rv ?x) = (rv ?y) + (rv ?z)  ))
)
```

Figure D.4: *An example of an RV rule in the database* `rvruledb.lisp`*. The* `MSI-1` *template connects the occurrence of the additions* $x = y + z$ *and* $\hat{x} = \hat{y} + \hat{z}$ *in the model of the first and second system, respectively, to the RV constraint* $RV(x) = RV(y) + RV(z)$*.*

constraints restricting the assignment of RVs to variables. This CSP is represented as a special kind of QDE in the QSIM language. Variables are discrete QSIM variables with possible values ∞ ($\Uparrow$), 0 ($\|$), and $-\infty$ ($\Downarrow$). The RV constraints are special QSIM constraints which have been defined for this purpose (see table 5.1 in van Raalte & de Jong [1997]). This representation has the advantage that the Cfilter algorithm in QSIM can be used to solve the CSP and thus obtain the valid comparative states at the pair of comparison. The functions `cs-generate-qde` and `cs-compute-comparative-states` in `comparative-states-pc` are used to generate and solve the CSP, respectively.

The RV constraints included in the CSP originate from the action part of RV rules whose condition part is satisfied by features of the qualitative models and behaviors being analyzed. An example of an RV rule is shown in figure D.4. The condition part refers to ADD constraints in the qualitative models, whereas the action part specifies an RV constraint which is transformed into a special QSIM constraint by `cs-deduce-constraints`. There are several kinds of RV rules. The RV rule shown in the figure is an `MSI` rule which considers model and state information at a single pair of comparison. In addition, the rule is domain-independent in the sense that it is applicable to any pair of qualitative models. In the report on the implementation a complete overview of the different kinds of RV rules is given, including a formal syntax diagram (van Raalte & de Jong [1997], ch. 4).

D.2 Implementation of KIMA

The algorithms of the methods making up the KIMA system have been implemented in Common Lisp in the same technical environment as CEC*. The KIMA implementation comprises about 4000 lines of code and is built on top of the implementations of QSIM and CEC* whose main functions are repeatedly called.

The data flow diagram of the KIMA system is shown in figure D.5. The user can choose from four different tasks: selection of a measurement base, conflict detection, conflict resolution, and error identification. In the case of conflict detection, the constant d_{limit} functioning as a threshold for statistical conflicts (definition 36 in chapter 5) is supplied by the user. For the tasks of conflict resolution and error identification the user specifies which measurements in the selected measurement base are to be compared. The candidate models and candidate experimental conditions are provided with the property measurements. Simulation of the experimental systems with the candidate models and candidate conditions is followed by comparative analysis of all possible combinations of behaviors resulting from the simulation. Some of the initial RVs required for the comparative analyses are obtained from the measured states, others are supplied by the user.

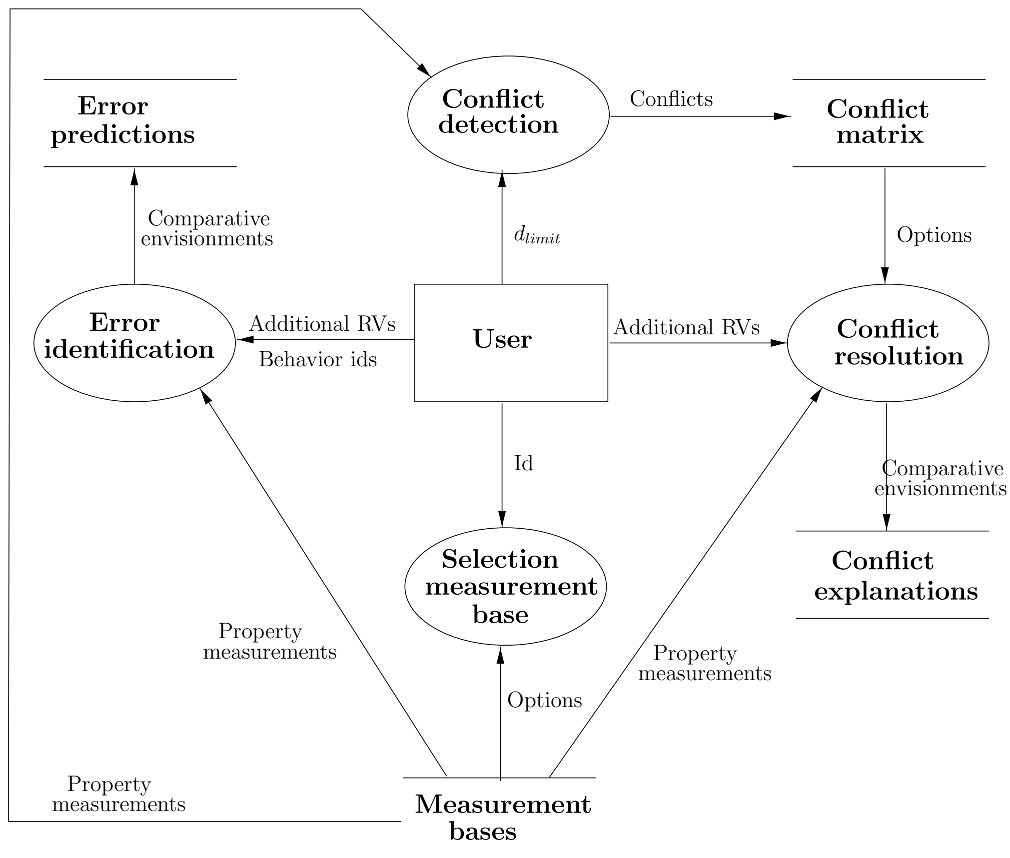


Figure D.5: Dataflow diagram of KIMA.

The KIMA implementation consists of the following modules:

`KIMA.lisp` Top-level functions of KIMA.

`KIMA-measset.lisp` Definitions of property measurements, properties, measured states, and other concepts introduced in section 5.1 as Lisp structures.

`KIMA-confdet.lisp` Functions for performing conflict detection.

`KIMA-confres.lisp` Functions for performing conflict resolution.

`KIMA-errid.lisp` Functions for performing error identification.

`KIMA-interface.lisp` Interface functions.

A set of measurement bases containing property measurements is the main input required for KIMA. A measurement base is represented as a Lisp file with the name `KIMA-data-*.lisp`, where `*` is an arbitrary identifying string.

The main KIMA function is shown in figure D.6. It is a simple loop allowing the user to choose between measurement base selection, conflict detection, conflict resolution, and error identification.

```
(defun KIMA ()
  (format t "========== KIMA ==========")
  (do ()
      ((format t "========== Main menu ========== 1. Conflict
                   detection~&2. Conflict resolution~&3. Error
                   identification~&4. Choose measurement base~&5.
                   Stop~&-------------------------~&(1-5) : "))
      (case (read)
            (1 (exec_conf_dect))
            (2 (exec_conf_res))
            (3 (exec_error_id))
            (4 (exec_choose_meas_prop))
            (5 (return t))
            (t (format t ‘‘Euuuh !? Can’t you count to 5 !"))
            )
      )
  )
```

Figure D.6: *The function* `KIMA` *in* `KIMA.lisp`*.*

The top-level function for conflict resolution is `conflict-resolution` (figure D.7). The global variables `*PM1*` and `*PM2*` contain structures representing the property measurements to be compared. They are used by the functions `derive_behaviors`, `eliminate_behaviors`, and `derive_comparative_behaviors` in order to derive qualitative behaviors by means of QSIM, rule out qualitative behaviors conflicting with the measured states, and derive comparative behaviors by means of CEC*, respectively. These functions are used in a different configuration for error identification, with a parameter value `'forward` instead of `'backward` in `derive_comparative_behaviors`.

```
(defun conflict_resolution ()
  (format t "Simulate experimental systems")
  (derive_behaviors *PM1*)
  (derive_behaviors *PM2*)
  (format t "Eliminate behaviors of experimental systems")
  (eliminate_behaviors *PM1*)
  (eliminate_behaviors *PM2*)
  (format t "Compare experimental systems")
  (derive_comparative_behaviors *PM1* *PM2* 'backward)
  )
```

Figure D.7: *The function* `conflict-resolution` *in* `KIMA-confres.lisp`.

Appendix E

Sample Traces of CEC* and KIMA

This appendix will show the output of the CEC* and KIMA programs when presented with a comparative analysis and conflict resolution example, respectively. The examples have been drawn from chapter 7. CEC* is given the question what might cause the lower fracture strength of a specimen described by QDE_{tt} in comparison with a specimen described by $QDE_{tteccsurf}$. KIMA tries to explain the conflict in example 2 of the case-study, that is, the conflict between two fracture strength measurements performed on systems with sets of candidate models $CM = \{\langle QDE_{tt}, QS_{tt}\rangle, \langle QDE_{ttecc}, QS_{ttecc}\rangle\}$ and $\hat{CM} = \{\langle Q\hat{D}E_{ttsurf}, \hat{Q}S_{ttsurf}\rangle, \langle Q\hat{D}E_{tteccsurf}, \hat{Q}S_{tteccsurf}\rangle\}$. The question answered by CEC* is a step in the process of resolving the conflict presented to KIMA.

E.1 CEC* trace

As explained in appendix D, CEC* needs a number of input files for its proper operation. In this case, we have the files `tf-quocr.lisp` and `tfeccsurf-quocr.lisp` containing the QDEs and simulation functions of the two experimental systems to be compared, the file `tf-quocr-ini.lisp` with alternative initial qualitative and relative values and additional CEC* parameter settings, and the file `tf-quocr-rv.lisp` with domain-dependent RV rules. I will show (excerpts of) the files `tf-quocr.lisp`, `tf-quocr-ini.lisp`, and `tf-quocr-rv.lisp` to give an idea of their contents (`tfeccsurf-quocr.lisp` is omitted because it is largely similar to `tf-quocr.lisp`).

```
;;;; File: tf-quocr.lisp
;;;;
;;;;    Definitions of the model and simulation functions for a tension test
;;;;    applied to a material exhibiting pure brittle fracture (mode I).
;;;;    The model is identical to the one in tf-critstr.lisp, but a single
;;;;    crack shape parameter quocr (c/rho) is used.
;;;;    The model adopts a critical-stress perspective on fracture (p. 404-5 in
;;;;    Courtney).
;;;;
```

```
(in-package :QSIM)

;;;
;;; ** QDE definitions **

;;
;; QDE: tensile-load-stress
;;
;; Model for the build-up of a stress accompanying the gradual stretching of
;; the sample. Fracture occurs when the maximum stress at the crack tip
;; reaches the theoretical strength.
;;

(define-QDE tension-load-stress
        (text "Build-up of the applied stress in tension test")
        (quantity-spaces
          (smct          (minf 0 inf)            "maximum stress at crack tip")
          (sa            (minf 0 inf)            "nominal applied stress")
          (sc            (minf 0 inf)            "applied stress at crack")
          (sth           (0 inf)                 "theoretical strength         ")
          (eps           (minf 0 inf)            "strain                       ")
          (e             (0 inf)                 "Young's modulus")
          (f             (minf 0 inf)            "applied force")
          (lr            (0  inf)                "loading rate")
          (d             (0  inf)                "gage-section diameter
                                                  of sample")
          (a             (0  inf)                "gage-section area of sample")
          (l             (0  inf)                "length sample")
          (l0            (0  inf)                "initial length sample")
          (diffthms      (minf 0  inf)           "theoretical strength-maximum
                                                  stress")
          (diffl         (minf 0  inf)           "diffl=l-l0")
          (num           (0 2 2sqrt2 inf)        "dimensionless number")
          (quocr         (0 inf)                 "quocr=c/rho")
          (sqrtquocr     (0 inf)                 "sqrtquocr=sqrt{c/rho}")
          (scsqrtquocr   (0 inf)                 "scsqrtquocr=sc(sqrt{c/rho})")
          (gam           (0 inf)                 "surface energy per unit area")
          (a0            (0 inf)                 "interatomic distance")
          (game          (0 inf)                 "(gam)(e)")
          (gamea0        (0 inf)                 "(gam)(e)/a0")
        )
        (constraints
          ((D/DT         l lr))
          ((MULT         eps  l0 diffl))
          ((MULT         sa  a f))
```

```
          ((EQUAL       sa sc))
          ((MULT        eps e sa))
          ((M+          quocr sqrtquocr)   (0 0) (inf inf))
          ((MULT        sc sqrtquocr scsqrtquocr))
          ((MULT        num scsqrtquocr smct))
          ((ADD         diffthms smct sth))
          ((ADD         diffl l0 l))
          ((M+          d a)           (0 0))
          ((MULT        gam e game))
          ((MULT        gamea0 a0 game))
          ((M+          gamea0 sth)    (0 0))
          ((CONSTANT    sth))
          ((CONSTANT    num))
          ((CONSTANT    e))
          ((CONSTANT    lr))
          ((CONSTANT    quocr))
          ((CONSTANT    d))
          ((CONSTANT    a))
          ((CONSTANT    l0))
          ((CONSTANT    a0))
          ((CONSTANT    gam))
        )
        (layout (nil nil nil)
                (sth smct nil)
                (diffthms eps nil))
        (transitions
          ((diffthms    (0 dec)) -> t))
;         ((diffthms    (0 dec)) -> trans-crack-propagation))
        (unreachable-values (l inf))
)


;;
;; QDE: tension-crack-propagation
;;
;;      Model for the situation in which the cracks in the sample start
;;      propagating until the sample is broken. A sample is broken if the
;;      crack length equals the diameter of the sample.
;;

(define-QDE tension-crack-propagation
        (text "Propagation of cracks in tension test")
        (quantity-spaces
          (msct         (minf 0 inf)            "maximum stress at crack tip")
          (as           (minf 0 inf)            "applied stress")
          (ths          (0 inf)                 "theoretical strength")
```

```
    (e          (minf 0 inf)           "strain")
    (ym         (0 inf)                "Young's modulus")
    (f          (minf 0 inf)           "applied force")
    (d          (0  inf)               "diameter sample")
    (a          (0  inf)               "cross-section sample")
    (l          (0  inf)               "length sample")
    (l0         (0  inf)               "initial length sample")
    (c          (0  inf)               "crack length")
    (cv         (0  inf)               "crack velocity")
    (rho        (0  inf)               "crack radius")
    (diffc      (minf 0  inf)          "diffc=d-c")
    (diffl      (minf 0  inf)          "diffl=l-l0")
    (quocr      (0 inf)                "quocr=c/rho")
  )
  (constraints
    ((D/DT       c cv))
    ((MULT       e  l0 diffl))
    ((MULT       as  a f))
    ((MULT       e  ym as))
    ((MULT       quocr rho c))
    ((ADD        diffl l0 l))
    ((ADD        diffc c d))
    (((M + +)    quocr as msct)   (0 0 0))
    ((M+         d a)          (0 0))
    ((M+         msct cv))
    ((CONSTANT   ym))
    ((CONSTANT   as))
    ((CONSTANT   ths))
    ((CONSTANT   d))
    ((CONSTANT   a))
    ((CONSTANT   l0))
    ((CONSTANT   l))
    ((CONSTANT   rho))
  )
  (layout (nil nil nil)
          (sth smct nil)
          (diffthms eps nil))
  (transitions
    ((diffthms   (0 dec)) -> t))
  (unreachable-values (l inf))
)


...
...
```

```
;;;
;;; ** Simulation functions **
;;;

;;
;; Functions: simulate-tension-test-qsim, simulate-tension-test
;;
;; Description:
;;      Functions which create the intial state for the tension test with
;;      brittle fracture.
;;      The set of initial values is identified by the optional argument setid.
;;


...
...


(defun simulate-tension-test ()
    (cond ((null *INITIAL-VALUES*)
           (cec-error '() "No initial values specified"))
          (t (let ((initial-state
                    (make-new-state
                     :from-qde tension-load-stress
                     :assert-values *INITIAL-VALUES*
                     :text "Load the sample.")))
               (qsim initial-state)
               initial-state)))
)


;;;;
;;;; File: tf-quocr-ini.lisp
;;;;
;;;; Initializations for the pure brittle fracture examples.
;;;;

(in-package :QSIM)

;;;
;;; Definition of the initial value sets for QSIM simulation.
;;;

(setf *INITIAL-VALUE-SETS*
  '((v1  ;tf-quocr
     (d ((0 inf) std))
     (a ((0 inf) std))
```

```
     (l ((0 inf) nil))
     (l0 ((0 inf) std))
     (diffl (0 nil))
     (diffthms ((0 inf) nil))
     (e ((0 inf) std))
     (lr ((0 inf) std))
     (quocr ((0 inf) std))
     (num (2 std))
     (gam ((0 inf) std))
     (a0 ((0 inf) std)))
    (v2  ; tfsurf-quocr, tfdamsurf-quocr
     (d ((0 inf) std))
     (a ((0 inf) std))
     (l ((0 inf) nil))
     (l0 ((0 inf) std))
     (diffl (0 nil))
     (diffthms ((0 inf) nil))
     (e ((0 inf) std))
     (lr ((0 inf) std))
     (quocr ((0 inf) std))
     (num (2sqrt2  std))
     (gam ((0 inf) std))
     (a0 ((0 inf) std)))

...
...

    (v4  ; tfeccsurf-quocr, tfeccdamsurf-quocr
     (d ((0 inf) std))
     (a ((0 inf) std))
     (l ((0 inf) nil))
     (l0 ((0 inf) std))
     (diffl (0 nil))
     (diffthms ((0 inf) nil))
     (e ((0 inf) std))
     (lr ((0 inf) std))
     (quocr ((0 inf) std))
     (ec ((0 inf) std))
     (r ((0 inf) std))
     (xr (1 std))
     (num (2sqrt2 std))
     (gam ((0 inf) std))
     (a0 ((0 inf) std)))

...
...

))
```

```
;;;
;;; Definition of the initial relative value sets at the start of CEC
;;;

(setf *INITIAL-RV-SETS*
  '((p0)                ; Empty
    (p1  (d  =) (l =) (l0 =) (quocr =) (lr =) (e =) (gam =) (a0 =))
    (p2  (d  =) (l =) (l0 =) (quocr +) (lr =) (e =) (gam =) (a0 =))
    (p3  (d  =) (l =) (l0 =) (quocr -) (lr =) (e =) (gam =) (a0 =))
    (p4  (d  =) (l =) (l0 =) (lr =) (e =) (gam =) (a0 =))
    (p5  (d  =) (l =) (l0 =) (quocr =) (lr -) (e =) (gam =) (a0 =))
    (p6  (d  +) (l =) (l0 =) (quocr =) (lr =) (e =) (gam =) (a0 =))
    (p7  (d  =) (l =) (l0 =) (lr =) (e =) (gam =) (a0 =))
    (p8  (d  =) (l =) (l0 =) (quocr =) (lr =) (e =) (gam =) (a0 =))
    (p9  (d  =) (eps -) (l0 =) (lr =) (e =) (gam =) (a0 =) (sa -))
    (p10  (d  =) (eps -) (l0 =) (lr =) (e =) (gam =) (a0 =) (sa -) (ec =))
    (p11  (d  =) (eps -) (l0 =) (quocr +) (lr =) (e =) (gam =) (a0 =) (sa -))
    (p12  (d  =) (l =) (l0 =) (quocr =) (lr =) (e +) (gam =) (a0 =))
    (p13  (d  =) (l =) (l0 =) (quocr +) (lr =) (e +) (gam =) (a0 =))
    (p14  (d  =) (eps =) (l0 =) (lr =) (e =) (gam =) (a0 =) (sa =))
    (p15  (d  =) (eps =) (l0 =) (quocr +) (lr =) (e =) (gam =) (a0 =) (sa =))
    (p16  (d  =) (eps +) (l0 =) (lr =) (e =) (gam =) (a0 =) (sa +))
    (p17  (d  =) (eps +) (l0 =) (quocr +) (lr =) (e =) (gam =) (a0 =) (sa +))
    (p18  (d  =) (l0 =) (lr =) (e =) (gam =) (a0 =))
    (p19  (d  =) (l0 =) (lr =) (e +) (gam =) (a0 =))
    (p20  (d  =) (l0 =) (lr =) (e =) (gam =) (a0 =) (quocr +))
    (p21  (d  =) (l0 =) (lr =) (e +) (gam =) (a0 =) (quocr +))
    (np)
   )
)

;;;
;;; Definition of the initial relative value sets by means of rules.
;;;

(setf *INITIAL-RV-RULE-SETS*
  '((ir0))              ; Empty
)

(setf *SYS-FORMS*               '(("simulate-tension-test"
                                        (simulate-tension-test))
                                  ("simulate-tension-test-ecc"
                                        (simulate-tension-test-ecc))
                                  ("simulate-tension-test-sh"
                                        (simulate-tension-test-sh))
```

```
                                  ("simulate-tension-test-shsurf"
                                        (simulate-tension-test-shsurf))
                                  ("simulate-tension-test-grip"
                                        (simulate-tension-test-grip))
                                  ("simulate-tension-test-gripsurf"
                                        (simulate-tension-test-gripsurf))
                                  ("simulate-tension-test-surf"
                                        (simulate-tension-test-surf))
                                  ("simulate-tension-test-eccsurf"
                                        (simulate-tension-test-eccsurf))
                                  )
)

(setf *SYS1-FORM*               (second (first *SYS-FORMS*)))
(setf *SYS2-FORM*               (second (third *SYS-FORMS*)))
(setf *SYS1-INITIAL-VID*        'v1)
(setf *SYS2-INITIAL-VID*        'v1)
(setf *INITIAL-RV-ID*           'p1)
(setf *INITIAL-RV-RULE-ID*      'ir0)
(setf *CS-REASONING-DIRECTION*  'forward)

(setf *IM-TRACE* 1)



;;;;
;;;; File: tf-quocr-rv.lisp
;;;;
;;;; Specific model-dependent RV-Rules for the tensile fracture examples.
;;;; The rules are model-dependent in that the condition part of
;;;; the rules can refer to specific QDEs.
;;;;
;;;; Type of the rules specified is
;;;;    . Model and State Information rules on an interval between a specific
;;;;      pair of comparison and one of its predecessors (INT-MSI).
;;;;

(in-package :QSIM)

;;;
;;; Rules related to the Model and State Information on the interval
;;; between PCi and PCj, where PCj is a predecessor of PCi.
;;;

;; Various combinations of fracture models which all have the same RV constraints
;; for the stress build-up stage of brittle fracture.
```

```
(INT-MSI
 INT-MSI-TF-1
 :constraints
        (and (rv-var-qdir-equal-int-p 'L 'inc Previous Current)
             (or (rv-state-sequence-int-p '(tension-load-stress)
                                          '(tension-load-stress)
                                          Previous Current)
                 (rv-state-sequence-int-p '(tension-load-stress)
                                          '(tension-load-stress-ecc)
                                          Previous Current)
                 (rv-state-sequence-int-p '(tension-load-stress)
                                          '(tension-load-stress-surf)
                                          Previous Current)
                 (rv-state-sequence-int-p '(tension-load-stress)
                                          '(tension-load-stress-eccsurf)
                                          Previous Current)

...
...
                 (rv-state-sequence-int-p '(tension-load-stress-surf)
                                          '(tension-load-stress-gripsurf)
                                          Previous Current)
                 ))
 :result
        (( IF (( (rv TIME Interval) = minf ))
           THEN (( (rv L) <  (rv L Previous) + (rv LR Previous) )))
         ( IF (( (rv TIME Interval) = 0 ))
           THEN (( (rv L) =  (rv L Previous) + (rv LR Previous) )))
         ( IF (( (rv TIME Interval) = inf ))
           THEN (( (rv L) >  (rv L Previous) + (rv LR Previous) )))
        )
)
```

The above input files are used by CEC* in performing the explanatory comparative analysis mentioned in the introduction. The following output is produced by the program:

```
screwdriver% Cec
;;; Lucid Common Lisp/SPARC Solaris,  Version: 4.2.1
;;; Development Environment (DBCS),  Release Date: 25 December 1994
;;; Copyright (C) 1994 by Harlequin, Inc.,  All Rights Reserved
;;;
;;; This software product contains confidential and trade secret information
;;; belonging to Harlequin, Inc.  It may not be copied for any reason other
;;; than for archival and backup purposes.
```

```
;;;
;;; Lucid and Lucid Common Lisp are trademarks of Harlequin, Inc.  Other
;;; brand or product names are trademarks or registered trademarks of their
;;; respective holders.

> (im-cec)

+=============================================================================+
|                          = Parameters for CEC =                             |
|=============================================================================|
| * System Information *                                                      |
|=============================================================================|
| Simulation   | System 1                      | System 2                     |
|              |--------------------------------------------------------------|
|  Function    | NIL                           | NIL                          |
|  Value set   | V1                            | V1                           |
|-----------------------------------------------------------------------------|
|  Initial relative value set (Var/RV pairs) : P0                             |
|      Initial relative value set (RV rules) : IR0                            |
|        Processing direction of the OPC set : FORWARD                        |
|=============================================================================|
| * Trace Information *                                                       |
|=============================================================================|
| Modules      | *traced*                      | *not traced*                 |
|              |--------------------------------------------------------------|
|              | Ordered Pairs of Comparison   |                              |
|              | Rules for Relative Values     |                              |
|              | Comparative States            |                              |
|-----------------------------------------------------------------------------|
|                                 Trace level : 10                            |
+=============================================================================+

Use these settings? [Y/N] n

Change the system parameters? [Y/N] y

* Select the model:

    1. Tensile Test
    2. Cascaded Tanks
    3. Two Tanks filled by a single Pump
    4. Heat Exchangers
    5. Bath Tubs
    6. Vibrating Springs
    7. Sliding Blocks
    8. Tension Tests
```

```
    9. Bend Tests
    10. Tension Tests (single crack parameter)
    11. Bend Tests (single crack parameter)
    12. Tension Tests (Griffith)
    13. Bend Tests (Griffith)
    14. Bend Test and Tension Test

And the winner is ... 10
;;; Loading source file "/home/kbs/hdejong/qsim/models/tf-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfsurf-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfdamsurf-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfecc-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfsh-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfgrip-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfshsurf-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfgripsurf-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfeccsurf-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfeccdamsurf-quocr.lisp"
;;; Loading source file "model/tf-quocr-rv.lisp"
Adding RV-Rule INT-MSI-TF-1 of type INT-MSI-RULE
Adding RV-Rule MSI-TF-1 of type MSI-RULE
Adding RV-Rule MSI-TF-2 of type MSI-RULE
;;; Loading source file "model/tf-quocr-ini.lisp"

* Select an initialization function:

    1. simulate-tension-test
    2. simulate-tension-test-ecc
    3. simulate-tension-test-sh
    4. simulate-tension-test-shsurf
    5. simulate-tension-test-grip
    6. simulate-tension-test-gripsurf
    7. simulate-tension-test-surf
    8. simulate-tension-test-damsurf
    9. simulate-tension-test-eccsurf
    10. simulate-tension-test-eccdamsurf

And the winner is ... 1

* Select the initial value set for simulation:

    1. V1 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...
    2. V2 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...
    3. V3 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...
    4. V4 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...
    5. V5 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...
```

```
    6. V6 ((D ((0 INF) STD)) (A ((0 INF) STD)) (DH ((0 INF) STD ...
    7. V7 ((D ((0 INF) STD)) (A ((0 INF) STD)) (DH ((0 INF) STD ...
    8. V8 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...

And the winner is ... 1

Use the same model and behavior for the second system? [Y/N] n

* Select an initialization function:

    1. simulate-tension-test
    2. simulate-tension-test-ecc
    3. simulate-tension-test-sh
    4. simulate-tension-test-shsurf
    5. simulate-tension-test-grip
    6. simulate-tension-test-gripsurf
    7. simulate-tension-test-surf
    8. simulate-tension-test-damsurf
    9. simulate-tension-test-eccsurf
    10. simulate-tension-test-eccdamsurf

And the winner is ... 9

* Select the initial value set for simulation:

    1. V1 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...
    2. V2 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...
    3. V3 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...
    4. V4 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...
    5. V5 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...
    6. V6 ((D ((0 INF) STD)) (A ((0 INF) STD)) (DH ((0 INF) STD ...
    7. V7 ((D ((0 INF) STD)) (A ((0 INF) STD)) (DH ((0 INF) STD ...
    8. V8 ((D ((0 INF) STD)) (A ((0 INF) STD)) (L ((0 INF) NIL) ...

And the winner is ... 4


* Select the initial relative value set (Var/RV pairs):

    1. P0 NIL
    2. P1 ((D =) (L =) (L0 =) (QUOCR =) (LR =) (E =) (GAM =) (A ...
    3. P2 ((D =) (L =) (L0 =) (QUOCR +) (LR =) (E =) (GAM =) (A ...
    4. P3 ((D =) (L =) (L0 =) (QUOCR -) (LR =) (E =) (GAM =) (A ...
    5. P4 ((D =) (L =) (L0 =) (LR =) (E =) (GAM =) (A0 =))
    6. P5 ((D =) (L =) (L0 =) (QUOCR =) (LR -) (E =) (GAM =) (A ...
    7. P6 ((D +) (L =) (L0 =) (QUOCR =) (LR =) (E =) (GAM =) (A ...
```

```
    8. P7 ((D =) (L =) (L0 =) (LR =) (E =) (GAM =) (A0 =))
    9. P8 ((D =) (L =) (L0 =) (QUOCR =) (LR =) (E =) (GAM =) (A ...
    10. P9 ((D =) (EPS -) (L0 =) (LR =) (E =) (GAM =) (A0 =) (SA -))
    11. P10 ((D =) (EPS -) (L0 =) (LR =) (E =) (GAM =) (A0 =) (S ...
    12. P11 ((D =) (EPS -) (L0 =) (QUOCR +) (LR =) (E =) (GAM =) ...
    13. P12 ((D =) (L =) (L0 =) (QUOCR =) (LR =) (E +) (GAM =) ( ...
    14. P13 ((D =) (L =) (L0 =) (QUOCR +) (LR =) (E +) (GAM =) ( ...
    15. P14 ((D =) (EPS =) (L0 =) (LR =) (E =) (GAM =) (A0 =) (S ...
    16. P15 ((D =) (EPS =) (L0 =) (QUOCR +) (LR =) (E =) (GAM =) ...
    17. P16 ((D =) (EPS +) (L0 =) (LR =) (E =) (GAM =) (A0 =) (S ...
    18. P17 ((D =) (EPS +) (L0 =) (QUOCR +) (LR =) (E =) (GAM =) ...
    19. P18 ((D =) (L0 =) (LR =) (E =) (GAM =) (A0 =))
    20. P19 ((D =) (L0 =) (LR =) (E +) (GAM =) (A0 =))
    21. P20 ((D =) (L0 =) (LR =) (E =) (GAM =) (A0 =) (QUOCR +))
    22. P21 ((D =) (L0 =) (LR =) (E +) (GAM =) (A0 =) (QUOCR +))
    23. NP NIL

And the winner is ... 12

* Select the initial relative value set (RV rules):

    1. IR0 NIL

And the winner is ... 1

* Select the direction of processing the OPC set:

    1. FORWARD
    2. BACKWARD

And the winner is ... 2

Change the trace parameters? [Y/N] y

Enter trace level (currently 1): 1

Switches for tracing system modules:

    1. Ordered Pairs of Comparison            (ON)
    2. Rules for Relative Values              (ON)
    3. Comparative States                     (ON)

    Q. Quit

And the winner is ... q
```

```
+==============================================================================+
|                           = Parameters for CEC =                             |
|==============================================================================|
| * System Information *                                                       |
|==============================================================================|
| Simulation   | System 1                     | System 2                       |
|              |---------------------------------------------------------------|
|  Function    | (SIMULATE-TENSION-TEST)      | (SIMULATE-TENSION-TEST-E ...   |
|  Value set   | V1                           | V4                             |
|------------------------------------------------------------------------------|
|  Initial relative value set (Var/RV pairs) : P11                             |
|      Initial relative value set (RV rules) : IR0                             |
|        Processing direction of the OPC set : BACKWARD                        |
|==============================================================================|
| * Trace Information *                                                        |
|==============================================================================|
| Modules      | *traced*                     | *not traced*                   |
|              |---------------------------------------------------------------|
|              | Ordered Pairs of Comparison  |                                |
|              | Rules for Relative Values    |                                |
|              | Comparative States           |                                |
|------------------------------------------------------------------------------|
|                                 Trace level : 1                              |
+==============================================================================+

Use these settings? [Y/N] y

Initial relative values (var/RV pairs)...
    <D 0>
    <EPS MINF>
    <L0 0>
    <QUOCR INF>
    <LR 0>
    <E 0>
    <GAM 0>
    <A0 0>
    <SA MINF>

Initial relative values (RV rules)...

Initial simulation values for (SIMULATE-TENSION-TEST)...
    <(D ((0 INF) STD))>
    <(A ((0 INF) STD))>
    <(L ((0 INF) NIL))>
    <(L0 ((0 INF) STD))>
```

```
    <(DIFFL (0 NIL))>
    <(DIFFTHMS ((0 INF) NIL))>
    <(E ((0 INF) STD))>
    <(LR ((0 INF) STD))>
    <(QUOCR ((0 INF) STD))>
    <(NUM (2 STD))>
    <(GAM ((0 INF) STD))>
    <(A0 ((0 INF) STD))>

Variables in TENSION-LOAD-STRESS likely to chatter are NONE.
Run time: 0.080 seconds to initialize S-0.
Variables in TENSION-LOAD-STRESS likely to chatter are NONE.
Run time: 0.100 seconds to simulate 4 states.

* Simulation resulted in 1 behaviour:
        D       : Display results
        Number  : Selected behavior
        Q       : Quit
Your choice: 1

Initial simulation values for (SIMULATE-TENSION-TEST-ECCSURF)...
    <(D ((0 INF) STD))>
    <(A ((0 INF) STD))>
    <(L ((0 INF) NIL))>
    <(L0 ((0 INF) STD))>
    <(DIFFL (0 NIL))>
    <(DIFFTHMS ((0 INF) NIL))>
    <(E ((0 INF) STD))>
    <(LR ((0 INF) STD))>
    <(QUOCR ((0 INF) STD))>
    <(EC ((0 INF) STD))>
    <(R ((0 INF) STD))>
    <(XR (1 STD))>
    <(NUM (2SQRT2 STD))>
    <(GAM ((0 INF) STD))>
    <(A0 ((0 INF) STD))>

Variables in TENSION-LOAD-STRESS-ECCSURF likely to chatter are NONE.
Run time: 0.080 seconds to initialize S-3.
Variables in TENSION-LOAD-STRESS-ECCSURF likely to chatter are NONE.
Run time: 0.150 seconds to simulate 4 states.

* Simulation resulted in 1 behaviour:
        D       : Display results
        Number  : Selected behavior
        Q       : Quit
```

```
Your choice: 1
* Generating OPC
* Behaviour first system : (S-0 S-1 S-2)
* Behaviour second system: (S-3 S-4 S-5)

* Graph representing the Ordered Pairs of Comparison:
*  Root: (PC-0)
*  Leaf: (PC-1)
*  Nodes (2):
      (PC-1 NIL (PC-0))
      (PC-0 (PC-1) NIL)

Do you want to cheat? [Y/N] n
**Collecting constraints for PC-1 with rules at one PC
**Collecting constraints for PC-0 with rules at one PC
**Collecting constraints for PC-0 with rules at an interval between two PCs
**Overview of the relative values of variables

===================================================================+
          |SMCT| SA | SC | STH| EPS|  E |  F | LR |  D |  A |  L |
===================================================================+
CS-PC-1-0 |  = |  - |  - |  = |  - |  = |  - |  = |  = |  = |  - |
CS-PC-0-0 |  = |  = |  = |  = |  = |  = |  = |  = |  = |  = |  = |
===================================================================+


===================================================================+
          |LO  |DIF.|DIF.| NUM|QUO.|SQR.|SCS.| GAM| AO |GAME|GAM.|
===================================================================+
CS-PC-1-0 |  = |  = |  - |  + |  + |  + |  - |  = |  = |  = |  = |
CS-PC-0-0 |  = |  = |  = |  + |  + |  + |  = |  = |  = |  = |  = |
===================================================================+


* Graph representing the Comparative Envisionment
*  Root: (CS-PC-1-0)
*  Leaf: (CS-PC-0-0)
*  Nodes (2):
      (CS-PC-1-0 (EDGE-0) NIL)
      (CS-PC-0-0 NIL (EDGE-0))
*  Edges (1):
      (EDGE-0 (CS-PC-0-0) (CS-PC-1-0))
*  Labels (1):
      (EDGE-0 ((RV-TIME-PC-0-PC-1 (MINF IGN))))
* Paths in graph: (1):
    (CS-PC-1-0 ((RV-TIME-PC-0-PC-1 (MINF IGN))) CS-PC-0-0)
```

```
(NIL)
>
```

The output of CEC* corresponds with the comparative envisionment shown in figure 7.7.

E.2 KIMA trace

KIMA uses as input a measurement base with property measurements. The file `KIMA-data-tens.lisp` contains the property measurements in the tension test examples. Some relevant excerpts from this file are shown below. It shows the definition of the property measurements compared in example 2 of chapter 5, the definition of the fracture strength property, and the definition of the experimental system of one of the property measurements. The definitions of the measured states and measured values are omitted.

```
;;;; File: KIMA-data-tens.lisp
;;;;

(in-package :QSIM)

(setf *PROP-MEAS-BASES* `((mbtt (,pmtt-1 ,pmtt-2 ,pmtt-3 ,pmtt-4
,pmtt-5 ,pmtt-6 ,pmtt-7 ,pmtt-id1 ))))

;;;;; Property measurements

(setq pmtt-1
      (
       make-property_measurement
       :description "pmtt-1 "
       :type "strength tension test"
       :experiment-type "tension test"
       :exp_system estt-1
       :property p-1
       :measured_states s-50
       )
      )


...
...


;;;; Properties

(setq p-1
      (
```

```
        make-property
        :description "strength"
        :definition_state 'qs-end
        :definition_quantity 'SA
        )
       )


;;;; Experimental systems


(setq estt-1
       (
        make-exp_system
        :description "alumina sample"
        :exp_system_structure '(simulate-tension-test simulate-tension-test-ecc)
        :exp_system_conditions '(v1 v3)
        :type '("/home/kbs/hdejong/qsim/models/tf-quocr"
         "/home/kbs/hdejong/qsim/models/tfecc-quocr") ;;; file names
        )
       )

...
...
```

Now, KIMA is started and applied to example 2 of the case-study. Four comparative analyses need to be performed (figure 7.7), but I will only show the output resulting from the comparison of QDE_{tt} and $QDE_{tteccsurf}$, that is, the same analysis as performed with CEC* in the previous section. This allows one to gain an insight into the interaction between KIMA and CEC*.

```
screwdriver% Cec
;;; Lucid Common Lisp/SPARC Solaris,  Version: 4.2.1
;;; Development Environment (DBCS),  Release Date: 25 December 1994
;;; Copyright (C) 1994 by Harlequin, Inc.,  All Rights Reserved
;;;
;;; This software product contains confidential and trade secret information
;;; belonging to Harlequin, Inc.  It may not be copied for any reason other
;;; than for archival and backup purposes.
;;;
;;; Lucid and Lucid Common Lisp are trademarks of Harlequin, Inc.  Other
;;; brand or product names are trademarks or registered trademarks of their
;;; respective holders.

> (load 'kima)
;;; Loading source file "KIMA.lisp"
```

```
;;; Loading source file "KIMA-measset.lisp"
;;; Loading source file "KIMA-confdet.lisp"
;;; Loading source file "KIMA-confres.lisp"
;;; Loading source file "KIMA-errid.lisp"
;;; Loading source file "KIMA-aggr.lisp"
;;; Loading source file "KIMA-interface.lisp"
#P"/home/kbs/wit/KIMI/KIMA.lisp"
> (kima)


========== KIMA ==========


========== Main menu ==========

1. Conflict detection
2. Conflict resolution
3. Error identification
4. Choose measurement base
5. Stop
-------------------------
(1-5) : 4



========== Choose measurement base ==========


* Choose type of test

 1. Four-point bend test
 2. Tension test
 3. Four-point bend test (extended crack parameters)
 ----------------
(1-3) : 2

 Load models, initial conditions, simulation functions and RV rules

;;; Loading source file "KIMA-data-tens.lisp"
;;; Loading source file "/home/kbs/hdejong/cec/model/tf-quocr-ini.lisp"
;;; Loading source file "/home/kbs/hdejong/cec/model/tf-quocr-rv.lisp"
Adding RV-Rule INT-MSI-TF-1 of type INT-MSI-RULE
Adding RV-Rule MSI-TF-1 of type MSI-RULE
Adding RV-Rule MSI-TF-2 of type MSI-RULE
;;; Loading source file "/home/kbs/hdejong/qsim/models/tf-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfsurf-quocr.lisp"
```

```
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfgripsurf-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfshsurf-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfeccsurf-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfgrip-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfsh-quocr.lisp"
;;; Loading source file "/home/kbs/hdejong/qsim/models/tfecc-quocr.lisp"

* Select measurement base :

    1. MBTT (pmtt-1  pmtt-2  pmtt-3  pmtt-4  pmtt-5  pmtt-6  pmtt-7  pmtt-idsurf)

And the winner is ... 1


========== Main menu ==========

1. Conflict detection
2. Conflict resolution
3. Error identification
4. Choose measurement base
5. Stop
-------------------------
(1-5) : 2



========== Conflict resolution ==========


Enter trace level (currently 1): 1

* Select property measurement 1 :

    1. pmtt-1  (((285 3) MPa))
    2. pmtt-2  (((210 2) MPa))
    3. pmtt-3  (((275 3) MPa))
    4. pmtt-4  (((195 2) MPa))
    5. pmtt-5  (((270 3) MPa))
    6. pmtt-6  (((225 2) MPa))
    7. pmtt-7  (((175 2) MPa))
    8. pmtt-idsurf ((NIL unknown))

And the winner is ... 1

* Select property measurement 2 :
```

```
    1. pmtt-1  (((285 3) MPa))
    2. pmtt-2  (((210 2) MPa))
    3. pmtt-3  (((275 3) MPa))
    4. pmtt-4  (((195 2) MPa))
    5. pmtt-5  (((270 3) MPa))
    6. pmtt-6  (((225 2) MPa))
    7. pmtt-7  (((175 2) MPa))
    8. pmtt-idsurf ((NIL unknown))

And the winner is ... 2

Simulate experimental systems


Variables in TENSION-LOAD-STRESS likely to chatter are NONE.
Run time: 0.060 seconds to initialize S-0.
Variables in TENSION-LOAD-STRESS likely to chatter are NONE.
Run time: 0.060 seconds to simulate 4 states.
Variables in TENSION-LOAD-STRESS-ECC likely to chatter are NONE.
Run time: 0.050 seconds to initialize S-3.
Variables in TENSION-LOAD-STRESS-ECC likely to chatter are NONE.
Run time: 0.090 seconds to simulate 4 states.
Variables in TENSION-LOAD-STRESS-SURF likely to chatter are NONE.
Run time: 0.100 seconds to initialize S-6.
Variables in TENSION-LOAD-STRESS-SURF likely to chatter are NONE.
Run time: 0.070 seconds to simulate 4 states.
Variables in TENSION-LOAD-STRESS-ECCSURF likely to chatter are NONE.
Run time: 0.060 seconds to initialize S-9.
Variables in TENSION-LOAD-STRESS-ECCSURF likely to chatter are NONE.
Run time: 0.080 seconds to simulate 4 states.

Eliminate behaviors of experimental systems

Compare experimental systems


...
...

Model system 1 : SIMULATE-TENSION-TEST
Behavior system 1 :  (S-0 S-1 S-2)
Model system 2 : SIMULATE-TENSION-TEST-ECCSURF
Behavior system 2 : (S-9 S-10 S-11)


* Select initial relative values (Var/RV pairs) :
```

```
   1. P0 NIL
    2. P1 ((D =) (L =) (L0 =) (QUOCR =) (LR =) (E =) (GAM =) (A0 =))
    3. P2 ((D =) (L =) (L0 =) (QUOCR +) (LR =) (E =) (GAM =) (A0 =))
    4. P3 ((D =) (L =) (L0 =) (QUOCR -) (LR =) (E =) (GAM =) (A0 =))
    5. P4 ((D =) (L =) (L0 =) (LR =) (E =) (GAM =) (A0 =))
    6. P5 ((D =) (L =) (L0 =) (QUOCR =) (LR -) (E =) (GAM =) (A0 =))
    7. P6 ((D +) (L =) (L0 =) (QUOCR =) (LR =) (E =) (GAM =) (A0 =))
    8. P7 ((D =) (L =) (L0 =) (LR =) (E =) (GAM =) (A0 =))
    9. P8 ((D =) (L =) (L0 =) (QUOCR =) (LR =) (E =) (GAM =) (A0 =))
    10. P9 ((D =) (EPS -) (L0 =) (LR =) (E =) (GAM =) (A0 =) (SA -))
    11. P10 ((D =) (EPS -) (L0 =) (LR =) (E =) (GAM =) (A0 =) (SA -) (EC =))
    12. P11 ((D =) (EPS -) (L0 =) (QUOCR +) (LR =) (E =) (GAM =) (A0 =) (SA -))
    13. P12 ((D =) (L =) (L0 =) (QUOCR =) (LR =) (E +) (GAM =) (A0 =))
    14. P13 ((D =) (L =) (L0 =) (QUOCR +) (LR =) (E +) (GAM =) (A0 =))
    15. P14 ((D =) (EPS =) (L0 =) (LR =) (E =) (GAM =) (A0 =) (SA =))
    16. P15 ((D =) (EPS =) (L0 =) (QUOCR +) (LR =) (E =) (GAM =) (A0 =) (SA =))
    17. P16 ((D =) (EPS +) (L0 =) (LR =) (E =) (GAM =) (A0 =) (SA +))
    18. P17 ((D =) (EPS +) (L0 =) (QUOCR +) (LR =) (E =) (GAM =) (A0 =) (SA +))
    19. P18 ((D =) (L0 =) (LR =) (E =) (GAM =) (A0 =))
    20. P19 ((D =) (L0 =) (LR =) (E +) (GAM =) (A0 =))
    21. P20 ((D =) (L0 =) (LR =) (E =) (GAM =) (A0 =) (QUOCR +))
    22. P21 ((D =) (L0 =) (LR =) (E +) (GAM =) (A0 =) (QUOCR +))
    23. NP ((D =) (L0 =) (LR =) (E =) (GAM =) (A0 =) (QUOCR +) (EPS -) (SA -))

And the winner is ... 21

* Select initial relative values (RV rules) :

    1. IR0 NIL

And the winner is ... 1

Initial relative values (var/RV pairs)...
    <D 0>
    <L0 0>
    <LR 0>
    <E 0>
    <GAM 0>
    <A0 0>
    <QUOCR INF>
    <EPS MINF>
    <SA MINF>

Initial relative values (RV rules...
* Generating OPC
* Behaviour first system : (S-0 S-1 S-2)
```

```
* Behaviour second system: (S-9 S-10 S-11)

* Graph representing the Ordered Pairs of Comparison:
*  Root: (PC-2)
*  Leaf: (PC-3)
*  Nodes (2):
      (PC-3 NIL (PC-2))
      (PC-2 (PC-3) NIL)

Do you want to cheat? [Y/N] n
**Collecting constraints for PC-3 with rules at one PC
**Collecting constraints for PC-2 with rules at one PC
**Collecting constraints for PC-2 with rules at an interval between two PCs
**Overview of the relative values of variables

===================================================================+
           |SMCT| SA | SC | STH| EPS|  E |  F | LR |  D |  A |  L |
===================================================================+
CS-PC-1-0  |  = |  - |  - |  = |  - |  = |  - |  = |  = |  = |  - |
CS-PC-0-0  |  = |  = |  = |  = |  = |  = |  = |  = |  = |  = |  = |
===================================================================+


===================================================================+
           |LO  |DIF.|DIF.| NUM|QUO.|SQR.|SCS.| GAM| AO |GAME|GAM. |
===================================================================+
CS-PC-1-0  |  = |  = |  - |  + |  + |  + |  - |  = |  = |  = |  = |
CS-PC-0-0  |  = |  = |  = |  + |  + |  + |  = |  = |  = |  = |  = |
===================================================================+


* Graph representing the Comparative Envisionment
*  Root: (CS-PC-3-0)
*  Leaf: (CS-PC-2-0)
*  Nodes (2):
      (CS-PC-3-0 (EDGE-1) NIL)
      (CS-PC-2-0 NIL (EDGE-1))
*  Edges (1):
      (EDGE-1 (CS-PC-2-0) (CS-PC-3-0))
*  Labels (1):
      (EDGE-1 ((RV-TIME-PC-2-PC-3 (MINF IGN))))
* Paths in graph: (1):
    (CS-PC-3-0 ((RV-TIME-PC-2-PC-3 (MINF IGN))) CS-PC-2-0)
```

. . .

. . .

The output of KIMA in the above table is readily seen to be equal with the CEC* output in the previous section and figure 7.7.

References

H.M. Collins (ed.) [1981], "Knowledge and Controversy: Studies of Modern Natural Science," special issue of *Social Studies of Science* 11, 3–158.

S. Addanki, R. Cremonini & J.S. Penberthy [1989], "Reasoning about assumptions in graphs of models," in *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence, IJCAI-89*, N.S. Sridharan, ed., Morgan Kaufmann, San Mateo, CA, 1432–1438.

L.K. Alberts [1993], "YMIR: An ontology for engineering design," University of Twente, Ph.D. thesis, Enschede, the Netherlands.

F.I. Baratta [1984], "Requirements for flexure testing of brittle materials," in *Methods for Assessing the Structural Reliability of Brittle Materials, ASTM STP 844*, S.W. Freiman & C.M. Hudson, eds., American Society for Testing and Materials, Philadelphia, 194–222.

B.A. Barry [1978], *Errors in Practical Measurement in Science, Engineering, and Technology*, John Wiley & Sons, New York.

D. Berleant & B. Kuipers [1997], "Qualitative and quantitative simulation: Bridging the gap," *Artificial Intelligence* 95, 215–256.

P. Borst, H. Akkermans & J. Top [1997], "Engineering ontologies," *International Journal of Human-Computer Studies* 46, 365–406.

L.M. Bouter [1994], *Meta-Analyse: Controleerbaar en Reproduceerbaar Literatuuronderzoek als Basis voor Rationele Beslissingen in de Gezondheidszorg*, Amsterdam University Press, Amsterdam.

W.E. Boyce & R.C. DiPrima [1992], *Elementary Differential Equations and Boundary Value Problems*, John Wiley & Sons, New York, 5th ed.

C.-T. Chen [1970], *Introduction to Linear System Theory*, Holt, Rinehart and Winston, New York.

C. Chiu & B. Kuipers [1992], "Comparative analysis and qualitative integral representations," in *Recent Advances in Qualitative Physics*, B. Faltings & P. Struss, eds., MIT Press, Cambridge, MA, 313–328.

D.J. Clancy & B. Kuipers [1997], "Static and dynamic abstraction solves the problem of chatter in qualitative simulation," in *Proceedings of Fourteenth National Conference on Artificial Intelligence, AAAI-97*, AAAI Press/MIT Press, Menlo Park, CA/Cambridge, MA , 125–131.

H.M. Collins [1992], *Changing Order: Replication and Induction in Scientific Practice*, University of Chicago Press, Chicago, 2nd edition.

R. Courant [1959], *Differential and Integral Calculus,* 2 vols, Blackie & Son, London.

T.H. Courtney [1990], *Mechanical Behavior of Materials*, McGraw-Hill, New York.

J.P. Coutures & M.H. Rand [1989], "Melting temperatures of refractory oxides: Part II, lanthanoid sesquioxides," *Pure & Applied Chemistry* 61, 1461–1482.

W.E.C. Creyke, I.E.J. Sainsbury & R. Morrell [1982], *Design with Non-Ductile Materials*, Applied Science Publishers, London.

L. Darden [1991], *Theory Change in Science, Strategies from Mendelian Genetics*, Oxford University Press, New York.

R.W. Davidge [1979], *Mechanical Behaviour of Ceramics*, Cambridge University Press, Cambridge.

R. Davis & W.C. Hamscher [1988], "Model-based reasoning: troubleshooting," in *Exploring Artificial Intelligence: Survey Talks from the National Conferences on Artificial Intelligence*, H.E. Shrobe, ed., Morgan Kaufmann, San Mateo, CA, 297–346.

D. DeCoste [1991], "Dynamic across-time measurement interpretation," *Artificial Intelligence* 51, 273–341.

E. Dörre & H. Hübner [1984], *Alumina: Processing, Properties, and Applications*, Springer-Verlag, Berlin.

D. Dvorak & B. Kuipers [1991], "Process monitoring and diagnosis, a model-based approach," *IEEE Expert* 5, 67–74.

B. Falkenhainer, A. Farquhar, D. Bobrow, R. Fikes, K. Forbus, T. Gruber, Y. Iwasaki & B. Kuipers [1994], "CML: A compositional modeling language," Knowledge Systems Laboratory, Stanford University, Technical report KSL-94-14, Stanford.

B. Falkenhainer & K.D. Forbus [1991], "Compositional modeling: Finding the right model for the job," *Artificial Intelligence* 51, 95–143.

A. Farquhar [1994], "A Qualitative Physics Compiler," in *Proceedings of the 12th National Conference on Artificial Intelligence, AAAI-94*, AAAI Press/MIT Press, Menlo Park, CA/Cambridge, MA, 1168–1174.

A. Farquhar, B. Kuipers, J. Rickel, D. Throop & Qualitative Reasoning Group [1993], "QSIM: The program and its use," University of Texas, Technical report UT-AI-90-123, Austin, TX.

U. Fayyad, D. Haussler & P. Stolorz [1996], "Mining scientific data," *Communications of the ACM* 39, 51–57.

K.D. Forbus [1984], "Qualitative process theory," *Artificial Intelligence* 24, 85–168.

K.D. Forbus [1987], "Interpreting observations of physical systems," *IEEE Transactions on System, Man, and Cybernetics* 13, 350–359.

K.D. Forbus [1988], "Qualitative physics: Past, present, and future," in *Exploring Artificial Intelligence*, H. Shrobe, ed., Morgan Kaufmann, San Mateo, CA, 239–296.

K.D. Forbus [1989], "QPE: A study in assumption-based truth maintenance," *International Journal of Artificial Intelligence in Engineering* 3, 200–215.

K.D. Forbus & B. Falkenhainer [1992], "Self-exploratory simulations: Integrating qualitative and quantitative knowledge," in *Recent Advances in Qualitative Physics*, B. Faltings & P. Struss, eds., MIT Press, Cambridge, MA, 49–65.

P. Fouché & B. Kuipers [1992], "Reasoning about energy in qualitative simulation," *IEEE Transactions on Man, Machine, and Cybernetics* 22, 47–63.

A. Franklin [1986], *The Neglect of Experiment*, Cambridge University Press, Cambridge.

P. Galison [1987], *How Experiments End*, University of Chicago Press, Chicago.

P. Gärdenfors [1988], *Knowledge in Flux, Modeling the Dynamics of Epistemic States*, MIT Press, Cambridge, MA.

D. Gooding, T. Pinch & S. Schaffer [1989], *The Uses of Experiment: Studies in the Natural Sciences*, Cambridge University Press, Cambridge.

J.R. Griesemer [1992], "The role of instruments in the generative analysis of science," in *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences*, A.E. Clarke & J.H. Fujimura, eds., Princeton University Press, Princeton, NJ, 47–76.

T.R. Gruber & G.R. Olsen [1994], "An ontology for engineering mathematics," in *Proceedings of the Fourth International Conference on Principles of Knowledge Representation and Reasoning, KR-94*, J. Doyle, P. Torasso & E. Sandewall, eds., Morgan Kaufmann, San Francisco, CA, 258–269.

U.G. Gupta, ed. [1991], *Validating and Verifying Knowledge-Based Systems*, IEEE Computer Press, Los Alamitos, CA.

I. Hacking [1983], *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*, Cambridge University Press, Cambridge.

I. Hacking [1992], "The self-vindication of the laboratory sciences," in *Science as Practice and Culture*, A. Pickering, ed., University of Chicago Press, Chicago.

W. Hamscher, L. Console & J. de Kleer, eds. [1992], *Readings in Model-based Diagnosis*, Morgan Kaufmann, San Mateo, CA.

C.G. Hempel [1966], *Philosophy of Natural Science*, Prentice Hall, Englewood Cliffs, NJ.

J. Hlaváč [1982], "Melting temperatures of refractory oxides," *Pure & Applied Chemistry* 54, 681–688.

R.G. Hoagland, C.W. Marschall & W.H. Duckworth [1976], "Reduction of errors in ceramic bend tests," *Journal of the American Ceramic Society* 59, 189–192.

J.E. Hunter & F.L. Schmidt [1990], *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*, Sage Publications, Newbury Park.

Y. Ishida [1989], "Using global properties for qualitative reasoning: A qualitative system theory," in *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence, IJCAI-89*, N.S. Sridharan, ed., Morgan Kaufmann, San Mateo, CA, 1174–1179.

Y. Iwasaki, A. Farquhar, R. Fikes & J. Rice [1997], "A web-based compositional modeling system for sharing of physical knowledge," in *Proceedings of the 15th International Joint Conference on Artificial Intelligence, IJCAI-97*, M.E. Pollack, ed., Morgan Kaufmann, San Francisco, CA, 494–500.

Y. Iwasaki & A.Y. Levy [1994], "Automated model selection for simulation," in *Proceedings of the 12th National Conference on Artificial Intelligence, AAAI-94*, AAAI Press/MIT Press, Menlo Park, CA/Cambridge, MA, 1183–1190.

H. de Jong [1992], "Inconsistency resolution in knowledge integration," University of Twente, Memorandum UT-KBS-92-38, Enschede, the Netherlands, M.Sc. thesis.

H. de Jong [1996], "Comparative Envisionment Construction: A Method for the Comparative Analysis of Dynamical Systems," University of Twente, Memorandum UT-KBS-96-14, Enschede, the Netherlands.

H. de Jong, N.J.I. Mars & P.E. van der Vet [1993], "Consistency of scientific knowledge bases," in *Working notes of the AAAI-93 Workshop on Validation and Verification of Knowledge-Based Systems*, A.D. Preece, K. Bellman, J. Combs, C. Culbert, D. O'Leary, L. Miller, N. Sizemore & R. Stachowitz, eds., AAAI Press, 64–65.

H. de Jong, N.J.I. Mars & P.E. van der Vet [1996], "CEC: Comparative analysis by envisionment construction," in *Proceedings of the 12th European Conference on Artificial Intelligence, ECAI-96*, W. Wahlster, ed., John Wiley & Sons, Chichester, 476–480.

H. de Jong, N.J.I. Mars & P.E. van der Vet [1998], "Computer-supported analysis of experiments in materials science," to be presented at the *International Congress on Discovery and Creativity*, Ghent.

H. de Jong & F. van Raalte [1997a], "Comparative analysis of structurally different dynamical systems," in *Proceedings of the 15th International Joint Conference on Artificial Intelligence, IJCAI-97*, M.E. Pollack, ed., Morgan Kaufmann, San Francisco, CA, 486–491.

H. de Jong & F. van Raalte [1997b], "Comparative envisionment construction: A general method for the comparative analysis of dynamical systems," Centre of Telematics and Information Technology, University of Twente, CTIT Technical Report 97-06, Enschede, the Netherlands.

H. de Jong & A. Rip [1995], "Ontdekkingspraktijk in plaats van logica: Wetenschappelijk ontdekken in computerondersteunde ontdekkingsomgevingen," *Kennis en Methode* XIX, 131–159.

H. de Jong & A. Rip [1997], "The computer revolution in science: Steps towards the realization of computer-supported discovery environments," *Artificial Intelligence* 91, 225–256.

H. de Jong, V.J. de Wit, N.J.I. Mars & P.E. van der Vet [1998], "Model-based identification of systematic errors in measurements," in *Working Notes of the 12th International Workshop on Qualitative Reasoning, QR-98*, AAAI Press, Menlo Park, CA, to appear.

J. Kalagnanam, H.A. Simon & Y. Iwasaki [1991], "The mathematical bases for qualitative reasoning," *IEEE Expert* 5, 11–19.

H. Kay & B. Kuipers [1993], "Numerical behavior envelopes for qualitative models," in *Proceedings of the 11th National Conference on Artificial Intelligence, AAAI-93*, AAAI Press/MIT Press, Menlo Park, CA/Cambridge, MA, 606–613.

W.D. Kingery, H.K. Bowen & D.R. Uhlmann [1976], *Introduction to Ceramics*, John Wiley & Sons, New York, second edition.

J. de Kleer [1992], "Qualitative physics," in *Encyclopedia of Artificial Intelligence*, S.C. Shapiro, ed. #2, John Wiley & Sons, New York, 1149–1159, 2nd edition.

J. de Kleer & J.S. Brown [1984], "A qualitative physics based on confluences," *Artificial Intelligence* 24, 7–83.

T.S. Kuhn [1970], *The Structure of Scientific Revolutions*, The University of Chicago Press, Chicago, 2nd edition.

T.S. Kuhn [1977], "The function of measurement in modern physical science," in *The Essential Tension: Selected Studies in Scientific Tradition and Change*, University of Chicago Press, Chicago, 178–224.

B. Kuipers [1986], "Qualitative simulation," *Artificial Intelligence* 29, 289–388.

B. Kuipers [1994], *Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge*, MIT Press, Cambridge, MA.

J.G. van der Laan & P.W. Bach [1987], "Mechanische beproeving van keramische materialen door middel van 3- en 4-punts buiging," The Netherlands Energy Research Foundation, Memo 87-117, Petten, the Netherlands.

B. Latour [1987], *Science in Action: How to Follow Scientists and Engineers Through Society*, Harvard University Press, Cambridge, MA.

B.R. Lawn & T.R. Wilshaw [1975], *Fracture of Brittle Solids*, Cambridge University Press, Cambridge.

W.W. Lee & B.J. Kuipers [1988], "Non-intersection of trajectories in qualitative phase space: A global constraint for qualitative simulation," in *Proceedings of the Seventh National Conference on Artificial Intelligence, AAAI-88*, Morgan Kaufmann, Los Altos, CA, 286–291.

R. Leitch & Q. Shen [1993], "Prioritizing behaviours in qualitative simulation," in *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence, IJCAI-93*, J. McDermott, ed., Morgan Kaufmann, San Mateo, CA, 1523–1528.

C.M. Low & Y. Iwasaki [1992], "Device modelling environment: an interactive environment for modelling device behaviour," *Intelligent Systems Engineering* 1.

A. Makarovic [1991], "Parsimony in model-based reasoning," University of Twente, Ph.D. thesis, Enschede, the Netherlands.

A. Malik & P. Struss [1996], "Diagnosis of dynamic systems does not necessarily require simulation," in *Working Notes of the Tenth International Workshop on Qualitative Reasoning, QR-96*.

N.J.I. Mars, H. de Jong, P.-H. Speel, W.G. ter Stal & P.E. van der Vet [1994], "Semi-automatic knowledge acquisition in Plinius: An engineering approach," in *Proceedings of the Eighth Banff Knowledge Acquisition for Knowledge-based Systems Workshop*, B. Gaines & M. Musen, eds., Banff, 4–1–4-15.

C.W. Marschall & A. Rudnick [1974], "Conventional strength testing," in *Fracture Mechanics of Ceramics. Volume 1: Concepts, Flaws, and Fractography*, R.C. Bradt, D.P.H. Hasselman & F.F. Lange, eds., Plenum Press, New York, 69–92.

P.P. Nayak [1995], *Automated Modeling of Physical Systems*, Springer Verlag, Berlin.

R. Neches, R. Fikes, T. Finin, T. Gruber, R. Patil, T. Senator & W.R. Swartout [1991], "Enabling technology for knowledge sharing," *AI Magazine* 12, 37–55.

M. Neitzke [1992], "Modeling physical systems with relative descriptions of parameters," in *Proceedings of the Tenth European Conference on Artificial Intelligence, ECAI-92*, B. Neumann, ed., John Wiley & Sons, Chichester, 683–684.

M. Neitzke [1997], *Relativsimulation von Störungen in technischen Systemen,* Dissertationen zur Künstlichen Intelligenz, Bd. 154, Infix, Sankt Augustin.

M. Neitzke & B. Neumann [1994], "Comparative simulation," in *Proceedings of the 12th National Conference on Artificial Intelligence, AAAI-94*, AAAI Press/MIT Press, Menlo Park, CA/Cambridge, MA, 1205–1210.

R.C. Newnham [1975], "Strength tests for brittle materials," *Proceedings of the British Ceramic Society* 25, 281–293.

T. Ohji [1988], "Towards routine tensile testing," *International Journal of High Technology Ceramics* 4, 211–225.

P. Polak [1979], *Systematic Errors in Engineering Physics*, Macmillan Press, London.

A. Pos, H. Akkermans & J. Top [1998], "Automated model revision," *IEEE Expert*.

G.D. Quinn [1991], "Strength and proof testing," in *Engineered Materials Handbook. Volume 4: Ceramics and Glasses*, S.J. Schneider et al., ed., ASM International, 585–598.

F. van Raalte & H. de Jong [1997], "Comparative envisionment construction: Implementation, version 1," University of Twente, Memorandum UT-KBS-97-02, Enschede, the Netherlands.

H. Radder [1984], *De Materiële Realisering van Wetenschap: Een Filosofiese Visie op de Experimentele Natuurwetenschappen, Ontwikkeld in Discussie met Habermas*, VU-uitgeverij, Amsterdam.

O. Raiman [1986], “Order of magnitude reasoning,” in *Proceedings of the Fifth National Conference on Artificial Intelligence, AAAI-86*, Morgan Kaufmann, San Mateo, CA, 100–104.

J.K Ravetz [1996], *Scientific Knowledge and its Social Problems*, Transaction Publishers, New Brunswick, NJ.

P.J. Rousseeuw & A.M. Leroy [1987], *Robust Regression and Outlier Detection*, John Wiley, New York.

W.J. Rugh [1996], *Linear System Theory*, Prentice Hall, Englewood Cliffs, NJ, 2nd edition.

E. Sacks [1990a], “A dynamic systems perspective on qualitative simulation,” *Artificial Intelligence* 42, 349–362.

E. Sacks [1990b], “Automatic qualitative analysis of dynamic systems using piecewise linear approximations,” *Artificial Intelligence* 41, 313–364.

C. Schut & B. Bredeweg [1996], “An overview of approaches to qualitative model construction,” *Knowledge Engineering Review* 11, 1–25.

J. Shrager & P. Langley, eds. [1991], *Computational Models of Scientific Discovery and Theory Formation*, Morgan Kaufmann, San Mateo, CA.

D.J. deSolla Price [1984], “Of sealing wax and string,” *Natural History* 93, 48–56.

P.-H. Speel [1995], “Selecting knowledge representation systems,” University of Twente, Ph.D. thesis, Enschede, the Netherlands.

G.L. Squires [1985], *Practical Physics*, Cambridge University Press, Cambridge, 3rd edition.

W.G. ter Stal [1996], “Automated interpretation of nominal compounds in a technical domain,” University of Twente, Ph.D. thesis, Enschede, the Netherlands.

P. Struss [1988a], “Global filters of qualitative behaviors,” in *Proceedings of the Seventh National Conference on Artificial Intelligence, AAAI-88*, Morgan Kaufmann, Los Altos, CA, 275–280.

P. Struss [1988b], “Mathematical aspects of qualitative reasoning,” *Artificial Intelligence in Engineering* 3, 156–169.

P. Thagard [1988], *Computational Philosophy of Science*, MIT Press, Cambridge, MA.

E.R. Tufte [1983], *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, CN.

J.W. Tukey [1977], *Exploratory Data Analysis*, Addison-Wesley, Reading, MA.

P.E. van der Vet & N.J.I. Mars [1993], “Structured system of concepts for storing, retrieving, and manipulating chemical information,” *Journal of Chemical Information and Computer Sciences* 33, 564–568.

P.E. van der Vet & N.J.I. Mars [1998], “Bottom-up construction of ontologies,” *IEEE Transactions on Knowledge and Data Engineering*, to appear.

P.E. van der Vet, P.-H. Speel, W.G. ter Stal, H. de Jong, F. van Raalte & N.J.I. Mars [1994], "Plinius intermediate report," University of Twente, Memorandum UT-KBS-94-35, Enschede, the Netherlands.

J.B. Wachtman [1996], *Mechanical Properties of Ceramics*, John Wiley & Sons, New York.

D.S. Weld [1989], "Automated model switching: Discrepancy driven selection of approximation reformulations," University of Washington, Technical Report 89-08-01, Seattle.

D.S. Weld [1990], *Theories of Comparative Analysis*, MIT Press, Cambridge, MA.

D.S. Weld [1992], "Reasoning about model accuracy," *Artificial Intelligence* 56, 255–300.

D.S. Weld & J. de Kleer, eds. [1990], *Readings in Qualitative Reasoning about Physical Systems*, Morgan Kaufmann, San Mateo, CA.

M.N. Wise, ed. [1995], *The Values of Precision*, Princeton University Press, Princeton, NJ.

V.J. de Wit & H. de Jong [1998], "Implementatie van KIMA," University of Twente, Memorandum UT-KBS-98-06, Enschede, the Netherlands.

S. Xia & N. Smith [1996], "Automated modelling: A discussion and review," *Knowledge Engineering Review* 11, 137–160.

Index

Curriculum Vitae

Hidde de Jong was born on July 7th 1968 in Delft, in the Netherlands. He completed secondary school, VWO, at the Carmel Lyceum in Oldenzaal, after which he went to study at the University of Twente in Enschede. Between 1986 and 1994 he obtained M.Sc. degrees in Computer Science, Management Science, and Philosophy of Science, Technology, and Society, all three of them *cum laude*.

In 1994 he started to work as a Ph.D. student in the Knowledge-Based Systems group, headed by Professor N.J.I. Mars, at the faculty of Computer Science of the University of Twente. His research focussed on the development and use of qualitative reasoning techniques for the model-based analysis of measurements in science and engineering. Apart from that, he maintained a strong interest in the potential role of computers in making scientific discoveries.

He hopes to continue research on advanced computational techniques for computer-supported scientific discovery.