A source modelling system and its use for uncertainty management

Bokma, Albert Franz

A Source Modelling System and its Use for Uncertainty Management

Albert Franz Bokma

Abstract

Human agents have to deal with a considerable amount of information from their environment and are also continuously faced with the need to take actions. As that information is largely of an uncertain nature, human agents have to decide whether, or how much, to believe individual pieces of information. To enable a reasoning system to deal in general with the demands of a real environment, and with information from human sources in particular, requires tools for uncertainty management and belief formation. This thesis presents a model for the management of uncertain information from human sources. Dealing, more specifically, with information which has been pre-processed by a natural language processor and transformed into an event-based representation, the model assesses information, forms beliefs and resolves conflicts between them in order to maintain a consistent world model. The approach is built on the fundamental principle that the uncertainty of information from people can, in the majority of situations, successfully be assessed through source models which record factors concerning the source's abilities and trustworthiness. These models are adjusted to reflect changes in the behaviour of the source. A mechanism is presented together with the underlying principles to reproduce such a behaviour. A high-level design is also given to make the proposed model reconstructible, and the successful operation of the model is demonstrated on two detailed examples.

Abstract

Human agents have to deal with a considerable amount of information from their environment and are also continuously faced with the need to take actions. As that information is largely of an uncertain nature, human agents have to decide whether, or how much, to believe individual pieces of information. To enable a reasoning system to deal in general with the demands of a real environment, and with information from human sources in particular, requires tools for uncertainty management and belief formation. This thesis presents a model for the management of uncertain information from human sources. Dealing, more specifically, with information which has been pre-processed by a natural language processor and transformed into an event-based representation, the model assesses information, forms beliefs and resolves conflicts between them in order to maintain a consistent world model. The approach is built on the fundamental principle that the uncertainty of information from people can, in the majority of situations, successfully be assessed through source models which record factors concerning the source's abilities and trustworthiness. These models are adjusted to reflect changes in the behaviour of the source. A mechanism is presented together with the underlying principles to reproduce such a behaviour. A high-level design is also given to make the proposed model reconstructible, and the successful operation of the model is demonstrated on two detailed examples.

A Source Modelling System and its Use for Uncertainty Management

Albert Franz Bokma

March 1993

Submitted for the Degree of Doctor of Philosophy

School of Engineering and Computer Science

University of Durham

Supervisor: Dr. Roberto Garigliano

To my parents

*"Die Vergangenheit dient dazu,
der Zukunft besser gerecht werden zu können."*

Richard von Weizsäcker
2 September 1990

Acknowledgements

My foremost thanks go to my parents who have so generously supported me in many ways and without whom all this would not have been possible.

I am also very grateful to the Division of Computer Science, SECS, for making my stay so pleasant. My special thanks go to my supervisor, Roberto Garigliano, who has been a great source of advice and support. It has been very enjoyable to work with him, and he often kept me on course, when I was about to lose orientation and had enough foresight to persuade me that the end was in sight, when I still could not quite make out the light at the end of the tunnel.

There are many who have supported me over the years in my studies and I am grateful for all of them. The work has been very rewarding, but also long and arduous, especially towards the end, and the encouragement they have given has been much appreciated. My thanks therefore go to Simon, Kevin, Ashley and especially Russell who kindly volunteered to proof-read the thesis.

The support of the Science and Engineering Research Council (AY 1989-90) is also gratefully acknowledged.

All work presented in this thesis is the author's and the views expressed do not necessarily reflect those of the University of Durham or SERC.

Table of Contents

Overview

In our every-day experience we continually have to deal with pieces of information from human sources, and most of that information is of an uncertain nature. In this thesis a model for the management of information of that kind is presented and whose fundamental purpose is to assign beliefs.

There are a number of models and systems presently available, which are able to assess the impact of propagation of uncertain information on systems of belief, once initial levels of uncertainty are given. The current field is characterised by a considerable number of approaches using uncertain information, while there is a distinct lack of approaches to assess these initial levels of uncertainty, needed to start that process. The emphasis of this model is on assessing initial levels of uncertainty and assigning beliefs as opposed to the propagation of uncertainty during reasoning.

In contrast to many approaches which deal with specialised domains and situations, the purpose of our model is to be domain-independent and to assess information with the help of a principled system of general heuristics.

Our model is based on the principle that the believability of uncertain information can be assessed to a great extent by considering the source of information and the context in which the information was given. To this end, the system maintains models about the sources it is acquainted with, which are then used in the belief formation process. There is also an important feedback loop between the information and the source models, to allow them to be adapted to reflect changes in the behaviour of sources.

There is also a classification mechanism in operation, whereby sources are seen as members of a particular class. As a consequence, a lot of useful information about the properties of a source can be inferred from the typical properties of that class. These properties can be used as defaults in the absence of more concrete evidence.

Another major approach to uncertainty management from human sources, assesses the uncertainty of information through a qualitative analysis of the arguments, or endorsements, that are given in their support. This approach appears to be complementary with our proposed model for a number of reasons:

- the endorsement approach is particularly suited for situations where decisions have to be based on the arguments. While being precise and elaborate, the process is difficult and costly and should be employed in those situations where there is time and when the situation is sufficiently important.
- the source control approach is particularly suited for situations which are not decided on the basis of the weight of argument, but by source considerations. It should be able to deliver results more quickly although sometimes less accurate and therefore suited for situations where precision is not the primary concern, or where a fast approximation is better than a precise answer which comes too late.

The source control model is best suited to operate with input pre-processed by a natural language analyser. To function optimally, there needs to be a close cooperation between the source control model and the natural language analyser. While the natural language analyser needs to know whether the information it is given is believable, the source control model needs information about the context of the information and the sources involved to correctly assess its credibility. An experimental natural language system has been built and can be described as a conversational system [LOL92]. It is capable of syntactic, semantic, pragmatic and dialogue analyses and uses a conceptual graph as its memory. This system is capable of supporting the source control model.

There are a number of objectives in our approach:

1) to build a model which is general and not domain-specific and to make the model applicable to information from human sources in realistic situations.
2) to make the system pragmatic, concentrating on making the model practical and useful. To forego precision which cannot be achieved in this large domain or which is not necessary for the purpose.
3) to make the approach practical, to give a design and thus making the model precise, traceable, reconstructible as well as implementable.

In the process of abstracting the heuristics we have chosen a wide and varied set of examples and looked for common ways of assessing the uncertainty in them and the way beliefs tend to be formed. In the process we have identified a set of general heuristics which appear to be domain independent and which do not change significantly when applied to different examples.

To give a simplified example, consider the situation where John, a friend of the system, tells the system that he needs to buy a motorbike to go to work with. He says that he has had a look at a motor-bike at a local second-hand motorbike dealer who recommended one of his motorbikes to him as being sound and cheap. When he went to the dealer his friend Paul came along as well and said that the motorbike the dealer pointed out was looking good. John is however still unsure and is asking the system for help.

The information ip1 from dealer1 can be described as:

```
ip1:
 object: Motor-bike1 is sound and cheap
 source: dealer1
 certainty: high

dealer1:
 ability:
  expertise: motor_mechanics=high
  reasoning: average
  judging sources: average
 interests: selling motor-bikes
```

dealer1 -> clients(John)
helpfulness: high
trustworthiness: low

The dealer strongly recommends the motorbike as being a good buy. He definitely has the expertise to judge the quality of motor-bikes but there is a problem. He has an interest in selling motor-bikes. Also, when the dealer is dealing with clients he is not considered to be very trustworthy. John is in the role of a potential client and in that case the dealer cannot be trusted. Consequently the system fundamentally cannot believe the information, unless there is a good reason to convince the system that the dealer is giving truthful information. In this case no such indication is given.

Alternatively, John could go by what his friend Paul says, who is cautious in giving a recommendation and says that he thinks that the motor-bike maybe a good buy.

ip2:
object: Motor-bike1 looks good and may be a good buy
source: Paul
certainty: medium

Paul:
ability:
expertise: none(default)
reasoning: average
judging sources: average
interests: none(default)

Paul -> friends(John)
helpfulness: high
trustworthiness: high

The system can create the link to connect John to being a friend of Paul and find that Paul is trustworthy to his friends as well as generally helpful. Paul's cautious approach signals to the system that Paul will probably be disinclined to accept responsibility for his recommendation and when the system examines Paul's abilities it finds that Paul does not appear to have any expertise. Although Paul can be trusted he is lacking the necessary expertise to give a reliable recommendation. This is reflected in his reluctance and cautious attitude.

Although the system can believe the information the level of belief is only medium.

Unfortunately, neither piece of information is strong enough to help John with his problem and the system might ask John whether he knows anybody who can give reliable expert advice. John may then respond that the RAC does valuations for a fee and that they are trustworthy. Having had the valuation on the motorbike John then comes back with the following result:

```
ip3:
 object: Motor-bike1 is a good buy
 source: RAC valuer
 certainty: high

RAC Valuer:
 ability:
  expertise: motor_mechanics
  reasoning: average
  judging sources: average
 interests: fees, reputation

 RAC -> clients(John)
  helpfulness: high
  trustworthiness: high
```

The system will find that the valuer of the RAC is highly trustworthy and helpful, as well as having the necessary expertise. As the valuer also considers the motorbike to be a good buy the system can now safely recommend to John to follow the RAC's advice.

This gives a brief overview of the criteria employed by the system in its analysis of information. This should also give an idea of the basic premiss of source control, namely that source considerations enter and influence the analysis process.

Chapter 1

Methodological Introduction

1.1 Introduction

In the present thesis we are concerned with a model for the management of uncertainty of information from human sources. As human beings are faced with a considerable amount of information of uncertain nature and with the need to form beliefs about the information, so the purpose of our model is to deal with a situation of that kind.

Before we embark on the more technical treatment of the proposed model in the chapters to follow, we first need to clarify some methodological issues with respect to our position in the current field of research in AI and the limits of our approach. This is followed by a section to explain controversial terminology and a section on how to read the thesis.

1.2 The Dilemma of AI

The gap between the sheer complexity and sophistication of human behaviour and the inadequacies of current tools to describe and model them has characterised research in AI since its beginnings. As a consequence of the realisation of the enormity of the task and the desire to introduce scientific, principled approaches, researchers have typically concentrated their efforts, by restricting the domain, or by simplifying the problem, or both. This is reflected in the development of research into two different directions:

- research in formal approaches
- research in heuristic approaches

1.2.1 Research in Formal Approaches

Much effort has been invested in various fields, to produce a formal basis in which results can be proven and properties demonstrated as well as providing an excellent tool for comparison between rival theorems. This task has already proven to be extremely difficult and cumbersome for small problems. Despite their obvious results formal approaches seem to be until now unsatisfactory for two reasons:

- lack of wider applicability
- doubts about their usefulness

Despite their exactness, provability and generality, formal theories tend to apply only to very simplified situations. In order to manage the considerable task of formalisation, formal theories have had to strip their domain of much detail, by restricting the parameters and making simplifying assumptions, thus leaving only the bare core. Considering the domain we are interested in, it is very doubtful if such models could cope with the added complexity.

Even if a formal theory can be produced for 'real-world' situations it is uncertain whether it would be particularly useful in practical application. It is likely that a formal model for more realistic scenarios, if feasible, would be unwieldy in its use, very demanding in space and time, and deliver results which would be more precise than is necessary for our purposes. Formalisations have been very useful for more abstract situations but there are usually difficulties in applying them to common, everyday situations where practical approaches have worked rather more successfully.

1.2.2 Research in Heuristic Approaches

The other school of thought has taken a pragmatic approach, by starting from more realistic scenarios to explore the heuristics employed by humans. Due to the size and complexity of the problem, the researchers have often concentrated on more specific examples, in the hope that the lessons learned from them could be used in a wider field. However, there are also difficulties involved with this approach:

- problem of specificity
- tendency towards ad hoc heuristics
- difficulty with their reconstruction

Using examples which are very specific and follow a common pattern, there is a danger that the examples are chosen to fit the model rather than the model to fit the examples. This would throw serious doubt on the generality of the solutions.

Another problem with this approach is that the heuristics might be fitting only the particularities of the problem. We are interested in results which can be used successfully in 'real-world' situations and we suspect that many heuristics lack the power to cope with these situations. Care has to be taken not to end up with heuristics which are *ad hoc* and not principled enough to be applicable to wider domains.

Unfortunately, in many cases where more general claims are made and more complex models are presented, the designs of the solutions referred to are usually not given. It is therefore impossible to reconstruct a clear and complete set of requirements and design by which the results can be shown to be repeatable.

Somehow, this often leaves the suspicion that the claims are more extravagant than the reality.

There is also a third group of researchers who have taken a position in between the two extremes. Faced with the unfeasibility of formalising a complex class of problems and the tendency towards ad-hoc-ness, a more general and principled approach to the design of heuristics systems has been taken, in the attempt to avoid the inherent problems of these opposite positions.

1.3 The Chosen Method

We are interested in a rich, 'real-world' domain which requires general heuristics and which is too complex, we believe, to be amenable to a suitable formalisation, at least for the time being. For these reasons we aim for the middle ground, between the two extreme positions:

When selecting a class of problems, it is important that the choice is effected by virtue of an independent and external principle and not according to whether the elements will fit the particular model. For our present task we choose the class of *everyday information from human sources*, which is both natural and general. We choose a large domain and start with a wide and varied set of common, everyday examples and try to build a theory which has the explanatory power to describe and analyse the phenomena in terms of general heuristics.

Some simplifications are still unavoidable: we examine only *literal meaning input*, i.e. text without metaphors, humour, emotional undertones or strong contextual meaning. We expect that a natural language processor has eliminated these in the pre-processing phase and that the information, which is the input to the envisaged source control system, can be taken to be literal. This considerably reduces the range of input considered, but leaves in our domain the kind of information which is more relevant under a belief analysis point of view.

As it is impossible to test our model on all elements of this potentially unlimited class, we have attempted to choose examples from a wide selection in order to demonstrate that our model can cope with them. Consequently, we hope that we have preserved a strong element of generality in them. Furthermore, the heuristics employed do not change significantly when applied across these examples, thus reinforcing our confidence in their generality.

It is not feasible to formalise the model in its present complexity, if one wants to preserve its ability to deal with the class of situations we are interested in. That is, we are not attempting to prove properties such as consistency or complexity, nor do we give axioms to describe the system and derive theorems from them. Nevertheless, we do give a complete design which makes the model precise and implementable, and shows it to be principled. This also has the advantage that the claimed results can be reproduced.

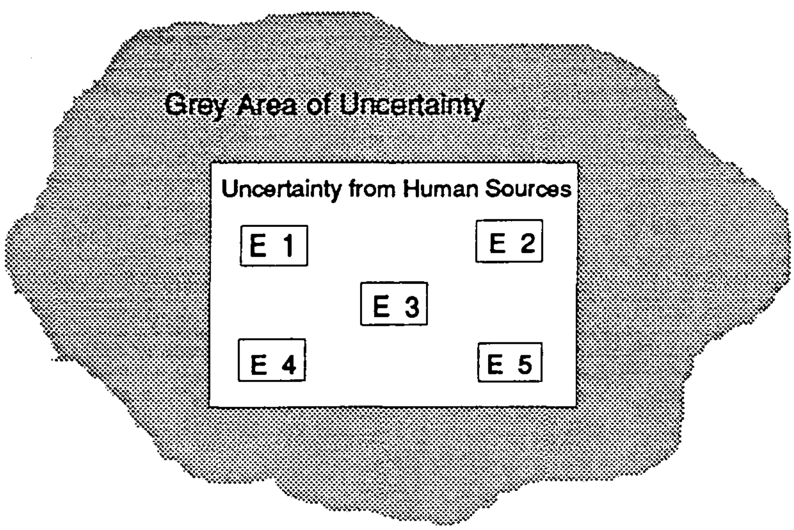


Figure 1: Grey Area of Uncertainty

1.3.1 Criteria For Success

It is important to point out that in AI we introduce a model which must simulate a behaviour, but not the mechanism by which this behaviour was originally achieved. As recently defined, "Artificial Intelligence is the field of research concerned with making machines perform tasks which are generally thought of as requiring human intelligence" [BEA89]. This means that we are not primarily concerned with copying the mechanics, as psychologists might be, but by reproducing the behaviour. In our case this means the behaviour of managing uncertain information, as exemplified by the specific instances which we describe in Chapter 2.

The important point is that the model is copying a behaviour of dealing with information. When we lay claims to a behaviour and that behaviour is objectively defined, then the successful proof is dependent on establishing that we can reproduce that behaviour. Thus, if there is a man X who does Y and we claim that we have a model which will behave like that then the proof is in demonstrating that this can be achieved. It should be noted that we do not lay claim to a reasonable model modelling a reasonable man with a reasonable behaviour, but that there is a claimed response which can be demonstrated to be reproducible by the model. As a matter of fact, we do consider the behaviour shown in our example to be 'reasonable' and common: this is a point which could be interesting to argue, but is neither an assumption nor a conclusion of the present work.

1.4 The Limits of the Approach

Although the approach is based on a large class of situations, represented by a varied collection of examples, there are a number of limits to the situations it should deal with. These limits can be stated as follows:

- accuracy
- human sources
- flow of information
- contextual situations

Accuracy: The model uses generalisations in its analysis, which are based on considerations of source and circumstance. The result will therefore not be guaranteed to be truth preserving and accurate, and in those situations where the accuracy of the results is not crucial, the model will be suitable. Alternatively, if a precise answer is required and there is enough time, an approach which analyses purely the arguments behind a piece of information would be more suitable.

Human Sources: The model is geared to deal with information from human sources, by using information about the sources and a considerable amount of contextual information in the evaluation process. It is therefore not suited for information coming from non-human sources, where other strategies will be more appropriate.

Flow of Information: The model will function best in situations where there is a flow of information, rather than just sporadic input. Its machinery is designed to abstract information from patterns in the behaviour of sources and to use this in the maintenance of indices in the adaptive mechanisms. If the system is only dealing with one source, or the information is rather sporadic, the indices may take a long time to adjust and other approaches may be more suitable.

Contextual Situations: The source control model would be most suited to operate in a natural environment, dealing with everyday situations, as the source control approach uses a lot of contextual information in the construction of the various models, using considerations like groups, classes, advantages and risks which are either declared or more often inferred by the model.

1.5 Terminology Issues

In this section we discuss how we deal with terminology. We adopt the following schema: terms which are technical, but standard in the relevant literature, are concisely defined in the Glossary. Technical terms which are used only by a particular author are briefly explained in the text at their first mention, and also in the Glossary for reference; similarly for terms introduced by us which do not seem to cause any controversy. Finally, terms which are used in a technical sense, but about which there is no precise agreement in the community, are discussed in the rest of this section.

1.5.1 Controversial Terms

Uncertainty: this is probably the fuzziest of all terms in this field; according to the various authors and applications, it is taken to mean imprecision, inaccuracy, ignorance, ambiguity, guess, chance or preconception. It is even used to signify a very precise and accurate numerical value. For us it is mainly a synonym for *all information coming from human sources, to which a belief must be attached*. We also use it for indicating the heuristic nature of the process by which the system decides such beliefs.

Belief: we use it to indicate the level of support that the system has assigned to a piece of information. This is summarised by a (qualitative) index, but its real (operational) meaning is given by a set of endorsements qualifying it. In other authors or contexts, it is taken to mean a precise numerical value, the set of endorsements without any absolute value or the whole set of facts dependent on the piece of information. We also use the term 'belief' in a specific technical sense, as one of the indices in our model of the human source.

1.6 The Logical Progression of the Thesis

This work is organised according to a very precise plan, in which every part performs a specific role and function:

The Overview has the role of giving the reader a quick taste of the whole work: it can be considered as an extended abstract. The usual short abstract cannot give enough information to provide a general feel of the work. As we believe that it is very important (especially in a new field such as AI) to define the method used early on, it follows that a general overview, external to the logical progression of the body of the thesis, is needed at the beginning.

Chapter 1 (the present one) covers these methodological issues. It is thus a very important chapter, since it defines *explicitly* what we set out to do and why we do it in a particular way, allowing the problems of general method to be discussed separately from those of specific content. This section (1.6) on the logical progression of the work can only be understood once the preceding discussion of methodology in AI, and the description of the method used here, have been fully examined. This section, in turn, is crucial to understand the flow of argument in the subsequent chapters. We therefore recommended that the reader should refer to this section when starting a new chapter, or whenever in doubt about the function of a particular section in the general argument.

Chapter 2 defines with considerable precision the problem area, so as to allow, in Chapter 6, for a more objective measure of the degree of success in solving our problem. Following our chosen method of analysis, we start by discussing the general, fuzzy area of management of uncertainty. We then cut it down to a more precise and manageable subset, using *external criteria*, which are both natural and precise, and can be justified without any reference to our solution. Simplifying assumptions are then added, which reduce noticeably the complexity of the phenomena to model, but still leave inside a class of realistic, common and important behaviours. This space is then sampled for very different cases, to obtain finally a small set of precisely described behaviours to be modelled.

Chapter 3 puts our problem area in the context of existing work which has attempted to solve, even partially, this or related classes of problems. We also examine work which was not applied to this domain, but it might be, and work which has no real relevance, but it could be thought (often because of controversial terms) to have some. We would like to make here the following point: under the strict perspective of the flow of argument, only the work which has direct relevance to our own (either because we build on it, or because we criticise it) needed to be mentioned in this chapter; the rest of the material has been presented for academic reasons of completeness. The reader wanting to get quickly to the substance of our work can safely ignore the main text of this chapter at first reading, relying on the sections' conclusions to provide the background information actually needed later on.

Chapter 4 presents the principles of the source control approach in full, and is therefore the most central chapter of the thesis, setting forth the requirements which the design needs to conform to. Our task is to model the observable situations described at the end of Chapter 2 and not the mechanics by which that behaviour is obtained in human agents: however, we give here also the intuitions and justifications behind our choice of those principles. We think that this might help the reader in accepting them before their full function is shown, and would also give a glimpse of the mental processes by which we have determined them. A set of heuristics, consistent with the general principles, are then given. Their formal counterparts in the next chapter form the backbone of the whole design. These heuristics are domain independent, require only basic information from the natural language environment or the reasoning unit, and are equally applicable and useful in all the examples we set out to model. We are thus adhering to our method, which prescribed exactly such properties in a set of heuristics, to avoid problems of ad-hoc-ness.

Having presented a principled solution to the task, we then introduce in Chapter 5 a high level design. The design is clearly based on the work in the previous chapter, and is detailed enough, we are confident, to allow a competent programmer to reconstruct it. This satisfies the requirement of our method that results should be precisely defined and verifiable. This applies to software systems in the same way as physical experiments or mathematical theorems. To this end, we think that an implementable design is the right level of detail, as it is precise enough to be experimented with, without the amount of details of a real program, which might hide unwarranted short-cuts essential to the program's behaviour. Nevertheless, we believe that the reader has the right to demand evidence that the intended situations have indeed been modelled, without having to apply them to the design. To this end, we present, at the end of Chapter 5, two examples from our initial set modelled according to our design.

These two examples are shown at different levels of details, and in different contexts, so as to make them even more convincing. In the first example, which deals with a 'classical' belief-formation situation, internal data structures are presented in more detail. The second one has been framed as a question-answering problem, to show how our system would be useful in that situation too, and how it should relate to a reasoning unit of present abilities. Various possible scenarios have been considered. Given this added complexity, the exposition has been kept at a higher level of abstraction than for the first example. This satisfies the final demand of our method and allows for a precise evaluation of the results achieved.

The reader less inclined towards the design details, or the one already convinced by the previous chapter, may prefer to skip the design part of Chapter 5 at the first reading. We recommend, however, that at least one of the examples be inspected, as they represent the final link in this programme of research.

Chapter 6 deals with an analysis of the complexity and stability of the Source Control Mechanism. The complexity analysis has been carried out in terms of the decision process, and is based on a structural representation of the SCM provided in the appendices. While the complexity analysis shows the feasibility only in terms of the computational cost, we also briefly consider a potential application of the SCM, which serves to show its feasibility in terms of providing solutions to realistic engineering problems. Furthermore through a formal analysis of the properties of different forms of modification, the stability of the mechanism has been examined to determine the conditions for potential modifications and refinement within the present framework. The results of the chapter will be of particular interest to potential implementers of the SCM, to establish its computational feasibility for a given application, as well as showing its flexibility to allow for refinements.

In Chapter 7 the whole progression is summarised, so as to offer, at a glance, a comparison at each stage between what was required and what has been achieved. Reasons are presented as to why the thesis is believed to form a complete unit, and hence why this is a good point to stop. However, much research could be done by expanding on the themes presented here, and we give a few pointers from where we foresee that the most promising developments may arise.

Then there is a glossary of uncontroversial technical terms, which is followed by references and a more general bibliography.

Finally there are three appendices, the first two of which which give a structural representation of a simplified complete version of the SCM for the purposes of the complexity analysis. The SCM has two major components, namely belief formation and source reevaluation and the first appendix shows the decision trees of the belief formation component, while the second specifies the decision trees of source reevaluation. This is followed by an appendix containing the details of a current proposal of a project in which the SCM is to be applied as part of a decision support system.

Chapter 2

The Problem Area

2.1 Introduction

In this chapter we consider the general, fuzzy area of uncertainty, we describe the particular aspect we want to address and we determine the boundaries of the problem area, before embarking on a solution. In the following section we first look at uncertainty in general and the way it features in everyday experience, before looking at uncertain information from human sources in particular. At the end we present a set of examples, which are instances of the problems we want to address and which we propose to solve with our model.

2.2 On the Nature of Uncertainty

In our everyday experience we are continuously faced with situations where the majority of information is to some degree uncertain. We are also continuously forced to take decisions of one form or another, and hence there is a need to assess the uncertainty of pieces of information before being able to use them as a basis for action. There are a number of factors which contribute to create uncertainty when one is confronted with information from people, and this problem is reinforced by the need to take decisions.

One way of looking at the problem of uncertainty is by considering the different ways in which uncertainty is caused. This includes not just the uncertainty that can objectively be established, but also uncertainty which is perceived to be a problem and features in the decision taking process. These causes can be put under six major headings:

- changing environment
- problems of communication
- problems of complexity
- source of information
- background theories
- reasoning

The world we live in is continuously changing. A particular flower on my rose-bush was only a bud a few days ago, is now in full bloom and will be fading in a week's time. The red sports-car parked next to my car had disappeared a few hours later. The milk in my fridge was fine until yesterday but has now gone sour. These are examples of the ways in which change is part and parcel of our everyday experience. In fact there are only very few things not subject to change, like laws of mathematics or some physical constants like the speed of light and that bodies in free fall accelerate with the square of the distance and so on. As a consequence pieces of information will in most cases be uncertain in that the situation they describe may no longer exist, due to a change of the environment.

Also, a bad communication process may cloud or distort the actual message, thereby adding an element of uncertainty to the information received through this process. The way the message is being put may be ambiguous or imprecise or even fragmented. Thus for example when we follow a conversation in a bar with loud background music, the information we get is like pieces of a mosaic where we have to fill in the gaps of the information we did not actually hear but can infer from the sentence structure, mode of speaking, context and so on. Likewise some teachers are better than others at teaching a particular subject due to the way they get the message across.

Uncertainty is also introduced by way of complexity. A considerable amount of information would require a substantial amount of expertise to completely understand and assess. The information is therefore uncertain in the sense that we do not completely understand all the intricacies of the problem. An aeroplane, for example, is a very complex piece of machinery and apart from more general statements which can be made by everybody, recent crashes and accidents have shown that even experts need to make lengthy enquiries to establish how accidents could have come about.

Another major cause of uncertainty stems from the source of information. Sources typically tend to be reliable only to a point and differ in their reliability. Thus instruments tend to have a margin of error which affects the data they produce. In a similar way human sources, whether single or compound, differ in their behaviour, thus making it difficult to gauge how reliable the information received from them is. Unlike any other source of information, however, people are capable of manipulating the information before passing it on: be that by added reasoning, influence of personal beliefs, or sheer deceit.

Background interpretations are involved in many scientific contexts. The problem of uncertainty becomes particularly acute when the observations can only be made with the help of background theories and the correctness of the information depends on the correctness of that theory. Thus phenomena in modern particle physics cannot be directly observed but are inferred after lengthy calculations and assumptions about the detectors and their accuracy. Background theories also feature in everyday situations, such as in politics where the same set of data can be used to argue quite contrary positions, dependent on the particular political persuasion of the beholder.

Finally, another cause of uncertainty stems from reasoning. Although the basic data may be sound and defensible, plausible reasoning techniques introduce an element of uncertainty as their results, although useful and often convincing, are not strictly valid. Two examples of that are analogy and induction. Inductive arguments conclude from a number of instances, and an assumption of homogeneity on a class of which the instances are members [GAL89], that there is a general law or principle according to which most members of the class behave like the instances. Analogy, by contrast, jumps from a statement about a term to making the same statement about another term based on how much the terms are alike in the context of the statement. Such reasoning techniques are widely applied but only hold to a degree and counter-examples can be found. Uncertainty therefore arises when we consider the possibility that applications may be fallible.

2.3 Information from Human Sources

We have now outlined a classification of uncertainty based on causes. It would of course be possible to discuss the general problem of uncertainty in much greater detail; however, our present objective, as specified in the initial step of our method, is to mark out a specific problem area. The analysis by causes shows that the sub-class of uncertainty originated by human sources is a well defined, interesting and common class with peculiar properties of its own.

Human beings consciously or subconsciously tend to put their interpretation of the data they receive, partially in response to their recognition of the uncertainty involved. There are a number of different aspects about human sources which seem to add to the problem of uncertainty and which can be distinguished under five major headings:

- indication of conviction
- reliability

- trustworthiness
- conflicts
- situation dependency

Sources may indicate their conviction or degree of belief in the information they provide, thus indicating their assessment of the uncertainty of the information. Thus, for example, a friend may say that he was quite sure that he saw my car in town last night. It is also a common phenomenon that human beings are fallible in that respect. A shop assistant in a supermarket might have told me where I can find canned vegetables, but I subsequently find out that they are somewhere else. This adds another dimension to the problem of dealing with uncertainty.

It appears that this indication of belief can take a number of forms: implicit in the context, indication about the strength of belief (for and possibly against) or the supporting evidence or reasons for the belief. Thus, for example, someone in my department sent a message to everybody about a walk in the Lake District indicating that the weather should be warm and sunny according to the weather forecast. He aired a moderate belief in favour of the sun turning out and quoting the weather forecast and hence the Met. Office as his reason for his belief.

There is also a problem with the reliability of sources. While some sources may be very competent and provide information which can generally be trusted, others may appear very poor in their performance. This performance may also vary within single sources. Thus we may get a brilliant physicist who has established a reputation in his field while being unreliable when it comes to practical, everyday matters. Equally, it is not uncommon to find that arts and social science students are totally at sea when faced with computers.

Another serious problem can arise with trustworthiness of sources. An old lady across the road was visited by a woman who introduced herself as being from the social services to check on her well being, only to find out that while she was talking to the old lady an accomplice went through the lady's possessions to rob her of her pension money. Even in more subtle cases we may find that when buying second-hand merchandise, the salesmen may try to over-charge and not admit to faults and problems they are aware of. Especially in important situations when there are potential gains to be made one may have to be careful not to put one's trust in what we are told without good reasons.

It is also not uncommon to find that information we get is in conflict with some other information we may already have. Thus my financial adviser may advise me on one occasion that I should buy shares in a major chemical company, only to tell me to sell them a few weeks later. Similarly we can find that two politicians may argue reasonably coherently for completely opposite positions on the same problem.

It is not untypical to find that uncertainty can have different impacts on different situations. If someone has been witness to a serious accident they may be readily prepared to make statements on what they seem to have experienced, especially to the press, but in the subsequent proceedings in the witness box of a court they will have to make much more careful statements. The uncertainty may therefore vary with different situations where a different level of accuracy is required. Similarly, cases of mutual understanding may have an influence on what kind of statements, or accuracy of statements, are made on the basis of what is expected.

These different situations show that there are a number of ways in which human sources may compound the problem of uncertainty in general as described in the previous section. Although both accounts are not necessarily exhaustive they highlight the general problem of uncertainty.

The different kinds of problems we have described show that uncertainty is an issue in a wide area of our experience. It also appears that there are two main classes of uncertainty, namely those which are more objective and can be assessed in a scientific way as well as those types which are more subjective and dependent on the observer. Even though the latter kind of uncertainty may not always be scientifically verifiable it is none the less significant as human beings will have to take decisions on the basis of their limited perception and knowledge. It is also important to note that although the categorisation may give the impression that there are clear divisions to the problem, in fact the uncertainty involved in most situations tends to be a combination of factors. The root of the problem cannot always easily be identified and consequently we are dealing with a grey area.

Considering that human beings continually have to deal with new information and there is a lack of time, it is not possible to stop at every point and start to conduct scientific experiments. Therefore, in many situations pragmatic techniques seems to be required. Artificial Intelligence has tended to consider uncertainty in a more statistical sense, usually with the help of various forms of probabilistic techniques. The problems described so far do, however, also suggest that there are different kinds of uncertainty involved when we get information from human sources in common everyday situations, which may be born of subjective perceptions, limited understanding or incomplete knowledge, but which are nonetheless significant as they affect real problems and real decisions. Looking at the size of the problem of uncertainty we can only start to conceive the amount of work which remains to be addressed in Artificial Intelligence.

2.4 Selection of a Class

One area which is of significant interest, is the uncertainty involved in information from human sources. There is only a limited number of different ways in which information reaches us in everyday situations, such as direct observation through our senses, or with the help of instruments, like measuring devices. Information from human sources occupies a prominent place. At the same time this problem has not been addressed widely in current research, but if we want to build systems which can cope with a real and general domain, a solution will be required.

The general area of uncertainty as we have described it, is not very clear-cut and we need to look for criteria for extracting a more well-defined section. The uncertainty involved in information from human sources is a good candidate for a number of reasons:

- it is a clear, well-defined and natural class
- it is a prominent class
- it is an important class

The problem of uncertainty in general covers a very large, fuzzy and diverse area as we have seen from the discussion in Section 2.1. Following our method, we need to cut out a more clear-cut subsection for a number of reasons: firstly, in order to assess whether a potential solution can be judged to be successful one needs clear criteria to define what the solution is supposed to cover and to check whether that has indeed been achieved. Secondly, the radical differences, outlined above, in the possible kinds of uncertainty makes it doubtful whether a single approach can be found which covers them all. Finally, the sheer size of the general problem makes it next to impossible to address it in a single attempt.

The uncertainty involved in information from human sources provides a suitable criterion where candidate members can be easily judged to fall into the class or not. It will also be intuitively plausible that this class is a natural class and describes therefore a class of behaviours which is not chosen because it happens to suit our model. Information from human sources is also significantly different from, say, information from instruments or other non-human sources, and due to the rather different nature of that type of information it requires special treatment.

The class of information from human sources is also a prominent class. Everyday experience will easily convince us, that, in fact, a great proportion of information comes from identifiable human sources and that that presents a wide variety. Providing a solution to this problem should therefore make a significant advance on the problem of uncertainty management in general. This should also help considerably to enable systems to deal with a real domain, through the fact that it may give a handle on a major proportion of information which can be expected in this situation.

Uncertainty about information from human sources is also an important class as it is a crucial starting point. Although progress has been made on some of the other areas of uncertainty, the traditional dilemma has been how to deal with the propagation and effects of uncertainty in databases of various kinds, while being unable to automatically establish initial levels of uncertainty needed to set the process in motion. One cannot get started without them, and a solution to the problem of information from human sources should be instrumental in establishing a significant amount of automatization at these initial levels.

2.5 Some Simplifications

While we want to deal with uncertainty from human sources in general, there is a need to make some simplifications with respect to this class of behaviours for a number of reasons. The overall domain is very complex and includes a number of different problems which have no direct connection to our problem or are of such a fuzzy nature that no clear criteria of success can be produced. Such a restriction should therefore not take anything away from a solution to the problem itself.

Information from human sources is extremely varied and at closer inspection reveals itself to be a mixture of components. We have already hinted at the fact that in many situations there may be much more to the information conveyed than is contained in the sentences uttered, as is the case when there is a great deal of mutual understanding and non-explicit communication involved. There is also the tone of conversation and other non-verbal acts of communication. To successfully deal with these areas, a natural language processor would require to analyse the discourse on a more general level and also to capture non-verbal, allegoric and emotive aspects. These aspects are a matter for a natural language processor to deal with and we expect that in the transformation of information to a machine representation these will be eliminated and that our system therefore only deals with literal meaning input. This simplification is necessary to be able to judge the results of our approach against the criteria of success.

Consequently, we propose to deal only with what is contained in the verbal, plain English part of the information. The other simplification is that the information is assumed to be serious. To deal with joking, humorous and other non-serious conversation is usually dependent on a great deal of culture and upbringing. Even experienced foreigners, for example, find it extremely difficult to grasp the British humour especially if that is of a more subtle nature, which leaves the suspicion that a great deal of local, domain specific knowledge is required. We expect that this has been taken care of by a natural language analyser and we think that this simplification does not impoverish the domain significantly, as most realistic situations are still covered, such as when sources of information are lying or trying to deceive the system.

2.6 The Examples

So far we have described the general field of uncertainty and the different kinds of problems involved. We have then proceeded to cut a particular class of behaviours out of this rather fuzzy area which we propose to model, and finally we needed to make some simplifications about the behaviours we will address.

2.6.1 Criteria for Selection

Since it is impossible to demonstrate that the model we propose will work successfully on every element of this potentially unlimited class, we need to select a number of examples as representatives of the whole class: we could then model them in more detail. The selection is intended to address the main categories of the domain and to provide a good coverage of the area. On a general perspective it appears that there are three main categories:

- problems with isolated pieces of information
- problems with multiple pieces of information from the same source
- problems with multiple pieces of information from different sources

In the simplest case we might be dealing with single pieces of information which have no significant connections to other information and can therefore only be considered on their own, or that prove to be unsustainable on their own.

Alternatively, there may be a connection with information supplied by the same source on a previous occasion, leaving the problem of which of the two pieces of information to believe.

Finally, there may be cases where there are multiple pieces of information from different sources, again creating uncertainty as to what to believe as a result.

Apart from the general types of situation the second consideration is to select examples which are from a wide variety of subjects to demonstrate how the model is fundamentally domain-independent. In the following we present five major examples which, we claim, can be reconstructed with the help of our model:

- the second-hand motorbike example
- the copied assignments example
- the financial advice example
- the squash racket example
- the John Wayne example

2.6.2 Buying a Motorbike

Suppose we want to buy a second-hand motorbike to go to work with. We go to a local shop to have a look at some motorbikes and the dealer asks me how much I wanted to spend, tries to sell me one of his bikes, arguing that the bike is in good shape and would be a real bargain. Suppose also that a friend of mine came along as well and said that the motorbike the dealer pointed out was looking good. If I am however not an expert on these matters, I will still be unsure whom to believe and what to do - should I buy the bike or should I not. I may put off the decision for a while but in the end I will have to take a decision either way.

2.6.3 Copying Assignments

Suppose a teacher is assessing assignments and finds that two of them are virtually identical. A colleague, asked for his considered opinion, agrees that they seem to be copied and the teacher therefore calls in the two students separately and asks them whether they agree that the assignments are remarkably similar and whether they can give an explanation. Suppose the student with the better assignment agrees that they are indeed similar but that he did not copy from the other while the other student denounces that they are not very similar and that he certainly did not copy from the other. It is a very tricky situation as there is no concrete evidence to prove what has happened and one needs to investigate carefully who is lying and who speaks the truth. If we believe the colleague then at least one of the students is a liar and if we believe the students then the colleague and our own judgement must be mistaken.

2.6.4 Financial Advice

Supposing we were to invest a considerable sum of money in the financial markets, but because we do not have sufficient expertise we go to see a financial adviser. Asking a financial adviser at my bank, he tells me that I should buy ICI shares as they have developed a new environment-friendly coolant for fridges which should give them a major share of the market. Going to see another adviser I am told that I should not buy ICI because of take-over rumours but rather buy shares in Abbey National as they will get a good share of the mortgage market with their new all-inclusive home-moving package.

2.6.5 Buying a Squash Racket

Suppose I need a new squash-racket and a friend recommends a particular model which he heard had received good reviews. It is difficult to decide which model to buy, because there are so many of them and there is so much hype about the latest technology some of which may make a difference to my play, but am I really good enough to get the benefit of it and if I do is it worth the extra expense as opposed to a simpler model. When I see him again a week later, he insists that it is an excellent racket as he bought one himself and felt it was a great improvement on his old one. I may believe that he feels that the racket is making a difference but perhaps he is only excited by the fact that he has a new racket which has kindled his enthusiasm. He was talking about the reviews which I could have a look at, but can I be bothered to go through lengthy technical enquiries considering that I am not a physicist to appreciate the finer implications of the technology.

2.6.6 The John Wayne Example

This example is an extract from the film, "Cowboys". John Wayne plays an old rancher who has been forced to hire children to help run his ranch, because his cowboys have joined the gold rush. Three men come along, asking for a job. He asks where they have worked. "All over the place" they say, "You name it". "No, you name it", Wayne replies. They mention a few names, and he asks them when they last saw one of those mentioned. "Last month", they say, and he tells them that that man had been dead for three years. They are caught out, and admit that they are just got out of jail and saw the names on a list. "I have no use for you", he tells them. "You don't want us because we are ex-prisoners?" they challenge him. "I never hold a man's past against him, but I cannot stand liars", he replies.

2.7 Conclusions

The purpose of this chapter has been to look at the problem domain of uncertainty in general and to select a more well-defined section which we propose to model. The selection of a clearly defined sub-class is necessary in order to establish that the proposed model does cover that sub-class successfully. The selection criteria themselves have to be independent of the model itself to ensure that the class wasn't chosen just because it happened to fit a pre-determined model. We also had to apply a few simplifications which reduce noticeably the amount of details to be modelled, but leave inside the core of the domain. As a model cannot be tested on every member of a class of that size it is necessary to select examples to demonstrate that a prospective model works in order to establish clear criteria for success.

We have described the general grey area of uncertainty and selected the class of uncertain information from human sources. We have given arguments to suggest that that class is both well-defined and satisfies the constraint of being independently justifiable. We also have given justifications for the simplifications and provided a set of examples which is both varied in type and from a wide selection of subjects. It is also important to note that the line of argument has not made reference to the model to be constructed and it remains to demonstrate that the model we propose in the remainder of the thesis can deal with these examples. Out of this class we have chosen examples which serve as representatives of the whole. We shall demonstrate in the following chapters that our model can deal with them.

Chapter 3

Related Work

3.1 Introduction

As we have seen in the previous chapter, uncertainty is an important factor in the evaluation of the kind of information we tend to receive in everyday situations. We have looked at the problem from various angles and decided to work on a model for uncertain information from human sources. An inevitable consequence of uncertainty in that context is that we have to deal with conflicting information. There are a number of problems to be addressed:

- uncertainty introduced by the possibility of change and the difficulty in its prediction
- uncertainty born out of our limited understanding and knowledge and the need to make assumptions and estimates
- the possibility of having to deal with and resolve conflicting information

A number of techniques and skills are required to deal with these problems. We need techniques for uncertainty management: especially, we need techniques which can deal with human sources. We also need techniques to deal with conflicting information to achieve conflict resolution.

In the last few years there has been an increased interest in the management of uncertainty at various levels. At the same time, the available literature bearing on the problem is still not very plentiful, at least not on the specific problems we are considering. Despite a strong resemblance in terminology there are vast differences in the goals and intentions involved and there are, alas, only a handful of pieces of research which are directly connected to the present problem.

Within the relevant field of current research there are a number of different directions which we want to consider in more detail. These different areas can be put under the following headings:

- plausible reasoning
- truth maintenance
- conflict resolution

A number of interesting developments have been made in the area of plausible reasoning, which has been concerned with the management of uncertainty in general as well as with the management of uncertain reasoning techniques.

Truth maintenance is a well established and widely known discipline, which has been looking at the problem of ordering the beliefs of inferentially and logically connected belief systems of various kinds. There have been a number of recent departures from the traditional Boolean view of the world.

There are also a number of contributions which have addressed the problem of conflict, the necessity to restore consistency, and the need to decide between conflicting courses of action. Despite a common terminology there are a number of very different kinds of conflict addressed, not all of which have to do directly with information.

Apart from these three areas we shall also have a brief look at theories of probability. Representations of uncertainty of information in AI have by and large been based on various probabilistic theories and we therefore shall look at probability theories on a general level.

Finally, we shall revisit one approach from the plausible reasoning community which deals with a similar problem situation to ours and is therefore presented in more detail, after having been put into perspective by comparing it with other approaches.

In the previous chapter we have looked at the problem we want to address and in this chapter we need to look at the current developments in that light. As we have not yet presented the principles of our model, the discussion in this chapter is restricted to more general comments and we shall discuss the finer points which directly relate to our work as we present the principles of our model in next chapter. Another reason why we shall not go into great detail is that by and large there is no direct connection to our work, with the notable exception of [GAR86] and [GAL89], from which the present project takes its starting point, and [GAB88a], [GAB88b] and [BGA90] which record interim results.

3.2 Plausible Reasoning

The term plausible reasoning has been used in AI to describe two different kinds of reasoning:

- plausible reasoning techniques
- uncertainty management

Human beings display a remarkable ability to draw implicit information out of the explicit information they are given. Thus, plausible reasoning techniques are used to draw inferences by techniques like analogy and induction, when no explicit links are available. These reasoning techniques are uncertain as they cannot guarantee to deliver correct results but are good approximations and a valuable method to exploit existing information. For an in-depth formal model of analogy, see Long [LON87].

Plausible reasoning, however, is also used to refer to reasoning techniques employed in the assessment of uncertain or inexact information. This uncertainty has been represented in different ways, like probabilities, possibilities, fuzzy sets and belief functions and as a result we get an inexact kind of reasoning usually referred to as approximate or plausible reasoning.

The first category can be considered as a form of reasoning management as plausible reasoning in this field is a matter of controlling uncertain reasoning techniques. In the second category plausible reasoning is a form of uncertainty management, dealing with imprecise and uncertain information. Both types of reasoning techniques can deliver incorrect conclusions but are at the same time very powerful in allowing the user to significantly transcend the limitations, say, of classical logic.

For an overview of some of the approaches we would like to refer to Berenji and Lum [BER87], O'Neill [NEI87] and Backer, Van der Lubbe and Krijgsman [BAV88], who give an account of the current architectures of plausible reasoning systems and other related methodologies. Most of them deal mainly with probabilistic approaches and an exposition of non-probabilistic reasoning can be found in Horvitz, Heckerman and Langlotz [HORV86]. For a comparison of formalisms for dealing with uncertainty we suggest Pearl [PEA88a][PEA88b], although there the emphasis is particularly on non-probabilistic formalisms.

Looking at recent developments, there appear to be three different interests: there are a number of researchers interested in everyday human plausible reasoning in the attempt to model this very successful reasoning process in the field of AI. Secondly, there are a number of projects addressing the problem of uncertainty and reasoning management on a more technical level. Finally, we discuss a very distinctive approach to uncertainty management, which we consider separately as it appears to be particularly relevant to our method.

3.2.2 Plausible Reasoning Techniques

Although, as we have said, plausible reasoning techniques are not fail-safe, there is a definite advantage to be gained by using them. Human agents often use inductive arguments, based on a number of observed instances and an assumption of homogeneity [GAL89], to draw conclusions about general rules. Likewise, analogy is rather common too; if understood in a more scientific way it means jumping from a statement about one term to making the same statement about another term based on how much the terms are alike in the context of the statement. Thus if I knew two teachers who were very much alike in their character and I believed that character has an effect on teaching methods, I could probably say that both teaching methods would be much the same.

Work which has been carried out on the subject of reasoning management is much more sparse than that on uncertainty management; we will consider three pieces of research: Stefik [STE81], Collins and Michalsky [COM89] and Dontas and Zemankova [DOZ88].

Stefik's particular interest lies in a planning approach, which integrates and extends the least-commitment and heuristic strategies. There is a need for a system to make intelligent guesses and Stefik resorts to plausible reasoning techniques to compensate for limitations of the knowledge base. The planning approach has a layered control structure separating the planning problem from the planning process and is consequently called meta-planning; this has the advantage that the system can reason about its own performance. The approach has been implemented in MOLGEN, a KBS which plans gene cloning experiments in molecular genetics.

The most important and most widely known work is that by Collins and Michalsky [COM89] who, on a more theoretical level, present a general analysis of plausible human inference patterns, such as analogy and induction, together with some parameters of conditional likelihood, typicality and similarity. They give a formal representation of human plausible inference patterns which are widely used in everyday reasoning. A framework is established which combines the plausible inference patterns with different certainty parameters.

A practical application of Collins' and Michalsky's work has been presented by Dontas and Zemankova [DOZ88] with their implementation of APPLAUSE. In this system, human knowledge is represented as objects in hierarchies, while construction and creation of links represent the learning process. This link construction corresponds to the ability to draw inferences when no direct links between the objects concerned are available. The process uses generalisation, specialisation, similarity and dependencies as well as the tool of confidence parameters.

A predominant reason for the use of plausible reasoning techniques is the need to bridge gaps in the knowledge base which could not be bridged otherwise with the help of traditional methods. Although, for instance, the results of reasoning achieved solely with the precepts of classical logic can be guaranteed to be true (provided the correctness of the premises), their usefulness is rather limited. Plausible reasoning techniques are much more powerful and allow for useful conclusions to be drawn which could not be drawn otherwise, but at the same time there is the obvious trade-off of validity versus richness, as plausible reasoning techniques cannot be guaranteed to deliver correct results.

Both Stefik, and Dontas and Zemankova acknowledge the need to draw inferences to fill important gaps in their knowledge bases, which could not be filled by more classical methods, with the help of plausible reasoning techniques. Collins and Michalsky think that these reasoning techniques are a dominant feature of human reasoning and consequently they establish a model for them based on human responses to everyday problems. Acknowledging that the results of such reasoning are less than completely reliable, they also model the uncertainty surrounding such consequences and the varying degree of belief human beings display. The emphasis is, however, on how to reason about an existing stock of information rather than how to assess the uncertainty of information as and when it first reaches the system, which is the task we want to address with our model. Consequently, it seems that the connection between the approaches is more on a complementary basis.

3.2.3 Uncertainty Management

As plausible reasoning techniques are less than completely reliable, their use introduces uncertainty into the system. At the same time, uncertainty is a common feature in a number of situations and not just a product of plausible reasoning. There are two fundamentally different representations which have been applied. By far the most popular approach is based on uncertainty measures represented by quantitative parameters. The second method, which has attracted less attention, is based on qualitative measures modelled upon human methodologies for uncertainty management.

The work on uncertainty management with non-probabilistic human plausible reasoning methods appears to follow in the footsteps of [RES76]. According to his theory of plausible reasoning, the credibility of a piece of information depends on the reliability of the source supporting it. The model will therefore reduce the credibility of the information to the degree of reliability of the source concerned (which is known to the system). He argues that the use of probability indices for formulae is problematic, as combinations of probabilities in chains of reasoning will cause a rapid deterioration of the original levels of probabilities. If we have two premises with $p = 1/2$ the result of an inference depending on both premises will only have a probability of $p = 1/4$ and so on. Rescher draws attention to this effect, arguing that this makes the result of any realistic chain of inference so low as to be virtually useless. He claims that his theory of plausibility can overcome this limitation with his new plausibility measure. He adopts the 'pars deterior principle', where the strength of a conclusion is at least as strong as its weakest premise.

An application of Rescher's approach can be found in Bestougeff and Ligozat [BES87], who have implemented a plausible reasoning module which is an extension of an expert system able to handle first order predicates and equipped with forward and/or backward chaining. Their system is based on an adaptation of Rescher's approach, where each fact has a plausibility index attached allowing easy computation of the plausibility of any deduction.

Alternatively, Popchev and Zlatareva [POZ87] propose a non-monotonic inference mechanism which can deal with incomplete and uncertain information, based on human schema for plausible reasoning. Their understanding on human reasoning schema is, however, built on probabilistic methods. The authors suggest that problem-independent plausible inference mechanisms can be implemented. A description of the most important plausible reasoning schema is also given.

Uncertainty Management has been a subject in various areas of Artificial Intelligence. There are a number of projects which address the subject in the field of truth maintenance as we shall see in the following section. There are also a few pieces of work which address the subject from the angle of plausible reasoning, like Prade [PRA85], Farreny and Prade [FAP85], Quinlan [QUI85], Shao-Hung [SHH86], Paass [PAA86] and Hori and Sheu [HOS88]. For a review on different approaches we would like to refer to Prade [PRA85] who discusses various techniques through a unifying framework.

In the same article Prade introduces the concept of degrees of truth and discusses truth qualifications of propositions. He considers two strongly related concepts - uncertainty and imprecision of propositions. In addition the case of multiple sources is also being considered although the model is not elaborated in great detail. On the same basis, Farreny and Prade [FAP85] deal with uncertainty management in rule based systems. Uncertain facts or rules are represented by probabilities. At the same time the authors use Rescher's method for manipulating information from multiple sources and the pars deterior principle.

On a more specialised issue, Quinlan [QUI85], criticises the usual approach of associating a validity measure with each fact or rule and to compute the validity of any deductions. This can be inappropriate for some problems, particularly when the evidence is not internally consistent and proposes a new approach based on finding consistent subsets of the evidence in question.

Shao-Hung [SHH86] has produced a uniform formalism to account for plausible and causal reasoning. Basing it on the Dempster/Shafer theory of evidence, which features prominently in current plausible reasoning systems, he produces an extension of the theory to achieve a uniform way of accounting for the causation aspect as well as the certainty aspect of an inference.

Paass [PAA86] again deals with assigning a measure of probability to rules and facts in a rule based system, measuring their validity. His particular interest is in a probability propagation algorithm which does not require that probabilities be given for each proposition. This method estimates the joint distributions by the maximum likelihood approach and assesses the uncertainty of the results.

Finally, for a more technical application, Hori and Sheu [HOS88] consider the management of uncertainty in the design of expert systems, for the purpose of trouble-shooting complicated systems. Their method is again based on the Dempster/Shafer theory of evidence. The paper describes an algorithm which calculates basic probability assignments from a-priori probabilities and the statistics of a target system.

With the exception of Rescher, and Bestougeff and Ligozat the approaches to uncertainty management are by and large put on a probabilistic or quasi-probabilistic basis. Leaving aside the specific aspects the various projects address, Quinlan and Paass use probability theory in their respective models, whereas both Shao-Hung and Hori and Sheu use the Dempster/Shafer theory of evidence which is a probability bounding method adapted from Bayes Theorem, which is by far the most popular form of probability theory used to represent uncertainty. Only Prade, and Prade and Farreny are the exception as they claim to use probabilistic measures of uncertainty while using Rescher's non-probabilistic methods of uncertainty assessment and manipulation.

While Rescher's approach appears to make significant progress to overcome some of the problems of probabilistic methods, there are nevertheless also some shortcomings. Although he can deal with assigning initial levels to the plausibility of statements through his source models, the indices involved are expected to be given and correct and the model has no way of re-assessing and changing them in the light of contrary evidence. He is however at least able to assign values to given statements which is more than most probabilistic methods can do. Incidentally, considering the approach Prade and Farreny take and the fundamental differences between truly probabilistic and plausibilistic approaches it seems doubtful how they can be reconciled on a theoretical level, an issue which they do not explore explicitly. As far as our direction of research is concerned, we agree with the criticisms made by Rescher on the use of probabilistic measures for capturing and propagating beliefs. Although there are certain areas of application such as statistics, where probabilistic methods are more suitable, the majority of types of uncertainty as described in the previous chapter do not seem accessible with a probabilistic approach.

3.2.4 Evidential Reasoning

Another direction of research which developed in response to the felt inadequacies of probabilistic indices of belief is that pursued by Cohen [COH85], and is based on endorsements as representing the arguments for and against a particular piece of information. The model is entirely non-numeric and thus distinguished from the other two schools of plausible reasoning.

He questions the representational adequacy of numbers used to hide the qualitative reasons behind a belief in a proposition, arguing that many of our beliefs are based on one kind of evidence or another and may also depend on the importance of the uncertain situation and what the information would be used for. His model analyses the evidence or endorsements in a qualitative way, to decide whether the information is believable enough for the purpose. There are many kinds of uncertainty which he wants to represent in his system and reason about, with the help of heuristic knowledge. In a comparison between the endorsement approach and a Bayesian model, he argues that numeric approaches fail to capture vital properties of the evidence.

Strictly numerical approaches have limitations as the arguments behind a proposition may need to be analysed given the particular situation. Although Cohen deals with information from human sources, his approach does not have a source model along the lines of Rescher's proposition. There seems to be much common perspective in both views and the two models appear to be complementary to a certain degree.

A different approach to evidential reasoning in expert systems has been proposed by Baldwin [BAL87] and is based on a probabilistic understanding of uncertainty. He argues that man's knowledge consists of statements which cannot be guaranteed to be true and which are often phrased in imprecise language. Uncertainties therefore need to be properly modelled, which leads him to use inductive, abductive, analogical and plausible reasoning methods with an emphasis on the strength of evidence. In his system a conclusion does not logically follow from premises but is supported to a certain degree by evidence.

Baldwin's [BAL87] can therefore be counted in both the uncertainty management as well as reasoning management camps. His fundamental premise that a conclusion does not always logically follow from premises, but is supported to a certain degree by evidence, is however typical of evidential reasoning systems.

3.2.5 Conclusion

Cohen [COH85] makes the important point that the AI community has over years made many elaborate efforts of avoiding to deal with uncertainty by a process of carefully engineering uncertainty out of the problem domain. The subject of plausible reasoning has made the refreshing attempt to actually deal with uncertainty rather than to avoid it. Also, with the introduction of human-like reasoning methods a significant step has been made to increase a system's power to elicit implicit information from the knowledge base.

As our interest is in dealing with uncertain information from human sources in everyday situations, we shall be concerned primarily with assessing the uncertainty and believability of pieces of information as we receive them. We are interested in establishing initial levels of belief as opposed to what other information can be drawn out of an existing stock. This implies that our interests in plausible reasoning is mainly concentrated on uncertainty management and evidential reasoning rather than plausible reasoning techniques. We agree with Rescher's argument about the limitations of probability theory for the particular domain and therefore adopt a non-probabilistic approach. At the same time Cohen presents compelling arguments against purely numeric measures irrespective of what these numbers represent.

Rescher's most significant contribution is the idea that source considerations have an influence on the certainty or plausibility of the particular piece of information in question, yet it appears that his model would profit from a source control mechanism which can adapt the perceived reliability of the source to changing circumstances, to produce an intuitively more appropriate response.

Cohen provides a good argument for his stance that numeric probabilistic certainty indices are representationally inadequate and that the belief in a proposition depends on a qualitative analysis of the actual arguments supporting it. We would however propose two main areas of improvement for the endorsement model. Firstly, as with Rescher's approach, Cohen's endorsement model could profit from a source control mechanism especially since a large proportion of information we tend to analyse is not given with specific arguments. It is this kind of information Cohen finds difficult to represent. Secondly, we think that some index representing the strength of belief would be useful to run in parallel to the full endorsement analysis as Cohen himself admits that the full analysis is time consuming and not all problems warrant it. This would enable the system to provide quick solutions to unimportant problems and at the same time it would allow the system to analyse the underlying endorsements if the situation so requires. There are also many situations in which the supporting evidence is simply not available, while the source of information is almost always known. We think that a marriage between these two approaches would be profitable in many respects.

3.3 Truth Maintenance

3.3.1 Introduction

The basic motivation behind truth maintenance is the need for belief revision. Many problems involve the need to make choices about how to proceed with less than perfect information and truth maintenance provides a machinery that allows the consequences of assumptions to be determined and the set of beliefs revised, if necessary. Truth Maintenance Systems (TMSs) are house-keeping sub-systems of reasoning systems. The problem solver passes the inferences it makes to the TMS, which in turn organises the beliefs of the problem solver. As the TMS has no access to the semantics of the information, it is usually the responsibility of the problem solver to ensure correctness of information.

[DOY78] is said to have coined the name Truth Maintenance System and since that time work has developed in the field in two related but separate directions:

Justification based TMSs (JTMS) are direct descendants from Doyle such as [ALE80] and [GOO84] as well as Thompson [THO79]. The main protagonists in the area of Assumption based TMSs (ATMS) are de Kleer ([KLE85],[KLE86a-c]) and Martins & Shapiro ([MAR83],[MAS84],[MAS86],[MAS88]).

These two types differ in the way dependencies between the data are recorded. While the JTMS records only the immediate relation between a datum and its support, in the ATMS all the hypotheses on which a datum ultimately depends are recorded with each datum. Doyle's [DOY79] JTMS maintains a node for each datum associated with the justifications provided by the problem solver. Nodes either have a justification and are thereby currently believed or have no current justification and are not believed. By contrast, de Kleer's ([KLE85] and [KLE86a]) ATMS is based on manipulating assumption sets, which allow the system to work quickly and efficiently with inconsistent data. It allows multiple contexts where each context is a consistent subset of assumptions and inferences drawn from them. As a result the process of backtracking, which can be a severe problem in TMSs, is substantially reduced and as a useful by-product the architecture allows multiple potential solutions to be considered simultaneously.

In general, the TMS provides two services to the problem solver, truth maintenance and dependency directed backtracking. Truth maintenance is required when a node which is currently disbelieved is given a valid justification. The TMS must therefore calculate the status of other nodes which are in any way connected or affected by the change. Dependency Directed Backtracking is required when a node which has been declared as valid is found to cause a contradiction. The TMS finds the set of assumptions on which the justification for the contradictory node depends. One of the assumptions from this set (culprit) is retracted. If this does not succeed in forcing the contradiction out, the process is repeated.

Research in truth maintenance has been very plentiful and the following are examples of a very rich field. Drummond, Steel and Kelleher [DRS87] provide a brief overview of current approaches to Truth Maintenance Systems: as the ATMS is the most popular form of TMS, it is discussed in greater detail. Shadbolt gives a review of TMSs and ATMSs and describes a number of applications using these methodologies. Bigham in turn draws a comparison between the TMSs of Doyle [DOY79], McAllester [ALE80] and de Kleer [KLE86a]. The comparison is however more specific, mainly in terms of examples from diagnosis of electronic system faults and maintenance in industrial applications. We shall review the following topics:

- truth maintenance systems issues
- TMSs, plausibility and uncertainty
- truth maintenance systems applications
- truth maintenance and questions of logic

3.3.2 Truth Maintenance Systems Issues

There are several works concerned with specific problems of the classical architectures of truth maintenance systems, which are either due to their inherent shortcomings or due to the need of adapting TMSs to the needs of new application areas. There are traditional problems of efficiency connected with the incorporation of new information into a database, and with the process of consistency recovery once contradictions have been discovered. As far as new requirements are concerned, there have also been problems with modelling actions and the adaptation to distinctly non-monotonic domains.

The incorporation of new information into the system can become quite complex and time-consuming, especially once the database has reached a considerable size and the new information necessitates numerous changes to the system of beliefs. Support propagation immediately follows the incorporation of new information and thus it has been suggested by Petrie [PET86] that, due to a large amount of repetition involved, this process could be significantly speeded up by parallel processing.

Backtracking is in fact the more serious problem and it is an issue which has specifically been addressed by Dechter [DEC87], who has developed a support propagation and diagnostic model which finds solutions to contradictions, and which necessitates a minimum of changes to the database. Another difficulty with backtracking is that it has been beset by problems of dead ends; Dechter has addressed that too [DEC86], suggesting that this highly inefficient process can be significantly improved by recording the reasons for dead ends: an information which can later be used by the control regime to ensure that the system will not fall into the same trap again.

Another proposal for improvements to the backtracking process has been made by Rodi [ROD89]. His approach can find all solution states, while avoiding backtracking, by focusing on the regularities of the solution space. Unfortunately, its time complexity is potentially exponential, thereby restricting its usefulness to small databases of up to 100 beliefs. With a view to using it on larger problem spaces Rodi, however, suggests that the problem can be contained by computing only partial solutions.

TMSs, and particularly ATMSs, are powerful tools for searching a space of alternatives: they have traditionally been dedicated to inferential problem solving and belief formation. They do not have facilities for modelling actions which are often the natural outcome of the decision taking process. Morris and Nado [MOR86] have suggested such an approach based on an ATMS, but their findings do also indicate that this will require a more elaborate treatment of contradiction. Solutions to contradictions have more often than not been dictated by expedience rather than problem specific reasons; actions, due to their strong connections to the external world, impose a number of restrictions to the consistency recovery process.

Another reality of everyday experience is the nature of change. TMSs have usually been based on more classical types of logic and consequently held a non-monotonic view of the world. This implies that TMSs are concerned with maintaining a consistent database in the face of a changing world. Even though

De Kleer's ATMS allows for multiple contexts, the fundamental principles of truth maintenance in general are based on a strongly monotonic view of reality, therefore insisting that justifications have to be monotonic. This in turn implies that non-monotonic justifications can only be introduced at a higher level. Dressler [DRE87][DRE88] deplored this shortcoming and, to enable a system better modelling of a changing reality, has proposed a modified model for an ATMS which allows for non-monotonic justifications at the same level as monotonic ones.

These examples show some of the areas of current development in truth maintenance. Although the more technical problems addressed are obviously specific to truth maintenance systems, they give us at least a flavour of the issues involved. At the same time it is interesting to note that researchers like Dressler, and Morris and Nado have tried to put truth maintenance into a wider context and to free them from some of their traditional restrictions.

3.3.3 TMSs, Plausibility and Uncertainty

Plausible reasoning and uncertainty management is at the heart of our own research project and some work in that direction has also been done in truth maintenance. At the same time, this is only a marginal issue in the field and there has been less research there than in other areas of truth maintenance.

The work done on the introduction of measures of uncertainty to truth maintenance can be categorised along similar lines as the work already discussed in the section on plausible reason. Again a majority of work is based on probabilistic or quasi-probabilistic interpretations of uncertainty, with a notable exception based on Rescher's model for plausible reasoning.

Like the introduction of non-monotonic justifications by Dressler, the introduction of uncertain justifications causes a number of problems, which are connected to the fundamental differences between traditional forms of logic and the new forms of logic needed to deal with uncertainty and non-monotonicity. Falkenhainer's [FAL87] general-purpose belief maintenance system, for example, tries to manage probabilistic beliefs in a similar way as usual TMSs manage boolean beliefs. He considers his system as a generalisation of a truth maintenance system in which usual boolean beliefs are also catered for. D'Ambrosio [AMB87] addresses the same issue in a completely different way, namely with numeric certainty estimates where assumptions are used to represent beliefs in uncertain facts. The new method derives numeric truth values from numeric truth estimates of assumptions. This has the advantage of improved management of dependent and partially independent evidence, and the ability to query the certainty of a proposition from different perspectives. In [AMB88], D'Ambrosio extends this new approach into a hybrid reasoning scheme which combines numeric with symbolic methods for uncertainty management. This hybrid system bases its symbolic techniques on an adapted version of an ATMS and combines it with numerical methodologies from an adaptation of a Dempster/Shafer theory of evidence.

The approach of Falkenhainer is probabilistic and tries, like Dressler, to cause a minimum of changes to the classical architecture of truth maintenance systems. D'Ambrosio, by contrast, goes a few steps further by trying to incorporate a full model for uncertainty into a truth maintenance system. While Falkenhainer's approach is a proper probabilistic one, D'Ambrosio adopts an approach using the Dempster/Shafer theory of evidence. He is careful to provide a solution which can deal with dependent evidence, which is one of the major shortcomings of classical probability theories.

To give an example of a quite different approach, Fangqing Dong and Nakagawa [FAN88] have addressed the introduction of uncertainty by using the concept of 'possible nogood' and 'possible good', enabling the system to make hypotheses about a current situation where the information is insufficient for reasoning.

Finally, Katai and Iwai [KAI89] introduce a pluralistic evaluation of belief plausibility based on the plausibility index introduced by Rescher [RES76]. They compare their approach with the usual probabilistic ones to explore the properties of their approach, especially those of Rescher's indexing of beliefs. The authors also show that this new methodology can effectively be utilised in TMSs which deal with beliefs in complex and dynamically changing situations. Since Rescher's theory is basically monotonic, the authors adaptation is mainly to do with extending it to cope with non-monotonic situations. This approach incorporates a full extended model for plausible reasoning into an ATMS. Incidentally, the model can be implemented quite easily, since changes which were made on a boolean basis before, are now made on the basis of pair-wise comparison between two plausibility estimates and the new datum has only an impact if it is stronger than the other propositions involved. Therefore the extension to the ATMS in this fashion proceeds without major changes to the fundamental design of TMSs.

A number of researchers have seen the need to enable truth maintenance systems to deal with uncertain information. This has not been without problems, as it significantly complicates the truth maintenance process and architecture. Falkenhainer has therefore proposed a way in which probabilistic beliefs can be handled in a similar way to boolean beliefs. D'Ambrosio in turn uses assumptions in an ATMS to represent beliefs in uncertain facts and thereby avoiding complicated changes to the architecture by moving uncertainty to a different level. Fanqing Dong and Nakagawa's approach, likewise, tries to avoid costly changes with the introduction of possible good and possible nogood thereby enabling their system to engage in hypothetical reasoning. The motivations of these researchers seems to have been different in each case. Whereas Falkenhainer concentrates on introducing uncertainty to truth maintenance with a minimum of disturbances, d'Ambrosio is interested in introducing the whole machinery of probabilistic uncertainty management into the ATMS. Although uncertain facts are represented by the common entities of assumptions, he is able to derive numeric truth values from uncertainty estimates of the assumptions. He also addresses the problem of dependent evidence which can mar probabilistic approaches. Katai and Iwai's approach by contrast is quite different from the other three as it is based on Rescher's model for plausible reasoning and is non-probabilistic by nature. They have introduced the whole machinery needed for plausible reasoning into the TMS.

3.3.4 Truth Maintenance Systems Applications

Truth maintenance systems have found a wide-ranging area of applications, reaching from electronic testing to medical diagnosis, computer aided tutoring and system security.

The testing of electronic devices as to their proper functioning is an ideal area of application, especially for the traditional versions of TMSs, as the domain is well constraint and also strictly logical. Pau's [PAU85] system for testing and monitoring of electronic circuits is one example of this. It has knowledge of the architecture of the device and can accordingly apply stimuli and compare the expected response with the actual response to see if there are any faults. Alternatively, an example of the use of truth maintenance systems for diagnostic problem solving on a more general level has been accomplished by Provan [PRO88]. To maintain efficiency in the use of complicated diagnostic problems, existing ATMSs are, however, in need of methodologies to rank competing solutions and restrict their search space to explore only likely solutions. Consequently, he proposes an extension of an ATMS based on the Dempster/Shafer theory of evidence which can rank competing solutions and restrict itself to exploring only the most likely ones.

A similar problem of how to deal with competing explanations or diagnoses has been addressed with truth maintenance by Rake and Smith [RAS88]. They have worked on an automation of the process of extracting data from cardiac angiogram images. In particular they have an interest in resolving ambiguous features and producing a unique explanation of data that can have a number of different interpretations. Widman [WID89], by contrast, has an interest in the use of expert system reasoning about dynamic systems by semi-quantitative simulation. He uses a complex model with semi-quantitative parameters, representing the cardiovascular system as an example to illustrate the method. The mathematical side includes first-order differential equations and numerical integration by standard methods. This is paired with the symbolic power of causal inference methods. A common database is used and truth maintenance helps to resolve conflicting explanations.

Finally, two applications dealing with computer systems are presented in [VEA88] and [IKM89]. Venkat, Rangan and Ashany [VEA88] have been dealing with the process of securing communication in distributed systems for industrial applications. Concepts like the theory of belief, the theory of evidence, belief reasoning with uncertainty and truth maintenance have been applied to this security issue. Ikeda, Mizoguchi and Kakusho [IKM89] have been working on an intelligent computer-assisted instruction system, ICAI, which includes a student model and a tutoring strategy module. The student module models what the student does or does not understand. An ATMS is incorporated to deal with a student's inconsistent answers.

These examples show that truth maintenance has found a wide area of applications. Truth maintenance systems are however sub-system of an overall problem solving system and their degree of involvement in the problem solving process varies. Thus in applications like [PAU85] and [PRO88] the TMS plays an integral role in the process of testing of circuits and the establishment of causes of failure. The two medical applications, by contrast, involve the TMS to a lesser degree. Their main task is to model the cardiovascular system: truth maintenance is only used to resolve ambiguous or conflicting interpretations of the data-set.

Similarly, in [VEA88] truth maintenance is used to check consistency of beliefs and in [IKM89] to resolve inconsistencies of students' answers. Though the work of the TMS is important, in both cases the TMS is not directly involved in the main task, fulfilling an auxiliary role.

These examples of applications of truth maintenance seem to suggest that they work particularly well on problems which are more technical and well defined, and where there is a relatively small amount of uncertainty involved. Also, in the application to potentially large problem spaces, Provan's experience suggests that there is a need to cut down on the exploration of solutions in order to maintain an acceptable level of efficiency. Finally, it is intriguing that though the emphasis is usually on considering contradictions a nuisance to be eradicated (JTMS) or to be resolved by generating alternative views of the world (ATMS), Ikeda actually applies the ATMS precisely to the purpose of sorting out inconsistencies in the student's answers. The only application which is concerned with uncertainty is that of Provan, but the uncertainty is not integral to the truth maintenance process: it is only used to steer the process of exploring likely solutions.

3.3.5 Truth Maintenance and Questions of Logic

There has also been a significant interest in discussing logical matters relating to truth maintenance systems. This includes issues like their theoretical underpinning's, their relation to other methodologies and their shortcomings and possible extensions, the issue of non-monotonicity, interests in unification of formalisms, formal semantics and demonstrations of common semantics between formalisms. For a comparison of truth maintenance systems and other major approaches to evidential reason like rule-based systems, Bayesian networks, Dempster/Shafer formalisms and non-monotonic logics we would like to refer to Pearl [PEA88a][PEA88b]. Pearl focuses particularly on the fact that they are systems dealing with uncertainty.

The issue of uncertainty and modelling of beliefs of active processes have already been considered by McDermott and Doyle [DER80]. They argue that non-monotonic logic are very important in modelling beliefs of active processes. Given that the information to be processed is incomplete, it must be possible to make and subsequently revise assumptions in the light of observations. Presenting the motivations and history of such logics they also discuss the relation to truth maintenance systems.

The issue of unification of formalisms has been addressed by Ginsberg [GIN88], who claims that first-order theorem provers, ATMSs and formal systems such as default logic or circumscription can be captured by a uniform formalism. He tries to demonstrate that a default reasoner can be implemented in this way and that it can be combined with ATMS methods to form an incremental default reasoning system. In this model consistency checks need not be made before drawing tentative conclusions, but beliefs can be adjusted when a default premise or conclusion is overturned in the face of convincing contradictory evidence.

On a different slant, there has also been interest in formal semantics for truth maintenance. Brown [BRO88] has been working on a mathematical logic which is equipped with an underlying model theory and has been applied to characterise precisely some well-known models of truth maintenance. The characterization is claimed to be precise in that it gives meaning to truth maintenance in terms of formal logic, where each characterising logic corresponds to a particular truth maintenance system and vice-versa.

These articles have been concerned with various logical issues of truth maintenance systems, their provability and relation to other methodologies dealing with similar problems. Although these issues are important for putting truth maintenance on a firm foundation and as a means for precise comparison of different methodologies, they do not directly touch on our own research interests. Even though Pearl, for example, has been interested in uncertainty, the emphasis has been on how to deal with known uncertainties and how they are affected during the propagation of justifications rather than on how to assess uncertainties initially, which is our particular interest. As far as a formalisation is concerned, we doubt if it could be properly and usefully carried out at present for a domain of such complexity as that of natural language information from human sources about real situations.

3.3.6 Conclusion

Truth maintenance system have made a significant impact and have found a wide area of application. At the same time their design is very specific and thus making them rather inflexible to adapt to the kind of requirements we are considering. Their basic assumption is that data are inferentially connected and usually of a boolean nature. The strength of the model lies in exploring these logical connections and in finding wholly consistent sets of beliefs, and to adapt the system of beliefs in the advent of new information. The model will consequently work at its best when applied to problems of that nature.

Experience has however shown that many of our everyday problems are not of that kind in that uncertainty and changeability are a prominent feature. This accounts for contributions like [MON86] who have addressed the need of modelling actions as well as beliefs, [DRE87] and [DRE88] who tried to adapt the ATMS to allow for non-monotonic justifications to take account of the changeability of certain situations and the contributions on uncertainty by [FAL87], [AMB87], [AMB88], [FAN88] and [KAI89].

Of these adaptations the non-monotonicity is perhaps the easiest to achieve as it preserves the main features of the TMS. The introduction of uncertainty, by contrast, is much more complicated as it violates the precepts of the traditional types of logic which dominate the support propagation and backtracking process. This necessitates the introduction of a whole new machinery, to determine the reliability of the result of chains of reasoning and how to deal with potential conflicts.

The typical areas of application of truth maintenance strengthens this view. The problem of testing electronic circuits is perhaps the one most closely related to the classical architecture of truth maintenance. Provan tackles a typical problem of efficiency and introduces an extension based on evidential reasoning to avoid wasteful computation of unlikely solutions. Rake and Smith's problem of ambiguity in the possible interpretation of abnormalities in cardiac angiograms already suggests that there is uncertainty about the correct diagnosis, although they still seem to treat these different interpretations as if there was no uncertainty involved. This may perhaps not be surprising as another precept of truth maintenance holds that it is the problem solver's responsibility to ensure correctness of information and that it is the purpose of the TMS proper to order the beliefs of the problem solver accordingly.

Considering our problem of dealing with a changing and uncertain world in general and with information from human sources in particular, there are a number of important implications. Firstly, the domain is very large and a prospective system will have to deal with large amounts of data. We also want to deal specifically with assigning degrees of belief to information as it reaches the system rather than focussing on how beliefs are propagated in inferentially connected belief spaces. While TMS traditionally focus on the latter we want to redress the imbalance and concentrate on the former. Furthermore, on a general note, we feel that the majority of information we are likely to deal with are not completely connected inferentially but rather of a more atomic nature and if there are connections they are likely to be implicit rather than explicit and have to be inferred first. Consequently, there are good grounds for focussing on the former which has not been widely addressed, but which is essential if we want to generate systems with some degree of autonomy.

3.4 Conflict Resolution

Considering that in an environment where we have to deal with uncertain information from a variety of sources the situations will arise where various pieces of information may be at odds with each other, we also have to face the task of resolving conflicts in order to maintain a coherent and consistent view of the world.

Unfortunately, unlike plausible reasoning or truth maintenance, conflict resolution, as a term, has not acquired a strict meaning and consequently we are dealing with a variety of issues. As a consequence pieces of work which appear, through the use of a common terminology, to apply to the same subject may have little to do with each other. In the current field there seem to be three different directions of research:

- In rule based systems, conflict resolution is about dealing with conflicting rules.
- In robotics, specifically during the planning stage which precedes action, there can be planning conflicts which need to be resolved.
- Planning conflicts are also an issue in multiple agents systems.

3.4.1 Conflict Resolution in Rule Based Systems

Typically, rule based systems are made of a collection of rules which are used on the data supplied to the system. The rules usually have particular conditions attached to them under which they can be applied to data; these conditions are commonly called 'triggering conditions'; any particular rule will remain inactive until its triggering conditions are fulfilled by the data and/or control structure. When the conditions are fulfilled the rule becomes active: this is called 'firing'. When fired, a rule can either deliver a direct result or in turn trigger another rule and so on. When there are no more rules to be fired the result is output. These rules, defined by an expert, are often assumed to be consistent, an assumption which is not always warranted, as experience shows [ION89]. The available literature in this area deals primarily with the problem of how to resolve conflicts between incompatible rules which have, for some reason, been fired together. A small review of different methodologies can be found in [LAU84], where a number of different approaches for a control cycle over the operation of knowledge bases are considered.

The methodologies that have been proposed in the area can be divided into those where the system lets the user decide which rules to adopt and which to drop, and those where the system will decide for itself which rules should be discarded. Cases where the system will decide for itself seem to be more popular than those where the user is asked to assist the conflict resolution. Ogami, Nishiyama and Kakusho [OGN81] and Morris [MOR86] address the issue of conflict resolution in production systems where the system either uses heuristic or domain- specific rules [OGN81], or uses meta-level reasoning in the form of meta-rules or meta-knowledge [MOR86]. By contrast, White and Sykes [WHS86] propose to resolve conflicts among rules by user preference, arguing that the users may have preferential knowledge specific to the user and/or the situation they are in, and that this should be taken into account in the process of conflict resolution.

Another quite different approach has been proposed by Fayyad, Voorhis and Wiesmeyer [FAY88] in STAC, a system which can learn control information. They propose that connections between rules should be established each time a task has successfully been performed. These connections have a strength associated with them, which encodes a history of the success/failure rate of sequences of rule firing during problem solving. The strength of the connection is subsequently used to guide the inference engine in two ways: the selection of the next rule to consider and as a basis for conflict resolution.

The interesting point is that the rules are generally assumed to be correct: it is only some combinations which are thought to be problematic. Conflict resolution consequently becomes the process of getting out of a tricky situation, whereas the actual conflict might suggest that something is wrong with the rules. Amongst different strategies, the one in [WHS86] does not appear to go beyond the immediate problem of rule conflict either, as it simply asks the user to give an assessment of the correctness of the rules involved. The only exception can be found with [FAY88], where the success rate of the rules is monitored, and this knowledge used to guide the system. Consequently, [FAY88] appears to be the only model which attempts a long term re-assessment of the conflicting rules, to improve the system's performance rather than opting for a quick solution to the immediate problem, which would leave the system just as vulnerable to the same problems in the future.

It is incidentally interesting to note that conflicts are by and large seen solely as a nuisance, a problem to be overcome. Conflicts should rather be considered as an opportunity to learn, as they suggest that something is wrong and needs to be sorted out. It is therefore a stimulus to re-assess the contributing factors and to learn from the situation, in order to produce a better response in the future. Only in [FAY88] this aspect seems to have been perceived.

The situation where we have to deal with uncertain information from different sources is bound to give rise to conflicting information, but the problem there is rather different from the ones just examined, as conflicting rules are about conflicting strategies whereas we are dealing with conflicting information and the approaches used to deal with conflicts will have to be quite different.

3.4.2 Conflict Resolution in Robotics and Technical Applications

Conflict resolution is also an issue in the field of robotics and similar technical applications of expert systems; examples can be found from a wide variety of applications including navigation, medical diagnosis and scheduling.

Thus, Balakrishnan, Mahapatra, Nayak, Poulose and Krishna [BAM86] deal with Air Traffic Control, where conflict resolution methodologies are employed to avert potential aircraft collisions during route planning. On a similar issue of collision avoidance and route planning Gilmore, Semeco and Eamsherangkoon [GIS85a] [GIS85b] have been working on autonomous vehicles which can find their own way through natural terrain. In both systems conflict resolution occurs when the planning algorithm proposes a strategy which conflicts with the physical contingencies of the terrain. Conflict resolution is a matter of finding a plan of action which avoids collision with obstacles.

To mention two further projects briefly, Birman [BIR82], for example, has been working on a diagnostic tool for electrocardiogram analysis, using a rule-based learning algorithm to perform wave interpretation to check for cardiac abnormalities. Conflicting interpretations may occur from time to time and a simple conflict resolution mechanism will resolve them into a single unambiguous one. Finally, Shaw [SHA86] gives an example of conflict resolution in the scheduling of flexible manufacturing systems. Again we are dealing with the problem of generating plans which do not conflict with the contingencies of the real world or with each other.

The conflicts addressed in these projects are either planning conflicts or interpretation conflicts rather than conflicts about uncertain information. It is interesting to note that all projects assume the data they are given to be correct, an assumption which will probably be warranted in most cases: it is however the problem we are most interested in.

3.4.3 Conflict Resolution in Multiple Agents Systems

Unfortunately very little work has been done on multiple agents, which is a subject much more closely related to our line of research. There are a number of different definitions of what agents are. Agents could be systems or system components which are to some degree independent and autonomous, like vehicles in Adey's [ADE88a] navigation simulator or independent system components as in a distributed AI system as in [DOR88], [ORG88] and [HER88]. Alternatively, there are also systems which have a world model made of concepts of agents external to the system. Examples of this interpretation are human agents as in [BEL88] and Galliers [GLL87], [GLL88] as well as [COS88], although Connah, Shiels and Wavish include simulated human agents as well. Not all of these, however, do specifically cover conflicts and conflict resolution: [BEL88], [HER88], [COS88] and [DOR88] address multiple agents without explicitly dealing with conflict and methodologies for conflict resolution. In fact, Adey and Galliers appear to be the only two researchers specifically interested in this problem.

Adey [ADE88a] describes a system to control multiple vehicles in a training simulator for aircraft or marine navigation, where vehicles may be required to act as independent agents. Conflicts occur when plans of these vehicles are failing due to a new traffic situation etc. and need to be repaired. The system is rule-based and a decision has to be taken when conflicting rules are fired together. The author uses meta-rules to resolve conflicts. This is very much akin to [MOR86] except that multiple agents are involved in this case.

By contrast, Galliers [GLL87], [GLL88] concentrates on the system's interaction with human agents, working in the area of cooperative planning frameworks. She works with a model of multiple, autonomous agents and proposes to use dialogue for cooperative multi-agent planning. Conflicts between the agents will be resolved by way of negotiation. Galliers builds on the work of Cohen and Levesque [COL87a], [COL87b] who have devised a formal theory of rational interaction as a basis for communication. Gallier's article describes some aspects of a computational model for multi-agent dialogue. This incorporates cooperative as well as conflicting agents. Galliers works towards the implementation of a cooperative system which will use dialogue to negotiate and resolve differences. She recognises that the real world is riddled with conflicting situations born out of constantly changing and unpredictable environments, and that a system should include dialogue to negotiate and potentially remove conflict and achieve cooperation.

Galliers describes a number of properties that agents enjoy in her system: agents have preferences which are a relationship between beliefs and goals; autonomous agents have control over their beliefs and the adoption of goals (mental states); they act autonomously on the basis of their own preferences, and if there is an existing contradictory goal, then the only condition for being able to drop one goal is if the agent prefers the other in the light of the current circumstances. Furthermore agents have interests, which are types of goals which the agent not only believes achievable, but which will eventually be achieved. It is also acknowledged in this system that agents have only partial control over the outcome of their actions, as agents are not isolated but operate in a social environment. Dialogue is therefore part of a game of strategy, and can be used to further an agent's interest by inducing in the other agent beliefs and goals which will not conflict with the first agent's interests. Finally there are also postures which are an agent's attitudes towards another agent's beliefs and goals, and which can be one of cooperation, indifference or conflict.

Adey's multiple agent simulation is very much akin to the kinds of conflict and conflict resolution we found in [BAM86], [GIS85a] and [GIS85b] as it is part of the planning and re-planning process, although with the added feature of multiple agents. It is a conflict of plans rather than of information, such as the one we are primarily interested in. Galliers' model, although could be interpreted as a sophisticated model to repair plans, is primarily a model of motivation, and strategy to further them. Through the model of motivation, it leads on to a model of manipulation used to further each agent's motivations and hence goes considerably beyond mere planning. Whereas Adey and similar problems are plan orientated, Galliers is motivation orientated. As our interest is primarily in sorting out conflicts between conflicting information of uncertain epistemic status our problem is quite different. At the same time it appears at least from the outside that there is a certain amount of complementarity between our aims and that of Galliers as multiple, at times conflicting agents are as much part of observable reality as is uncertain information. A truly autonomous general reasoning system which is expected to deal with the real world much in the same way as a human agent, will have to be able to model intentions and motivations of other agents as well. This, however, goes beyond the limits of the present enquiry.

3.4.4 Conclusions

Conflict resolution in rule based systems is a matter of deciding between conflicting rules. These rules are by and large assumed to be correct, and. the solution to the conflict does not take into account previous performance. [FAY88] is the only exception in this field, since success/failure rates are monitored and thereby the system may improve its performance. At the same time even this model is very introspective as it tries to learn about its own performance rather than to learn about the correctness of its world model. Such correctness is the very issue we question and will try to provide a solution to.

Similar differences are manifest with respect to the technical applications of section 3.2.3. We are not currently considering planning problems, which feature prominently in that area. Furthermore the nature of their data from radar, vision, electrodes etc. can be classed as information from instruments, which we are not concerned with. Most of that information is also more easily verifiable, whereas our particular concern is to deal with uncertain and incomplete information.

3.5 Probability

3.5.1 Introduction

Probability theory has found a widespread application in the AI community, and in particular in those areas concerned with the modelling of uncertainty or the modelling of belief. It has been suggested [COH85] that the reason for this lies in the simplicity and ease of use rather than in the representational adequacy. It is also quite intriguing to find that most contributions happily use probability theory without ever stating what kind of interpretation or definition of probability they employ, a consideration which can have important implications. In the following we shall look at probability theory on a more theoretical level rather than of examples of its use (which can be found in articles mentioned before), as our doubts about the representational adequacy of probability theory for our purposes will obviously affect any approach built on it.

Probability is said to affect our decisions in everyday situations to a considerable degree, and is often taken as a guide of life. At the same time, in everyday situations, people may disagree on what is, or is not, probable given the same information and consequently there must be different interpretations of what probability is. Kyburg [KYB61] is interested in finding a viable definition of probability which can be considered as rational.

The property of rationality has been closely associated with the use of probability. Clearly, if one wants to be rational one would be expected not only to be consistent but also devoid of self-contradiction. Likewise it would be irrational to reject conclusions supported by evidence and hence when they are very probable. Kyburg maintains that we are particularly rational when our degree of belief in a given statement is precisely the degree to which it is supported by evidence. Conversely it is irrational to ignore the evidence and believe what is not warranted by the evidence. There are three major types of probability theory:

- frequency theories of probability
- personalistic theories of probability
- logical theories of probability

3.5.2 Frequency Theories of Probability

Frequency theories interpret probability statements as being about relative frequencies or ratios of classes, or alternatively as an abstract property of certain sequences of events obeying the laws of probability calculus. Thus the connection with the world is through derived relative frequencies of classes which can be either observed or contemplated.

Thus Venn [VEN86] states the probability that an A will be a B as a limit of the relative frequencies of B's among A's as the number of A's is increased without limit. In von Mises formulation the sequence of A's is referred to as a collective and is subject to the condition that the B's occur randomly in the sequence. Randomness in turn is defined objectively without reference to anything anyone does or does not know. Wald [WAL37] demonstrates that, according to this definition, there are random sequences with limits.

Kyburg maintains that this interpretation of probability is completely objective. A given sequence is or is not a collective regardless of what anyone does or does not know or believe. Similarly, the probability that an A is a B has a certain value p regardless of what anyone knows or believes. Note that there may be no value p where there is no limit of relative frequency of Bs among As. Kyburg draws attention to the fact that no finite collection can qualify as a collective and it is doubtful, even if there were infinite sequences of natural events, whether they could qualify.

On other interpretations objective theories can be derived for finite classes. Russell [RUS48] and Neyman [NEY50] have defined probability as a proportion in a finite class. Thus the probability that a card in a 52 card deck is an ace is 1/13 as 1/13th of the cards are aces. Alternatively, Braithwaite takes probability as an abstract characteristic of 'selections' from a population. Thus that As are Bs is to be interpreted as a class ratio in a model.

Another definition of probability is given by Cramer [CRA46]: Whenever we say that the probability of an event E with respect to an experiment 'e' is equal to P the concrete meaning of this assertion is that given a series of repetitions of 'e' it is practically certain that the frequency of E will be approximately equal to P. Given the result of an experiment we can decide what probability statements to accept on the basis of a decision technique. Thus the mathematical probability can be interpreted as the conceptual counterpart of the relative frequency with which the random variable takes on a value.

When we want to quantify the concept of whether a proposition is believable, we judge whether the statement is believable or not, but not whether the event described in the statement is probable because of some intrinsic property. This suggests that the frequency theory is only applicable to certain uses, such as scientific and statistical contexts, and a number of frequency theorists would agree with that position. Kyburg draws attention to the fact that there is no obvious way to provide a frequency interpretations of existential statements like "it is very probable that Cesar crossed the Rubicon" or that "it is improbable that there is life on Jupiter". There are many types of problems for which a frequency interpretation would not be meaningful and Kyburg's examples seem to reinforce that view.

3.5.3 Personalistic Theories of Probability

The basic idea of personalistic theories (first suggested by Ramsey [RAM31] as a supplement to a frequency theory and followed by Savage [SAV54] and de Finetti [FIN37]) is that probability statements are statements about actual degrees of belief. Thus some statements are certain and others believable only to a certain degree, dependent on the subjective evaluation of the understander.

This raises a number of problems: If probabilities become a matter of personal preference they will lack objectivity and hence there can be no comparison between different evaluations of probability and no methodology to settle disputes. Savage himself agrees that it would be unreasonable to expect that any two persons would be bound to agree on the probabilities given the same evidence.

Another implication of the subjective view is that the views need not even be rational. Thus Kyburg argues that a fundamentalists belief in creation in the face of evidence is permissible, even though most people would consider it irrational.

Finally, it is very difficult to determine the exact degrees of probability to be ascribed to statements. This problem has been addressed by some researchers and various techniques were proposed. According to Ramsey it is possible to discover the degree of probability by considering a hypothetical sequence of bets. The value of probability is thus determined by the highest odds one would offer in the betting process. Kyburg argues that due to the diminishing marginal utility of money the technique is difficult to apply, although as the sums become smaller and smaller it becomes more plausible.

It appears that the personalistic theory does not give us a clear interpretation of scientific or statistical situation like coin-tossing or population analysis. Attempts have been made to deal with such situations but Kyburg argues that these attempts are contrived and indirect and generally not true to life. He furthermore deplores the fact that the theory can scarcely handle general statements without having to take recourse to the angelic hypothesis, where an omniscient angel knows the next state and can make his bets accordingly.

Although statements about personal belief can be handled without problem the theory is problematic as it is too liberal to be reasonable and at the same time too strict as it appears intuitively unplausible that the perceived probabilities could be given such a precise value. Finally there also appears to be no overall framework to settle disagreements as people are free to disagree at leisure. The personalistic theory consequently appears to be too vague to be useful.

3.5.4 Logical Theories of Probability

Probability can also be defined as a part of logic and understood in that way it provides a framework of rational belief not merely confined to self-consistency. Rules can be laid down to express the probability of a given statement relative to the given evidence, resulting in a real number which is determined on logical grounds alone. Consequently, probability is to be considered as a logical concept and probability statements are logically true if they are true at all.

Keynes [KEY21] was the first to suggest such a theory, arguing that probabilities are not subjective but objective relative to a body of knowledge. Statements are therefore not probable because we think they are, but what is probable or not has been objectively fixed through the facts given in the body of knowledge and is hence independent of our perception and opinion. This also implies that experience has no influence on the probability of a given statement and the given evidence.

Although it may appear that the probability is irrevocably fixed through the logical relationships, it is not uncommon to observe that experience may yield new evidence which should be included and which may influence the probability of the statements in question. Kyburg argues that this is precisely the kind of concept of probability needed for a rational guide of life, especially when we speak, amongst other things, of the high probability of scientific inferences.

There is, however, one problem with Keynes' definition as it considers probability to be primitive. This means that probability cannot be defined in terms of the other logical primitives. Therefore probability has to be known intuitively in the same way as we recognize intuitively that if 'q follows p' and 'p' then 'q'. Though the latter can be established as a formal rule (p,p->q .: q), Keynes is unable to provide anything like this for his probability relations; he cannot provide formal rules which eliminate the necessity of intuition. Yet this is the kind of thing we need if we are to demand that two rational beings faced with the same evidence will agree on the probability of a given statement. Even Keynes admits that the ability of rational beings to intuit probabilities varies, some being more apt than others.

Carnap [CAR51] manages to eliminate some of the shortcomings of Keynes' theory. He defines probability in terms of conventional logical concepts which does not require any extra-logical reference for their definition, nor any extraordinary intuition for their application. Carnap does not attack frequency theories but opts for peaceful co-existence. While frequency theories are considered appropriate for many applications in scientific and statistical problems, the logical theory of probability is considered appropriate in other areas.

Carnap proposes a special language to deal with probability which, it has been argued by Kyburg, is only able to express very simple sentences and would not be adequate for the needs of, say, modern science. This is complicated by the requirements of completeness and independence of the information concerned, a requirement which is impossible to fulfil in many cases. Logical independence requires that there are no connections between objects and properties. This is a severe restriction as one cannot even say that an object is red or green considering the logical truth that if an object is red it cannot at the same time be green. Neither can relational predicates be used like a>b>c as this implies that a>c. The requirement of completeness has even more severe implications than the requirement of independence. It demands that any two individuals differ only in a finite number of independent respects and that a system of predicates be taken which is sufficiently comprehensive for expressing all the qualitative attributes in the given universe to which the logic formulated for the language is to be applied. Kyburg argues that that has ontological implications, and sounds like the old-fashioned postulate about the uniformity of nature. This would deny the possibility to add new names of properties as they become apparent in the process of scientific investigation. One may also experience difficulties in imagining that there is an infinite number of individuals which differ only in a finite number of ways.

3.5.5 Conclusions

If we consider our earlier discussion about the kinds of uncertainty we are dealing with in everyday situations, it seems doubtful whether any of these different kinds of probability theory will be able to help in the task we want to address. There may be notable exceptions where for instance frequency theories are applicable, especially when the nature of the problem is of a scientific-statistical kind. However, when we deal with the problem of whether Cesar crossed the Rubicon or whether the milkman has already delivered the milk this morning then such an approach seems dumbfounded. The personalistic theories seem to account better for the fact that people tend to have degrees of conviction about the data they hold, but then this theory appears to allow for any point of view, however absurd, and clearly something more principled is needed. The basic idea of logical theories that the probability of a given statement is a function of the supporting evidence appears to be much more appealing, but as we have seen, there are still a number of formal problems which stand in the face of common, everyday experience. Kyburg's argument that logical theories can be interpreted to produce probabilities of differing reliability may be bad news for logical theories, but are good news for us. Thus, for example, it will not be uncommon to find situations where human agents have taken decisions on the basis of some evidence and where the decision may have turned out to be wrong as new evidence came to light. Although the first decision may be wrong in hindsight, we may agree that we would have reacted in the same way and that the first decision was reasonable, given the limited evidence. Although this falls short of absolute and objective measures, it nevertheless suggests that the human agent proceeded in a principled manner in his or her evaluation, which was the best that could be done given the situation.

The degree of conviction in a piece of information does not usually depend on relative frequencies of classes of events as is represented by frequency theories of probability. It will be intuitively plausible that in fact most of the information we get in everyday situations is not statistical in nature for such a definition to apply. Thus Cesar's alleged movements are not easily captured by a frequency definition, and statements of that nature are the ones which our model would be dealing with. There is also another subtle but important distinction. Probability theories would be interested in the exact amount of probability. Given the sparse historic evidence one may question whether an exact probability can be ascribed at all to speculations about Cesar's excursions as the meaning of the statement is that though historians may favour the idea that Cesar in fact crossed the Rubicon they do not know for sure and probably never will.

Personalistic probability statements are said to be statements about actual degrees of belief, dependent on the subjective evaluation of the understander. This may have some resemblance with the problem we try to address, but although personalistic models try to account for actual degrees of conviction they do not give any help in establishing the respective levels of conviction for any given case. Ramsey's proposal to discover the degree of probability by considering a hypothetical sequence of bets tries to address this issue, but, although his idea underlines the need for consistency between beliefs and actions, it is difficult to envisage the bet experiment being useful in practice. Furthermore, the observation that the marginal utility of money in this process influences the evaluation of probability shows that the evaluation is dependent on the financial situation of the understander. Although the personalistic theories try to account for actual strength of belief in human agents the use of the theory is severely restricted through its lack of objectivity and the impracticality of the methods used, an important consideration when we consider a practical implementation.

Logical theories of probability provide a framework of rational belief not merely confined to self-consistency. Thus the probability of a given statement relative to the given evidence results in a probability which is determined on logical grounds, if they are true at all. This concept of probability obviously avoids the objections which have been made about personalistic theories and their lack of objectivity.

At the same time this objectivity is only relative to the body of knowledge. If the body of knowledge is complete the probabilities will be objectively true whereas if the body of knowledge is incomplete, and in parts uncertain, the probabilities will also be uncertain. It is an interesting phenomenon that the reliability of probability statements increases with the amount of evidence. We think that this is an indication that incomplete knowledge issues in only partially reliable probability statements. Although that may create problems for the logical theory, we do not think that this is a strange phenomenon but rather the logical implication of assessing the probabilities relative to a body of knowledge. When we consider human reasoning we find that we often have to take decision on the basis of an incomplete and uncertain body of knowledge. This is in most cases a quite reasonable thing to do.

A far more serious problem is the restriction in Keynes' definition, which requires probability to be primitive and therefore to be known intuitively. He cannot provide formal rules which eliminate the necessity of intuition which is needed if any two rational beings are to agree on the probabilities, given the same evidence. Even though Carnap manages to eliminate some of the problems of Keynes' theory, the requirements of completeness and independence of the information concerned pose a requirement which is impossible to fulfil in many cases. We think that especially in the everyday situations we try to model, these two requirements cannot be met.

Kyburg thinks that as long as external influences act on the application of a logical theory of probability it is impossible to completely systematise probability and therefore one is in a way no better off than with the frequency theory as a guide of life. We think that this view is not necessarily justified if we keep in mind that the original claim was only that the probabilities are established relative to a body of knowledge. Although we shall not use probabilities in our model, the implicit assumption is that human beings have a particular degree of belief in a proposition based on some evidence or endorsement. To demand that body of endorsements and evidence should be complete and sufficient is often unrealistic and opinions have to be formed despite incomplete information. Indeed we consider it a virtue of human reasoning not to get stuck in the face of incomplete and uncertain information but still to come to a reasonable decision. This may not please the purist who wants a neat and tidy solution and objective measures, but in order to dig the proverbial cart our of the mud a shovel is much more useful than a silver spoon.

Finally, although we have addressed only the three major schools of probability theories, there have been many hybrids developed in order to solve various problems. Thus Bayes' Theorem and the Dempster/Shafer theory of evidence have found wide acceptance in the AI community amongst others. The basic problems however remain the same and we hope that it has generally become clear that despite initial appearances there are significant differences between probability theories in general and the kind of approach we need to pursue for our particular problem. We agree with Cohen's argument that beliefs in proposition, statistical cases apart, are to be based on a qualitative analysis of the arguments involved rather than by a quantitative approach. Kyburg argues that a reasonable response is to have a strength of belief commensurate with the known probability. As his exposition shows, this objective probability is often very hard to come by, if not impossible. Cohen proposes an elegant alternative where beliefs are reasonable if they are backed by adequate arguments or endorsements for the decision at hand. This model is highly intuitive, but unfortunately presents, up to now, serious problem when it comes to deliver results.

3.6 Model of Endorsement

3.6.1 Introduction

In this section we want to have a closer look at Cohen's model of endorsement [COH85] which we briefly mentioned in the sections on plausible reasoning and on probability. As the model is perhaps the closest of all approaches to the principles which underlie our own method for uncertainty management, we will consider it in more detail. Since the model of endorsement is built on a qualitative, non-numeric analysis of evidence for and against arguments, it is therefore a definite departure from the usual probabilistic or quasi-probabilistic approaches which dominate this field. Cohen argues that the numeric representations of strength of evidence hide important aspects of the evidence which may be crucial to the reasoning process and the strength of the conclusions. The argument Cohen produces for his view is twofold:

- dependence on the type of arguments
- dependence on the purpose of the information

He thinks that a numeric value hides different kinds of evidence which cannot be meaningfully represented in this fashion. The qualitative differences between different kinds of evidence make some preferable to others: thus in most cases eyewitness testimony is preferable to circumstantial evidence, direct evidence is preferable to indirect evidence, corroboration is preferable to contradiction and inference is preferable to assumption. He replaces the numerical certainty estimates by the actual evidence supporting a datum, inference or conclusion.

Apart from the evidence, the strength of our belief may also depend on the importance of the uncertain situation and the utility of the evidence is judged in this light. If the decision is just a matter of forming an opinion then less evidence is needed than if the decision has more serious implications. To go with Cohen's example, King Solomon's decision on the fate of the baby that two women lay claim to has serious consequences and substantial evidence is needed, whereas his courtiers can easily make bets about his decision as little depends on them. This implies that the purpose of the information needs to be taken into account when its uncertainty is considered and again numeric approaches do not have the ability to take this into consideration.

3.6.2 The Problem of Dealing with Uncertainty

Cohen deplores the vehement adherence of current approaches to numerical methods, especially in areas where they seem inappropriate. He thinks that these quasi-probabilistic numerical methods for reasoning about uncertainty are often adopted less for their advantages than for the lack of better methods.

There are a number of different ways in which researchers have tried to deal with uncertainty and amongst them there are two main approaches which dominate the field:

- engineering uncertainty out of the problem domain
- dealing with it with probabilistic approaches

In some cases it is easier to engineer uncertainty out of the problem domain and therefore to avoid having to deal with it. One way to do this is by making assumptions such as presuming noise-free data or assuming only relevant data or making a closed world assumption. Another approach for discounting uncertainty is diversification, the process of reducing the impact of uncertainty by spreading the risks. These are basically low-risk strategies where uncertain outcomes of recommendations might be reduced by suggesting a course of action with less gain but higher security. Although these approaches may work quite well for some cases they do not help us when we have situations where we cannot avoid having to deal with the problem.

One of the most widely used approaches which actually deal with uncertainty is Bayes's theorem, which is a simple mathematical method for updating the probabilities of a hypothesis given some evidence. Assuming a set of hypotheses and their relations, plus a basic assignment of probabilities, Bayes' theorem propagates and maintains the probabilities of the hypotheses. Nevertheless there are a number of well-known problems with this approach, such as that it requires huge amounts of data, that initial probabilities have to be provided, that events are independent and that the approach does not distinguish uncertainty from ignorance. Cohen's most fundamental objection is that numerical beliefs fail to capture everything one needs to know about the uncertain situation. He thinks that one should not only know how a particular conclusion is derived and how much it should be believed but also more fundamentally why it is to be believed. The Dempster/Shafer theory of evidence, which is an extension of Bayes' theorem, remedies the criticism that uncertainty and ignorance cannot be distinguished in the Bayesian model by allowing the assignment of a degree of belief to a set of formulae, thereby leaving it open exactly how probable each formula in the set is.

In the model of endorsement, endorsements are the arguments for one's belief in a particular piece of information, where the certainty of a hypothesis is dependent on its strongest endorsement. As it is important to decide what to do if the needed evidence is lacking, one can try either to reduce uncertainty through procuring extra information or finding ways to discount or reduce the uncertainty. Unlike other models, Cohen's system asserts hypotheses even if the preconditions are not satisfied, but it expresses its mistrust in the endorsements of the conclusion. In the event that the conclusion were needed for reasoning the system should embark on a resolution task aimed at increasing the believability of the hypothesis, for example by procuring further evidence.

As endorsements have two different roles in Cohen's model, representing reasons for believing or disbelieving a hypotheses and whether a hypothesis is believable enough for a particular purpose, they can also serve as a guide for problem recovery when there is insufficient evidence for a particular purpose. In that case they can be used to select a strategy for gaining more evidence. Provided the criteria needed for a conclusion to be believed are given, the system will ensure that no recommendations are made which do not satisfy these criteria.

Cohen reinforces his point with the example of King Solomon's proverbial dilemma, which shows that one's certainty in a result should depend on what the result is wanted for. Cohen insists that the believability of proposition changes, and may be adequate for one purpose but not for another; changing goals affect the usefulness of conclusions. The believability of a conclusion may also change with new evidence for or against it. This does however have the side-effect that endorsements which are non-numeric cannot be easily ranked or compared. Cohen admits that this may be inconvenient, but that one can at least establish a partial ordering and he thinks that to compare two completely different and unrelated hypotheses is meaningless anyway.

3.6.3 Comparison of Numeric and Endorsement Approaches

Cohen draws a comparison between numeric methodologies and the model of endorsement with the help of an example from anthropology [WAL78], about three different hominid fossil remains which have given rise to a number of theories about whether they belong to the same species.

The various arguments used for and against such claims are then used in a comparison between the endorsement approach and a Bayesian approach.

Cohen reproduces the method by which the original argument is represented in a Bayesian way, but in order to derive the probability of each hypothesis, it is necessary to assess the prior probabilities and the conditional probabilities for all hypotheses, a data requirement which Cohen considers to be unrealistic. Unfortunately, the Bayesian theorem requires a lot of data, although a design can be made which requires only prior probabilities and likelihood ratios: Cohen still criticises the fact that numbers have to be made up and plugged into the design for it to work.

To present an endorsement-based analysis of the evidence, the example has to be reorganised in the form of inference rules. These rules can then be used in the evaluation of evidence. Thus the question whether two pieces of evidence against a hypothesis outweigh the evidence in favour depends on the quality and kind of evidence. Arguments may be too general or may be based on assumptions which are not necessarily warranted and so on.

Cohen is particularly interested in the representational adequacy and ease of construction, and admits that the endorsement representation lacks facilities to represent degrees of belief that will immediately show wether the evidence against one hypothesis is more damaging than that against another. Cohen also concedes that the endorsement approach does not easily capture arguments which are about likelihood, an aspect which can easily be represented by the Bayesian approach, since it is an argument from probability.

Cohen admits that the inability or unwillingness to specify the relative weight of endorsements can limit the usefulness of the endorsement-based approach. The probability approach will immediately be able to give preference to an argument over another according to the relative strength of evidence without, however, being able to say why it does decide that way. The advantage of the probability approach is achieved at the cost of allowing the user to state beliefs without justifications. Cohen concludes that a hybrid approach between the endorsement and probability approaches would prove more satisfactory. The probability approach has the great advantage of making degrees of belief explicit, whereas in the endorsement approach it is easy to align evidence pro and con a hypothesis and to distinguish different kinds of evidence, but it can be difficult to express how much one is believing something. Whether a proposition is certain enough to warrant action depends on a quantitative action threshold for numerical approaches, whereas for the endorsement approach the action threshold is a qualitative one.

As far as ease of construction is concerned Cohen claims that the endorsement-based approach has a clear advantage. Cohen suggests that the need for a sophisticated Bayesian design results from trying to recast evidence in terms of probabilities. The numerical approach may be much more natural for expressing degrees of belief, but this means relinquishing information which we may need at a later point of time; the information contained in the endorsements is not just intended for justifications alone but is also used during reasoning. Finally, reasoning with numbers is reasoning with hidden and implicit information. To make this information explicit requires that efforts are made to maintain it and this is the cost of the endorsement approach. Cohen does however agree that some kinds of reasoning about uncertainty will not warrant the extra effort.

3.6.4 Implementation of SOLOMON

The model of endorsement, according to Cohen, has been implemented in a system called SOLOMON, which is able to deal with uncertainty in this manner, and is intended as a tool for building rule-based expert systems for domains where reasoning about uncertainty is necessary. In SOLOMON, uncertainty is represented in terms of its effect on reasoning, where the properties change with the style of reasoning and with the availability, type and quality of evidence.

Built as a backward-chaining problem solver, it proves assertions represented in first-order predicate calculus. In order to solve a problem it collects all rules which conclude about it and then tries to prove the rules' preconditions. This is done by generating a task and deciding whether it is worth running it, in order to screen out tasks which do not contribute to the system's certainty in its goal. If the system runs out of ways to prove a task it may generate a resolution task; the conclusion may be too general, but in conjunction with another conclusion from another task it may be strong enough. A resolution task may look for the same conclusions and put them together. Conventional backward chaining mechanisms would have stopped when the original goals and sub-goals were generated and exhausted; these mechanisms would have had to reject the conclusions as being too weak in that case. SOLOMON's resolution tasks go further and allow of an attempt at corroboration.

3.6.5 Conclusions

Uncertainty is present in a considerable amount of information we deal with in everyday situations and, considering the amazing human ability to deal with this problem, Cohen proposes to model this behaviour with a semantic approach. His method works by analysing the arguments for and against the data, together with the potential reliability of the reasoning process built on uncertain information. This approach is one of the most promising recent developments in research about uncertainty management and has made a considerable contribution for a number of reasons:

- It is the first significant departure from the usual numeric methodologies for treating uncertainty, proposing a semantic rather than a syntactic treatment of the arguments behind a proposition.
- It has thrown doubt on the representational adequacy of numerical certainty estimates as hiding qualitatively different evidence.

- It has shown that a qualitative representation can be built, although he does not provide convincing ways to handle it.

Despite the intuitive appeal, the approach is not without problems. His model of endorsement is thorough but at the same time expensive and, by his own admission, some cases may not warrant such elaborate treatment.

Unfortunately, with respect to the problem of endorsement proliferation over chains of inference, the point has to be made that long chains of inference will make the system come to a grinding halt, although it is doubtful whether in the majority of problems the chains of reasoning will, in fact, be very long. This will of course depend on the kind of applications, but it may create a serious problem, especially when one considers the flood of information a system might have to deal with in a real-world environment. Cohen himself admits that there is a need for steps to be taken to stop the proliferation of endorsements, but he lacks the right heuristics to do so. Similarly, he does not provide credible heuristics for relative weighting of endorsment lists, short of a rather vague appeal to pairwise comparisons.

Finally, despite the fact that Cohen presents the model as a general purpose system, it is still strongly dependent on domain specific rules. This makes the whole model less general than it may appear at the outset and less applicable for those areas where these domain specific rules are unavailable as yet.

3.7 Conclusions and Perspectives

The management of uncertain information in general, and from human sources in particular, is a largely unexplored field of research in its own right. At the same time it also touches on a number of current and established areas, although that connection is rather tenuous. The reasons for this is that the goals have been very different and the use of common terminology tends to obscure this fact.

Thus a great deal of work in conflict resolution by and large addresses conflict from a system design perspective. Measures of expedience and economy, which assume correct data, based on a more or less mechanic response, are the most commonplace techniques: these may be justifiable given the nature of the problems involved. The only notable exceptions which are somewhat more interesting from our perspective are Fayyad [FAY88] and Galliers [GLL87], [GLL88].

Fayyad has seen the importance of enabling a system to learn control information so that it could improve its performance as time progresses. We, likewise, endorse this design precept although the particular domain we try to deal with is obviously different from that of Fayyad. In that respect, Galliers' work is more closely related, as it specifically considers an environment of multiple human agents. Her model does however not deal specifically with the management of uncertain information, but is concerned with a model of motivations and strategies to achieve goals which are the result of motivations.

Current work in plausible reasoning addresses itself to the impact of uncertain reasoning techniques, on one side, and to the management of uncertainty in inferential databases and expert systems, on the other side. Both aspects will finally need to be included in a comprehensive approach to deal with uncertainty, but we are at present only interested in the latter. Considering the current trends in research in both areas the emphasis has been by and large on dealing with the propagation of uncertainty in system of beliefs, rather than how these uncertainties are initially derived. The one notable exception to this rule is Rescher's model for plausible reasoning, which assesses prior plausibilities based on its knowledge about the reliability of the respective sources. This is an important starting point, akin to our approach, which likewise wants to initially assess uncertain information. The one disadvantage of Rescher's model is that his source models are static and depend on being given prior levels of reliability of its sources. Rather than providing a final solution, this approach only eliminates the problem by transferring it onto a different level and a more adaptive method would be required for a truly autonomous system.

A system, faced with a continual flood of new information which needs to be incorporated into the existing set of beliefs, will very likely find itself confronted with conflicting information from time to time. In order to maintain a consistent view of the world there will always be the need for good consistency recovery techniques. This subject has been addressed by the truth maintenance community and other researchers interested in conflict resolution. Unfortunately, the world model generally adopted by truth maintenance systems is very simplified, as propositions are usually boolean and part of inferentially connected belief-spaces. We consider this to be insufficient in a real-world environment.

In the conflict resolution community, conflicts are more often than not resolved on the basis of expedience and economicity rather than real world considerations. Again this will not be sophisticated enough for our requirements. In general, the impression generated is that conflicts are by and large considered as a problem to be detected and eradicated. While we agree that it is a problem to be addressed, a conflict is at the same time a useful trigger to indicate that something is wrong with the world model and therefore an opportunity to learn about it.

One of the most widely used measures of uncertainty in AI is built on considerations of probability. Of these, Bayes' theorem features most prominently and provides an easily applicable method for the propagation of probabilities in inferentially connected belief-spaces. It does however depend on being given a complete set of initial probabilities. These initial values are often difficult to obtain and it is those our project specifically addresses. Our short excursion into probability theories also shows that they are fraught with conceptual difficulties. It is also remarkable to see that most researchers using probability never seem to state which school of probability theory they subscribe to, and why.

This has also been noticed by Cohen [COH85] in a most promising and innovative approach to uncertainty management. He argues that probability theories have probably found unquestioned use because of practical considerations and lack of better alternatives. He proposes in turn a qualitative rather than a quantitative approach. We too question the adequacy of approaches based on probabilistic or quasi-probabilistic considerations. While Rescher's model for plausible reasoning is a good starting point, it is falling short in a number of respects and will require modification if it is to be applied to uncertain information from human sources. We would like to combine the basic principles with Cohen's model of endorsement. Rescher's model needs to be enriched to become an autonomous system which can order its own affairs, while Cohen's model is too ambitious and vague for the majority of everyday decisions and needs to be complemented by a quick and yet considered methodology to survive in a real-time, real-world environment. We therefore suggest that a possible solution to the particular problem we are considering will lie somewhere between these two approaches.

Chapter 4

Principles of Uncertainty Management Through Source Control

4.1 Introduction

Having stated the problem area we are concerned with in the present enquiry in Chapter 2 and having examined current developments of research in Chapter 3, in the present chapter we will discuss the principles of source control before we can embark on providing a design for a system which can deal with the problem we described. As can be seen from the dichotomy between Cohen's and Rescher's approach, there are two basic approaches to deal with uncertain information from human sources, namely by analysis of the arguments or by analysis of the source. As we favour the latter strategy, we now need to analyse the principles involved in this very successful human way of solving problems.

As the task ahead is rather large and complex the following overview shows the subdivisions of the chapter:

- the source control mechanism in the context of current research
- constraints on the behaviour

- the strategy of the source control model
- concepts and definitions
- principles of information evaluation
- principles of conflict resolution
- principles of the enquiry
- principles of source re-evaluation

In the following section we first put the source control approach into the context of current research. In the subsequent sections we shall follow our methodological strategy by defining the fundamental constraints on the behaviour of the proposed model, followed by an explanation of how the behaviour can be produced and controlled. We then briefly present some of the basic concepts involved before examining the principles of belief formation and conflict resolution in detail. Finally, we describe the principles which govern the re-evaluation of the source models used in the source control model (SCM). This, in turn will give us the requirements on which the design can be built to demonstrate that the model can actually be reconstructed, to produce the behaviour we initially claimed.

4.2 The Source Control Mechanism in the Context of Current Research

The goal of the present project is to provide a mechanism for the formation and maintenance of beliefs about uncertain information of a general nature, as obtained through human sources. Given a general reasoning system operating in a natural language environment to model a human agent capable of natural language understanding and interaction with other human agents, such a general system would require a sub-system to deal with the uncertainty of information and to form and maintain beliefs. The purpose of a source control model is to control the belief formation process about information from human sources, and thereby order the beliefs of a general reasoning system.

In order to embed such a model within the current field of research, we need to consider the requirements, which are the result of our definition of the problem of uncertainty. This task of uncertainty management is influenced by three different considerations:

- the nature of uncertainty
- the nature of human sources of information
- the management of systems of belief

In order to arrive at a reasonable evaluation of the uncertainty one needs to consider the nature of uncertainty. As we have seen, there are a number of potential causes which contribute to the problem. We need to consider the potential cause of uncertainty, the problem the information may be used for and the impact of the uncertainty of the respective piece of information on other information it may have a relevant connection to.

Human sources of information account for a major proportion of information we usually deal with. Human sources also display a number of interesting, and at times unfortunate, properties which need to be taken into account. Thus sources vary in their competence and willingness to provide reliable pieces of information, a factor which has to be taken into account when evaluating the uncertainty of information from them.

The process of information evaluation cannot be considered in isolation, but includes the incorporation into a system of beliefs. This is not always without problem as the new potential belief can be in conflict with other beliefs, and in the interest of self-consistency action will be required. Thus, in the case of conflict decisions have to be taken as to what to believe or not to believe, whether this affects other beliefs or our opinion of the sources involved.

Uncertainty appears to be the product of a variety of factors and a reality which has to be dealt with. This recognition has led researchers in a variety of areas of AI to introduce various representations for different reasons. Truth maintenance is concerned with ordering beliefs and maintaining the consistency of systems of belief. Traditionally they have tended to be built on Boolean beliefs connected in an inferential manner and to determine the impact of a propagation of beliefs which are the result of the addition of new information. The introduction of degrees of uncertainty instead of Boolean indices has not been greeted with overwhelming enthusiasm, which can probably be put down to the fact that the mechanics of truth maintenance depend on propagation algorithms closely built on the precepts of classical logic and propositional calculus which cannot easily be changed to accommodate degrees of uncertainty. The same problem appears to apply to the introduction of non-monotonicity of beliefs [DRE87/8], although their introduction causes more theoretic than practical problems. Since change is a reality in our everyday experience the SCM, too, will need to accept a non-monotonic view of the world, but since it is not built on classical logic, the associated problems do not occur.

There is also another difference between truth maintenance and the SCM. Whereas truth maintenance is principally concerned with logically interconnected systems of belief, the SCM deals primarily with a system of beliefs where the constituent members are more or less atomic. Although some of these beliefs may be interrelated, not all of them are likely to be connected in the way which is typical of TMSs.

Despite considerable difficulties, measures of uncertainty have been introduced by [FAL87], [AMB87/8], [FAN88] and [KAI89]. Apart from [FAN88] who works with the concept of "possible goods and nogoods" which circumvents the introduction of uncertainty by delegating it to the realm of hypothesis, the remaining camp splits into two, namely [FAL87] and [AMB87/8] who advocate a probabilistic representation of uncertainty and [KAI89] who proposes a representation built on [RES76]. The advantage of probabilistic representation lies partially in the fact that there are relatively straightforward algorithms available for the propagation of probabilities. [COH85] puts the apparent popularity of probabilistic methods down to their ease-of-use and the lack of better alternatives, while questioning their representational adequacy. On the other side, [KAI89] uses an adapted version of Rescher's plausibility index, thus basing his model on a consideration of sources of information and their reliability. This is also the general line taken by the source control model.

Research in plausible reasoning has been concerned with the management of uncertainty in a number of different contexts. Our particular interest is, however, restricted to that subset concerned with the management of uncertain information as opposed to the management of uncertain reasoning techniques. In this field, the SCM is especially concerned with belief formation about information from human sources. Consequently, the subset of plausible reasoning concerned with the enlargement and exploitation of existing sets of beliefs with the help of uncertain reasoning techniques are of less interest. This is not to diminish their value, as such techniques are the source of much successful human reasoning indispensable for a general reasoning system, but they are outside the scope of the present project.

Similar to other areas of research, with respect to representations of uncertainty, the same two basic approaches are evident: [SHH86], [PAA86] and [HOS88], for example, propose probabilistic measures of uncertainty, whereas [BES87] and [POZ87] suggest uncertainty measures built on human methods for plausible reasoning of [RES76]. Finally, the most interesting has been proposed by [COH85], advocating a qualitative analysis of the arguments which support a given belief. His argument about the representational inadequacy of probabilistic measures of uncertainty and the importance of taking into account the qualitative differences between reasons is endorsed by the SCM, however with two differences: Firstly, although we agree that a quantitative index cannot fully represent qualitative differences of these reasons, we further question the meaning of probability in general as an adequate representation as understood in the sense of Chapter 2 and as argued in agreement with [KYB61]. Secondly, there is a significant amount of information we have to deal with which does not warrant the elaborate treatment proposed by the model of endorsement. When the information is not important or when, as a matter of urgency, a quick solution is required, a faster and more efficient methodology will be necessary. The SCM proposes to fill this gap and the relationship to the model of endorsement is consequently envisaged to be symbiotic as both approaches appear to be highly complementary; while the model of endorsement provides a thorough qualitative analysis of arguments, which requires time and effort, the SCM proposes to provide a quick, efficient and reasonable decision when time and effort cannot be spent.

[RES76] and his school are built on a numeric index of uncertainty but radically different from probabilistic considerations. Deploring the inadequacies of probabilistic methods they advocate a plausibility index which makes the uncertainty of propositions dependent on the known degree of reliability of the source in question. The SCM endorses this view in principle, although there are a number of important differences: whereas the model of plausible reasoning deals mainly with inanimate sources, the SCM specialises in human sources. The model of plausible reasoning is static and dependent on being given prior reliability levels of sources, whereas the SCM is a dynamic and adaptive mechanism deriving much of its strength from that feature. Rather than depending on being given reliability indices for sources the SCM will develop its own indices from scratch and maintain them adapting as time progresses, by learning about its sources and their behaviour.

The model of endorsement is devoid of such an explicit source model, deciding beliefs solely on the basis of the subject matter, namely the arguments for and against a particular proposition. Unfortunately, the model is dependent on a considerable amount of domain specific knowledge and we argue that in a considerable number of situations we may not have sufficient expertise, but still have to take decisions about the credibility of the information. We also argue that human beings quite successfully adopt the strategy of forming their beliefs by source considerations in these situations, which we attempt to model. We believe that this is a strong argument for bridging the gap between the basic motivations of Rescher and the model of endorsement with the SCM.

Apart from providing a model for belief formation this should also prove to be a profitable strategy to address problems of conflict resolution, which has been an issue with truth maintenance as well as a separate conflict resolution community. As truth maintenance is about maintaining self-consistent systems of beliefs, the elimination of contradictions is of primary concern. As we have seen, this is achieved by an identification of the conflict set and the choice of a 'culprit' whose elimination will hopefully restore consistency. TMSs differ in the strategy for choosing culprits, which can be achieved purely randomly, by user choice or on the grounds of expedience such that the elimination of the culprit causes a minimum of disturbance to the existing belief system. Incidentally, the term 'culprit' is very revealing as it suggests that the culprit is chosen not necessarily for its guilt but for other reasons. Indeed, the focus of the majority of TMSs on internal consistency rather than consistency with reality seems to reinforce this suspicion. The primary objective of the source control mechanism is to keep in touch with reality and a solution to conflicts will therefore need to be oriented on external reality rather than mere internal consistency.

The majority of work in conflict resolution itself is however of an altogether different nature, such as conflicts amongst competing rules or conflicts in planning rather than conflicts of beliefs and information, which are our concern. About the only exception can be found with [GLL87] [GLL88] who has addressed conflicts between competing strategies of different agents. Her model shows how to devise strategies and influence other agents so as to achieve one's goals and avoid potential conflicts of interests or goals. The model is significant in that it provides a simple and elegant method to model motivations and actions important for a general reasoning system which has to survive in a highly competitive world. Such a system would be relevant to the SCM as a tool to detect and determine motivations, which could be used in the assessment of trustworthiness of sources (as well as their reliability). As Galliers model is about motivations it could feed the SCM with the necessary information to assess a source's trustworthiness in general as well as in particular, but is not currently available in a suitable form. The remainder of research in conflict resolution is about local solutions to local problems rather than the more general, long term objectives of the SCM and the connections are therefore largely insignificant.

Finally, there is also the subject of multiple agents, which is of considerable importance to the SCM since it has to deal with a number of different sources and more significantly with the problem of ordering its beliefs when conflicts occur between pieces of information supplied by different sources. Current research on this problem is particularly meagre although there has been a recent surge in interest in multiple agents [ALV88]. Despite a similar terminology the connections are, however, tenuous as a great deal of research uses the term either to describe systems constructed from a number of largely autonomous components or if the world model identifies multiple agents then these are largely used to describe non-human agents like vehicles in a simulator.

Galliers, as we have seen, uses multiple human agents in her world model where each agent has its interests and goals and may compete with other agents in the attempt to further its goals. The conflicts are however about goals and motivations rather than about uncertain information which need to be addressed in a different way. Conversely, the other pieces of research about multiple agents in the context of conflict resolution deal with competing agents in problems of routing and the need for replanning to avoid conflicts in the sense of collisions.

Although Cohen acknowledges multiple sources of information his interest is to decide conflicts by argument and sources as such are thus only of secondary importance. Rescher, by contrast, shows a stronger interest in this subject but since his model lacks adaptive features and hence also the relevant interest to explore, his model decides arguments amongst competing sources by strength alone.

If uncertainty in the sense of Chapter 2 is to be modelled in a real-world scenario, multiple source conflicts are inevitable and methodologies to deal with such situations are therefore needed. On one side, Rescher's answer to the problem lacks the necessary sophistication and on the other side Cohen acknowledges his difficulty in the face of conflicting propositions from different sources with complex endorsements. The SCM thus proposes to presents a significant advance on the subject, which has not been covered hitherto. This establishes the position of the SCM in the following way:

1) The SCM will be built on a non-probabilistic understanding of uncertainty. The mechanism will use a coarse-grained index to express strength of argument and conviction which represents the arguments behind it in an iconic fashion. This allows the mechanism to use the index for most operations while enabling the model to look at the actual arguments should that be necessary.

2) The SCM proposes uncertainty management through source control, taking its starting point from Rescher and Garigliano. The treatment is however significantly more complex and tries to bridge the gap to Cohen to share the task of uncertainty management where problems which do not warrant an elaborate treatment in Cohen's model, or where the expertise required is not available or where the problem is sufficiently urgent to make an endorsement analysis unfeasible.

3) The SCM wants to make a significant advance on the problem of conflicting information, based on an elaborate treatment of conflicts with an adaptive model which learns from the advent of conflicts rather than suffering from it.

4) The SCM should operate as a sub-system of a general reasoning system. Being closely linked to it along the lines of Garigliano, the SCM is designed to maintain a system of beliefs representing the systems view of the world. It could be paired with a model of endorsement and could benefit from a model of motivations like that of Galliers to maintain its views about the trustworthiness of sources as opposed to their ability.

4.3 Constraints on the Behaviour

Having put the source control mechanism into the context of the general field of current research, we now need to describe the basic constraints on the model in order to put it into context with rational agency.

Looking at the problem of forming beliefs in a situation where we are faced with different sources providing us with a great variety of different information there are a number of competing motivations which rational agents should expect to control. These constraints can be summarised under five headings:

- maintaining consistency in the face of contradiction
- maintaining interaction with the environment
- improving one's performance
- active problem solving
- ability to take decisions under poor conditions

Being continuously faced with new information, we are bound to find ourselves in contradiction from time to time and as life has to go on we will have to take steps to recover from contradiction to regain a consistent world model. Since the source control mechanism is faced with the same situation of having to form beliefs, it has to take decisions and re-establish consistency in its system of beliefs, however serious that contradiction may have been and thereby display a certain degree of robustness. It is also important that the criteria employed in deciding contradictions try to orientate themselves on reality rather than consistency merely within the world model, as the latter strategy would mean that the agent would start to live in a fantasy world which would hardly be a desirable behaviour to model.

Obviously the easiest way to maintain consistency is to stop interacting with the environment once the world model is free of contradiction. Unfortunately this would result in some form of solipsism and produce a pathological and useless system and therefore in order to produce a reasonable response of the system, the maintenance of interaction with the environment is paramount. This needs to go beyond the mere exchange of, say, speech acts designed to keep up the appearance of communication, but requires that communication causes a reasonable response in the recipient. This implies that if information is to be rejected it is done for a different reason than just because it would disturb the consistency of the system of beliefs.

Another behaviour which is expected of a rational agent is a certain ability to learn from past experiences in order to improve one's performance for the future. If a system stopped being able to learn and started to fall into the same trap time and time again it could be accused of becoming senile. It is therefore important that a system of the type we are interested in has some ability to learn about its environment and to improve its performance. Even though it may not be perfect in the beginning it would at least redeem itself as time progresses, which is not an unreasonable behaviour to expect from human agents.

There is also a certain amount of general curiosity evident in rational agency. The entire history of scientific investigation is an example of the insatiable desire of rational agents to find explanations. When we are encountering problems with evaluating information we are usually not content to see that there is a problem and to aim at a quick solution, but we usually want to find out the true root of the problem. We do not like to be puzzled and if there are problems we do not just want solutions but also an explanation as to why the problem arose. We therefore will want the source control model not just to be passive but have some strategies to explore.

The information we get in everyday situations is often not just uncertain but also incomplete. It is also an important property of human agents to be able to form beliefs in a variety of different circumstances and to deal with information of almost arbitrary complexity. Often we find situations where the information will have gaps, but human agents have a considerable capacity to deal with information despite obvious deficiencies. This flexibility to perform under poor conditions will therefore also be required from a computational model which is to deal with realistic situations. Since we are dealing with human sources, we also have to consider that the source may lie to the system or try to deceive it in other ways. The prospective system will therefore also have to try to distinguish truth from falsehood. This process is not easy and a successful solution will require a rather active engagement in belief formation and revision in order not to just acknowledge problems but also to attempt to find out where the roots lie to find an appropriate solution.

These fundamental requirements of the system need to be translated into actual behaviour. Viewed from this level, the system's primary objectives and hence behaviour can be summarised in the following way:

- formation of beliefs about data
- formation of beliefs about sources
- resolving conflicts in the system of beliefs

The most important task the system has to perform, is to form beliefs about uncertain information from human sources. As sources typically differ in their competence and cooperation, a great deal can be learnt about the information by looking at the source. In order to do so, the system needs to generate and maintain models for each of the sources involved. These source models record various properties of the source, which could have an influence on the information from that source, and help in the evaluation process. To make this work, there is a need for an important feedback mechanism where the source model is used in the process of evaluating information from that source. Information gained from that analysis can then be fed back and in conjunction with other information from that source and other information from the existing system of beliefs be used to look for patterns in the behaviour of the source or indications about its properties. The result of this analysis can subsequently be used to modify and improve the source model. Finally, the belief formation process may be complicated if that new piece of information is in conflict with other beliefs already held by the source control model. In that case the mechanism will have to decide on a solution and whether that may have an effect on the source model.

4.4 The Strategy of the Source Control Model

The constraints provide a framework for the behaviour but do not give us an indication as to how, in general, this behaviour can be achieved with a source control model.

From the introduction we can see that there are two different ways of evaluating uncertain information:

- we can analyse the arguments which are presented for the information
- we can make our belief dependent on our opinion of the source

Cohen adopts the first strategy while we pursue the second. Although the second line is perhaps less accurate and less objective, it nevertheless is both a very powerful strategy which gives results much faster and can deal also with cases where there are no explicit arguments or where we are unable to understand them. For example, if I go to my GP because of a complaint I have, and I get a prescription I will have to put my trust in his recommendations as I do not have enough medical knowledge to challenge his decision. He may even give me the arguments why I should do what he requires but I still may not understand them.

It is a common phenomenon that as we make acquaintances and get to know people we learn where their strengths and weaknesses are and whether we fundamentally can trust them. The model we build about their behaviour is revised as new evidence emerges and is applied to any new information we get from them. Thus, to stay with our example, I may follow the prescription and find that it works very well and the next time I have to see my GP this experience will reinforce my confidence in his competence and I will be less likely to question his instructions.

Given that we get a piece of information and a source model, there are a number of things we try to do:

- we try to show that the information is compatible with the source model
- if there is a problem we try to weaken the information to fit the source model
- in the process we try to do a surface analysis and not to go into a full analysis of the arguments but use pragmatics and shallow heuristics first

Source models represent our considered opinion of the source and therefore encapsulate our expectations of the source's behaviour. As rational agents want to have a grasp on reality which means not being surprised by events which do not fit with their world model the first reaction when faced with new information is to see whether that fits with the source model. If there is no problem, then the process of belief formation is relatively straightforward.

Considering that source models are an important tool for our grasp on reality and to evaluate information we have a strong interest to make sure that they are correct and as human beings tend to prefer an ordered and stable world model they also do not like to have to change them. Thus, if there is a clash between the information and the source model, the usual reaction will be to consider the source model to be stronger than the information and therefore to try to weaken the information to fit the source model rather than to change the source model to fit the information. This is perhaps not surprising as source models tend to be the product of a long term experience with the source.

As we usually do not have much time to spend on each instance of belief formation, this suggests a layered approach whereby we try to do a shallow analysis of the information, just enough to be able to abstract the general type and properties. As long as we do not encounter difficulties or if the situation is not too important we will be able to deliver a quick solution. If there is a problem, however, or if the case is serious, rational agents will want to find an explanation.

This suggests that a source control mechanism can be organised in the following way:

- Importance Analysis
- Information Evaluation
- Conflict Resolution
- Enquiry
- Source Model Re-evaluation

In order to steer the source control mechanism to behave in a reasonable fashion it is necessary that the mechanism can make distinctions between what is important and what is not, in order to decide how much effort should be spent on any particular case. Thus if there is a problem with a trivial matter then one should not lose too much time over it and alternatively if the case is more significant one may want to get to the root of the problem. An importance analysis will therefore at the beginning need to find out whether there are any indications to take an interest in the matter to start the analysis process. Given that there are problems becoming apparent later on, one needs to decide whether it is worth carrying on, or indeed whether it is important to find an explanation for a case which has the potential of breaking a strong source model.

During the information evaluation process the model needs to determine the properties of the information and compare them with the respective parts of the source model to see whether the information is consistent with the model. Given that there are no obvious problems the process may quickly come to a decision as to whether to believe the information and how much. Alternatively, if there are problems, then given there is still enough interest then the process will have to try to get to the root of the problem and find an explanation or at least suggest where a solution can be found. In the evaluation process generally the information and source considerations go hand in hand.

Supposing that at the end of the entire process the information has not been completely rejected, the next step will be to try to fit the information into the system of beliefs. This is the second dimension to the source control mechanism. In the first dimension the information needs to be accommodated in the source model and here the information needs to be consistent with the established system of beliefs. If there are no direct connections to other beliefs the information can be introduced without problems. Conversely, one will have to re-examine the respective pieces of information and the sources involved to see whether there is a problem with the information or whether there is a problem with the source models. A shallow examination may reveal that the nature of the information is such that a difference in opinion between the sources involved is possible or that the information may have a 'limited shelf life'. Again the same considerations about the importance of the situation apply and in serious cases the mechanism will want to find an explanation for the problem as well as a solution. The mechanism will only be able to hold one opinion and therefore the mechanism has to decide what to believe as a result. A change of opinion is not that significant but if it is a matter of having to acknowledge that the system has wrongly assessed something important, the matter is more serious.

Finding an explanation to a problem amounts to being able to unify apparently conflicting items in the overall world model of the mechanism. Not to be able to find an explanation seems to suggest that the mechanism may be losing its grip on reality, and that is a serious situation which needs to be resolved. As we have said earlier, the mechanism may not be able to find an explanation but may nevertheless have an idea where a solution can be found. This information can be used in the process of starting an enquiry whereby the system actively tries to resolve the problem by seeking assistance from other facilities of a general reasoning system or from sources directly. As this is a rather complex and costly process, one needs to consider carefully whether that expense can be justified.

Finally, at the end of all analysis the question will continuously arise whether the source models are adequate or whether the previous analysis suggests that changes have to be made. Incidentally, the discussion of what happens when problems arise may have created the impression that only negative information has an effect on the mechanism. It is important to note that positive as well as negative information will affect the source model re-evaluation. While a particular problem may suggest that the opinion about a particular aspect of the source may have been too optimistic, an unexpected piece of information may also suggest that the source perhaps has expertise in a field the mechanism was previously unaware of and may prompt the mechanism to enquire whether the source actually has the suspected expertise.

We have explained the considerations which drive the source control mechanism and presented an organisation which performs in that manner and which also satisfies the constraints. The conceptual model of the mechanism can thus be represented in the following way:

This is the first decisive step towards the solution. After briefly discussing some fundamental concepts, we shall detail the principles which govern the actual process of information evaluation, conflict resolution, enquiry and source model re-evaluation.

4.5 Concepts and Definitions

The source control model takes its starting point from [GAL89], where, amongst other things, a model for conflict resolution through source control was presented. The model proposed there suggests assessing information through an adaptive model of sources. Although the basic concept of source control is preserved, the source control model which is the result of the present project is radically different from that described in [GAL89]. A number of intermediate results along the road of evolution have been presented in [GAB88a], [GAB88b] and [BGA90].

In [GAB88a] an analysis of the factors involved in the assessment of uncertainty and belief formation is presented together with a number of suggestions for solutions. Subsequently [GAB88b] [BGA90] the model was substantially enhanced and extended to deal with multiple sources which had not been addressed previously. The aspect of multiplicity of sources has since then enjoyed an increasing popularity in current research.

There are a number of concepts and definitions which need to be introduced before we can proceed to explain the functionality of the source control mechanism:

- uncertainty and belief
- classification
- source models
- information models
- strength of conviction
- conflict of information

After giving definitions of the basic concepts we can then present the principles of information evaluation together with a description of the functionality of source models. In the final step the principles of conflict resolution are explained.

4.5.1 Uncertainty and Beliefs

Being faced with information of indetermined uncertainty and the necessity of forming beliefs, the assessment process is difficult, but nonetheless important, as the information may influence our decisions and actions. Thus, if the weather forecaster tells me that there will be rain this afternoon I may have to decide whether to take along an umbrella. Human beings are generally able to form opinions about the reliability of information and their suitability as a basis for action and it is the purpose of the source control model to model this successful behaviour as a building block towards the development of a general reasoning system. On the most basic level the task is described in the following diagram:

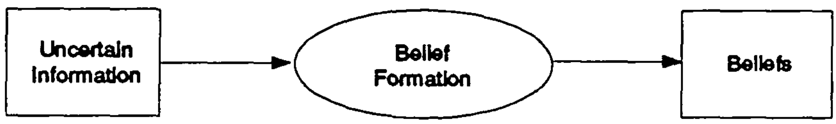


Figure 2: Context of Belief Formation

As a consequence, a model dealing with this problem takes as its input information of indeterminate certainty, performs a process of belief formation, and produces, as output, beliefs. Before we consider how this is to be achieved, we need to consider what we understand by uncertainty and belief.

UNCERTAINTY

Considering the complexity of the uncertainty problem as described in Chapter 3 and the restriction of the area we try to model, we are dealing with a subset of uncertainty. This subset can be described in the following way with respect to the divisions made in Chapter 3:

- Changing Environment: The SCM is designed to be domain independent and a great deal of very specific knowledge about the respective domain is required to determine whether a change could have taken place. If the general reasoning system is able to produce information to that effect, the SCM will be able to take that into consideration in the evaluation process.
- Problem of Communication: The SCM deals with the event-based representation of information which is the result of pre-processing by a natural language processor. The information is taken from the

conversation between system and source and we assume that there are no specific problems of communication. To deal with that problem would require a special theory of communication and is beyond the scope of the present project.

- Source: Information can come from a variety of sources which need special treatment depending on their properties. The SCM deals with information from human sources and the model is based on a machinery specifically designed to deal with information of that kind. Consequently the definition of uncertainty has to be seen largely in this context.
- Background Interpretation: To deal with background theories in general requires a great deal of scientific and domain specific reasoning. This would be more along the line of uncertain reasoning techniques we are not concerned with. However, the concept of interpreting information from human sources with the help of a source model can be considered as a background theory, in which case the SCM also touches on this category.
- Reasoning: This, again, is the realm of uncertain reasoning techniques which should be treated separately by a more appropriate approach, be it with management of uncertain reasoning techniques or an endorsement approach. The SCM is however able, given a general classification, to make general decisions. For example, given a long chain of reasoning on uncertain information and with uncertain techniques usually no strong claims can be made about the conclusions and if such claims are made the model can express its doubts about the information. The ability of the SCM in this area will therefore be very restricted.
- Accuracy, Precision and Robustness: These categories are not just dependent on the information itself but also on what the information is to be used for and is therefore very context sensitive. This requires a separate theory of action. The SCM is however primarily concerned with forming beliefs about information rather than what the information is to be used for and uncertainty in this sense is not covered by the SCM.

- **Importance:** Considerations of accuracy, precision and robustness will also feature in an analysis of importance, although importance is a wider issue than just that. Although the SCM does not include a theory of action the issue of importance features in a more basic form as the system may have to economise its resources to accord a very detailed analysis only to cases which warrant the treatment. The SCM will therefore have to find justifications to perform an exhaustive analysis.
- **Spreading of Uncertainty:** The uncertainty of a piece of information may not just be restricted to itself but it may spread to other information connected to it. To establish whether and how far the uncertainty will spread requires a considerable amount of reasoning to see whether there is a connection and whether it is relevant. The SCM is designed to attribute beliefs to single pieces of information rather than the propagation through a system. At the same time the SCM also considers situations where the information to be evaluated has a direct effect on individual beliefs held by the system.

This definition of the goals of the SCM excludes a number of issues which should be dealt with elsewhere and concentrates on the task of establishing initial beliefs about information from human sources. This perspective does however have a considerable potential as information from human sources have not been dealt with in this form before and will enable a system to exploit this area. Also, given that there are still many areas which need a great deal of specific heuristics and have not been solved we can indirectly draw on the expertise of sources. Although first hand information is better than second hand, second hand information is still better than no information at all.

This may appear to be rather pragmatic, but this is needed since the world is of such a complex and unpredictable nature that exhaustive, scientifically principled evaluations are either not possible or we do not have the time or leisure to pursue them. We therefore have to make do with suitable approximations and simplifications.

This state of affairs implies that we can hardly ever claim to have immutable and absolute knowledge. More often than not, we have to make do with beliefs which we form about the information we are confronted with.

Those adopting probabilistic measures of uncertainty in AI tell us much about the propagation of beliefs in connected belief sets, but few researchers actually tell us what their definition of probability is and hence their definition of uncertainty to which they seem to equate it. Be this as it may, Kyburg [KYB61] has shown that there are conceptual difficulties with all major schools of probability. The main problem is however that the advocates of probabilistic theories in AI by and large assume that initial probabilities are given, whereas we are interested in the very establishment of initial uncertainties. We cannot assume that they are always provided, and if so, whether they are correct. Compared to our original division of the problem of uncertainty, probabilistic theories have no concept of source but tend to side with reasoning, background theories and the spreading of uncertainty in chains of information .

Cohen, on the other hand, argues that uncertainty is a function of the arguments behind a proposition and the purpose for which it is to be used. Uncertainty thereby becomes relative to the purpose and the arguments. This implies that, with respect to our original classification, the model of endorsement is primarily concerned with reasoning as the cause of uncertainty and also with respect to the problem, as uncertainty is also evaluated with respect to the purpose the information is to be used for. Although Cohen also deals with chains of information he acknowledges that his model can at present only deal with relatively short chains due to a problem with the rapid increase in complexity which may bring the model to a grinding halt.

According to Rescher's model the uncertainty or plausibility of a piece of information is the product of the plausibility claimed by the source and the systems knowledge about the reliability of the course. The "pars deterior" principle then implies that the result of reasoning performed on pieces of information with given plausibilities is at least as strong as its weakest premise.

This makes the uncertainty of information dependent on the supporting source and he also deals with a model for reasoning roughly based on probabilistic principles, although with important modifications. This can be considered as some form of background theory by which levels of uncertainty are determined and since the model is also about the propagation of uncertainty of chains of information the problem of size is also addressed. As opposed to the model of Cohen, there is no concept of uncertainty in the context of the problem the information is to be used for.

This shows that the main approaches differ greatly in the aspects of the problem of uncertainty they cover. Each approach is specialised to deal with the problem they are addressing and consequently, despite initial appearances, there are limits to a meaningful comparison.

BELIEF

Since the SCM is primarily about belief formation we also need to consider what we understand by belief. The term belief has furthermore been used in a number of different ways in AI and we need to clarify our position.

Considering the problem of uncertainty it follows that there are limits placed on our capacity to correctly assess the uncertainty. As a result we form beliefs which are the subjective response to the perceived uncertainty which can be defined in the following way (at least as far as the SCM is concerned):

A Belief is a conviction in the correctness of an uncertain piece of information. The strength or quality of the conviction is dependent on two aspects:

- an analysis of the arguments for and against a belief in the information
- dependent on our opinion of the source's competence and trustworthiness to provide reliable information

These beliefs may vary in strength, depending on the evidence supporting them and may vary as new evidence emerges. Beliefs therefore have to be seen relative to the evidence supporting them.

This stands in contrast to beliefs as perceived by a large section of the truth maintenance community who deal with Boolean, monotonic beliefs. Although there is a notion of evidence for a belief in a particular proposition by way of justifications and assumptions, the proposition is categorically believed if there are valid justifications and denied if there is contradictory evidence. There is also a strong element of monotonicity as a 'change of mind' will lead to a contradiction leading to an emphatic denial of the offending propositions.

Apart from approaches which remedy the monotonicity problem [FAL82] but which leave the other problems untouched, there have been developments to introduce uncertainty by way of probabilistic approaches. These, like probabilistic approaches by the plausible reasoning community allow for the representation of degrees of belief as probabilities. As has been argued before, those advocating probabilistic approaches appear to treat belief and probability as identical.

Rescher's plausibility theory does give an explicit definition. As he modifies the source's strength of assertion in a proposition by the system's knowledge about the degree of unreliability of the source, he seems to support the idea that beliefs are determined with the help of a source modelling technique.

Finally, Cohen's model does not explicitly define belief, but relies on the operational semantics of the model. Consequently, there is no apparent distinction between belief and uncertainty, but both rely purely on an evaluation of the evidence for and against, together with the intended purpose of the information.

The SCM adopts the position that a belief and its intensity or degree of conviction are dependent both on one's opinion of the source and the arguments for and against. Which approach to take in the particular case will depend on the situation, because if there is no time to carefully analyse the arguments or if the situation is not important enough it may be advisable to use a source model approach whereas in the converse case an analysis of the arguments may be more appropriate.

4.5.2 Classifications

From the concept of source control we can also detect an important phenomenon. Human beings have a strong tendency to try to understand their environment by way of establishing classifications and to make sense of new items through trying to fit them into the system. Biology and archaeology are typical examples of this approach as they try to make sense of new finds but the mechanism is also evident in other areas. Thus when I first go to my GP with my complaint, I may not know him, but I will have a reasonable idea of what to expect. He will have a great deal of medical expertise, that he knows how to deal with people, that he will act with integrity and so on. In effect I will ascribe a number of properties to him because I know that he belongs to the class of doctors and that doctors have certain attributes. Again as I get to know him I will revise some of my initial expectations, but probably not a great deal.

The consequence of this phenomenon is that we are dealing with three kinds of knowledge as far as source models are concerned:

- default properties due to class membership
- actual properties due to experience
- default properties due to ignorance

The default properties due to class membership may not be exactly right but are a good approximation in the absence of properties we know exist through our actual experience with the source. Alternatively, if we do not have any specific information from either we can still expect average capabilities under the assumption that if there were special areas of expertise we would have heard about them.

4.5.3 Representation of Sources

From the previous section we have seen some of the aspects involved in the source models. We can now describe the source model required for this type of information evaluation. There can be great difference between sources with respect to their propensity to provide reliable information which needs to be reflected in their representation. At a general level, there appear to be two considerations:

- the source's *ability* to provide good information, as opposed to
- the source's *trustworthiness* to give the best information it is capable of

A source may be willing to give the best information available, yet being incapable through lack of expertise or lack of sound reasoning. At the same time a source may be very competent at evaluating information and conveying it correctly but be unwilling to do so because vested interests are involved. For example, a car salesman may be a competent mechanic but be disinclined to be completely open about the cars he wants to sell, whereas a friend may want to be as helpful as possible yet lack the expertise to give a competent evaluation.

The source's ability seems to be a combination of a number of general factors as well as areas of expertise. This can be categorised in the following way:

- expertise
- reasoning
- judging sources
- experience

It is not uncommon to observe that sources may have, apart from a general competence, expertise in certain subjects. Thus, the motorbike dealer will in all likelihood know a great deal about mechanics. This is however different from more general abilities like being able to follow and construct chains of reasoning as the teacher can be expected to handle well. Again, both the dealer and the teacher will probably be a good judge of people, as their work involves a great deal of interaction with people. Finally, most sources will usually be able to competently handle information from their experience.

Consequently, in order to evaluate information appropriately we need to model the source's abilities with respect to those categories. At the same time there are also a number of considerations to be made about the source's trustworthiness:

- trustworthiness
- helpfulness
- interests and beliefs

Whether we can trust sources is a very important consideration no matter how appealing the information may sound. In contrast to a source's ability which will apply generally, it seems that trustworthiness is very much a matter of special relations between the source and other agents or classes of agents. Apart from the peculiarities of the character of the source this is strongly influenced by interests and fundamental beliefs. Thus the dealer has a strong interest of selling bikes for a financial gain and there is therefore a special relationship between the dealer and clients where the trustworthiness will not be particularly high. In parallel to trustworthiness there is also a certain level of helpfulness, indicating the source's willingness to give information, whereas the trustworthiness indicates the inclination of the source to misinform.

We can now establish a basic model for the car salesman as:

CarSalesman1:
Ability:
Expertise: mechanics=high
Reasoning: average
Judging Sources: high
Experience: high

Interests: selling
Beliefs: ?
dealer -> client :
Trustworthiness: low
Helpfulness: high

One also needs to keep in mind that not all of the values will be known and that a great deal of information may be defaulted from class membership, in this case to the class of salesmen.

4.5.4 Representation of Information

From the discussions of Chapter 3 we can see that information from human sources is typically a complex package, containing various items which can divided in the following way:

- the *message* or cognitive content
- the *qualifications* of the message

Each act of communication is designed to convey a certain *message* from the source to a recipient. At the same time, there may be more or less explicit *qualifications* attached to it which describe the source's relation to the message and its implications. For example in "I heard that John bought a car, "John bought a car" can be considered to be the message whereas "I heard that ..." is an indication about the source's claim to the message, namely that it is hearsay. In this instance the qualifications of the message are poor and no strong claims are made.

It is the task of a natural language analyser to process natural language input in this way and not the responsibility of the SCM, which deals with information only once it has been transformed into events. There are, however, a number of aspects about the *qualifications* of the message which are relevant for the purpose of evaluating its credibility:

- indications of the strength of belief
- arguments for and against the proposition
- indications of whether responsibility is assumed for the correctness of the information
- whether the source would have a vested interest and would gain an advantage out of the system believing the information.

As we said before, the information we typically have to deal with can vary greatly in complexity and in some situations little more than the actual message is conveyed in an act of communication. Alternatively, it is also not uncommon to observe that sources may indicate in some way their conviction in the information and perhaps even give the reasons for their belief. Thus if I plan to go on a hike with a friend of mine and he tells me that he heard in the weather forecast that there may be rain on the day and that I should bring some waterproof clothing, then he seems to express a moderate belief that there may be rain and quotes the weather forecast as the reason for his belief. However, to find that the source also indicates whether responsibility is assumed for the correctness of the information or whether it would derive an advantage if we were to believe and act on the information. With the case of the motorbike dealer these considerations can be very important.

The package of information can therefore be represented in the following way:

```
Information1:
    message: John bought a car
    arguments : hear say
    conviction: average
    responsibility: questionable
    advantage: questionable
```

It is important to keep in mind that there are great variations in the amount of extra information supplied with a particular message, ranging form the complete set of auxiliary information to the bare message. While explicit arguments are sometimes given, they are often omitted and may be assumed or implicit from the context of the conversation. This situation does not appear to create much problem for human agents who can operate in a variety of situations. The source control model therefore has to be able to operate in a similar way, whether information is sparse or plentiful.

4.5.5 Strength of Conviction

As described earlier, the source may indicate its strength of conviction and considering the implausibility of anyone being able to give precise values to this strength of conviction a more simplified index is more appropriate. We therefore use an index to take one of five different values which represent the actual arguments:

- nil
- low
- average
- high
- top

There are also limits on what can be consideration to be reasonable levels of the conviction for (c) and conviction against (d) which we call the rule of inverses. If both c and d are provided by the source, the level of c implies an upper bound on d not to exceed the inverse of c; which can be described as:

c		d
nil	→	top
low	→	high
average	→	average
high	→	low
top	→	nil

Thus if c is 'high' d cannot be expected to exceed 'low'. One cannot reasonably insist to be highly convinced about a proposition while insisting to be highly convinced about the proposition being wrong.

Thus, if the upper bounds are reached, the source can be considered to claim to be completely informed about the arguments for and against, whereas below the bounds there is room for reasonable doubt as it is implicitly acknowledged that there may be arguments the source is unaware of. At the same time the relative weight between c and d establishes the source's weight of conviction. Consequently, there are two lessons to be drawn:

- the 'spare capacity' between the given levels of c and d and the upper bound indicates the degree of uncertainty remaining
- to exceed the upper bound can be considered unreasonable and throws doubt on the credibility of the piece of information

If the source only provides c but not d then, obviously, that problem does not arise.

4.5.6 Conflicts of Information

Given uncertainty and the remarkable human capacity to go wrong, we need to briefly describe our concept of conflicting information before we address the principles to deal with it in practice. There appear to be three different situations in which conflicts can occur:

- pieces of information can be inconsistent in themselves or malformed in some way
- the same source may previously have given a piece of information with the same message which may be at odds with the new piece of information

- a different source may previously have given a piece of information with the same message which may be at odds with the new piece of information

The SCM is interested in detecting and dealing with conflicts in this sense, wanting to resolve conflicts of information in order to regain a consistent view of the world and to learn about the behaviour of its sources in order to improve the systems belief formation and consistency recovery capabilities.

Conflicts, as considered in other areas are of a quite different nature, dealing predominantly either with rule conflicts or planning conflicts. In either case, the conflicts of competing rules or inconsequent plans are generally seen as a nuisance to be overcome and are usually solved by predetermined, straightforward strategies. Except for [FAY88] and perhaps [GLL88] there is no interest in learning about the conflict in order to avoid falling into the same trap again, or improving one's response as time goes on. The SCM considers conflicts not as exceptional but commonplace and as an opportunity and trigger to re-assess, learn and improve one's performance.

4.5.7 Action Point

The entire information evaluation process is dominated by the consideration that the given piece of information may be required as a basis for action. Although the question of whether or not to take action on the basis of the information is not the domain of information evaluating or uncertainty management, its distinct possibility influences the process. There are two points which have to be considered:

- the consequence of action or inaction
- the type of action

The consequences of action can be as serious as the consequences of inaction. If I had stepped on a stone during a hike, as I was advised, I could have broken my neck. by the same token, I could have risked my life ignoring a warning that the building I was in, was on fire. This demonstrates that although in most cases a cautious rather than an exuberant response to information may be prudent, it may at the same time limit our capacity to act.

This suggests that there is a point at which the system's connection will be strong enough to act upon - this is defined as the *action point*. In fact it appears that this action point varies dependant on the purpose or implications of the action considered. In order just to air an opinion it may suffice that the system's conviction is just above the medium level, or even less if the system was to say "that it had been told that ...", while giving the piece of information and carefully disassociating itself from it. Alternatively, as a basis for any significant action a stronger conviction is required and for important decisions the system needs to be strongly convinced about the information and a confirmation from an independent source may be required.

The implication of these considerations are that pieces of information considered to be highly reliable have to be examined with more care and attention than those of low credibility. While the question of whether or not to act on the information is not to be taken by the source control model, care has to be taken about forming an adequate opinion. If the system is too conservative about the credibility of information it may prevent the system from making proper use of the information. At the same time a too liberal attitude may allow the system to act on potentially unreliable information.

4.6 Principles of Importance Analysis

Given a piece of information, we first have to decide whether we want to take an interest. There are a number of different ways in which the information can be significant and the source control mechanism will have to do a quick check to see whether there are any indications to that effect:

- the source is important to the system
- the information is substantial
- the system is interested in the subject of the information

The basic idea is that the mechanism will need to find out on a very superficial level whether any of these considerations produce an interest. If the source is trustworthy and competent then the mechanism will want to be helpful and will need to take an interest. Failing that the mechanism may still take an interest as the information from the levels of conviction may suggest that the information has at least potential to produce a strong belief by the mechanism. Thirdly, even if the source or the information may not appear to be that appealing, the subject of the information may be of interest to the system. Thus, if the system can generate some form of interest, it will want to start to evaluate the information more closely.

The importance analysis also plays a role later on, if there are any problems with the information, to decide whether there is any point in carrying on with any analysis which may get very complicated and time consuming while there are other more important issues to consider. This will become particularly significant if there are reasons to suggest that a full enquiry is required to resolve the problem and the potential gain has to be weighed against the cost. The mechanism may therefore at that point start to loose interest and take a decision at that point.

From this we can abstract a number of general principles which govern the considerations of importance:

- that we do not spend much time and effort on information which is not of interest

- that we may have to reconsider whether to pursue problems if the information is not important enough
- that we may take an interest in information not because the information is interesting, but because we want to find out something about the source
- that if the matter is important one may want to get some confirmation if there are any doubts about the information

4.7 Principles of Information Evaluation

Given that we take an interest in the information, we have to decide whether or not to believe it, and if so, what the basis of our belief is. Do we believe it without being strongly committed to it or do we believe it to the point that we are prepared to act on it. Thus when the motorbike dealer confidently tells me that the bike I am considering to buy is in good shape and a real bargain, do I believe the information at all, and if so, do I believe it strongly enough to actually buy the bike?

From what we have seen from the basic principle of source control, there are two aspects we have to balance when we consider how far we are prepared to believe a piece of information.

- properties of the information
- properties of the source

On one side we need to consider whether the information is well-formed and on the other side we need to consider whether that is compatible with the source model, whether the source is trustworthy and competent enough to support the piece of information.

There are three ways in which the source model is used in the process:

- to assess the information
- to modify information
- to suggest ways to resolve problems

The primary purpose of the source models is to assess whether the information is compatible with the expectations of the source model or at least acceptable given certain conditions. Once that can be established, the source model is also used to modify the information, either to fill gaps in the information or to modify the information to fit the source model if a claim is made by the source which is too strong considering the level of ability ascribed to the source. Finally, we also seem to use our opinion of a source when there is a problem to find out more. Thus for example if our friend made a very strong claim in very technical terms then that will not fit our source model as we do not expect him to have expertise in mechanics, but if we ask him we may find out that unbeknown to us he has some qualifications in that subject.

We now need to look at the considerations which have to be made with respect to the information and the source. Although we shall introduce them separately, in the actual process of analysis these two go very much hand in hand.

4.7.1 Information Considerations

From the concept of information we see that there are a number of factors involved in each case and in the process of evaluation we do not just tend to consider whether the piece of information is well-formed, but also what the particular constellation of the qualifications tells us about the potential credibility of the information as a whole. Recalling our representation of pieces of information as:

Information1:
message:
arguments:
conviction:
responsibility:
advantage:

we need to explore the following:

- the relative strengths of conviction
- whether responsibility is assumed
- whether the source would derive an advantage if we were to act on the information.

To consider the arguments themselves, given their availability, can be quite complex and perhaps should be dealt with by an endorsement analysis if one wants to decide the matter entirely on those grounds. There may be situations when we have to look at the arguments, such as in an enquiry or when the analysis has come to a halt. Usually, it seems, the first step to take is to examine the levels of conviction. From what we discussed earlier we can see that it is important that there is no conflict between these levels as this would throw immediate doubt on the information. Given that we do not encounter any problems there, we also have to look at the overall level of conviction as this will give an indication as to whether we are dealing with a potentially interesting piece of information. Thus if I am given a piece of information with a low degree of belief then in the absence of an indication of conviction to the contrary, the information may amount to a weak opinion which will have little consequence and there may be little point in going into a thorough analysis. There may however be notable exceptions where there may be other motivating factors to carry on. Thus if the information is part of a serious discussion where opinions are being exchanged then even a weak opinion may be significant.

Alternatively, if the level of conviction is higher the information is more likely to be of significance and there are other indicators which need to be checked as well, such as whether the source, given its strong conviction assumes some sort of responsibility for its utterance. In other words, am I given guarantees and assurances that I will not unwittingly act on the information only to find out that I was misinformed, which could have a number of unpleasant consequences. The source may or may not give such an indication, consequently leaving us with three possible situations:

- responsibility categorically denied
- responsibility questionable
- responsibility accepted

The level of responsibility may have an important impact on the credibility of the levels of conviction. If responsibility is categorically denied it implies that the source would probably not have acted on the information itself, and therefore does not recommend us to act either. In that case we may have doubt about the veracity of the information and we will consequently be disinclined to attach a high belief to the information.

In the most common situation, it will be questionable whether the source assumes responsibility for the information. In that case the system will have to decide whether this implies that responsibility is denied or can be assumed. Although the safest option is to assume that it has been denied and to reduce the credibility of the information to avoid to act on doubtful information, this has the disadvantage that a great proportion of information would become useless as it cannot be acted upon. A more successful strategy is to consider the source's pattern of behaviour in that light, to determine whether the source is reliable enough and usually tends to accept responsibility. In that case responsibility can be ascribed and the credibility of the information can be preserved which would otherwise have had to be reduced. It seems that in this particular case the source model is not used to assess the information but to suggest whether certain assumptions can be made.

Positive assumption of responsibility by the source for strongly believed pieces of information will tend to reinforce our belief in the credibility of the information.

If the credibility of the information is not diminished for other reasons, one may be inclined to believe the information to the degree indicated by the source. On the other hand, to assume responsibility for the reliability of a piece of information whilst indicating only a weak belief in it, seems counter-intuitive. Since the belief in the information is too weak to be useful there are no risks involved.

Finally, there is also the possibility that the source could derive an advantage out of our acting on the piece of information. This concerns pieces of information which are claimed to be highly veracious. In this case one may need to satisfy oneself that there are good grounds for believing the information or to form only a weak opinion about it. If the source also accepts responsibility for the correctness of the information one may be inclined to give more credibility to the information, but care has to be taken that the guarantees are commensurate with the possible damage one might suffer in wrongly putting ones trust in the information. Considering the example of the motorbike we may be inclined to buy if, for example, the salesman offers a comprehensive guarantee or 'money back if not satisfied' since we could potentially lose a large sum of money on a motorbike which constantly breaks down.

From this analysis it appears that there are three basic principles involved in what we consider to be a credible piece of information, at least as far as the information itself is concerned:

- that the information is generally well formed and no excessive claims are made
- that the source accepts responsibility for important information and especially if strong claims are made and the source claims to be convinced
- that sources which want to act with integrity will accept responsibility and declare their advantage, unless that is obvious

4.7.2 Source Considerations

The results of the information analysis need to be seen in the context of the source model. As we have said earlier, the main motivation is to whether the information is compatible with our expectations about the source and there are two aspects to be considered:

- the source's trustworthiness
- the source's ability

Analysis of Trust

Given that there is at least some potential for credibility in the information, the first reaction, it seems, is whether the source is fundamentally trustworthy, because if the source might intend to mislead or otherwise be insincere then the credibility of the information is very much in doubt. This consideration is a very delicate one, as common experience shows that trust, once broken down, is very difficult to regain. For example, people with criminal convictions tend to find it very difficult to gain employment as the lack of trust on behalf of the employer is so deep-rooted that they do not consider a good working relation to be conceivable.

Coming back to the example of the motorbike dealer, he has a strong interest in selling and since we are in the position of a client, the dealer cannot be trusted very much. It also appears that he does not accept responsibility for his claims. If we were to believe in the information, our belief would be only weak. In a situation where the trust is higher one could be inclined to assume that the source would assume responsibility, but another complicating factor is that since the relationship is strongly dependent on an interest of selling this also leads us to conclude that there is an advantage involved if we were to act on the information. We can therefore modify the information to reflect that advantage. This situation makes it very difficult to consider that the source might assume responsibility. As there are a number of factors involved which are against believing the information, our belief will stay low. We may also add the reasons for the decision as an endorsement to the information, for future reference.

Supposing the situation was sufficiently important to try to find a solution and to raise our belief, the endorsement will show that the only way in which one can raise the belief is to ask for guarantees or for independent confirmation of the dealer's recommendation.

So far we have been dealing with problematic situations. As already mentioned, if the trustworthiness is higher then one could consider that the source would assume responsibility as long as there are no indications that there is a hidden advantage. In that situation one also needs to pay attention to the level of helpfulness, because if the source was expected to be helpful and if the information was asked for and only reluctantly given, the source is not likely to accept responsibility. Alternatively, if the expected helpfulness is low and the information was volunteered, then, again, there may be a hidden advantage by the source. Consequently, the safest situation to ascribe responsibility is when the source was as helpful as expected.

As a result, a number of fundamental principles can be identified in the considerations of trustworthiness which can be summarised as follows:

- trustworthiness can be very context sensitive, dependent on special relations, and may be strongly influenced by fundamental interests and beliefs of the source
- that if a source cannot be trusted it is very difficult to believe what it says unless there is independent confirmation, or guarantees for the correctness of the information are given
- conversely, if a source can be trusted and there are no vested interests involved, then it is easier to raise expectations that responsibility can be assumed
- even though some special relationships which carry a poor level of trustworthiness may not be evident, if the information touches on

fundamental beliefs and interests which are connected to strong motivations, then an advantage could be inferred

Analysis of Ability

Alternatively, if the source can be trusted and the information is well formed, the credibility of the information from that point onwards appears to be dependent on our opinion about the source's abilities. As we have already mentioned, the ability index includes a number of factors to reflect different aspects of the source's behaviour and we need to identify the appropriate index.

Apart from differences in expertise in certain subjects, there also seem to be a number of more general abilities such as the ability to judge sources. Someone who is very gullible and who puts his trust too easily in what people tell him is clearly not very good at judging sources. Another prominent feature is reasoning, and not many people who have not had much formal training in this area will be able to easily follow and reproduce long chains of reasoning. Finally, one category which most people will probably find easy is the ability to correctly evaluate one's experiences. Looking back at the different categories of ability, it will also be plausible that a source will not perform equally well on all of them and it is therefore necessary to decide which category the information is born out of.

To consider some examples, in the case of the motorbike, judgements on their qualities will require expertise in mechanics rather than particular abilities in judging sources or general reasoning. While we can trust our friend, we will not be able to rely on his recommendations because he does not have the relevant expertise. If our friend were to make strong claims about the qualities of the motorbike we would not be able to believe him as his abilities suggest that that is not possible, but since he only airs a weak opinion we can believe it even though that may not solve our problem of whether to buy the bike. This suggests that there is a limit, whereby we can believe information only to the degree that we consider the source capable of. The case with the dealer is different as he has the expertise, but due to the problems with trustworthiness we cannot assume that he will accept responsibility and therefore we cannot fix a strong belief to his recommendation. Had he given a guarantee or if it was not him who sold the bike we could have believed him to the degree we think him capable of.

Another interesting situation could arise when our friend, despite our expectations that he does not have expertise in mechanics makes a rather strong and technical statement. This may suggest that the friend either does not know what he is talking about or that he actually has more expertise than we expected and that may be worth to find out for example by asking him.

Alternatively, if the source makes a strong statement while refusing to accept responsibility, then we are dealing with a counter-intuitive situation. Whether the source has the necessary level of ability or not becomes immaterial, because the information can fundamentally not be believed or believed only to a low degree. The source is in fact saying that the information should be believed and not believed because it refuses to take responsibility. In that case we will probably consider it prudent to reduce our conviction to a low or moderate level.

Conversely, given that there are no such problems and the information is well formed it seems plausible to believe the information if the source is considered to be sufficiently qualified to make the statement.

Finally, it should also be noted that there is a difference between information in the source model which is the result of actual experience with the source as opposed to information derived from a classification. Information from classification can more easily be changed than information from actual experience, because class properties are only typical whereas actual experience may show that the source behaves different in reality.

Given that the source can be trusted to some degree, there are a number of fundamental principles involved in an analysis of ability which can be summarised:

- there are clear differences in the ability of sources, and a correct assessment of information from them requires that the different abilities are taken into account, usually depending on the predominant type of the information
- if there are problems which require expertise in a subject and the source does not seem to have relevant expertise, then one would not expect that strong claims would be made by the source
- the perceived abilities of the source in a particular area appear to be a limit to the credibility of that type of information
- while strong claims by the source may be unacceptable while no obvious expertise is ascribed to the source, it also seems that such statements could raise suspicions that the source actually may have an expertise we are not aware of

4.8 Principles of Conflict Resolution

So far, pieces of information have been considered in isolation. From the description of the strategy of the source control mechanism we have seen that the process of reconciling the information with the source model is the first of two dimensions. Having found a solution to that problem, the information also needs to be accommodated in the system of beliefs. If there are no direct connections then the new belief can be incorporated and the case ends at this point. As the system's system of beliefs is built up, the majority of new information will typically have no connection with the beliefs held by the system. As the system of beliefs is growing, they situation will arise where the new piece of information touches on a belief held by the system. The system will therefore have to decide what position to take on the basis of the new information and its previously held belief.

Trying to put some categorisation on the problem which will help us to explore a solution in an organised fashion, we can again identify two different dimensions. There are four possible situations to be considered:

- both pieces of information are roughly identical
- the new piece of information reinforces the old
- the new piece of information weakens the old
- the new piece of information contradicts the old

This has to be seen against two different types of cases which can occur in each situation when:

- both pieces of information (originally) are from the same source
- the system's belief is based on a different source

Although the problem in each of the four situations is the same as far as the system of belief is concerned, single-source and multiple-source differ considerably in their implications and the treatment they require.

4.8.1 Single-Source Conflict Resolution

It appears that in this situation the same principle as in the initial information evaluation process applies; we have to try to reconcile the situation with the source model. The four situations obviously differ in their implications and the first two should not be a cause of concern and only the second two may have further implications, either for the system of beliefs or an ensuing source re-evaluation.

If both pieces of information are roughly identical, it amounts to a situation where the same thing is said twice on different occasions. This is the most simple case as it will not have much consequence for the belief held by the system other than to indicate that the source is still subscribing to its previously aired opinion. Apart from that, it seems that reiterating one's position does not make that position any stronger unless more, or different reasons are given in its support.

Alternatively, we could be faced with a situation where the source reinforces its previously held position, which is interesting, as it might allow the system to reinforce its belief and might make the difference for the system to be able to act on the information. The system will have to take two considerations into account how credible the new piece of information is with respect to the source model, its ability and trustworthiness and whether the change of the levels of conviction contravenes the limits of determination. Therefore there are three possibilities which may prevent the system to raise its belief:

- the source is not considered sufficiently competent
- the source is not considered trustworthy
- the limits of determination

If the source is considered sufficiently qualified and trustworthy to make such a strong statement and there are no problems with the information as such, then the source control mechanism would probably not object to raising its belief in response.

If any advantage was to be had by the source influencing our beliefs and actions that would mostly come from making convincing statements and avoiding any hint of contradiction or change of mind and therefore one has to be especially vigilant in these situations as a source might try to warm us to a particular idea before making strong statements.

If there are no problems with the previous two considerations, the only problem that may still arise is when the previous information was highly determined thus indicating that the source claimed to be well informed and now seems to be changing its mind. In that case we may have to be careful whether to believe the source that much, because in either case the source has made a mistake somewhere.

As previously discussed, if the levels of conviction for (c) and conviction against (d) do not add up to full determination then the source implicitly acknowledges that apart from its present position, there is a certain amount of uncertainty involved about the information. The source thus agrees that there may be more arguments, unknown to it, which, may influence the levels of conviction. This amount of uncertainty determines the limits of manoeuvrability of c and d. Supposing the source held a belief B with c=average and d=low, it could raise c to c=high without having to go back on d. A violation of this principle would imply that the source would have to go back on its previously held position, thus acknowledging that it was wrong in the first instance.

The first two cases of the source reiterating its position or raising its belief in a previous statement are usually considered favourably by the system, given that the source is both capable enough to support the claim and can also be fundamentally trusted. The third and fourth case are a cause for concern. The situation is much more serious as the system may have acted on a piece of information which was badly supported. In the cases of weakening and contradiction two factors have to be weighed against each other:

- the source model is inadequate
- there is a change in the environment

In the case where the new piece of information is weaker than the one previously aired by the source, the source still holds onto its previous opinion but with less conviction. As far as the systems belief is concerned this means that the support for the belief has been withdrawn. Since there is no other support for the belief the system may have no choice but to weaken its opinion. The other consideration to be made is whether that could be expected from the source.

Usually, responsible agents will give some sort of explanation if they change their mind about something substantial, and which would indicate that there was a change in the environment. If the information was important to the system, it may have to try to find an explanation for where the problem originated, or if there is a problem at all.

If it was the source's fault and the source was not expected to act in this manner then this may have to be considered by the source re-evaluation process. There may also be a long term effect on the source if it is in the habit of making strong claims, only to change its mind subsequently. As the whole process can be quite complex and requires to looking closely at the information itself, a decision may have to be taken at a surface level, especially if nothing important is at stake.

The most serious situation arises when the source is contradicting itself by supplying a new piece of information which subscribes to the opposite viewpoint to the one previously held. In contrast to a case where there is a weakening of information, contradictions are very sudden changes and tend to indicate that there is a fundamental break with the previous information. How significant the effect on the system is dependent, in part, on the level of contradiction. If the contradiction is between two weak beliefs, the case can be considered to describe a change of mind. The system would not have acted on the information anyway, and can now decide whether to change its mind as well or summarily withdraw its support and refuse to be drawn to one side over the other. As the situation does not seem to be important from an information point of view the case can be stopped there except that the mechanism may make an endorsement to indicate the unreliability of the information and to make a record of the occurrence for the purpose of source re-evaluation.

The situation is much more serious when information with high degrees of conviction are affected, as the system might have acted on the information. Since this is a case of self-contradiction of a single source, the system is likely not to believe either side or only form a weak belief either way, dependent on its general faith in the source. If the mechanism does not reject the information outright, it would tend to side with the latest information as the support for the previous information has now been withdrawn by the source. As this situation is potentially quite serious, the system will want to find an explanation for the contradiction, and to see whether this means that either the source was at fault or the source model was wrong or there was a change in the environment.

As the conflict affects pieces of information which are substantial, the mechanism will be prepared to try to find an explanation.

If the source cannot be trusted, and there may be a chance that there are interests involved, then we may be inclined to reject both pieces of information unless we take a strong interest in the information itself. If we are inclined to find out more then we will need to see whether the subject of information is likely to change, which implies that the source acted correctly if the new information reflects the new state of affairs, in which case it can be believed. In order to determine that, the system may have to seek independent confirmation which would require an enquiry.

Alternatively, the source may have made a mistake, in which case we need to see whether the source model allows for that. We therefore need to look at the type of ability the information was modified with. If the respective index is sufficiently low, then we did not expect the source to be very competent. Alternatively if the index is high then either there has been a change of state in the subject of the information or our assessment of the source's ability may be wrong. If, for example, it is a matter of personal experience, then it is very difficult to check independently and we can only acknowledge that the source now feels differently about it. If it is a matter of reasoning then the source control mechanism could ask the system to check the reasoning for flaws. If it is a matter of judging sources or expertise, then we could ask the source to give an explanation.

Given that it appears that it was the source's fault, the system will have to consider whether this kind of behaviour could have been expected of the source. If the system knew about the source's poor competence and took a cautious approach before then this situation only goes to reassure the system's opinion of the source. In the converse case the system mistakenly put its trust in the source and information, it may have to reconsider its opinion about the source, depending on whether the problem is about trustworthiness or ability, as will be described in section (4.10).

Given that there is no problem with trusting the source, the single-source conflict resolution is dominated by whether the source in question is self-consistent. If it is within the spare capacity of the levels and degree of determination to move from the old position to the new then the source is perfectly at liberty to do so and the system will have little choice but to follow suit unless it has some independent information bearing on the situation. In the converse case it will depend on the severity of the change of position and obviously whether this could have been expected of the source. In severe cases this may therefore require a reassessment of the system's opinion of the source. A source which is in the habit of changing its position unexpectedly is not very useful to the system which looks for a more stable, predictable behaviour.

From this discussion we can abstract the following principles:

- if a source weakens or contradicts its previous position, this implies that the support has been withdrawn for the old information and a new belief will have to be based on the new information, taking into account the conflict situation
- in contrast to more sudden and sharp changes in a position, if the change is smaller and more subtle and there is no problem of trust, then one will be inclined to believe the new piece of information
- in those situations where there is a case of severe weakening, one will have little choice but to accept the new information, although one may still be inclined to keep the occurrence in mind when reassessing one's opinion of the source
- substantial problems should not occur often, and if they are rather unexpected, one will be inclined to ask for an explanation and whether there has been a change in the environment

- if we are dealing with a reinforcement of a position and the source's ability is considered to be good enough and it can be trusted, then one will be inclined to believe the new information
- while outright contradictions are a definite cause for concern, it seems that contradiction at a very low level of conviction appear to be more like a change of mind and if there are no interests involved, then one should follow the new information, especially if there is no time to find out if there has been a change in reality
- alternatively, if there is a serious contradiction and there is no trust in the source and no explanation available, then the safest option seems to be to reject both pieces of information, unless the information is significant enough to suspend judgement until such time as there is a possibility to find out what the problem is
- in those cases where a change of environment is plausible and the respective ability is not high, then one will be inclined to hold a weak belief in the new information
- conversely, if a change of environment is not plausible and the respective ability is not high, then one may be tempted to reject both pieces of information
- if a change of environment is plausible and the ability is high and the source can be trusted, then there seems to be little room to question the source's judgement and in the absence of confirming or discrediting evidence one will be likely to believe in the new information
- finally, if a change of environment is not plausible and the respective ability is not high, then it appears that the source model may be wrong and one will be inclined to try to start an enquiry if the case is sufficiently interesting
- also, while there may be a problem of trust with a substantial new piece of information, if responsibility is volunteered, then despite problems with the information it is feasible to cautiously believe new information provided that the guarantees are commensurate with the possible damage that may result out of putting one's trust in a bad piece of information

- therefore, if there may be a problem of trust with a substantial new piece of information and responsibility is not volunteered, then the only possibility to proceed is to enquire to get guarantees or seek independent advice

4.8.2 Multiple-Source Conflict Resolution

Single-source conflict resolution is relatively straightforward as there is only one source to blame when there is a source problem and contradictions occur. Multiple-source cases, by contrast, are much more complicated, but also have more potential. Thus if there is a conflict between the system's belief based on one source and a piece of information from another source and both sources are equally competent, then it is difficult to find out who is wrong. At the same time, if there is reasonable agreement between two or more sources than that can be very reassuring for the system's world model as well as a helpful source of information for the subsequent source re-evaluation.

Before we analyse the various situation that may arise we need to consider two interesting phenomena which influence the evaluation process:

- the solidity of information/beliefs
- the corroboration of beliefs

When presented with two pieces of information from different sources which are at odds with each other and where one source is considerably more competent than the other, the system will tend to side with the strong source. What is more, the system will also be disinclined to let the position of the stronger source be affected by the antagonism of the weaker source. We call this phenomenon the *solidity* of the information to describe the relative maintainability of the position of the stronger source (and the system's belief in it). The reason for this is that it seems counter-intuitive that the position of the strong source should be considered on the same level with that of the much weaker source. Thus, if a physicist deliberates on his subject it is unlikely that his information will be in much doubt if an inexperienced student challenges his position.

Similar to the solidity of information there is another phenomenon associated with the system's beliefs. Supposing that the system gets two compatible pieces of information from different sources it is likely to subsume the new piece of information into the system's belief held about the old one. This is addressed by adding the new source to the list of supporting sources and adding the new arguments for and against and recalculating new levels for c and d. As a result the system's belief is *corroborated* in the sense that it is supported from different sources and therefore more stable than from a single source. If a third source subsequently is at variance with that corroborated belief it will find it much more difficult to change the system's opinion than if the system's belief had been based only on a single source. The degree to which the system's belief becomes immutable or inert we call the *degree of corroboration* and is dependent less on the number of sources but more substantially on their ability. Thus a reliable source will still outweigh several sources known to be unreliable, whereas two sources will probably outweigh a third of equal reliabilities.

We now shall consider the principles of multiple source conflict resolution. As with the information evaluation and single-source conflict analysis the trust that can be placed in the information as well as the abilities of the source need to be taken into account and that together has to be offset against the solidity and corroboration of the system's view based on a different source or sources. Therefore there are again three possibilities which may prevent the system to change its belief:

- the source is not considered sufficiently competent
- the source is not considered trustworthy
- the opposing view is much stronger

If the source is considered sufficiently qualified and trustworthy to make such a statement and there are no problems with the information as such, then the source control mechanism would probably not object to believing the information in principle.

If there is no problem of trustworthiness, then the system will be prepared to believe the source to the limits of its ability, but it is difficult to imagine how an information with doubtful trustworthiness can survive against an opposing view where there are no such problems. At the same time if a problematic piece of information reiterates a the position of a different source then that may serve to dispel questions about the information.

If there are no problems with the previous two considerations, the only problem remaining is how the new piece of information relates to the existing position and how strong that existing position is compared to the new information.

Although we need to consider similar situations, there are nevertheless significant differences with respect to their meaning.

As with the single source case the situation where both pieces of information are roughly identical is considered favourably by the system. The implications of this situation are however different. Whereas a reiteration of the information by the same sources does not have any effect on the credibility of the information, a reiteration by a different source corroborates the system's belief as it suggests that the system is correct in its evaluation. This may also reflect favourably on the systems view of the source during the source re-evaluation process, like if there was a reason to belief, say, that the source may have some expertise in an area previously unknown to the system and there is independent evidence which might help to clarify the situation. As far as the strength of the belief is concerned this situation should not lead to any rise, although the new source will be added to the list of sources supporting the system's belief.

The system will also react positively to situations where the new piece of information reinforces the old, considering that the system is reaffirmed in its belief. Given that there are no objections to the information from the information evaluation process, the system may be prepared to raise its degree of conviction to the level of the new piece of information. Whether that can be done is dependent on the degree of corroboration and the relative strength of the supporting sources. If the sources are roughly equal in their ability, or if the latter piece of information is supported by a stronger source, the system will be inclined to raise its conviction whereas in the converse case it will hold on to its previous conviction. In either case the source of the reinforcing information will be added to the list of sources supporting of the system's belief. As with the previous case the verdict will be passed on to the source re-evaluation process for further analysis.

The last two cases are a cause of concern to the system, depending obviously on the severity of the problem. The system will have to weigh the different pieces of information against each other keeping in mind the possibilities that:

- the source models may be inadequate
- there may be a change in the environment

In the case where new information weakens the old belief of the system, the system may have to readjust its position. Again, the new piece of information needs to be considered on its own through a process of information evaluation and then has to be compared with the system's belief. The system then has to decide between two conflicting motivations, namely the desire to preserve strong (and hence useful) information and the inclination to take the safe option of reducing its belief and not to risk to rely on unsubstantial information. The decision which line to take is dependent on the degree of corroboration if the system's view is supported by a number of sources, or on the grounds of solidity of the sources involved. If the system's view is well corroborated and its sources stronger than the new source, it will maintain its belief at the present level. In the converse case the system may have to reduce its belief in line with the new piece of information or go for an intermediary position if the sources are evenly matched. In any case the new source will be added to the list of supports. If the matter is sufficiently important, we may have to look into the possibility that there was a change in the environment either by looking at the subject of information and estimating the likelihood that a change may have occurred or by trying to talk to the source to ask for clarification or by seeking independent advice.

Although cases of weakening are more like a nuisance, contradictions are a definite problem as the source might have mistakenly put its trust in a piece of information and its supporting source. If the contradiction is on a low level, then it could mean that we are dealing with different viewpoints and we may have to decide whether the information is important for other reasons in order to carry on in the analysis.

It is always reasonably straightforward to take a decision based on the strengths of the supporting sources but then the system does not learn much. On the other hand, to try to find an explanation is very complex as we are dealing with a variety of situations from two sources of similar competence, over two sources with different abilities and disabilities to situations where one source is competing against a number of sources. This dimension also has to be considered against the other dimension of differences in the composition of the pieces of information in question. If the situation is sufficiently important one may try to get further information and the case should perhaps be decided by a full endorsement analysis. This may not always be possible and it seems that at this point the considerations of corroboration and solidity are quite appropriate, as that makes for a decision mechanism which is less detailed but at least makes different pieces of information at least comparable.

When we are dealing with conflicts of pieces of information which carry a high degree of conviction and belief, we may therefore have to consider the new piece of information as a whole, against the degree of corroboration and the relative solidity. For stronger piece of information the system will have to decide whether it is worth trying to find an explanation or fix a belief depending on the relative weight of the viewpoints involved. Conversely, At it is difficult to decide between sources of equal strength: it is sensible in that situation not to take sides and to opt for a balanced view.

If we want to examine the case more closely, then there are a number of aspects which may lead to an explanation:

- ability versus trustworthiness
- context

If the problem is one which is purely a matter of ability, then we will have to proceed along the lines described so far, but the matter may be different if there are interests involved, in which case we may have to decide the matter on the basis of trust. On indication whether that may play a decisive role is whether the two pieces of information were completely independent or whether they actually occurred in the same context. If they came from a completely different context then it is unlikely that there was a hidden relationship which may cause one of the sources to try to deceive. A good example of this is the case with the copied assignments. The colleague asked about his opinion on whether the assignments were copied, can be easily believed as he has the expertise and no vested interests other than, presumably, his professional integrity and a sense of justice and reputation of the department. The matter with the students is however different as they have a strong interest that the work is considered to be their own. The situation may be quite obvious, if the student who copied from the other produces an assignment which is above his capabilities from past experience, as far as the teacher can make out, but the case could be more subtle and difficult to prove. Suppose the teacher was quite inexperienced and asked for help of how to proceed. As this is rather complex to determine it may require an enquiry.

It is important to note that with all these situations, the system, having reassessed and perhaps amended its belief about the piece of information, will add the new source either to the list of supporting or opposing sources. Together with the considerable degree of corroboration that may accumulate for either or both sides over a period of time this accounts for (and models) the common phenomenon that certain views will be much more reticent than others. Thus we may be easily inclined to change our mind over some common beliefs such as whether the rose bushes in my garden are likely to blossom within the next week as opposed to more fundamental ones like whether the sun will rise tomorrow. To unsettle well-corroborated views in the prevalent views of science will take a tremendous amount of arguments and proofs rather than a spurious belief from an isolated source whereas more trivial, everyday beliefs are much more subject to rapid change.

The following principles are at work in the multiple-source conflict resolution:

- given that there are no interests and advantages involved, the decision of what to believe in a case of conflict is largely a matter of ability
- alternatively, if there are interests and advantages involved, then one needs to pay attention to these interests and abilities and pursue an enquiry if sufficiently important
- supposing one is faced with a situation where both pieces of information are roughly equal, then the new source can be added to the list of other sources supporting the information. While this leaves the information largely untouched, the occurrence may be of interest in the process of reevaluating the source
- by contrast, if the new information reinforces the old, then one will have to decide whether to raise one's belief, although this may be dependent on the qualifications of the respective sources and pieces of information
- conversely, if the new piece of information weakens the old, then one will have to check the relative weight of sources and whether there are interests and advantages involved if they are in the same context:

  Thus, if there are no interests or advantages involved, then the matter may be decided on the strength of the sources involved.

  Alternatively, if there are interests and advantages, then one will be inclined to start to enquire unless that is not possible. In that process one has to carefully consider the relations between sources and therefore independent sources are likely to take precedence.

- as with single-source situations, if there is a contradiction and the beliefs are low, it looks like it could just be a change of opinion and if there are no interests one may just decide on the basis of strength of source or refuse to take sides
- finally, if there is a serious contradiction and the beliefs are high, then it will depend on whether are interests. If there are no interests then on can decide on the ground of ability, given that there is no likelihood that the environment may have changed.

  In the case where there is a possibility the environment may have changed, one will have to try to enquire further

  Likewise, if there are interests involved one should try to start an enquiry or if that is not possible, believe the party which is least likely to lie i.e. those with the least interests.

4.9 The Principles of the Enquiry

As we have seen from the previous sections, there are a number of points where an enquiry should take place. As enquiries are rather complex and involve getting deep into the information and situation they are also rather time consuming as one may have to try to engage in conversation with sources, some of which may not immediately be available, or the overall system may refuse such requests for other reasons such as that it is engaged in a conversation with someone else and does not have the time to help in the enquiry.

It will have become apparent that there is usually no strict borderline where in the analysis process the enquiry starts as all processing is inquisitive by nature. Usually, enquiry starts at the point where a solution cannot be found directly from the explicit parameters of the representations or is unsafe considering the problem and the system will have to find explanations and assurances to convince itself.

In the following we pick up from the examples we referred to an enquiry in the previous sections:

- the technical claims of the friend about the motorbike
- the request for a decision whether to buy the motorbike
- how to sort out the question of the copied assignments

At the same time, not all enquiry is potentially detrimental but may serve to confirm suspicions. Thus, for example, in the motorbike case, if our friend made a quite technical statement which we did not expect and we ask him to find out that he actually has some qualifications we were unaware of, then we can ascribe that expertise without further ado, as we can trust him sufficiently not to lie to us.

If the enquiry process is asked to give a solution to the question of whether to buy the bike or not, then the enquiry process can look at the information available, which is the weak statement from the friend which is trustworthy but does not have expertise and the statement from the dealer who has expertise but is fundamentally not trustworthy. In this situation what is needed is expertise in mechanics as the problem is a technical one. The problem with the dealer is that though he has expertise, he also has an interest which leads the mechanism to infer that he would have an advantage. The belief of the friend cannot be reinforced as he has no expertise and therefore one avenue would be to try to ask whether the dealer can give, say, a guarantee. Alternatively, if that is not possible then one could ask the source if he knew someone else with expertise in mechanics or the system might see whether it knew someone like that and might come across its model about the RAC valuation service and recommend the source to ask them. Given that the source could come back later to say that the RAC said that the motorbike was quite reasonable then the mechanism could agree with that.

Finally, if the enquiry is started in the case of the assignments, we have a situation where the colleague can be trusted, but his agreement confirms the teachers suspicion without proving what actually happened. The matter with the students is difficult as they have their own interests close at heart and the situation may not be very obvious. If one piece of work was clearly above the capabilities of the student the decision could be more straightforward, but that may not be the case. The information so far has two aspects, the evidence of similarity of the assignments and the suspicion that one copied from the other. Who copied from whom may be difficult to determine but it would be unlikely that the better student copied from the weaker student considering that students want to get high marks. In any case, both have an interest, and it seems best to use the evidence first and to ask the students whether they agree that the assignments are indeed very similar. If anyone has an interest and denies the evidence then they are very likely to lie. Given that both say that they agree that the assignments are very similar then one will have to ask whether they copied from the other. If both deny that and instead give the same textbook as their source then on have to see whether the assignments are very close to the textbook, but if they are quite different then both may be lying. If one says that he did the work himself and if that is feasible given the ability the teacher ascribes to the student but that his folder containing his work went missing for some hours then that opens the possibility that the other student copied the work without the other student's consent. One may also have to look at the relation between the two students, because if they are close friends or boyfriend and girlfriend they will be very helpful towards each other and may have connived. As can be seen from this, to try to find an explanation for a problem from multiple sources can be very complex an involve a considerable amount of effort.

As can be seen from these cases, the process becomes very complex very quickly and may take considerable time. In order to go into an enquiry must therefore be worthwhile as the case of the motorbike and the copied assignments demonstrate where we are dealing with a considerable financial commitment on one side and with people's futures and reputations on the other.

From what has been said it appears that there are a number of general principles which emerge from these cases:

- supposing that there is a problem of ability and no interests are involved, then one can ask the source or another competent source which knows the source in question and which is good at judging sources
- alternatively, if one is interested in raising the strength of belief the information, but not primarily in the source, then one may be inclined to ask another source which has good abilities for that type of problem
- one should not ask a source of whether it can be trusted, because if it cannot be trusted it is likely to lie about that and if the source was trustworthy the relation of mutual trust may break down because the source's loyalty is being questioned
- at the same time, if a source declares its beliefs and interests that can usually be believed, but denial of beliefs and interests do not imply that the source does not have them
- although one should not ask a source about its trustworthiness and there is a problem with trust but not with expertise, then one can ask the source for assurances or seek independent advice from someone who has that expertise
- incidentally, if there is evidence for a particular problem, but no trust, then one may be able to gauge the trustworthiness, because if the source denies the evidence, it is likely to be lying
- in order to get independent advice one needs to look at the ability required and who may have it or ask someone who might be able to point in the right direction
- given that there are a number of sources to talk to, the best strategy is to ask the one whose abilities are the best first, because then one is less likely to have to go back to sources after new information has emerged
- if there is a conflict between two sources about the same issue and there is no trust, then one should look at the relation between the sources and

whether there are common or competitive interests, before exploring the possibility that there is a problem of abilities

- similar to the enquiry, if one wants to find out who lies or speaks the truth, a good starting point is to look at who would lose most if the truth were known
- as it is unusual for people to collaborate unless they have common or mutual interests, the fact that they collaborate may indicate that there is a common interest which may not be explicit
- given the chance to interview a source find out as much as possible about the source first, so that the opportunity is not wasted and one may have to go back and be a nuisance by asking about the same problem twice
- try not to be a nuisance and do not switch between sources if avoidable as irritated sources will be less inclined to be helpful
- in the process of holding conversations a good strategy is try to establish common ground which gives an opportunity to increase mutual understanding, before talking about problems
- consequently, if one has to ask potentially irritating questions then it is best to do so after asking the less controversial questions first, as the conversation is likely to break down and the source will not be helpful after that

4.10 The Principles of Source Re-evaluation

So far we have been concerned with the information evaluation and conflict resolution process in which the system's model of sources is applied and beliefs are formed, revised and rejected. In the following, we shall look in detail into how the source models are generated and maintained. Before we shall examine the principles in detail, we need to briefly describe its relation to the information evaluation process and we also need to introduce the different aspects the source model is concerned with.

4.10.1. The Basic Constraints of Source Re-evaluation

The source control mechanism has two fundamental components, the source model and the world model, where the source model is the collection of all sources the mechanism had acquaintance with and where the world model is the set of all beliefs held by the system. This set of beliefs represents the world in the view of the system and the source model is instrumental in maintaining the connection between the world model and reality by placing the appropriate interpretation on the data. It is the combination of the world model and the source model that the system has its fundamental grip on reality and it is therefore important that the source models are carefully maintained, because they are the tool by which information can be assessed.

The interest of the information evaluation process is to produce reasonable beliefs given the constraints uncertainty and incompleteness of information as well as unreliable sources. The interest of the source model is to assess the ability and trustworthiness of sources needed for that purpose. As the diagram below shows, the information evaluation process feeds details about the information it is supplied with to the source model. In return, the source re-evaluation provides its evaluation about the sources known to the system by returning revised source models.

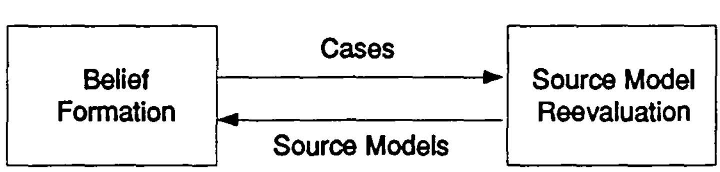


Figure 3: The Feedback Loop

Four Types of Analyses

There is also a marked difference between their mode of operation. Whereas the information evaluation process applies the source models for each piece of information, the source modelling process usually operates on a more long-term basis. How fast the re-evaluation takes to react depends on a number of factors:

- pattern analysis
- evidence from enquiry
- association with class
- revision of different defaults

Usually, a number of pieces of evidence need to be collected before a pattern in the behaviour emerges and can be translated into changes in the source model. The reason for this is that if a small sample is taken, spurious events and local abnormal behaviour may unduly influence the re-evaluation.

Alternatively, during an enquiry evidence can emerge to suggest that a source may have expertise in an area the source control mechanism was unaware of and the model may have to be changed immediately.

If the source for some reason may be associated with a new class it will need to inherit the properties for that class as far as they do not contravene the system's actual experience with that source. Alternatively, the system may encounter a new source and the source re-evaluation process will want to quickly generate a source model from average properties and class associations.

Finally, there is also a difference between indices which are the result of actual experience with the source and which are more difficult to change than those indices which are built on defaults. The easiest one's to change are defaults from ignorance once the system gets some evidence about the real behaviour of the source and then there are the defaults from classification which have a stronger basis and need more evidence.

Consequently, there are a number of general principles which need to be applied in the source re-evaluation process:

- considering that there may be no previous record of the source, we need to look at the type of source, whether it is presented by name or as a class-member (dealer, teacher, politician, GP etc.) to see whether a classification can be produced quickly, and used as a default
- once the source is known, we need to concentrate on whether there is a pattern emerging in the behaviour which should be reflected in the respective indices
- although one should be able to see a pattern in the behaviour of a source, if there is hard evidence from an enquiry, then that single incidence may be enough to change the source model if the source model was found to be incorrect
- as opposed to indices which reflect actual experience, if there is evidence to confirm or replace defaults then the defaults should more readily be replaced by indices labelled as originating from actual experience with the source

4.10.2. Maintenance of the Indices

In order to maintain the indices for a particular source model, the source control model needs to consider the following:

- how well the source model fits the new piece of information
- how well that piece of information fits with other, related information
- the source's track record

The re-evaluation process will need to consider firstly whether the information was in accordance with what could be expected from the source model. Thus if the information was about reasoning and the index for reasoning is high and the information was strong and correct then that reinforces the system's belief in a correct assessment of that index. If, however, the actual behaviour is different (i.e. better or worse) than the one expected, then one needs to weigh the strength of the evidence against the strength of the index. If the index is just a default by ignorance and the evidence is quite strong then we may be inclined to replace the old index with a level suggested by the information. It is, however, not very safe to build an index on the basis of a single piece of evidence unless it is marked as having to be treated with caution.

One way to get independent confirmation, once a variance with the source model is detected is through other information. Thus is a different source gave the same information, one can compare the performance and the levels of the indices in the source models and see at what level the index the other source is. Alternatively, the system may exchange its views about the source in question with a third source who knows that source well and use information from there.

Finally, perhaps the safest way is to keep record of the source's behaviour and to use a more statistical analysis although one has to make allowances for the fact that some evidence may be more substantial than another. Whereas the other two types of analyses may be reasonably quick, this method is much slower but less likely to go wrong. Indeed both types of analysis should go side by side as if there is any truth in the evaluation they should confirm each other over a longer period of time.

From this a number of principles are emerging:

- The consideration of the proportion of cases where the information does fit the source model gives a hint on whether the source model on the whole is appropriate.
- The other consideration is how the information fits with other information. That also has an effect on whether the source model appears to be appropriate or not.

4.10.3 Maintenance of Ability Related Indices

In the process of maintaining the ability index for each source, the source re-evaluation mechanism is concerned with the identification of four basic categories:

- the source's expertise in different areas
- the source's reasoning capabilities
- the source's competence in judging information
- the source's capabilities in handling its own experience

The source's areas of expertise are perhaps the part most affected by particular training as opposed to the other indices which are largely the product of life-experience. Information about expertise can come from the classes the source is associated with or can be deduced from the kind of statements the system gets from the source. Thus if the language used is very technical, using specific terminology the source may have expertise in that area which should be substantiated through communication with the source or other source's acquainted with the source in question.

The index about the source's ability to reason is a more general ability and not solely associated with a particular subject, although a source's sense of reasoning can be expected to be more elaborate in those areas. As reasoning can be learnt to a degree it can be expected to reflect the general standard of education of the source. The index expresses the source's ability to correctly follow and handle longer chains of reasoning and the system can gain information about the correctness of these chains of reasoning the source may convey to the system. The source control mechanism does not deal with the intricacies of chains of reasoning and they should be detected and checked by the general reasoning system.

The source's ability to judge sources of information describes the same situation the source control mechanism addresses, namely judging uncertain information from human sources. In order for the system to verify what the source's abilities are, depends on comparing the source's reactions with the system's evaluation and in order to do that it requires that the system knows the source's source of information as well as the basis of information on which the source built its belief. A full set of information to that effect is hard to come by and may come usually only from an enquiry by the system into how the source came to its belief. In most cases the system will only get the source's belief, which is a refined information, without a long winded explanation of how the source arrived at it from raw data.

Finally, the index about the source's experience records the source's ability to handle correctly one's personal experiences. This is perhaps the most difficult index to determine as it is not possible to experience other people's experiences. At the same time information based on personal experience tends to be very subjective by nature and usually no action is based on it and in those cases where that is necessary it has to be taken on trust. One way in which the index can be assessed is by looking at how much the source is inclined to use subjective experiences in support for supposedly more objective judgements. Subjective experiences should usually only be used for matters concerning oneself but not on matters of more general interest. Concrete information about the source's views and subjective perceptions can usually only come directly from oneself and one should not use other people's experiences as a basis for action. For example, to use someone else's experiences as a basis of one's beliefs suggests a disability to handle personal experience and provide one way in which the system can assess the source's capabilities in that area.

- supposing a source has expertise in an area it is likely to make stronger and more precise claims in that area than in other areas where it has less expertise
- therefore, if a source uses a lot of technical language then it is likely to have expertise in that area

- as opposed to the question of trustworthiness, sources are usually reasonably open about their areas of expertise and can therefore be asked about it
- if a source has expertise in an area it is less likely to make mistakes there and consequently when there are very few mistakes on a particular subject then there may be expertise in the area
- in the case where a source is in a professional classification then the expertise requisite for that profession is almost guaranteed to be evident in the source

4.10.4 Maintenance of Trust Related Indices

The trust related indices reflect the source's willingness to give good information as opposed to its capability. The willingness to supply information to the best of one's capabilities appears to be dependent on three major considerations:

- the source's fundamental beliefs
- the source's basic interests
- the source's special relationships

A source's trustworthiness and helpfulness are at least in parts dependent on one's interests and beliefs, but may also depend on the source's membership of certain classes. Whereas the source's beliefs and interests may influence the source's behaviour in general, the source may change in its behaviour if it is in a certain role. For example a salesman may display a certain behaviour if he is talking to people in a general situation and at the same time he may react differently when he is acting as a salesman. Accordingly, the beliefs and interests are more stable whereas the relations change with the context and depend on which role-model obtains between the source and who it is communicating with.

Maintenance of the Belief Index

The belief index reflects particularly strong general beliefs which may be explainable through the source's education and profession. Although that does not imply that the source is likely to deceive, some beliefs have an effect on the objectivity of the information. A particular political view may influence one's opinions of how the health service should be run, environmental issues, the judicial system, and so on. The detection of such beliefs can be very difficult unless one gets direct information when the respective subjects are being openly discussed, the system can otherwise only detect them through trend analyses which can be triggered through particularly strong expressions of the source on subjects which are contentious. Other than that, it may be ascribed through an association with a class.

- strong beliefs are often born out of ideologies or out of life experience and can be inferred from the source's social position or education
- beliefs do not tend to change with circumstances, but there tend to be situations where they apply while they do not touch other areas at all
- not all strong beliefs have an adverse effect on information and there are those which will compel the beholder to actually act with integrity
- since beliefs born out of ideologies are often the product of upbringing and education, they can be inferred from membership of social classes

Maintenance of the Interest Index

Leaving aside more platonic interests like hobbies and pastimes, we are concerned with more substantial interests which have strong motivations associated with them, like a source's profession and general goals in life. Again, it appears that indications to that effect can be obtained through the source's membership of classes where the principal properties are connected to strong goals, like financial gains, reputation and power which are particularly apt to give rise to vested interests.

- it appears that if a source is lying there may be a hidden interest which may be to do with the information or the situation
- interests are mainly born out of strong relationships, involvement of money, power or status
- interests may not be personal but be part of one's position or profession and if a source is acting in that capacity one may have to make distinctions
- when strong relationships are involved in the context of the conversation there may be interests in operation
- alternatively, it seems advisable to be very careful when there are financial gains involved as many people will lose their scruples at that point

Maintenance of the Trustworthiness Indices

While the fundamental beliefs and interests will be more stable, trustworthiness and helpfulness appear to be dependent on relations between the source and who it is communicating with and is therefore very context sensitive. Accordingly, there may also be a number of different sets of indices to reflect that. Some of them appear to be very general and applicable to most sources, like in situations were we are relating to friends we will be expected to be completely trustworthy and helpful. The same should be true if there is a relation between employee and employer or a witness and a judge, and so on. Alternatively, as we have seen in a relation between a dealer and a client the trustworthiness is expected to be low while the helpfulness is expected to be high. It is also plausible that as we get to know people we may quickly start to trust them and be helpful although the trust may collapse rather quickly and be difficult to regain if we have been let down. Indeed for that to happen does not take much evidence as may be required for some of the other indices of the source model.

- if the source has strong interests or beliefs which are the product of strong motivations, then if the information is about that it is not trustworthy

- if a person gets into a special relationship with a source the trustworthiness may change drastically
- if there is a particularly strong relation between sources then that will take precedence even when interests are involved
- if there is a strong relation of say friendship then the source will declare interests and beliefs more readily
- The expected helpfulness should be at a similar level to trustworthiness, except that in some situations the trustworthiness is high when the helpfulness is low, because the person may be trustworthy but is not allowed to give information
- Alternatively, there are situations where a person cannot be trusted, while one will expect him to be helpful, especially if the source will gain from the cooperation of another person, like in a businessman-customer relationship

4.11 Conclusion

Following on from our presentation of the current state of research, we have in this chapter presented and discussed the principles which should govern a source control mechanism in managing uncertainty and forming beliefs.

With a view to the specific aim of dealing with the management of uncertainty of information from human sources we first had to put the prospective source control mechanism into the context of existing work in that field. Following on from Garigliano's work, it appears that the natural habitat of such a model is between Rescher's model for plausible reasoning and Cohen's model of endorsement.

Subsequently we have described the constraints placed on the behaviour of the source control mechanism. In the task of forming beliefs and learning about sources and about the world, a system which wants to produce a realistic response has to be flexible. Thus it has to be capable of actively resolving conflicts in its world model, cope with varying conditions and maintain interaction with its environment.

Given the constraints it was necessary to define the basic strategy of the mechanism needed to produce the expected behaviour. The source control mechanism produces and maintains a world-model which comprises a system of beliefs and models of sources. Seen from an operational view, this implies that the model has to evaluate information and resolve conflicts with the help of source models and retrospectively assess whether the source models are adequate or are in need of revision.

This basic strategy in general, and the operational principles in particular, translate into the need for a number of different units required to translate these constraints and basic principles into an operational model. In this model an initial check of importance is required to determine how much interest to take in any particular piece of information. This is followed by an evaluation of the information to determine its believability. As the incorporation of information into an existing system of beliefs can cause problems, there is also a need for a mechanism to deal with conflicts. Since both processes can run into difficulties, it is also necessary to be able to steer an enquiry to find explanations and solutions to possible problems. Finally, as the source models are the primary tool for the belief formation process as a whole, it is important to have a review process by which changes can be made to source models in the light of new evidence.

In this chapter we have given a comprehensive presentation of the principles which govern the source control mechanism. We have shown that there are general principles involved in the task of belief formation, conflict resolution, enquiry and re-evaluation of source models which can be marshalled to produce the expected behaviour of a system which is to model this. In order to reinforce the view that this behaviour is reproducible and will behave in the indicated way we give a high level design and full examples in the next chapter.

Chapter 5

The Design for the Source Control Mechanism

5.1 Introduction

In the previous chapter we have explained the principles which determine the behaviour of the source control model, the information it processes and the sources it deals with. There, the emphasis has been to describe the model in terms of its external behaviour as it can be observed from the outside. In this chapter we complete the last steps in our methodology, by demonstrating first that the principles established in the previous chapter can be translated into a design, which ultimately can be transformed into an implementation. Secondly, we show that the mechanism works on two of the examples, and that we therefore have completed our original claim that we can devise a mechanism which can deal with them.

In this chapter we need to take a much more technical approach, to specify the various entities and functions needed to produce this behaviour. Therefore, we provide a high-level design of the model, which gives enough detail to be transformable into a more detailed design from which an implementation can be produced. To give a more detailed design here would not necessarily be more helpful, as that will usually have to be oriented on the particular implementation technology and that would also loose sight of the fundamental structure and functionality and to give a general design is therefore more adequate for our purposes.

In the following, we first give a brief description of how the source control mechanism would relate to other modules in a larger system; we consider a reasoning manager and an endorsement module. As the most likely application of the source control mechanism is as a sub-system to order the beliefs of a system capable of natural language processing, we also describe the likely interface that can be expected between the two.

Next, we explain the kind of data structures the source control mechanism requires for its task, including the event-based representation provided by the natural language environment.

Then we present a possible architecture model of the mechanism, which gives a high-level view of the mechanism. This also helps us to identify the major components of the system and to show how they are related.

In the sections which follow, we look at these components in turn and present their general design in rule form. Following the pattern laid out in the previous chapter we identify five components, with the addition of a control regime, required to steer the process of analysis.

Finally, we look at two full examples, chosen from the initial problem statement of Chapter 2. This will demonstrate that the mechanism can deliver the results initially claimed, and also has the advantage of showing the source control mechanism in operation, and will help the reader to get a feeling of how the mechanism functions in practice.

5.2 External Constraints

The source control mechanism is not designed to operate in isolation, but as a subsystem in the context of a general reasoning system to deal with the problem of belief formation about information from human sources. There are a number of important external constraints placed on the relations of the mechanism to other units in such an environment, which come under the following headings:

- the relation to an endorsement reasoning module
- the relation to a reasoning manager
- the relation to a natural language processor

Relation to an Endorsement Module: The source control approach is well equipped to deal with common, everyday situations, which are not usually exclusively decided on the grounds of the arguments involved, but where there are also considerations about the agents involved, their interests, motivations and relations. A pure endorsement approach would therefore not function appropriately in these situations. At the same time there may be situations where decisions have to be taken on the basis of the arguments and it would therefore be advantageous to divide the task between the two approaches and to transfer cases to an endorsement module if an analysis of the arguments is required.

Relation to a Reasoning Manager: In situations when the information is primarily the product of long chains of reasoning, the source control mechanism would want to enquire of a reasoning manager whether the reasoning is correct. For example, when the information is the product of reasoning, then the uncertainty needs to be judged according to different criteria, relating to the various techniques employed and what effect they have on the certainty of their conclusions. The source control mechanism is primarily geared to deal with initial levels of belief from human sources and the impact of application of reasoning techniques which are not truth preserving, is not strictly a part of that problem, except on a very superficial level. To be able to deal with more aspects of uncertainty, it would be advantageous if the source control mechanism could enquire into the quality of reasoning and the quality of the results.

Relation to a Natural Language Processor: Given a general reasoning system capable of natural language processing (which is able to generate the events which are the input to the source control mechanism) the primary purpose of the source control mechanism is to order and maintain the beliefs of such a system about information from human sources. Although the source control mechanism will be able to operate relatively independently, it will function more successfully if it can draw on the resources of a natural language processor, such as gaining information about the relation amongst data and information to help in the classification of sources as well as definitions of concepts. In order to operate at its full potential, a number of connections between these two components would be required.

Given that the source control mechanism (SCM) is closely integrated with a natural language processor (NLP), there are a number of different flows of information between the two:

Input to the SCM:

- Events: The SCM is working with an event-based representation and receives these events from the NLP and has to form beliefs about them.
- Source Models: Source Models are intended to be stored in the database of the NLP and the SCM receives the source models from the database appropriate to the events, which it uses in the evaluation process and may reassess during source re-evaluation.
- NLP Queries: The NLP may ask advice about how to improve weak beliefs.
- Requested Data: The NLP will return information in response to queries by the SCM.

- Enquiry Response: The System will return its response to requests for enquiry by the SCM.
- Stop Request: The NLP may run out of time and request an immediate decision.

Outputs from the SCM:

- Events: They are returned to the NLP with the beliefs attached to them.
- Source Models: They are returned for storage to the database. These source models are also changed by the SCM if appropriate.
- SCM Queries: The SCM may have to ask for clarification from the NLP on some aspects of the events.
- Recommendations: In response to NLP queries the SCM will return recommendations of how the belief in an event could be raised.
- SCM Enquiry: The SCM may ask the main system controller to gain more information by engaging in communication with sources.

5.3 Basic Data Structures and Basic Functions

There are three basic data structures the source control mechanism is concerned with:

- events
- source models
- case

5.3.1 Events

The natural language processor takes input directly from human sources and transforms the input into a set of events. The source control mechanism takes these events as its input from the natural language processor and attaches beliefs to them. A single piece of information given by the source may be analysed and translated into a series of connected events, some of which may be hypothetical and others real. The source control mechanism has the task to assign beliefs to the real events in this collection. Events are a collection of slots, which may vary in their composition and the following event is a typical example:

```
Event1:
 subject: John
 action: own
 object: desk1
 time: present
 source: Albert
 status: real
 certainty: high(default)
 belief: ?
```

5.3.2 Source Models

The SCM maintains the source models which are used in the information evaluation and belief assignment process and are adapted and refined according to changing circumstances by the source re-evaluation process. According to the principles, the following categories have to be represented and there are also defaults used for new sources in the absence of more concrete information:

- Ability - In the absence of concrete information it is assumed that the source displays average properties in reasoning and judging information, does not have any particular expertise and can confidently handle its experiences.
- Interests - If the system has no particular knowledge about the source's particular interests it is assumed that there are none.
- Beliefs - Again, if the system is not aware of particular strong beliefs it is assumed that there are none.

- Helpfulness - In absence of further information this is assumed to be average towards its general environment.
- Trustworthiness - Again, in the absence of specific knowledge it is assumed to be average.

There are two kinds of default which are employed in the source models, namely defaults born out of ignorance and defaults which are the result of a classification. Defaults from classifications are superior to ordinary defaults in that they are based on indirect knowledge. Through the knowledge that the source can be considered as a member of a class, a number of typical properties of that class can be attributed to its members and replaces defaults of ignorance. In the same way, factual knowledge will replace both kinds of defaults. Consequently, we can establish the default source as:

```
Source1:
   Ability:
      expertise: none(d)
      reasoning: average(d)
      judging information: average(d)
      experience high(d)

   Beliefs: none(d)
   Interests: none(d)

   Source1 -> system:
      Helpfulness: average(d)
      Trustworthiness: average(d)
   System -> Source1:
      Helpfulness: high(d)
      Trustworthiness: top(d)
```

5.3.3 Cases

In the evaluation process the SCM needs to keep records about the results of the examinations carried out and their results in order to avoid continuous re-examination at various stages. This requires that each stage in the analysis process can record its verdict and reasons as well as more general categories, which may be of interest later on. This will take a form like:

```
Case1:
   Event: event1
   Source: source1

   Importance Analysis:
      Source: yes
      Information: maybe
      System Motivation: no

   Information Evaluation:
      Determination: ok
      Problem of Responsibility: no
      Problem of Advantage: no
      Problem of Ability: no
      Problem of Trust: no
      Result: believe as given

   Conflict Resolution:
      Connection: yes
      Event: system-belief397
      Source: source4
      Same Context: no
      Type: reinforcement
      Problems of Trust: none
      Problems of Ability: none
      Result: reinforce system_belief and add source

   Source Model Re-evaluation:
      Classifications: clerical?
      Expertise: none(d)
      Reasoning: average (d)
      Judging Information: average(d)
      Experience: high(d)
      Beliefs: none(d)
      Trustworthiness: average(d)
      Helpfulness: average(d)
      Result: maybe clerical classification
```

A structure like this makes it easier to look up the general results and problems of a previous stage of examination as a guide for further analysis and as a means to take an immediate decision if necessary. Entries may also be replaced by results from further analyses and enquiries.

It should be noted at this point that the cases are available to the different components as they consider a piece of information and the results from previous steps of investigation can be used in directing the analysis process.

5.4 The Architecture Model

From the presentation of the principles in the previous chapter we can see that there are a number of clearly identifiable parts to the source control process. These parts appear to translate readily into separate components in the design. They are reasonably self-contained and are able to operate in relative isolation. They could also operate in parallel, subject to a controller passing on the cases from one stage to the next. Consequently, we can establish an operational architecture model with six components:

- control regime
- importance analysis
- information evaluation
- conflict analysis
- enquiry
- source re-evaluation

The source control mechanism needs to be controlled to adapt to the requirements of each case and changing circumstances. It needs to start the process, determine how much time and effort to spend on it and bring it to an end. The controller will also need to liaise with other system components and sources.

The importance analysis will try to see whether the situation is important for the system. That check should be fairly shallow to see whether there is a reason to go on, as to go deeply into it gets very complicated quickly. As a result the information will either proceed to the next stage or be expelled immediately. The importance will come from motivation of the system towards the source of the information or because of a system interest.

In the second stage the information is analysed to see whether it is well-formed and whether there are missing parts. If there is a problem then one needs to consider whether the case is worth pursuing further. If the problem can be solved more easily then it may be worth carrying on, but if the problem is more severe the case needs to be sufficiently important. In those cases it may be necessary to do an enquiry, but that depends on being sufficiently important.

In the third check we need to consider whether the information touches on other information already in the system. If there is a connection the system needs to see whether it is from the same source or a different one and whether there is a problem. If there are minor problems the system will try to carry on, but if there is a serious one than that again depends on the importance on whether the system will try to get an explanation or pursue an enquiry.

The enquiry is dealing with a variety of cases and tries to gain an explanation for problems by exploring the intricacies of the information and source either by exploiting the existing material more thoroughly or by trying to engage in interaction with the respective source or other sources which may be able to help.

Once the analysis and belief formation process has finished the results are passed to the source re-evaluation process, which has to consider the source models in the light of new information and to decide whether it is necessary to adjust them.

Considering this division and the relations between the components the following diagram shows a possible organisation for the source control mechanism. In this representation it is assumed that events, source models and cases are handed from one process to the next so that the receiving process has all the information available to operate. As this includes *cases*, this also implies that processes are able to draw on the findings of previous steps of analysis should that be required.

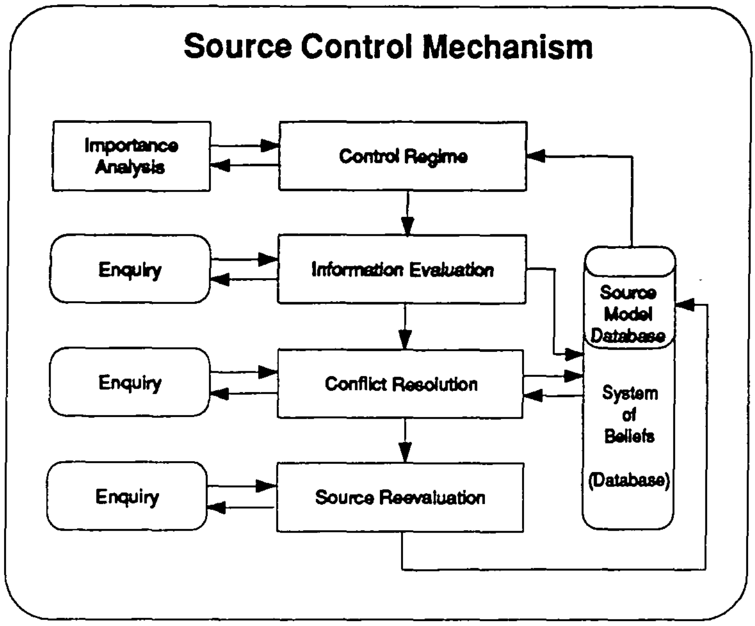


Figure 4: Components of the SCM

Having gained a first impression of the organisation of the SCM we now need to look at these components in more detail. It should be noted that this division follows the divisions of the previous chapter and in order to avoid unnecessary reiterations the material presented there will not be repeated here. We therefore recommend to refer back to the respective sections as they represent the requirements for the design in this chapter.

5.5 The Control Regime

The requirements of the source control mechanism are very versatile as can be seen from the fundamental constraints of the behaviour of the system. As the mechanism is rather complex, this also requires management of the actual process of analysis. The different tasks we identified and the principles that have emerged need to be translated into a flexible architecture which can adapt and respond to the semantics of the problem as well as to the type of situation.

The control regime controls the processing of the source control mechanism. This is achieved through a procedure where the control regime passes on *cases, events* and *source models* from one unit to the next so that each process has the necessary data to work on and can use the results of previous evaluations in its analysis. The control regime has a number of tasks to perform, which can be divided into two main areas:

- technical management of belief formation
- strategic management of the behaviour

The technical management of the belief formation process requires analysing the importance of the information and performing the processes of belief formation, conflict resolution and source re-evaluation to the extent warranted by the significance of the information or sources involved.

```
If there is a new case,
    then request an initial importance analysis

If the importance analysis shows that
    the information or source is interesting or the system takes interest,
       then request information evaluation
       else store information and evaluate when required

If the information evaluation shows that
    the interest in the information is greater than
    the problems with it,
       then request conflict resolution

If there are problems and the case is important,
    then try to do an enquiry by communicating with sources,
       else try to find an explanation by introspection

If there are problems and the case is not important,
    then else return result so far and store information

If after the conflict resolution there are problems involved,
       then try to find out more
       else return the result
```

If the case has been completely analysed,
then request a source re-evaluation

If the source re-evaluation is having problems
and the source is of above average importance,
then try to enquire to find an explanation

The strategic management will have to consider the system's major objectives and to pursue them in the current situation, such as whether it is currently communicating with single or multiple sources, whether it wants to continue communicating, whether there are demands placed on it and whether it wants to communicate with someone else. (How far this can be achieved depends on the general reasoning system as the source control mechanism is not entirely in control of the behaviour of the system.)

If the SCM is not communicating with ,
then try to communicate to get information or spend time
analysing unresolved problems

If the SCM is communicating,
then try to be helpful and maintain communication
and don't change sources unless it is necessary to

If the NLP requests information of the SCM,
and there is information in the system,
then try to find information and give answer
elseif there is no information in the system,
then report that there is no information
elseif the source presses for more information,
then try to produce it or suggest where it may be found

If there are problems with the information
and the information is important,
then ask the source for help
elseif there is a problem with the source,
then ask the source if that is feasible,
else try to find explanation and ask someone else

If there are problems and the information is important
and the SCM has to talk to other sources,
then try find out about them first by asking the current source

If there are problems and different sources need to be asked,
then start with the best source first
and proceed in order of trust and ability

5.6 The Importance Analysis

Embarking on a detailed analysis of a piece of information can be expensive especially when the situation is not straightforward. A vital step in the analysis process is therefore to determine whether the piece of information is worth the effort. An importance analysis is therefore a good starting point for the work of the source control mechanism. Given a source model and an event there are three ways in which importance can initially be established:

If the system has an interest in the subject of the information,
then the information is important to the system

If the information is strong and there is a connection to existing information,
then the information is important to the system

If the helpfulness of the system towards the source is high,
then the source is important to the system

The process of deciding whether a piece of information is important can be a considerable task, especially when that is not immediately obvious. Therefore, an importance check should initially be only very shallow, looking for obvious indicators to decide whether to take an interest. If there is a potential interest in the situation, then the system may examine the situation more thoroughly.

Although the importance check is shown at the beginning of the analysis process, it may also occur at a later point when there is a need for an enquiry. As enquiries are expensive, especially when they involve procuring information from sources not currently available to talk to, the system will have to determine whether it is worth pursuing the matter any further. The following rules are concerned with information which has been analysed already and where the importance analysis has to decide whether there is enough interest in the situation to make further enquiries:

**If there is a problem and the information and the source are not of interest,
then recommend not to enquire**

**If there is a problem and the effort required to solve it
is greater than the source or information warrants it,
then recommend not to enquire**

**If there is a potential problem with ability
and the information is of interest to the system,
then recommend further information analysis and enquiry**

**If there is a potential problem with ability
and the source is of interest to the system,
then recommend source analysis and enquiry**

**If there is a potential problem with trustworthiness
and the information is of interest to the system,
then recommend an enquiry**

**If there is a potential problem with trustworthiness
and the source is of interest to the system,
then recommend source re-evaluation and investigation**

**If there is a conflict with other information
and the source or information is of interest to the system,
then recommend further investigation**

During the analysis, the system needs to consider carefully whether the case is significant enough or whether the effort in producing a result is commensurate with the expected gains. It is therefore advisable to perform a quick analysis until obvious faults become apparent and to reconsider whether there is enough reason to go further.

5.7 The Information Evaluation Process

During the initial analysis of pieces of information, the source control mechanism needs to look at the information itself and at the source associated with it to see whether there are any problems or whether the information, on its own, can fundamentally be believed. There are two aspects to the analysis:

- the information analysis
- the source analysis

The information analysis looks at whether the information is well formed, whereas the source analysis looks at whether the information is credible considering the source involved. Both considerations are necessary to be able to make a decision on the credibility of the information and therefore have to be taken in conjunction:

5.7.1 Information Analysis

The information analysis process has to consider whether the information package is well-formed, whether there are missing parts in that package and what the information is about.

If the information is inconsistent, then it may have to be rejected unless there are reasons to do an enquiry. If there are missing parts, then it will depend on the claimed strength of the information whether it will be worth trying to fill the gaps. Fundamentally, there are three checks to be carried out:

- the relative strengths of conviction of the source
- whether there are indications to assume responsibility
- whether there are advantages involved

Checking the Strength of Conviction

The system needs to consider the degree of determination. If the information is under-determined then the source implicitly acknowledges that there is room for doubt, whereas if the information is almost fully determined then the source appears to indicate that it is well informed. If the information is, however, over-determined the source implicitly shows that it is not very competent in its handling of information.

```
if there is a conviction for the information
    but there is no conviction against,
    then determination -> ok
```

If there is a conviction for and a conviction against,
then
If conviction for + conviction against > top,
then determination -> overdetermined
If conviction for + conviction against = top,
then determination -> optimal
If conviction for + conviction against < top,
then determination -> underdetermined

Checking Responsibility

In the next step the mechanism has to determine whether responsibility is assumed for the information. It is unusual to find that sources will directly state whether they do, but it may be possible to infer from the context and the source model. The system will therefore look at the source's relation to the system and whether the source is trustworthy as a precondition it will then look at the level of helpfulness. If the information was asked for, and given to the system only reluctantly, then there is an indication that the source will be reluctant to accept responsibility. If the information was given readily, or volunteered, then the source is more likely to accept responsibility and the system can use this information as a default. Otherwise, the system has to go further into a semantic analysis.

If responsibility is assumed,
then Responsibility -> assumed

If responsibility is denied,
then Responsibility -> denied

If responsibility is not indicated and trustworthiness is low,
then Responsibility -> positive assurance required

If responsibility is not indicated and trustworthiness is high,
then
If Helpfulness is high and information readily given
then Responsibility -> expected
elseif information not readily given,
then Responsibility -> doubtful

If responsibility not indicated and trustworthiness is average,
then
If Helpfulness is low and information volunteered
then Responsibility -> suspect
elseif information asked for,
then Responsibility -> expected

```
If the source is asked and accepts responsibility,
    then Responsibility -> accepted
    else Responsibility -> denied

If a different source is asked
    and its ability to judge sources is high
    and its trustworthiness is high,
    and it agrees
       then Responsibility -> expected
       else Responsibility -> doubtful
```

Checking for Advantage

The source's advantage is usually not declared, especially when the source may try to deceive the person it is talking to. If there is no declaration of advantage, this information may have to be inferred, and there are two ways in which the system can do this:

- by examining the source model
- by getting the information from third parties

The first strategy involves going into an analysis of the context in which the information was given, to examine whether there are relations between the actors, where the source potentially could have vested interests. This needs to be compared with the kind of information to see whether the information falls into one of these categories:

```
If there is an interest of the source which includes the
    subject of the information,
       then Advantage -> possible
    else
       Advantage -> unlikely

If there is a special relation between the source and the
    system, or a source the system acts on behalf,
    and trustworthiness is low and the relation is built
    on an interest,
       then Advantage -> expected

If the information is relayed and the relaying source agrees that there may be an
    advantage,
       then Advantage -> expected
```

The last strategy can be applied when the information is being relayed and the relaying source can be questioned, or the source volunteers information about vested interests.

The latter situation will not arise often, and the source control mechanism will therefore have to rely mainly on the former strategy, which involves going into a source analysis, checking the source's relations and known interests for matches with the particular subject of the information and also by a comparison between the level of helpfulness and actual behaviour. If the source is more helpful than expected by the system and the trust is not high, then there may be a suspicion that the source may have an interest.

5.7.2 Source Analysis

Although the source models are maintained by the source control mechanism, they are integrated in the database of the NLP and are supplied to the source control mechanism with the information.

There are two different checks the source analysis has to perform on the given piece of information, to examine the information from a source model point of view:

- to analyse the trustworthiness
- to analyse the ability

In the information analysis process, the source model was used to derive missing parts and to fill gaps in the information package. The source analysis now has to take a step back, to consider whether the information can be believed and if so, how much, under the consideration of both trustworthiness and ability.

Apart from the considerations of the information analysis, which decide whether the information is acceptable in itself the source analysis will need to consider whether it can trust the source and whether the source is qualified to make the statement it did.

Checking Trustworthiness

In the process of checking the trustworthiness, the source control mechanism has to take into account the trustworthiness of the source, as well as whether responsibility is assumed and whether there are hidden advantages:

```
If the trustworthiness towards the system
    or, the source the system acts on behalf of, is low or average
    and the conviction is high
    and Responsibility is doubtful,
    and Advantage is at least possible
       then Problem of Trust -> possible
    elseif Responsibility is denied,
       then Problem of Trust -> yes
    elseif Responsibility is accepted or expected
       then Problem of Trust -> no

If the trustworthiness towards the system
    or the source acted on behalf of is high
    and the conviction is high
    and Responsibility is doubtful,
    and Advantage is at least possible
       then Problem of Trust -> possible
    elseif Responsibility is denied or there is an Advantage,
       then Problem of Trust -> yes
    elseif Responsibility is accepted or expected
       then Problem of Trust -> no

If the conviction is low and trustworthiness is low,
       then Problem of Trust -> possible
    elseif trustworthiness is high,
       then Problem of Trust -> no
```

Checking Ability

Having checked whether the source fundamentally can be trusted and having found that the source model and the information the source has provided are satisfactory, the source control model now has to assess whether the information is credible under the point of view of the source's abilities. In other words, is it competent enough to make such a statement.

The Ability index contains a number of indices and the system will need to apply the appropriate index. It is therefore necessary to determine the type of the information, whether it is born out of personal experience, reported from other sources, the product of reasoning or whether it requires expertise. The system will need to determine the dominant feature of the information.

In order to be able to assess this, the source control mechanism first has to find out what type the information is in order to determine whether the source is competent enough. These types need to be identified with some heuristics and shallow pragmatics.

If the information is originally from another source
    and the conviction is low
    and the ability to judge sources is at least low,
        then Problem of Ability -> no

If the information is originally from another source
    and the conviction is high
    and the ability to judge sources is high,
        then Problem of Ability -> no

If the information is originally from another source
    and the conviction is high
    and the ability to judge sources is low,
        then Problem of Ability -> yes

If the information is based on reasoning
    and the conviction is low
    and the ability to reason is at least low,
        then Problem of Ability -> no

If the information is based on reasoning
    and the conviction is high
    and the ability to reason is high,
        then Problem of Ability -> no

If the information is based on reasoning
    and the conviction is high
    and the ability to reason is low,
        then Problem of Ability -> yes

If the information is based on experience
    and the conviction is low
    and the ability to judge experience is at least low,
        then Problem of Ability -> no

If the information is based on experience
and the conviction is high
and the ability to judge experience is high,
then Problem of Ability -> no

If the information is based on experience
and the conviction is high
and the ability to judge sources is low,
then Problem of Ability -> yes

If the information requires expertise
and the conviction is low
and there is no record of that subject,
then Problem of Ability -> no

If the information requires expertise
and the conviction is high
and the ability in that subject is high,
then Problem of Ability -> no

If the information requires expertise
and the conviction is high
and the ability in that subject is low,
then Problem of Ability -> yes

Given that the information is well formed, the system will pay attention in particular to information with a certainty above the medium level and decide whether to reduce the credibility to the level of the respective index of ability. For sources which are considered to have a high ability, the system may therefore decide to maintain a belief at their present level:

If the conviction in the information is low
and there is a problem of trust,
then Belief -> nil

elseif the conviction in the information is low
and there is no problem of trust,
then Belief -> low

If the conviction in the information is high
and there is a problem of trust
and there is a problem of ability,
then Belief -> nil

If the conviction in the information is high
and there is a problem of trust
and there is no problem of ability,
then Belief -> low (not for use)

If the conviction in the information is high
and there is no problem of trust
and there is a problem of ability,
then Belief -> level of ability

If the conviction in the information is high
and there is no problem of trust
and there is no problem of ability,
then Belief -> high

5.8 The Conflict Analysis Process

After the information has been analysed in isolation, the conflict analysis process needs to consider how the information relates to other information available to the system. In that case, the system needs to decide whether important information, motivations or sources are involved, in order to determine how far to pursue the matter. There are two main categories to be considered:

- single-source conflicts
- multiple-source conflicts

These two categories differ considerably in their implications and requirements for their treatment. Therefore they need to be considered separately.

5.8.1 Single Source Conflicts

Single source conflicts have the advantage that when dealing with one source, and there is a problem with the source, it is obvious who is to blame whereas with multiple source problems that is still to be determined. Since there are fewer criteria to consider the whole process is much more simple, although a good solution may still be hard to find.

It is also important to note that there are significantly different implications with respect to what looks like being the same situation. Thus while a reiteration is of no great significance in a single source case it can be important in multiple-source situations and what is an unacceptable self-contradiction in a single-source situation is quite feasible in a multiple-source context even though that may not help the system in finding out what to believe.

If we remind ourselves from the discussions of the principles, there are four possible situations to be considered:

- both are roughly identical
- the new information reinforces the old
- the new information weakens the old
- both are in contradiction

Reiteration

If both pieces of information are roughly identical, the same statement is made twice on different occasions. This will not have much consequence for the belief held by the system other than to indicate that the source is still subscribing to its previously aired opinion. Reiterating one's position does not make that position any stronger, unless more, or different reasons are given in its support. In practical terms this means that more arguments are given for conviction or the source may explicitly assume responsibility, thus making the information more substantial while keeping the conviction at its previous level:

```
If the old conviction is equal to the new
    and the conviction in the new information is low
    and there is a problem of trust,
       then Belief -> nil

elseif the old conviction is equal to the new
    and the conviction in the new information is low
    and there is no problem of trust,
       then Belief -> low

If the old conviction is equal to the new
    and the conviction in the new information is high
    and there is a problem of trust
    and there is a problem of ability,
       then Belief -> nil
```

If the old conviction is equal to the new
and the conviction in the new information is high
and there is a problem of trust
and there is no problem of ability,
then Belief -> low (not for use)

If the old conviction is equal to the new
and the conviction in the new information is high
and there is no problem of trust
and there is a problem of ability,
then Belief -> level of ability

If the old conviction is equal to the new
and the conviction in the new information is high
and there is no problem of trust
and there is no problem of ability,
then Belief -> high

Reinforcement

Cases where the source reinforces its previous statement, may lead the system to reinforce its belief. The system will have to determine whether to raise its belief, but that is dependent on the solidity of the information:

If the new conviction is stronger than the old
and there is a problem of trust
and there is a problem of ability,
then keep old belief

If the new conviction is stronger than the old
and there is a problem of trust
and there is no problem of ability,
then Belief -> low (not for use)

If the new conviction is stronger than the old
and there is no problem of trust
and there is a problem of ability,
then Belief -> level of ability

If the new conviction is stronger than the old
and there is no problem of trust
and there is no problem of ability,
then Belief -> high

Weakening

If the new piece of information is weaker than the statement made previously, then the source still holds on to the same opinion, but with less conviction. This means that support for the strength of the belief has been withdrawn. Unless the system can find support for its belief somewhere else, it will have no choice but to reduce its belief on the basis of the new information:

If the new conviction is weaker than the old
and there is a problem of trust
and there is a problem of ability,
then Belief -> nil

If the new conviction is weaker than the old
and there is a problem of trust
and there is no problem of ability,
then Belief -> low (not for use)

If the new conviction is weaker than the old
and there is no problem of trust
and there is a problem of ability,
then Belief -> level of ability

If the new conviction is weaker than the old
and there is no problem of trust
and there is no problem of ability,
then Belief -> level of new conviction

Contradiction

Potentially the most serious situation arises when the source is contradicting its previous statement. How serious this is, is dependent on the level of contradiction and the importance of the situation.

If the contradiction occurred between pieces of information of low claimed certainty, the case can more easily be dismissed as the information was not substantially supported before, and can be considered to describe a change of mind. If the contradiction occurred at a higher level then there are two possibilities: Either the source did contradict itself, or there has been a change in the environment.

In order to find out whether there is a change in the environment, requires confirmation from the source or from another source who might know about that. Alternatively, it has to be decided whether the subject matter is such that a change is plausible, considering the time that has elapsed since the old information was first given:

If the contradiction is on a low level
and there is a problem of trust
and there is a problem of ability,
then both Belief -> nil

If the contradiction is on a low level
and there is a problem of trust
and there is no problem of ability,
then Belief -> low (not for use)

If the contradiction is on a low level
and there is no problem of trust
and there is a problem of ability,
then Belief -> low

If the contradiction is on a low level
and there is no problem of trust
and there is no problem of ability,
then Belief -> level of new conviction

If the contradiction is on a high level
and a change of environment is plausible
and there is a problem of trust
and there is a problem of ability,
then both Beliefs -> nil

If the contradiction is on a high level
and a change of environment is plausible
and there is a problem of trust
and there is no problem of ability,
then Belief -> low (not for use)

If the contradiction is on a high level
and a change of environment is plausible
and there is no problem of trust
and there is a problem of ability,
then Belief -> level of ability

If the contradiction is on a high level
and a change of environment is plausible
and there is no problem of trust
and there is no problem of ability,
then Belief -> level of new conviction

If the contradiction is on a high level
and a change of environment is not plausible
and there is a problem of trust
and there is a problem of ability,
then both Beliefs -> nil

If the contradiction is on a high level
and a change of environment is not plausible
and there is a problem of trust
and there is no problem of ability,
then Belief -> nil

If the contradiction is on a high level
and a change of environment is not plausible
and there is no problem of trust
and there is a problem of ability,
then Belief -> nil

If the contradiction is on a high level
and a change of environment is not plausible
and there is no problem of trust
and there is no problem of ability,
then Belief -> low

5.8.2 Multiple Source Conflicts

Compared with single source situations, multiple source conflicts are much more complicated and a number of additional criteria have to be taken into account. As we have seen from the discussion of the principles of multiple source conflicts, amongst other things the solidity of information/beliefs and the corroboration of beliefs play a prominent role in the decision process. Also whereas single-source conflicts are concerned more with the intricacies of the information, in multiple-source cases one tends to consider pieces of information at a higher level. Thus, if there are two pieces of information from different sources, it is more a matter of one position against another. If a more precise solution is required, then a more detailed analysis will have to be performed, guided by the particular problem in question.

When there is a conflict between two sources and one source is considerably more reliable than the other, the system will tend to go along with the strong source and the system will also be disinclined to let the position of the stronger source be much affected by the antagonism of the weaker source. Alternatively, if the system receives two compatible pieces of information from different sources, it is likely to combine the two pieces of information into the system's belief while adding the new source to the list of supporting sources and adding the new arguments for and against and reconsidering new levels for c and d.

Looking at the problem in more technical terms, there are again four different situations to be considered:

- both are roughly identical
- the new information reinforces the old
- the new information weakens the old
- both are in contradiction

Reiteration

When we are dealing with a reiteration of an existing belief by a different source, the new information will go to confirm the existing belief. If different sources supply the same piece of information, it has a corroborative effect on the system's belief, even though that does not raise the levels of belief. This basically means that the new source is added to the list of sources supporting the system's belief:

```
If the old conviction is roughly equal to the new
    and the level of convictions are low
    and there is a problem of trust,
        then Belief -> nil

elseif the old conviction is roughly equal to the new
    and the levels of conviction are low
    and there is no problem of trust,
        then add source to support

If the old conviction is roughly equal to the new
    and the levels of conviction are high
    and there is a problem of trust
    and there is a problem of ability,
        then Result -> check out the problem

If the old conviction is roughly equal to the new
    and the levels of conviction are high
    and there is a problem of trust
    and there is no problem of ability,
        then add source and check out the trust problem
```

If the old conviction is roughly equal to the new
and the levels of conviction are high
and there is no problem of trust
and there is a problem of ability,
then add source to list and check out ability problem

If the old conviction is roughly equal to the new
and the levels of conviction are high
and there is no problem of trust
and there is no problem of ability,
then add source to list

Strengthening

When two pieces of information reinforce each other, that case is very similar to the situation of identical pieces of information, and will differ only on the strength of the reinforcement. Whether the system can adopt the new piece of information, depends on the quality of the new piece of information and the quality of the new source. The precondition is that the new source can be trusted at least as much as the other source, then it is dependent on whether it has the relevant expertise if the information is of a kind where expertise is important. As there is a considerable increase in the possibilities between a single source and a multiple source conflict, the source control mechanism may have to take shortcuts for cases which are not very important and decide on the basis of the strength of sources.

If there is a case of reinforcing beliefs
and the new position is less solid than the old,
and investigation reveals that there is some substance
then keep old belief and add source,
otherwise keep old belief and do not add new source

If there is a case of reinforcing beliefs
and the old position is less solid than the new,
and there is a problem of trust
and there is no problem of ability,
then keep old belief and add source

If there is a case of reinforcing beliefs
and the old position is less solid than the new,
and there is no problem of trust
and there is a problem of ability,
then raise to level of ability and add source

If there is a case of reinforcing beliefs
and the old position is less solid than the new,
and there is no problem of trust
and there is no problem of ability,
then raise belief to level indicated and add source

If there is a case of reinforcing beliefs
and the old position is corroborated,
and the new position is more solid than the other source's,
then raise belief and add source

If there is a case of reinforcing beliefs
and the old position is corroborated,
and the new position is equally solid
and there is no problem of trust
then raise belief and add source

If there is a case of reinforcing beliefs
and the old position is corroborated,
and the new position is equally solid
and there is a problem of trust
then keep belief at present level and add source

If there is a case of reinforcing beliefs
and the old position is corroborated,
and the new position is less solid,
then keep belief at present level and maybe add source

Weakening

When new information weakens the old belief of the system, the aim of the system is to retain strong information. Whether that is possible, is dependent on the solidity of the information and whether there is any significant corroboration for one of the positions. If the system's view is well corroborated and its sources stronger than the new source, it will maintain its belief at the present level. In the converse case, the system will have to reduce its belief in line with the new piece of information or go for an intermediary position if the sources are evenly matched.

If there is a case of weakening of beliefs
and the new position is less solid than the old,
and a change of environment is not plausible,
then keep old belief and add source,
otherwise keep old belief and do not add new source

If there is a case of weakening of beliefs
and the old position is less solid than the new,
and a change of environment is not plausible,
and there is a problem of trust
then suspend belief and investigate

If there is a case of weakening of beliefs
and the old position is less solid than the new,
and there is no problem of trust or ability
then reduce belief and add source

If there is a case of weakening of beliefs
and the old position is corroborated,
and a change of environment is plausible,
and the new position is more solid than the other sources',
then reduce belief and add source

If there is a case of weakening of beliefs
and the old position is corroborated,
and the new position is equally solid
and there is no problem of trust
then marginally reduce belief and add source

If there is a case of weakening of beliefs
and the old position is corroborated,
and the new position is equally solid
and there is a problem of trust
then keep belief at present level and add source

If there is a case of weakening of beliefs
and the old position is corroborated,
and the new position is less solid,
then keep belief at present level and maybe add source

Contradiction

Contradictions are of definite concern as the system could have mistakenly put its trust in a piece of information and its supporting source and therefore could have taken action on the basis of a bad piece of information. This situation is particularly serious if the conflict is between two strong pieces of information and the system would have readily put its trust in both of them when considered on their own. Again the degree of corroboration and the relative strength of sources need to be included in the analysis.

If both beliefs are comparable in strength, the system may be inclined to be undecided, considering that there are arguments for both sides while keeping both sources as supports for its position. For stronger piece of information the system will have to reduce its belief substantially, or reject both positions unless there is a substantial difference between them. In that case the system will side with the stronger source or the more corroborated position.

If there is a case of contradiction of beliefs at a low level
    and there is no problem of trust,
        then add support and source to opposing side,
        otherwise keep old belief and do not add new source

If there is a case of contradiction of beliefs
    and the new position is less solid than the old,
    and a change of environment is not plausible,
        then keep old belief and add opposing source

If there is a case of contradiction of beliefs
    and the new position is less solid than the old,
    and a change of environment is plausible,
        then keep old belief and add opposing source ,
        otherwise add low belief and source to opposing side

If there is a case of contradiction of beliefs
    and the old position is less solid than the new,
    and a change of environment is not plausible,
    and there is a problem of trust
        then suspend and investigate

If there is a case of contradiction of beliefs
    and the old position is less solid than the new,
    and there is no problem of trust
        then suspend and investigate

If there is a case of contradiction of beliefs
    and the old position is less solid than the new,
    and a change of environment is plausible,
    and there is no problem of trust or ability
        then reverse belief, add old source and belief at low level on opposing side

If there is a case of contradiction of beliefs
    and the positions are equally solid,
    and there is no problem of trust
        then keep relative weight of beliefs, and add new source on opposing side

If there is a case of contradiction of beliefs
    and the old position is corroborated,
    and a change of environment is plausible,
    and the new position is more solid than the other source's,
        then reverse belief, add new position and source
            and reduce old belief to fit with the new position

If there is a case of contradiction of beliefs
and the old position is corroborated,
and the new position is equally solid
and there is no problem of trust
then marginally reduce belief and add new source
and supports on opposing side at most at average level

If there is a case of contradiction of beliefs
and the old position is corroborated,
and the new position is equally solid
and there is a problem of trust
then keep belief at present level and add source
and low belief on other side

If there is a case of contradiction of beliefs
and the old position is corroborated,
and the new position is less solid,
then keep belief at present level and maybe add source on opposing side

5.9 The Enquiry Process

There is usually no strict borderline where in the analysis process the enquiry starts and the enquiry usually involves doing more of the same, unless it means working out implications like in the example of assignments when the implications of copying have to be unearthed to give meaning to the incidence. Usually, enquiry starts at the point where a solution cannot be found directly from the explicit parameters of the representations, or is unsafe considering the problem and the system will have to find explanations and assurances to convince itself.

If the enquiry process is asked to give a solution to a question, then the enquiry process can look at the information available to find the main problem. If the problem is about the information itself, then the enquiry can generate a strategy to get the information required. Alternatively, if the problem is about the source, then the enquiry can use the source model to investigate the source and find an explanation or generate a strategy to find a solution.

As can be seen from these cases, the process becomes very complex very quickly and may take considerable time. In order to go into an enquiry it must be worthwhile. While there is a certain amount that can be done within the source control mechanism, there may be a requirement to request the overall reasoning system to engage in communication to procure further information. Whether that request is granted depends on the overall reasoning system:

If there is an isolated piece of information
and there is a problem of ability
and no interests are involved,
then one can ask the source or another competent source which knows the source in question and which is good at judging sources

If there is a problem with raising a belief
and there is a problem with trust in the source,
and the system is already talking to the source
then ask for assurances
else ask a different source with ability and which can be trusted

If there is a problem with trust,
then do not ask that source of whether it can be trusted but ask someone who is not involved

If the source declares its beliefs and interests,
and the source has no advantage in telling that,
then believe the existence of these beliefs and interests
else record these beliefs and interests as being possible

If a source is not highly trustworthy and denies beliefs and interests,
then the source may still have them

If there is evidence but no trust
and the source denies the evidence,
then the source is likely to be lying

If confirmation is required
and there is a problem of trust,
then ask independent source who has ability
or ask someone who might know where to get the information

If there is a conflict between two sources
and there is a problem of trust,
and there are common interests,
then look at the relation between the sources

5.10 The Source Model Re-evaluation

The source model re-evaluation maintains the source models for the sources known to the system. It operates either on the basis of accumulated evidence or in response to more significant or unexpected events.

The source model re-evaluation process is interested in assessing the ability and trustworthiness of sources. Once the belief formation process has assessed the information the results of that assessment (the *cases*) are sent to the re-evaluation process. In return, the source model is reassessed in the light of that new evidence and the ability and trustworthiness of the source are adjusted.

The re-evaluation process looks for emerging patterns of behaviour. Obviously, this can not be achieved from isolated pieces of information, and the process therefore operates over longer periods of time on the basis of accumulated evidence. The source re-evaluation model also aims at catching general trends in the source's behaviour and given more unexpected cases it tries to find explanations and to take appropriate action. In the following we shall consider the different aspects of the source model in turn.

5.10.1 Maintenance of the Ability Index

In the process of maintaining the ability index for each source, the source control mechanism takes into account the source's track-record in supplying information, as well as the evidence and information from enquiries and particular cases. In the identification of the level of ability of sources, the system has to try to assess and determine:

- the source's expertise in different areas
- the source's reasoning capabilities

- the source's competence in judging information
- the source's capabilities in handling its own experience

The source's expertise includes subjects which the source has had particular training in, as well as expertise in general, common knowledge and more practical experience. Information about areas of expertise can be deduced from the nature of statements or from various classifications.

The index about the source's ability to reason describes a very general aspect of source behaviour and is usually not limited to particular areas of the source's expertise. The index expresses the source's ability to correctly follow and handle longer chains of reasoning. To correctly determine the level of ability in this area requires the checking of information by an independent reasoning module or can be estimated from the standard and type of education and training.

The source's ability to judge information describes the source's ability to handle and evaluate information it receives from other sources. Ability in this area can be assessed by comparison with the behaviour of the source control model or by identifying types of expertise which involve dealing with people.

Finally, the index about the source's experience records the source's ability to judge and handle correctly its own, personal experiences. This ability does not depend on training but can only be inferred from the way information based on it is being handled. Sources should, however, usually be quite capable of handling their own experiences.

```
If there is a new case,
     and there is no connection to other information
     and there is no specific evidence,
         then add the type of case to the accumulated records

If there is a regular pattern in the records,
     then check whether that pattern can be explained in the source model

If there is a pattern which cannot be explained by the source model,
     and the index is built on long-standing evidence,
         then keep accumulating evidence and investigate
     elseif the index is not built on long-standing evidence,
         then weigh index against evidence and adjust index
```

**If the source makes strong technical claims
and the index does not record any expertise,
then investigate whether there could be a classification to explain it
else enquire with the source or sources who would know**

**If there is an opportunity to talk to a source,
then ask about its schooling, training and profession
and analyse the types of abilities required for that**

**If there has been an enquiry into the source,
and there is evidence of deficiencies or abilities
and the abilities are not reflected in the index,
and the evidence is stronger than the index,
then adjust the index
elseif the evidence is not stronger than the index,
then add evidence to records**

If a source is reporting about another source
and the reporting source is good at judging sources
and there is no problem of trust,
and the report is stronger than the index,
then adjust the index
elseif the report is not stronger than the index,
then add evidence to records

If a source is reporting about another source
and the reporting source is good at judging sources
and there is a problem of trust,
then investigate further or discard evidence

If there is evidence for a classification,
and classification can be explained by past performance,
then apply classification
elseif there is no past performance to judge against,
then apply classification
else investigate and record evidence

5.10.2 Maintenance of the Belief Index

The belief index records a source's particular strong beliefs which are, amongst other things, the result of education, profession and social circumstances. Strong beliefs can influence the objectivity of the source in these areas. The detection of such beliefs can be very difficult as they influence the source's behaviour indirectly. Unless the system gets direct information when the respective subjects are being openly discussed, the system can otherwise only detect them through trend analyses, which can be triggered through particularly strong expressions of the source about subjects which are contentious:

If there is a regular pattern in the records,
and that pattern is about opinions
and the source keeps reiterating its opinion
and there is direct evidence,
then add belief to list
elseif there is no direct evidence,
then record possibility of a strong belief

If there is evidence about education and social situation,
and that has strong beliefs associated with it,
then add beliefs to source model

If there is an association with a particular class,
and that class has strong beliefs associated with it,
then add beliefs to source model

If there is a belief and the source does not behave in accordance with it,
and there is no previous evidence for that belief,
then remove belief
else if there is previous evidence,
then investigate and add to records

If there is an opportunity to talk to a source,
then ask about its schooling and profession
and analyse the types of classes which may apply

5.10.3 Maintenance of the Interest Index

The source's behaviour may also be influenced by strong interests, which are based on strong motivations. Compared with beliefs, interests of the source can be slightly easier to determine, as they are often associated with particular classes. Among those, classes which have strong goals, like financial gains, reputation and power are particularly apt to give rise to vested interests and are significant as they tend to be concerned with important information.

If there is a regular pattern in the records,
and that pattern is about making strong claims while denying responsibility
and there is a connecting factor which implies some form of gain,
and there is direct evidence,
then add interest to list
elseif there is no direct evidence,
then investigate or record possibility of a strong interest

If there is an association with a particular class,
  and that class has strong interests associated with it,
    then add interests to the source model

If there is an interest and the source does not behave in accordance with it,
  and the interest comes from a classification,
    then remove interest
  else if there is previous evidence,
    then investigate the classification

If there are strong relationships involved,
  and there is a pattern of similar behaviour,
    then record interest for when these relationships
      and the subjects are involved

5.10.4 Maintenance of the Trustworthiness Indices

The trustworthiness indices are composed of a trustworthiness index describing the system's view of the source's honesty and integrity and hence inclination to mislead and misinform, together with a helpfulness index describing the system's view of the limit to which the source is inclined to be truthful. These indices tend to be very context sensitive, dependent on particular relations, or whether particular beliefs and interests are involved. Consequently, there may be a number of indices to represent specific situations.

The helpfulness index is unlikely to change much, and is used to gain information about possible advantages and potential vested interests of the source.

The trustworthiness index is more likely to be modified by the source re-evaluation process and usually starts on a high level, but may collapse rapidly if there is evidence to suggest that the source is cheating the system. The recovery of a collapsed trust index is very difficult, thus representing human experience that trust, once lost, is difficult to regain:

If there is no record of a breach of trust in the records,
and the number of recorded instances are significant,
  then general trustworthiness -> high
elseif there are minor problems with trust,
  then general trustworthiness -> average
else if there is a sudden problem with trust,
and the source is important,
  then investigate or record unresolved problem of trust

If there is a significant record of helpfulness,
and there is no obvious fracture,
then general helpfulness -> high
elseif there is no consistent record,
then general helpfulness -> average
else if there is a consistent record of being uncooperative,
then general helpfulness -> low

If there is a pattern of helpfulness or trustworthiness
associated with a particular subject,
and there is a strong interests associated with the subject,
and there is a relation associated with that interest,
then create that relation for the source model
and calculate trustworthiness and helpfulness from records

If there are strong relationships associated with the source,
and they are part of a class which has strong interests associated with it,
then add new relation and abstract typical behaviour from a semantic definition

5.11 Two Complete Examples

In the following we want to examine two examples in more detail. We have presented the principles of the source control mechanism in the previous chapter and given a design for the construction of the mechanism in this chapter. In this, the final step in our methodology, we demonstrate that the design of the source control mechanism can actually deal with the examples we gave in Chapter 2. Consequently we come back to our original claim that the mechanism is able to deal with the examples we stated at first. Given constraints of space, we give a full analysis of only two of them and expect that it will be clear from the analysis that the model is capable of dealing with all of them. Since they display a considerable complexity it should be evident that the mechanism will be able to deal with less complicated cases as well and that the strategies involved can be transferred.

As we want to exploit the particularities of the examples, and as we also want to give space to demonstrate different features of the source control mechanism, there is a difference of emphasis between the examples. In the first example we concentrate more on the technical side of applying the mechanism to an event-based representation. We also show the mechanism in the more passive role of an adviser who is trying to help one of its sources. The example is very versatile, as it shows well the different kinds of problems with information and as the problem is about forming beliefs on the basis of available information, it shows well the general mode of operation of the source control mechanism. In this example the intricacies of the information representation and task allocation are described, these are not dealt with in detail in the second example.

The second example is partially about forming beliefs, but it also deals with the system answering questions, which are not strictly a part of the source control mechanism but show how well the mechanism fits with other components in a general reasoning system and how it can help in the strategic procurement of information not available from the database. This has the advantage of showing the more active side of belief formation, which is less evident in the first example.

5.11.1 Buying a Motorbike

This example about buying a motorbike is a classic situation of belief formation and we will show the way in which the system will treat the information and form its beliefs. As the system will have to reconsider the position from time to time, there are three different strategies which can be pursued and the example shows how this is done. We will show how the SCM deals with the events it receives from the NLP and to which it has to attach beliefs. We will also try to show how the SCM and a NLP could co-operate in a closely integrated implementation.

Suppose, Albert tells the system that he needs to buy a motorbike to go to work with. He says that he has had a look at a motorbike at a local second-hand motorbike dealer who recommended one of his motorbikes to him as being sound and cheap. When he went to the dealer his friend Paul came along as well and said that the motorbike the dealer pointed out was looking good. Albert is however still unsure and is asking the system for help.

Supposing the system is presented with the following information from Albert:

Albert: I need a motorbike to go to work.

The natural language processor is the first to deal with new information from human sources. It transforms the natural language input into events and determines the type of information in order to know what to do with it. For example, questions are requests for information which do not require the assignment of belief, as are hypothetical events, which may be used in an argument but are not real and are not subject to belief. In our case, the NLP will identify the information from Albert as a set of connected real events (some of which are hypothetical) and identify the requirement of a belief on Albert's claimed need of a motorbike and send it to the SCM, labelled with certainty=high (by default). This can be represented in the following form:

Event1:
subject: Albert (named individual)
action: need
object: E3
time: present
source: Albert
status: real (default)
certainty: high(default)
belief: ?

Event2:
subject: Albert
action: go
destination: working place
instantiation: motorbike1
time: > Event1
source: Albert
status: hypothetical
cause: Event4

Event3:
subject: Albert
action: control
object: motorbike1
time: present
source: Albert
status: hypothetical
cause: Event2

Event4:
subject: Albert
action: work
time: > Event2
source: Albert
status: hypothetical

This is the point where the work of the SCM starts and the SCM now has to assign a belief to these events. In order to do this, it will first look for the sources involved and find that in this case Albert is the source of these events. It will therefore require the source model for Albert and consult the source model database to see whether such a source model exists. In our case it will find the following model:

```
Albert:
 Ability:
  expertise: computer science, philosophy, languages
  reasoning: high
  judging sources: average
  experience: high(default)
 Beliefs: none(default)
 Interests: none(default)]

 Albert -> System:
  Helpfulness: high
  Trustworthiness: high
 System -> Albert:
  Helpfulness: high
  Trustworthiness: high
```

The source control mechanism then has to find out whether the information is sufficiently interesting for the system before embarking on a more detailed analysis. If we remind ourselves from the design, there are three ways of doing this importance analysis:

- either by checking whether the information is important to the system
- or whether the information is well formed and consistent with other information
- or by checking whether the source or sources involved are important to the system

Proceeding by the first criterion, the SCM checks its motivation towards Albert by looking in the source model of Albert, to find that trustworthiness and helpfulness are high. The SCM will therefore take an interest in the information as it wants to be helpful to Albert and tries to assign a belief to the statement. Looking at Albert's relation to the SCM it will also find that the source can be trusted.

In the process of checking the type of information, it finds out that the information is about a need, which can be classed as being a subjective personal perception. As the source appears to be capable to correctly assessing its own experiences a medium belief is assigned by the system, as it is impossible to share other people's experiences objectively.

Then suppose that the NLP gets a second sentence from Albert like this:

Albert: I am considering buying motorbike1.

This sentence is a bit more complicated than it appears at first. To the NLP, 'considering' means that the source has not yet decided what to do, and the meaning is more like the question of whether the source should buy the bike, while the source is undecided. This implies that the information is translated into a hypothetical event where Albert buys the motorbike, and that is used in another event where the relation 'must' indicates motivation towards an action. The source is undecided, which is indicated by the certainty being medium/medium, thus indicating that the arguments for and against are evenly matched - the classical situation of a dilemma. The NLP will therefore produce the following set of events:

```
Event5:
  subject: Albert
  action: buy
  object: motorbike1
  source: Albert
  status: hypothetical
```

```
Event6:
  subject: Albert
  action: must
  object: E5
  source: Albert
  status: real
  certainty: medium/medium
  belief: ?
```

The SCM receives these events and as the information is from Albert, and as the system already determined to take an interest in information from Albert, it will also process this information. The information is about a matter of personal perception and the indication of conviction is not high and rather indecisive. The SCM will therefore have to take the information on trust and will agree with the source without being too strongly committed.

Then the NLP has to analyse the next piece of information from Albert about having seen the bike in a particular shop and about the dealer and his advice. This step is slightly more complicated as it requires that a connection is made by the NLP between the hypothetical bike which will get Albert to work and the bike the dealer recommends Albert to buy:

Albert: Dealer1 sells motorbike1.

Suppose the natural language processor produces the following set of events:

Event7:
subject: Dealer1
action: own
object: shop1
source: Albert
status: real
certainty: high
belief: high

Event8:
subject: motorbike1
action: is_in
object: shop1
source: Albert
status: real
certainty: high
belief: high

Event9:
subject: Dealer1
action: sell
object: motorbike1
source: system (endorsement)
certainty: high
belief: high

Event10: {universal}
subject: Dealer (bounded universal)
object: Customer (bounded universal)
action: sell
status: hypothetical

Event11: {instantiation}
subject: Dealer1
object: motorbike1
location: shop1
action: there_is
status: hypothetical

The NLP will need to connect these sentences, with respect to Albert and motorbike1, with a discourse analysis technique to make that inference. Since all the information is from Albert and the system believes Albert, the information will therefore generally be believed unless the source is reporting what another source has said. If that case arises then that information can probably be relied on having been conveyed properly, but has to be judged on its own merits.

It should now be clear, how the event structure works, and from now on we shall use a simplification, to concentrate on the problem solving aspect of the example.

Supposing the source provides the following information package to the system:

Albert: Dealer1 says that motorbike1 is sound and cheap

As this information is relayed by Albert and the system can rely on Albert relaying the information correctly the statement will be transformed into:

Dealer1: Motorbike1 is sound and cheap.

The SCM now has to assign a belief to this statement from him. Looking in the source model database it does not find a source model for the dealer and it therefore has to generate a default model based on some classification. As it knows about the class of salesmen in general and second hand dealers in particular together with the fact that the dealer in question is a motorbike dealer as can be inferred from the fact that he is selling a motorbike the SCM will generate the following model appropriate for the class:

```
Dealer1:
 Ability:
  expertise: motor mechanics(default)
  reasoning: average(default)
  judging information: average(default)
  experience: high(default)
 Beliefs: none(default)
 Interests: selling(default)

 dealer -> client:
  Helpfulness: high(default)
  Trustworthiness: low(default)
```

From the definition of selling and buying and the fact that Albert and Dealer1 are talking about the same bike, the system will also connect Albert to the client role in the dealer's source model as well as connecting the motorbikes mentioned.

As this information is from a different source the SCM has to reconsider whether it wants to take an interest in the information. Again there are different strategies to do this analysis:

- whether the source is important
- whether there is a motivation to examine
- whether the information is interesting in itself

Taking the strategy of looking at the source, we see that Albert is important to the system (system -> Albert Helpfulness: high) and that the system is willing to give good information to Albert. Now there is a need to find out whether the information is important to Albert. Albert is associated with the class of working people and it is important that Albert works. The motivation is connected to Albert in that Albert considers buying it. This is one approach which needs the general meaning of buying.

The faster way is to see whether motorbike1 is involved in any of Albert's motivation and find that Albert needs motorbike1 to go to work. Now one needs to find out how important working is and one needs to go to the script to find that getting to work is important and that the motorbike is instrumental to that purpose. If working is important, getting to work is important and a motorbike is instrumental for Albert. Thus the importance of the information is established.

Looking at the information from dealer1:

```
Event12:
 subject: motorbike1
 action: is sound and cheap
 time: present
 source: Albert
 status: real
 certainty: high(default)
 belief: ?
```

The SCM will find that the trustworthiness of dealer1 is low and so despite the dealer1's expertise and experience in motor-mechanics the SCM will remain very doubtful about the information.

One can also see problems purely from the fact that Albert is in the client role and that the trust he should have in the motorbike dealer is low. The problem with this approach is that it relies heavily on previous information, whereas the other approaches rely on present information. At the same time, the importance is difficult to work out and one should only go down this line if it is explicit and otherwise pursue another line.

In either case, the belief in the information from the salesman is low. Unless the responsibility assumed is high while the trust is low, the belief will be low. The usual way of gauging whether responsibility is assumed relies on the trust being high, as otherwise more concrete guarantees are necessary. In the script for exchange or selling in the NLP, there should be a slot to say that a suitable return for services is money. Thus if the warranty outweighs the possible losses, then we can rely on the information no matter what the dealer says. Alternatively, if the warranty does not quite reach that extent, then it at least give us an indication as to his intentions. In our case, in the absence of any guarantees, the SCM will have to assign a low belief to the dealer's statement.

Now supposing that Albert says:

Albert: Friend1 thinks that motorbike1 is looking good.

Again this would be linked to the previous events and the system would send a request to the SCM to form a belief about the following information:

Event13:
subject: motorbike1
action: is looking good
time: present
source: Friend1
status: real
certainty: medium
belief: ?

The system can again decide to check either the source, work out the consistency or check the importance. First, the system has to generate a source model for the friend from its knowledge about friends in general and the relation to Albert:

Friend1:
Ability:
expertise: none(default)
reasoning: average(default)
judging information: average(default)
experience: high(default)
Beliefs: none(default)
Interests: none(default)

Friend -> Friends:
Helpfulness: high(default)
Trustworthiness: high(default)

Assuming the system proceeds by source considerations, to work out the belief for the piece of information it will check the trustworthiness and helpfulness first. Both indices are high and the information is believable, but perhaps not very useful as the conviction attached by the friend is based on a subjective personal impression.

The link between Albert and Friend1 holds by virtue of the friendship relation (Friend -> Friends) and Albert's claim to be friends with Friend1. As the friend is important to Albert, this makes the information interesting to the system, which will try to attach a belief to the information. Again the system will have to decide which of the three strategies to pursue. Although, going straight to the source may somewhat pre-judge the information, we shall try this avenue for the moment. If the trust is low, there is no reason to believe the information, but instead we find a high level and proceed to examine the rest, keeping in mind that some values are defaulted whereas others are not. We also find that helpfulness is high. The helpfulness index has to be treated with care as it is very context sensitive and may be connected to fundamental beliefs and interests, none of which we are aware of at the moment.

The system may at this point ask Albert whether the information from the friend was volunteered or asked for, considering the expected helpfulness of the friend. The discrepancy between the system's expectations about the source's abilities and helpfulness and the actual behaviour can be significant, and the system can use this information to draw easy conclusions about responsibility and advantages. In the present case we would expect the friend to be helpful, but if the source was reluctant, we take it as an indication that the source is reluctant to accept responsibility. Typically, human agents do not state these indices explicitly but expect the other agent to pick up the signs themselves. Thus, if the source was urged to give information it did not want to give, the source's response will tend to be quite general and unspecific.

In our case the actual behaviour of the source and the expectations of the model correspond, and we can go along with the source model without needing to make further inferences; the source is generally very helpful and the information was almost volunteered. Now we have to check whether the source is able to give the information and to check its ability. First of all, we look at the expertise, because that is likely to be the hardest piece of information. If someone is an expert in a particular area, we are likely to know, or there may be a particular expertise as part of the class the source is associated with. Finally, we could ask Albert whether the source has the required expertise. In any case, if there is no specific mention of expertise, the model will carry a default in that slot.

Supposing we find out that the source has some degree of expertise with motorbikes then we can stop the inquiry and accept the information unless there is an explicit lack of responsibility. If the source did not have such expertise and the arguments employed were of a technical nature, then we would have to find out where it got the information from. One possibility is that it was the product of reasoning, in which case we have to ask from what kind of evidence it came from. As it cannot be expertise in this case it must be from somewhere else, or it is just an unsubstantiated belief based on faulty reasoning. It is important to note that once we have eliminated the possibility of expertise, all other possibilities are inferior, as intuitive belief or personal interests are insufficient to judge such a highly technical matter.

To embark on an inquiry is problematic as we are not in contact with the source and we cannot enquire directly. We could ask Albert whether he knew the origin of the information, but that is rather unlikely to be successful, because Albert in all likelihood would have told us where the information originally came from. Besides, the process would get too complicated if we asked Albert to go and find out, especially since the system was only asked for an opinion and not for an action. Such a response would therefore seem odd.

It appears that the only way is to analyse the information to see what kind of information it is, by using the semantics or very shallow pragmatics to see how the information needed may be gathered. The argument is that it is good looking, and that it is based on an aesthetic judgement close to a sensation. General heuristics will tell us that sensations are not likely to be reported from other sources unless one agrees with them. Since sensations do not come from reasoning or expertise, but obtained directly by experience, and it is expertise which is needed here, the information is not very credible.

The primary purpose of the source model is to store information about the source's strengths and weaknesses and to use this information to check whether the source has the credentials required for the information it supplies. The source model can, however, also be used to find out where the information could have come from.

The piece of information could also be the product of a long chain of reasoning, and those tend to come from the source itself as people generally find it difficult to remember other people's lines of arguments exactly, apart from the conclusions. It is therefore unlikely that someone reports someone else's chain of reasoning, or if one does, one will be likely to explicitly quote that person. The NLP will detect chains of reasoning and those can be checked by a reasoning module. As long as the SCM knows the abilities of the semantic parser it can ask for confirmation of its suspicions.

As it stands, the statement by the friend can be determined as being based on sensation and the heuristics are able to tell us that that is a matter of experience. And at this point the analysis is finished - the information can be believed as it is based on experience and the friend is able to correctly handle experience except that what can be believed is that he thinks the motorbike is good looking and that that is personal.

The SCM can stop at this point as this kind of information is not usually used in general reasoning or for actions. It will believe it but label it 'not for use'. In our case we could have saved a lot of work, because if we knew that we were dealing with sensations, we could have stopped there and defaulted to a normal belief. If the information was needed for action later on, one could have restarted the enquiry process.

Supposing, that the friend said the motorbike was a bargain and if the SCM looks up what a bargain is it would have found the definition that a bargain is something which is sold below its price. In order to be able to make that statement you must know what it was worth in the first place, which requires a certain amount of expertise. As with sensations, knowledge should not be taken from someone else without acknowledging the source. In our case the friend does not appear to have expertise, which puts the credibility of the statement in question. The source might be right, but unless we find better evidence we cannot put much trust in it. The belief assigned in that case is 'doubtful'. By default we assume that the source does not have the expertise, but we do not know for sure. The information is therefore suspended and labelled 'use only if corroborated'. The conclusion to make belief subject to corroboration may be a key to start an enquiry into whether the friend does have the expertise or not. This could be done by asking Albert to go and find out.

Alternatively, if the information from the friend was a strong piece of technical jargon we might assume that it does have the required expertise, because if the source does not, then the source would be a real fool (and we probably would know that). So we don't have to go into a deep semantic analysis to find out about that. Consequently, if jargon has been used and our default of 'no expertise' might have to change, whereas if we KNEW that he didn't have expertise then we would not believe the information and have to consider whether the source is a fool. Either the trust is high and the source is acting in good faith but is a fool or the trust is low and the source tries to manipulate and lie to us and is a danger. If it is a matter of jargon against defaults then the jargon is likely to win as it is a strong piece of information. This may also be a pointer to request a source model re-evaluation. We could also suspend the model and believe the info. If expertise is known to be lacking then one should not use the information and this therefore finishes the evaluation of the friend.

Now Albert asks the system whether he should buy the motorbike:

Albert: Should I buy motorbike1?

This will be represented as an event of type 'question' asking the system what to do. Questions are special in that they do not require to be believed and they are not strictly the domain of the source control mechanism. At the same time the SCM will want to be helpful and therefore treat them as a request for information about its beliefs.

The more cautionary the question is put, the less likely the source is to expect of the system to take responsibility for the recommendation. The system will want to act responsibly and not say anything unless it is reasonably sure. Although this does not involve the system staking money on it, the price is the system's reputation and good relations. Therefore it has a strong motivation to be careful.

Questions are not believed in the same way as information is believed and therefore they do not require that a belief is assigned by the SCM. At the same time, when the system has to answer them, the NLP has to consider the implications. In this case there is a piece of pragmatics involved, namely when it is that one buys a motorbike. This requires an analysis of what it means to buy things. Generally, one buys things which are in good working condition and that it is something one wants for one reason or another. Therefore one needs to find out what these reasons are. Looking back, we can see that Albert needs the motorbike to go to work, and that is a reason for buying. So now the question becomes something like 'if I buy the motorbike, will it be good enough to go to work?'. To work out this piece of reasoning is not part of the source control mechanism. It is the task of a semantic reasoner, although the source control mechanism will use the result in determining how to answer the question. In this case, good looks are not important, but reliability and cheapness are, and if it is a bargain, so much the better. We need to decide therefore whether the bike satisfies the needs of reliability and cheapness.

The system now has to try to prove any of the requirements of this list and therefore needs to examine what we know about the motorbike. Looking at the information we have, we cannot find any strong evidence to suggest to buy the motorbike, although there is some uncorroborated evidence and some doubtful evidence. It would be acceptable for a trivial opinion but the matter is important. As the system considers its relationship to Albert important, and the decision is important to Albert, the system does not want to give out information which is not good enough. On the other hand it does not have any information to suggest that the bike is not good. Therefore it can only say that on balance there is no good reason to buy the bike.

There is a dilemma between being helpful and being trustworthy, and the system has to find a balance between them. What we can do at this point, is to output the information but with a low responsibility and to give the reasons for our opinion:

System: Low positive belief for buying motorbike1,
but no responsibility taken.

If the system is pressurised by Albert it may suggest to try to get more information, or the system can launch into something more complicated like an enquiry to reinforce the information. It could ask whether the friend is an expert in vehicles, because if this is so than we can remove the doubt we had about the friend's information. If Albert says that he is an expert, then we can change the belief and the model to reflect that and recommend to abide by the friends' recommendation in the case where the friend claims that the motorbike is a bargain. If that is denied, we can still make changes to the source model, but the belief about the information stays low.

If no progress is made, the system might look for another source of corroboration. Suppose the system looks through its memory in order to find an expert on motor mechanics, and finds that it knows about the RAC and that the RAC operates a valuation service, available for a fee. Thus the system could recommend to Albert to get a valuation on the motorbike and to come back with an answer. Now suppose that Albert comes back with the following result:

Albert: The RAC valuer says that motorbike1 is sound and cheap.

This will lead to the following information being passed to the SCM:

```
Event3:
 subject: motorbike1
 action: is sound and cheap
 source: RAC valuer
 status: real
 certainty: high
 belief: ?
```

The SCM will then retrieve its generic source model for independent valuers, which looks something like this:

```
RAC Valuer:
 Ability:
  expertise: motor mechanics
  reasoning: average(default)
  judging information: average(default)
  experience: high(default)
 Interests: fees, reputation
 Beliefs: none(default)

 Valuer -> clients:
  Helpfulness: high(default)
  Trustworthiness: high(default)
```

The RAC Valuer is an ideal source, as far as the system is concerned, as it is trustworthy, helpful and has the necessary expertise to talk about motorbikes, as well as being completely independent. The system can therefore completely believe the information, although this is not the end as the system still needs to give a recommendation to Albert. The system then has to analyse the information to check whether the relevant connection has been made between the motorbike and motorbike1, as well as considering whether Albert does fit into the client position with the valuer. Both can be established and the connection is made by the NLE. Now the system has to analyse what the information says, with respect to the problem whether Albert should buy motorbike1 or not. Remembering from the analysis of buying, we see that the requirements are proper functioning and value for money. The general heuristics can then determine whether the stated properties fulfil the requirements.

In the case where the system does not know an expert, and neither does Albert, one strategy one could try is to find someone with personal experience. If that fails, then the system cannot do anything except getting Albert to gain the expertise himself through reading and courses. In the converse case, if Albert mentions the RAC then the system in turn might want to know about it for its own purposes and to find out what the RAC is and what they do etc. and ask Albert whether he can trust them. Then you find out that you have to pay them and them one needs to find out how much to see whether it is worth it. The system, knowing that it can trust Albert, can then engage in further dialogue to find out that the RAC are independent and trustworthy and that they operate a car valuation service for members and non-members at a certain fee. This information is then converted and a new source model is formed. At that point the system will pick up the processs as mentioned above, where the system knew already about the RAC. The system can then tell Albert to go to them and do what they say and that would conclude the incident.

The SCM can ask help of the NLP, although it has its own heuristics to make out fundamental distinctions. The beliefs are attached due to the checking of the model and the heuristics and endorsements are used in order to give the reasons why one does, or does not believe, and what to do in the respective situation. When questions have to be dealt with, that has usually nothing to do with the source control mechanism, except that the system might request information to see whether there are motivations to act. Once the source launches an appeal, it needs to be checked by the system against its motivations and therefore needs to ask the SCM what can be done. It is not the purpose of the SCM to go through the process of finding help, as this is a chain of reasoning. However, as new information arrives in it will be required of the SCM to analyse and assign beliefs to them.

5.11.2 Copied Assignments

In the second example, we deal with a situation which is very different in many respects from the example of the motorbike, thus demonstrating aspects which were not involved in that example. Since we have already in the previous example studied the way in which the information and source representations work, we shall concentrate in this example on the explorative aspect of the source control mechanism which is used to explore problems in the process of finding an explanation for them. This example also has the advantage of showing the kind of integration between the source control mechanism, natural language and reasoning modules which gets the best out of each module due to the fact that they cooperate in solving problems none of them could easily deal with on their own.

Suppose a teacher is marking assignments and finds that there are two assignments which are very similar, and which raise his suspicion that one of them was copied from the other. As he wants to get some confirmation for his suspicion he asks a colleague to look at them to see whether he agrees that they seem to be copied. Supposing also that the colleague agrees, the teacher therefore has to try to find out whether there is some substance to their suspicion.

Suppose the teacher, in essence, gives these two statements to the system:

Teacher1: Assignment1 from student1 is very similar to assignment2 from student2.

Teacher2: Assignment1 (a1) from student1 (S1) is very similar to assignment2 (a2) from student2 (S2).

In order to evaluate these statements the source control mechanism has to build two default models for teachers (of biology for the sake of the example):

```
Teacher1:
 Ability:
  expertise: biology=top(class-default)
  reasoning: high(class-default)
  judging information: average(default)
  experience: high(default)
 Interests: discipline, reputation
 Beliefs: justice(class-default)

 Teacher1 -> system:
  Helpfulness: high(default)
  Trustworthiness: high(default)


Teacher2:
 Ability:
  expertise: biology=top(class-default)
  reasoning: high(class-default)
  judging information: average(default)
  experience: high(default)
 Interests: discipline, reputation
 Beliefs: justice(class-default)

 Teacher2 -> colleagues:
  Helpfulness: high(default)
  Trustworthiness: high(default)
```

Suppose also that teacher1 subsequently asks the system whether they have copied from each other. As the source control mechanism is only concerned with generating beliefs and not with questions, the natural language environment will pick up the question and transform it into a special event which has no source attached and no status as it does not fall into categories like being hypothetical or real. The problem is therefore not to see whether the statement should be believed, but whether it can be proved:

Are assignment1 and assignment2 copied?

Now the source control mechanism, or a database manager, is faced with the problem of finding relevant information from the database. Supposing that there is no information in the database about that, one way to make progress is to do some semantic analysis on the action, which is copying. This will need to come from the natural language environment. Seeing that copying is an action which involves an original and the action of a human to produce an object which is identical or near identical to the original and which requires as a pre-requisite that one has access to the original. From a semantic analyser we get the following definition:

```
person1 has access to object1
      -> person1 copies from object1 to object2
            -> object1 = object2
```

Now some abductive reasoning needs to be applied, which will want to consider all the possibilities and to try to eliminate them, one by one, to get to a likely explanation. This is outside the scope of the source control mechanism and proceeds backwards from the fact that object1 = object2, and that there are a number of different possibilities to produce the result of near identical objects, which translates to assignments in our case:

```
S2 had access to a1 -> S2 copied from a1 to a2 -> a1 same as a2
S1 had access to a2 -> S1 copied from a2 to a1 -> a1 same as a2
S2 is friend of S1 -> S2 and S1 collaborated on a1 and a2 -> a1 same as a2
S2 wrote a2 and S1 wrote a1 -> coincidence -> a1 same as a2
```

The problem now is to try to find out which of these can be eliminated. In order to find out the system wants to talk to those sources who know, and that is S1 and S2. Also since the system has an interest in making good use of the chances to talk to sources it wants to be well prepared and as the system is still talking to the teacher1 who also knows S1 and S2 the system will want to find out about the sources following its interest to build source models quickly in order to be able to assess information from them. From the fact that they are students the source control mechanism will have built default source models for students and it can now ask teacher1 about the students abilities, trustworthiness and helpfulness. Suppose it can therefore build the following models:

Student1:
Ability:
expertise: biology=high
reasoning: average
judging information: average
experience: high(default)
Interests: getting good marks, showing expertise in biology
Beliefs: none(default)

Student1 -> Teacher1:
Helpfulness: high(default)
Trustworthiness: average(default)

Student2:
Ability:
expertise: biology=average
reasoning: average
judging information: average
experience: high(default)
Interests: getting good marks, showing expertise in biology
Beliefs: none(default)

Student2 -> Teacher1:
Helpfulness: high(default)
Trustworthiness: average(default)

Now the two teachers have provided some evidence for the fact that the assignments are very similar, which makes it possible that they have been copied. This can be seen from the definition of copying, where copying is one of the possible explanations for similarity of the results, which is assignments in our case. At the same time, the fact that the teacher asks for help suggests that he is not sufficiently certain, because otherwise he would not be likely to ask and since he is asking, we would not ask him the same question as it is pointless to ask the person who is asking you.

The source control mechanism will therefore want to ask someone else, such as the students who after all should know. In its preparation for meeting them, the source control mechanism will also want to find out about any relationships and whether there are advantages involved and who benefits the most. The teacher is therefore likely to say that they appear to be friends and that there are advantages, because copying does not attract good marks and that the weaker student would benefit most.

Consequently, given that the weaker student can be trusted less, because he would profit most and that the stronger student can be trusted more, the system would be inclined to ask the stronger student first, which also fits with the general principle of asking the strongest source first, as it is likely to know best. The source control mechanism will also have to take into account any relations between the two students, as the teacher1 suggested that they are friends.

There are a number of possibilities to consider which can be represented in the following way:

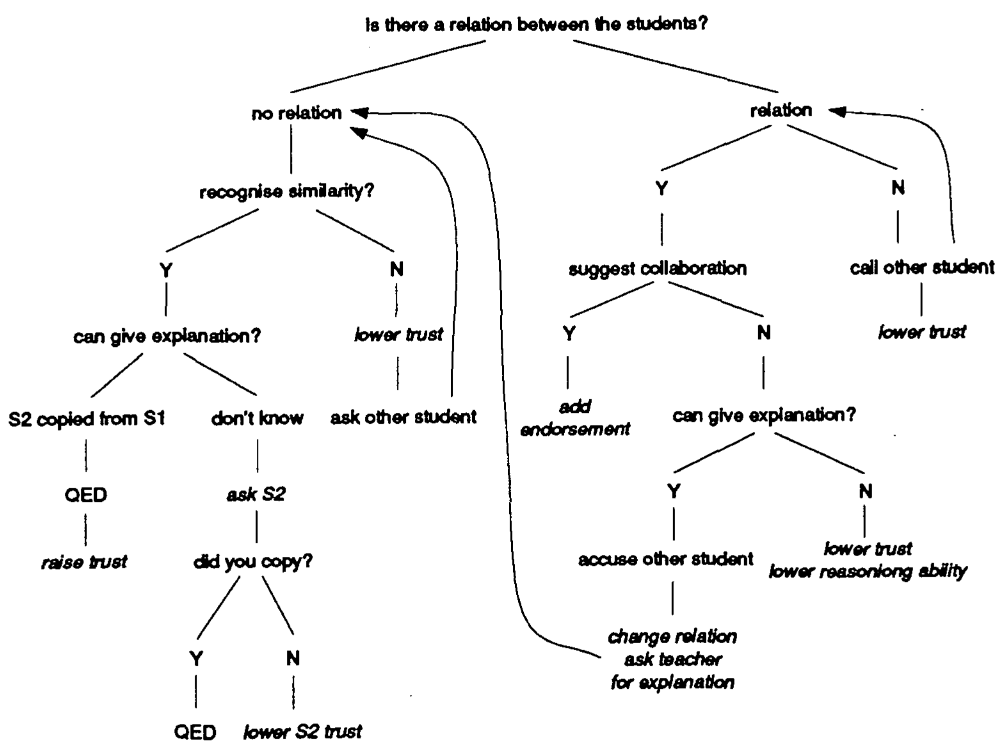


Figure 5: Reasoning of the Assignment Problem

Subsequently, talking to student1 the source control mechanism may want to start with the only evidence or endorsement it has for the case, namely the teacher's agreement that the two assignments are very similar:

System: Do you recognize that both assignments are very similar?

There are now two possibilities. Either the student will deny that there is any similarity, or he will accept it. If we consider for a moment that the student denies the similarity, then the system will lower its trust of the student straightaway, because he denies the evidence and proceed to call in the other student. If, on the other hand, the student accepts that the assignments are remarkably similar, then that information serves as an enforcement of the teachers opinions:

Student1: Yes, the assignments are similar.

The source control mechanism will in this situation build a record to collect the endorsements for and against an opinion of whether the assignments are copied. It should be stressed that the endorsements are not arguments for or against a proposition in the sense of Cohen, but the system uses source control endorsements which are reasons for believing or disbelieving.

So far we have two source endorsements, namely the teacher's opinion and the student's acceptance of similarity. Considering the definition of copying and the system's attempt to eliminate various alternatives and since there is an alleged relation of friendship between the students and that matches the requirements for collaboration, the system will therefore ask whether there has been any such collaboration:

System: Did you collaborate with Student2?

If the Student1 agrees that he did collaborate with Student2, then that reinforces the list of reasons for the belief that the assignments with a strong reinforcement and settles the matter because the similarity has been explained in the definition of copying by the satisfaction of the necessary preconditions for copying and the matter is settled in the view of the system, which can conclude that the assignments have been copied by a process of collaboration.

If, on the other hand, the student denies the collaboration then the system may exclude that possibility for the moment, but is still left with three further possible explanations for the similarity of the assignments. At the same time this is the only really likely explanation for getting similar assignments in a situation where the two agents are friends, which is an argument against the possibility that there was no collaboration and consequently there is a conflict. It will therefore ask the student for an explanation:

System: Can you explain why the assignments are so similar?

If the student responds that he cannot explain how the assignments could be so similar and there be no collaboration, then the system will lower the trustworthiness of the source because that response is not plausible and it will also be inclined to lower the reasoning index because the student is not able to keep together a coherent argument. The system then may suggest its opinion of the source and call in the other student.

Alternatively, assuming that the student proceeds to accuse the other student of having copied from him, this will give an explanation of why the assignments are so similar and add an endorsement for that side, but at the same time that throws serious doubt on the major assumption that there is a relation between the students, and the system will be inclined to ask the teacher why he thinks that there was a relation, considering that the student accuses Student2 of having copied from his assignment.

Given this situation the system will have to reconsider Student1's responses under the assumption that there is no strong relation between them. Student1 agreed that the assignments are similar and that Student2 copied from him, which is a possible situation under the definition of copying where copying does not require a special relationship but only access to the assignment to copy from. Student1's statement that Student2 copied from him therefore adds another endorsement for the conclusion that copying actually took place. The system can therefore proceed to call in Student2 to ask him:

System: Did you copy your assignment from Student1?

If Student2 denies the charge then there may be a stalemate, although the system will not be in a worse situation that a competent teacher would have been. In that case the system will reduce its trust in the source considering that that behaviour could be expected from Student2 as he is also the weaker student and has more to lose if he was found to have lied. If on the other hand the student admits that he copied, then the trust is restored as the student admits his action. At this point the case is solved as all four sources, the teachers and the students agree in their statements and the system can conclude that the assignments have been copied.

Although the procuring of information in response to a question is not strictly the domain of the source control mechanism but is done in collaboration between different units of a general reasoning system, the principle of finding an explanation for an important case is the same as used by the source control mechanism. This example therefore shows how an enquiry can be carried out successfully and thereby also demonstrates an active belief formation strategy typical of the source control mechanism as opposed to more traditional belief revision which is more passive and restricted to the information available at the time.

5.12 Conclusion

The system's overriding interest, apart from performing an adequate analysis of uncertainty in the information, is to gain strong reliable information with a high potential of being useful in any possible actions the system may consider. Although the source control mechanism is not concerned with taking action on information, it is instrumental to that end, and a realistic assessment of the ability and trustworthiness of sources is therefore paramount.

In this chapter we have given substance to our claim that the model we proposed works, and can actually be reconstructed from the high level design we provide. We have also shown that the model so described can actually cope with the examples we originally claimed. We therefore have closed the circle which started from the initial methodology, problem statement and claims to producing the model, following the methodology can proceed to deal with the problem by providing a solution to the examples.

In this chapter we have shown how the source control mechanism can be embedded in a general reasoning system designed to operate in a natural language environment. We have given the constraints and detailed the general interface. We presented the basic entities the mechanism operates on. We have suggested a possible architecture, comprising the following components:

- a Control Regime to operate the model
- an Importance Analysis Component
- an Information Evaluation Component
- a Conflict Resolution Mechanism
- an Enquiry Module
- a Source Model Re-evaluation Unit

The way these components have been organised, allows the system to provide a flexible response to different situations.

Finally, we have completed the last step in our methodology, by applying the model to two of the examples we stated in Chapter 2. In this chapter we have generated a design from the principles, and we have shown that the system we proposed is capable of dealing with the examples, as we originally claimed.

Chapter 6

Complexity and Stability

6.1 Introduction

In the previous two chapters we have described the principles of the Source Control Mechanism as well as giving a design. In this chapter we shall investigate two further issues which are of interest for potential implementations of the Source Control Mechanism, namely its complexity and stability:

- The complexity of an algorithm is a measure of the computational cost of using the algorithm, while
- the analysis of stability of an algorithm considers the restrictions on refinements.

Both these issues will be of interest to a potential implementer of the Source Control Mechanism. The complexity of the algorithm is an important consideration; an estimation of the expected cost can be made in order to establish whether it is computationally feasible for a particular application. Refining the algorithm for a more sophisticated response may be desirable for a potential application, but there are important conditions to be fulfilled. The stability analysis ensures that a refinement such as increaseing the levels of belief, maintains a connection to the original algorithm.

6.2 Foundations of Complexity Analysis

An algorithm is a method for solving a particular class of problems on a computer. The complexity of the algorithm is the cost of using the algorithm to solve problems.

Computation takes time and algorithms differ in the speed with which they solve their respective problem. This may be due to the intricacy of the problem solving process or because of the sheer amount of data to be processed. The generally accepted practice is to describe the complexity of the algorithm as a function of some measure of the amount of data required to describe the problem to the computer. In addition, some algorithms are designed to solve a class of problems. While the algorithm may be very efficient for some problems, its application may become unmanageable when applied to more difficult problems, and therefore it may only be feasible when used for problems up to a certain point. Consequently, it has become standard practice to also describe the complexity of an algorithm as the limit when the input n becomes arbitrarily large.

There are a number of conventions about the classification of problems and measures to assess the complexity of algorithms as well as a specific vocabulary, which we want to introduce at this point:

- categories of problems
- performance estimates
- classification of growth rates
- Quicksort: an example

Categories of Problems

As there can be a considerable variation in the complexity of algorithms they have traditionally been divided into categories. The general rule for distinguishing between *easy* and *hard* problems is that if the running time is at most a polynomial function of the amount of input data then the calculation is *easy*, and in the converse case it is *hard*.

Performance Estimates

Not all algorithms will display the same performance for each problem, but may vary in their performance depending on the particular properties of the problem they are solving. Thus, even though a problem may be hard, not all instances of it need to be hard. In order to establish the performance bounds as well as giving a sense to the typical performance that can be expected it is common practice to use *worst-case*, *best-case* and *average-case complexity estimates.*

Classification of Growth Rates

It is usual to classify the growth rate of algorithms by comparing them to different categories of functions. If a function grows faster than x^a for every constant a, but grows slower than c^x for every constant $c > 1$ it is said to be of *moderately exponential* growth. In general, a function f is of *exponential* growth if there exists $c > 1$ such that $f(x)$ grows at least as fast as c^x and there exists d such that $f(x)$ grows at most as fast as d^x. Beyond this there are functions which grow exponentially fast. There are also more moderate forms of growth, such as *linear* and *polynomial* growth. However, the growth ranges of most concern to computer scientists are between the *slow logarithmic* functions and those which are of *exponential* growth.

Quicksort: An Example

Suppose we are given an array $x[1],\ldots,x[n]$ of n numbers which we need to sort in non-decreasing order. Supposing we have an array with the following elements [17,12,8,4,6,5,23,47,79,65,36] the element 23 is in a special position as all numbers preceding it are smaller than 23 and all elements after it are larger. It is therefore a splitter in the array as it separates two distinct halves of the array. One can therefore sort each half separately and rejoin the two halves and make considerable economy. However, not every array has a splitter and one will have to do preliminary work to create one. This can be achieved by pre-sorting the array around a randomly created splitter, and by recursively splitting the sub-arrays. This is the principle that Quicksort is based on.

The complexity of the algorithm depends on two factors, namely which array we are sorting and how lucky we are in the random choice of splitting elements. The *worst case* occurs when the splitting element happens to be the smallest or greatest element in the array and therefore no real splitting occurs. In addition this needs to happen on each occasion we are trying to split the array. In that case one recursive call will be on an empty sub-array and the other will be on a sub-array of n - 1 entries. If $g(n)$ is the number of comparisons required and the splitting is required for an array of $n - 1$ elements, it can be shown that the cost can be in the *worst-case* as high as $(n(n - 1))/2$.

From this we can see that Quicksort is at worst quadratic, but its performance is usually much better and given a more suitable choice of splitters, it can be shown that the *average case* of the performance of Quicksort is about $2n \log n$ as $n \to \infty$. The *worst case* is similar to the performance of more primitive methods of sorting and the economy of Quicksort only becomes apparent when one considers the *average case.*

6.3 The Nature of the SCM Algorithm

As will have become clear from the description of the principles and design, the SCM is symbolic-qualitative rather than a numeric-quantitative. There are a number of properties which distinguish it from the typical examples of algorithms used in discussing complexity. In the following we shall look at these properties before embarking on the proposed analysis of complexity:

- The SCM is a heuristic approach which asks a number of questions and which takes decisions based on the answers to these questions; it generally does not perform numeric manipulations and there is no simple mathematical function that can be produced to describe it.
- The SCM is a multiple-component system, comprised of a number of relatively independent modules which deal with particular aspects of the information to be processed and which have been specially devised for their task.
- The SCM is highly multi-threaded: there are a number of possible paths the problem solving process can take, which differ in length and cost.
- The SCM is adaptive: the amount of analysis it receives depends on a number of factors: the merits of the information, the availability of related information, interest in the information (as determined by the system the SCM is connected to), and the SCM's interest in the source.

The quicksort algorithm is an example of the algorithms typically chosen to present complexity analysis and there are a number of interesting observations that can be made. Most algorithms usually considered are based on numeric or mechanical problem solving rather than on more rule-based approach typical of AI applications. Frequently, these algorithms are also single-threaded in the sense that they will deal with a single, uniform problem addressed with a single, uniform method. They are also frequently categoric, meaning that the algorithms will be given an input and will apply themselves without varying their strategy dependent on the properties of the particular input. In most cases the algorithms are considered in isolation and will only have the problem description as their input without taking into consideration the way these algorithms are used in a larger system. The SCM is heuristic, and proceeds by making a series of decisions to distinguish situations which are significant for semantic reasons rather than determined by the syntactic process leading up to them. The SCM also has a number of components which are different not only for their purpose but also different from a complexity point of view. They are connected with each other in a way that allows the SCM to respond to different situations. There are a number of possible routes the analysis may take, and which of the paths the analysis eventually takes depends in part on the information itself and in part on the system's interest in the information or the SCM's interest in the source. Therefore, the SCM does not display a categoric response, but varies its behaviour based on these additional considerations.

The number of different components involved and the fact that the actual behaviour of the SCM is partially dependent on these additional factors means that the actual cost will vary. Seen from this angle, this is not an unusual situation to arise in complexity analysis. As can be seen from the quicksort algorithm, its performance can vary considerably, and is dependent on the properties of the input string and the suitable choice of random splitters. At the same time the case is much more accentuated with the SCM, as there are a considerable number of components, not all of which may be required in each instance.

6.4 Types of Analyses

The nature of the SCM can be described as that of an intricate decision process. The most suitable way of analysing the complexity of the SCM therefore appears to be to describe the decision process and the different possible situations that can arise by way of a set of decision trees and to perform the following four types of analyses:

- best case analysis
- worst case analysis
- average case analysis
- likely case analysis

An analysis of the *best case* establishes the minimum cost of using the SCM; the mechanism will at least require that many steps. As we will see, this situation arises when the system has no interest in the information and no interest in the source and the information is soon rejected. This measure will therefore establish the lower bound of the complexity of the algorithm.

In addition, we will consider the ***worst case*** scenario. This is a conservative estimate which establishes the upper bound to the resources that could be consumed by some input. This situation arises, when, in the face of system interest in the information and the source, the information is fully analysed and where there are a number of issues which need to be clarified through enquiries.

Thirdly, we can consider the *average case* over all the situations distinguished by the system. This measure should be more realistic as the system is likely to face a mixture of cases in its application.

Finally, we will discuss a *likely case* scenario. The *average case* analysis looks at the algorithm in isolation, in a strictly mathematical way, considering the different possible outcomes. In order to give an approximation of the cost of using the SCM in a typical application environment, we shall also discuss the *likely case*.

In the following we will look at the best-, worst- and average-case analyses together, before we consider a likely case analysis.

6.5 Categorisations and Conventions

The SCM is not an algorithm based on calculation or mathematical manipulation. In essence, the SCM rule-based system and goes through a process of decisions in order to determine the belief to be assigned to each event. Therefore we describe the algorithm of the SCM with the help of a set of decision trees together with their connections.

As we discussed in the previous section, the SCM is a sophisticated mechanism and there are two distinctions we have to make in the process of giving an estimate of the expected cost:

Firstly, the SCM is designed to operate as a sub-system connected to a system which provides the input to the SCM, and the SCM deals with the uncertainty of the input and attaches beliefs. As a natural language processor is one of the most likely candidates the SCM would be connected to, we refer to such a system generically as NLP to distinguish it from the SCM. There are points during the analysis where the SCM interacts with the system it is embedded in, to ask for assistance or further information. What happens between the point where the SCM makes a request, and the point where it receives a reply, is not part of the SCM. A NLP may immediately supply the information or refuse the request or may take some time to produce the information requested. This is dependent on the way such a NLP is implemented. As this is not under the control of the SCM, the cost of system requests need to be distinguished.

Secondly, there are a number of decisions which are based on information which may, or may not be readily available from a NLP in a given application of the SCM. If available, the SCM could easily use the information or in the converse case the SCM may have to generate them. It is therefore difficult to establish the precise cost of these decisions, as they are implementation dependent. If the SCM was used, for example, with a system like LOLITA [LOL92], the organisation of its semantic net would make a considerable amount of information associated with an event readily available, but that kind of information may not be available and may need to be generated and we therefore need to distinguish decisions in this group.

On the decision trees and in the complexity analysis we distinguish three different kinds of nodes:

- Cat1
- Cat2
- Cat3

Cat1 are those nodes where the information is explicitly available to the SCM. They are a matter of a simple look-up or of recording a decision and should take a single step. Examples of this are to look at the level of a particular source index, or the existence of a particular endorsement. In the diagrams all nodes which are marked by a preceding dot are Cat1 nodes (i.e. .Situation Type?) as distinguished from other labels.

Cat2 nodes, which are denoted by a colon preceding their label (i.e. :Abil?), reflect the fact that the information is not immediately available but, is implicit and some preliminary work may be necessary to be able to decide. This may be relatively easy to achieve, such as determining whether there is a helpfulness problem, by considering whether the information was volunteered, or asked for, and whether the relation between the actual behaviour of the source and the helpfulness index suggests a problem. On the other side, to work out the ability type can be easy if the NLP can provide information about classifications of information types or may have to be worked out by the SCM and would therefore be more expensive. It will depend on the particular implementation environment and the availability of additional information but the process is still within the control of the SCM.

Finally, Cat3 nodes are those nodes which imply a system request, and are denoted by three dots preceding the respective label (i.e. .:Aresp?). Although occasional interaction with the main system may be required the architecture of the SCM is designed to use system requests as little as possible, as they can be expected to be expensive. What is involved in responding to these requests and how long it will take is outside the control of the SCM. As this is external to the SCM, and the cost cannot be determined categorically, the cat3 steps are distinguished from other nodes, thus making it clear what their impact is on the SCM.

There are important differences between these categories. Cat1 imply operations which are known to be simple. Cat2 decisions are those known to be more expensive and the precise cost will depend on the particular implementation. A categoric cost cannot be established, except that they are known to be more expensive than cat1 decisions. The distinction between Cat2 and Cat3 steps lies in the fact that while Cat2 decisions are under the control of the SCM and the SCM can influence their behaviour, Cat3 decisions lie outside the direct influence of the SCM and are external as opposed to cat2 decisions which are internal. The SCM can make requests and wait for a response, but cannot influence that process in any other way. Although Cat3 requests may be answered promptly in some cases, they are assumed to be the most expensive decisions and as a conservative measure we consider them to be incommensurate with cat2 and cat1.

As these decision points are different in category, we maintain this difference in the analysis by showing the type of each decision point in the diagrams, and by calculating the cost separately. The reader will notice that the results are presented as a 3-tuple like 4 - 2 - 1 where the first element shows the cost of 4 cat1 decisions, followed by 2 cat2 and 1 cat3 decisions respectively. This arrangement ensures that it can be seen at any point how many decisions are cat1, cat2 or cat3. Given that the magnitude of the separate categories can be determined for a particular application, the cost for the respective implementation can easily be established.

As a precise equivalence between the three categories cannot be established they are assumed to be incommensurate. This does not affect the average case analysis, but means that when we establish the best case we start by looking at the branches with the least amount of cat3 components and then look for the branches with the least amount of cat2 components followed by choosing the path with the least cat1 components. This implies that the best case is chosen as the best of a worst situation (where cat3 are always assumed to be more expensive than cat2). Similarly, to choose the worst case one needs to start by finding the path with the most cat3 steps, then the most cat2 steps and finally the most cat 1 steps. This is the principle applied in the calculations in the following sections.

Finally, to clarify some further conventions used in the diagrams. The meaning of the labels are given in the legend to be found preceding the respective diagrams in the appendices. Each separate diagram carries at the top node a label with its unique name. Diagrams which had to be split over consecutive pages are connected by a boxed node on the original diagram, which connects to the diagram by the unique diagram name shown within square brackets. This shows where the separate diagrams fit together.

If the original diagram continues after the boxed label, this means that after the component referred to has been executed, processing resumes on the path set out in the original diagram.

In the process of calculating the cost of the system, the cost of each major component of the SCM (i.e. Control Regime, Importance Analysis, Information Evaluation etc.) is calculated separately. The best and worst cases are selected according to the principle mentioned and the average cost of the component is calculated by adding the costs of all branches from the top node to the bottom of the tree, by maintaining the separation of categories and dividing each by the number of outcomes. The method of calculating the overall cost will be explained shortly.

6.6 The Components of the SCM

As stated earlier, the SCM is composed of a number of components which we need to introduce before examining the complexity of the entire SCM. At the top level the SCM splits into two major components:

- Belief Formation
- Source Reevaluation

These two components are in turn divided into a number of sub-components. This results in the following components for the Belief Formation mechanism:

- Control Regime
- Importance Analysis
- Information Evaluation
- Conflict Resolution

- Enquiry

In addition the Source Reevaluation mechanism has two levels, one of which acts notionally like a control regime while the other is a collection of analyses to be carried out, dependent on the distinction made in the control level:

- Source Reevaluation Control
- Source Reevaluation Analysis

We shall look at these components in turn, by describing the decision process by way of decision trees and by giving an analysis of their respective cost. After a brief consideration of the system architecture, we start by calculating the sub-component costs first, before calculating the overall costs.

6.6.1 The Architecture of Belief Formation

The belief formation process has a number of components and we need to clarify the architecture of the sub-system before considering the components separately. The following diagram gives an overall picture of the belief formation process to show how the different components of belief formation fit together.

Starting from the top, there are a sequence of analyses carried out and a number of decision points with a number of alternatives branching out into the different paths the analysis can take. Only one path is eventually carried out and once the bottom of the tree is reached, analysis stops and the result is recorded. In the course of analysis a select number of components are executed and these will be described in the following sections. The Control Regime which (apart from a small component to produce the Source Models) seems absent from the diagram, actually embodies this decision process, deciding which path the analysis will take and handing cases to other components to be processed. The different paths are identified from A to J.

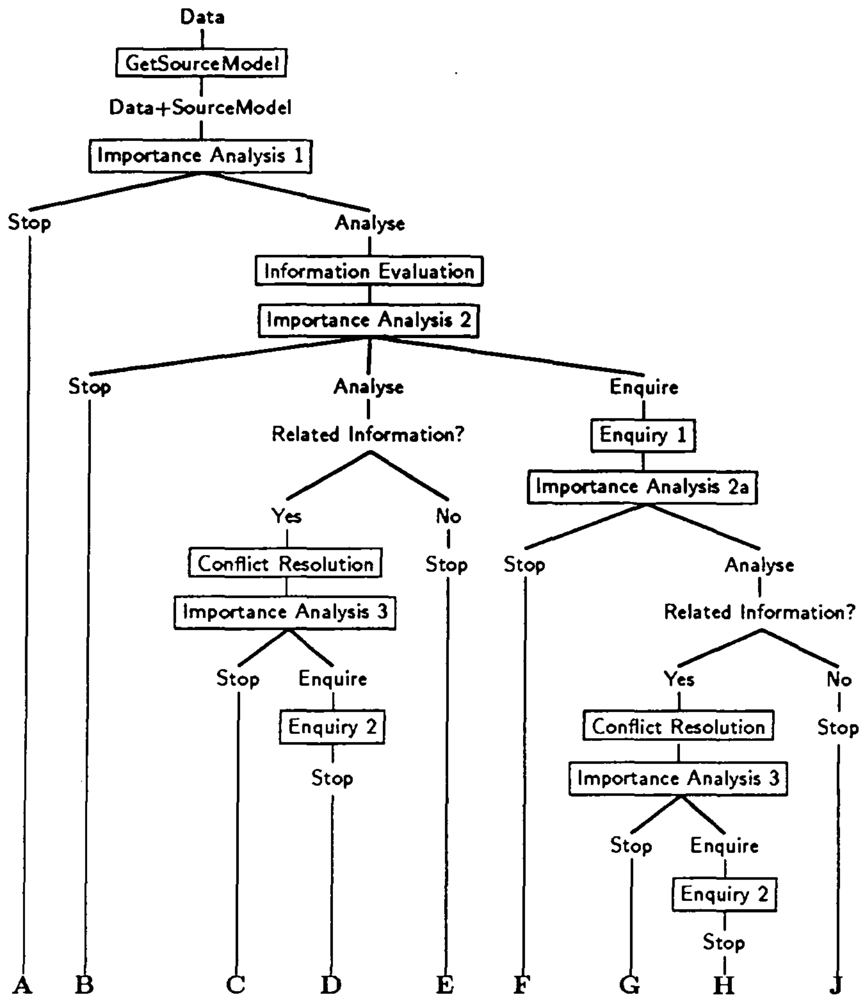


Figure 7: Overview of Belief Formation

In the following sections we look at the decision trees in Appendix A for the various components and calculate their respective cost before looking at the overall cost of belief formation.

6.6.2 Control Regime

The Control Regime is primarily a framework, coordinating the operation of the other components. The component is described in two decision trees, labelled [CtRg 1.1] which deals with the generation of the appropriate source model and [Cont Reg] which represents the main task of the Control Regime to steer the analysis process.

The boxed items on the main diagram refer to other components which are executed at the appropriate point and upon completion the Control Regime determines how to carry on the process of belief formation. The different paths are labelled from A to J. As the cost of the paths differ, and needs to be taken into account in the calculation of the overall cost, we give the costs separately for each path according to the conventions mentioned earlier. Note, that we calculate at this point the framework of the control regime *without* the costs of other components used along the path (the exception being [CtRg 1.1] which is part of the Control Regime), which will be added later, when we calculate the overall cost. The differences between best and worst cases are caused by the difference between the three paths distinguished in [CtRg 1.1] and the average is taken over these three paths and separately for paths A to J, as this is needed for the calculation of the overall cost discussed later:

Case	best case	worst case	average case
A	4 - 0 - 0	4 - 1 - 0	4.00 - 0.67 - 0.00
B	5 - 0 - 0	5 - 1 - 0	5.00 - 0.67 - 0.00
C	6 - 0 - 1	6 - 1 - 1	6.00 - 0.67 - 1.00
D	6 - 0 - 1	6 - 1 - 1	6.00 - 0.67 - 1.00
E	5 - 0 - 1	5 - 1 - 1	5.00 - 0.67 - 1.00
F	6 - 0 - 0	6 - 1 - 1	6.00 - 0.67 - 1.00
G	7 - 0 - 1	7 - 1 - 1	7.00 - 0.67 - 1.00
H	7 - 0 - 1	7 - 1 - 1	7.00 - 0.67 - 1.00
J	6 - 0 - 1	6 - 1 - 1	6.00 - 0.67 - 1.00

6.6.3 Importance Analysis

Importance Analysis is carried out at various points during processing, and depending on the merits of the information and the endorsements to process, this component recommends to the Control Regime whether to stop processing, whether to carry out further analysis or whether to recommend an enquiry. The component splits into four distinct sub-components and is described in four separate trees [ImpAn 1], [ImpAn 2], [ImpAn 2a] and [ImpAn 3]. As not all of them may be used during belief formation, each of them is calculated separately. Also since the choice of the overall paths A to J is dependent on the outcome of Importance Analysis, the calculation of cost is done separately for the different possible types of outcome of the respective decision tree:

component	best	worst	average
[ImpAn 1]			
stop	2 - 0 - 1	0 - 1 - 2	1.33 - 0.33 - 1.67
analyse	2 - 0 - 1	0 - 1 - 3	0.89 - 0.56 - 2.22
[ImpAn 2]			
stop	2 - 0 - 0	4 - 0 - 0	3.60 - 0.00 - 0.00
analyse	3 - 0 - 0	4 - 0 - 0	3.67 - 0.00 - 0.00
enquire	3 - 0 - 0	4 - 0 - 0	3.80 - 0.00 - 0.00
[ImpAn 2a]			
stop	3 - 0 - 0	4 - 0 - 0	3.37 - 0.00 - 0.00
analyse	3 - 0 - 0	4 - 0 - 0	3.90 - 0.00 - 0.00
[ImpAn 3]			
stop	2 - 0 - 0	4 - 0 - 0	3.43 - 0.00 - 0.00
enquire	4 - 0 - 0	4 - 0 - 0	4.00 - 0.00 - 0.00

6.6.4 Information Evaluation

Information Evaluation analyses pieces of information on an individual basis and recommends a belief. The decision process of the component is described in two diagrams, [Info Eval] and [InEv 1.1]. On the first diagram on the top level a distinction is made between information below or above the action point. On the right branch the next distinction is made on the question of whether responsibility is assumed by the source and two branches are shown distinguishing the case where responsibility is accepted and where it is denied. The second diagram deals with the case where responsibility is questionable, denoted by the label '?' and connects at that point. From the method previously described, the cost of the component can be described as follows:

Best Case:	3 - 1 - 0
Worst Case:	4 - 2 - 0
Average Case:	3.78 - 1.63 - 0.00

6.6.5 Conflict Resolution

Conflict Resolution deals with pieces of information which are related to other pieces of information already in the system's database. Two main categories are distinguished, namely whether the connection is with information from the same source or from other sources. The first diagram [Conf Res] distinguishes the case where there is a connection to information from the same source and cases where the sources are different. [CoRe 1.1] and [CoRe 1.1.1] deal with information from different sources. The third diagram connects to the second as it deals with the fourth case of contradiction in addition to the three cases of reiteration, weakening and strengthening. Only one of single-source or multiple-source sub-components is eventually executed, but as it cannot be determined *a priori* which case we are dealing with, the cost of conflict resolution is considered for both together:

Best Case:	3 - 1 - 0
Worst Case:	3 - 2 - 1
Average Case:	3.17 - 1.66 - 0.34

6.6.6 Enquiry

The Enquiry deals with cases, which, having received analysis, are found to have unresolved problems. From the Control Regime it can be seen that there are two different enquiries, which are distinguished by the particular situation they are examining. [Enquiry 1] is concerned with problem cases from the Information Evaluation Process, whereas [Enquiry 2] deals with cases as a result of Conflict Resolution.

[Enquiry 1] is described in three diagrams, the first of which deals with trust problems, whereas the other two [Enq 1.1] and [Enq 1.2] deal with ability problems. [Enq 1.1] deals with original information (i.e. personally from that source) whereas the third diagram [Enq 1.2] deals with information which has been reported from another source. As it cannot be determined in advance which situation the enquiry may have to deal with, their cost is considered together:

Best Case:	**4 - 0 - 0**
Worst Case:	**5 - 0 - 2**
Average Case:	**4.18 - 0.39 - 1.21**

[Enquiry 2] is again described in three diagrams, the first of which dealing with single-source self-contradiction problems, whereas the second [Enq 2.1] deals with multiple-source tie situations and the third [Enq 2.2] dealing with multiple-source contradictions. They are independent of each other and only one of them will eventually be required. The cost of them is therefore considered together as it cannot be determined beforehand which situation may obtain:

Best Case:	4 - 0 - 0
Worst Case:	4 - 1 - 2
Average Case:	4.19 - 0.73 - 1.09

6.6.7 The Aggregate Cost of Belief Formation

Having established the cost of the individual components we can now start to determine the overall cost of belief formation. In the following table the cost of individual components of the Source Control Mechanism are listed together (excluding the Control Regime):

COMPONENT	best case	worst case	average case
[Imp An 1]			
stop	2 - 0 - 1	0 - 1 - 2	1.33 - 0.33 - 1.67
analyse	2 - 0 - 1	0 - 1 - 3	0.89 - 0.56 - 2.22
[Imp An 2]			
stop	2 - 0 - 0	4 - 0 - 0	3.60 - 0.00 - 0.00
analyse	3 - 0 - 0	4 - 0 - 0	3.67 - 0.00 - 0.00
enquire	3 - 0 - 0	4 - 0 - 0	3.80 - 0.00 - 0.00
[Imp An 2a]			
stop	3 - 0 - 0	4 - 0 - 0	3.37 - 0.00 - 0.00
analyse	3 - 0 - 0	4 - 0 - 0	3.90 - 0.00 - 0.00
[Imp An 3]			
stop	2 - 0 - 0	4 - 0 - 0	3.43 - 0.00 - 0.00
enquire	4 - 0 - 0	4 - 0 - 0	4.00 - 0.00 - 0.00
[Info Eval]	3 - 1 - 0	4 - 2 - 0	3.78 - 1.63 - 0.00
[Conf Res]	3 - 1 - 0	3 - 2 - 1	3.17 - 1.66 - 0.34
[Enquiry 1]	4 - 0 - 0	5 - 0 - 2	4.18 - 0.39 - 1.21
[Enquiry 2]	4 - 0 - 0	4 - 1 - 2	4.19 - 0.73 - 1.09

Figure 8: Cost of Components of Belief Formation

As discussed earlier, the Control Regime introduces Importance Analyses at strategic points which determine whether further processing will be carried out. The following table shows the recommendations of the respective sub-components of the Importance Analysis, how many cases are stopped, how many are further analysed and how many cases should receive an enquiry:

COMPONENT	stop	analyse	enquire
[Imp An 1]	3	9	-
[Imp An 2]	5	6	5
[Imp An 2a]	8	10	-
[Imp An 3]	7	-	12

Figure 9: Decisions of Importance Analysis

The diagram for the Control Regime shows alternative ways in which a belief formation process can happen and which components will be required in the process the SCM goes through. The following diagram contains three further pieces of information necessary to determine the average complexity of Belief Formation. Apart from labelling the different routes from A to J, at each branching point the number in square brackets at the parent node describes the total number of cases distinguished and the numbers in square brackets at each child node describe how many of those cases will follow that path (i.e. [Imp An 1] distinguishes 12 cases, 3 of which are stopped and 9 go on to further analysis). At the leaves of each branch are indicated the probabilities for arriving at that point, representing the relative likelihood of going down that path.

In the subsequent table are contained the aggregate costs of running the components required for the path in question, together with the relative probability of having to deal with this particular case. Thus, for dealing with a case J the following components are involved: [Cont Reg], [Imp An 1], [Info Eval], [Imp An 2], [Enquiry 1] and [Imp An 2a]. The cost of path J is calculated by adding the average costs of the components involved to the average cost of path J from the Control Regime, and likewise for the best- and worst case.

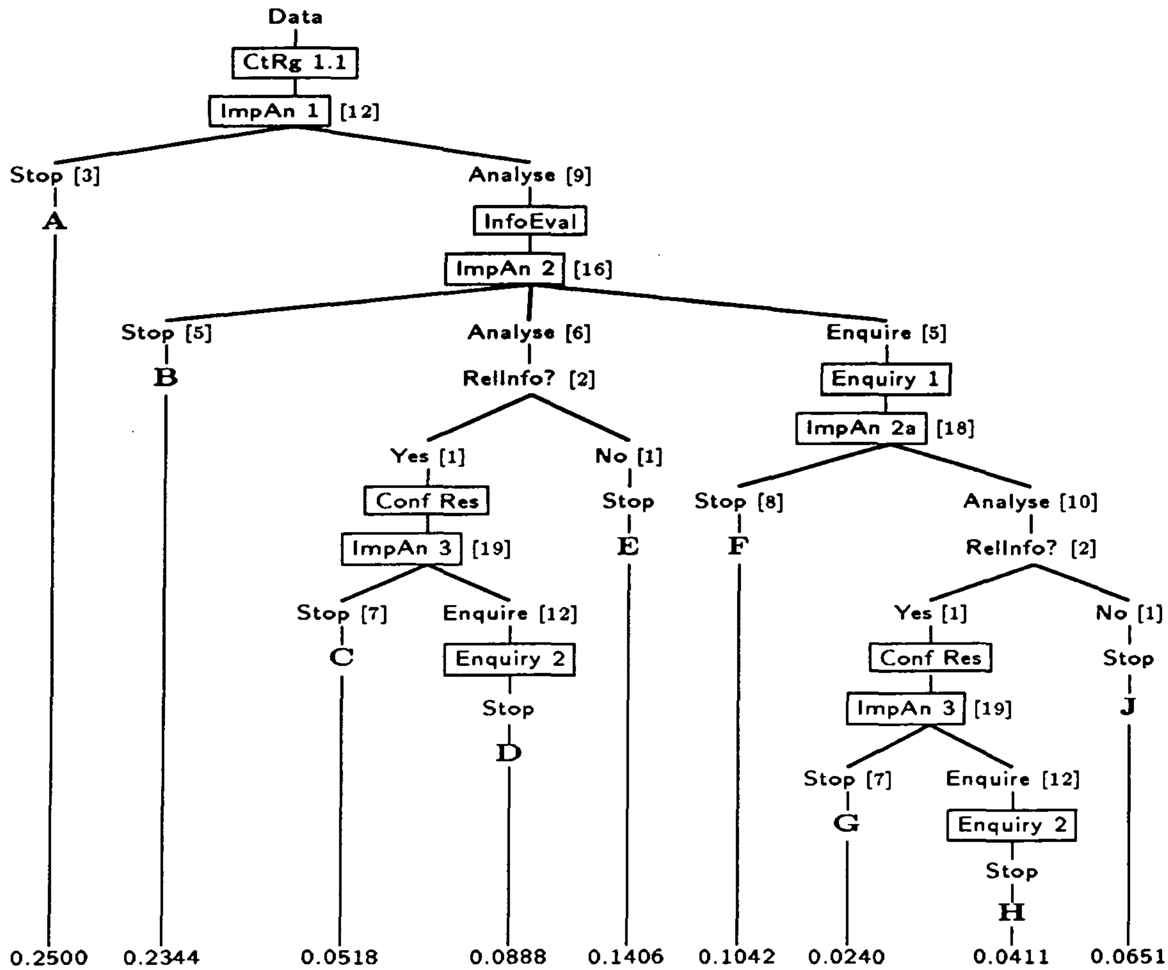


Figure 10: Probabilities of Paths

The probability for each particular path is the product of the ratios of cases for each branching point along the path. Thus, for example, the probability of path J is $p = (9/12 \cdot 5/16 \cdot 8/18) = 0.1042$.

Path	best case	worst case	average case	probability
A	6 - 0 - 1	4 - 2 - 2	5.33 - 1.00 - 1.67	p = 0.2500
B	12 - 1 - 1	9 - 4 - 3	13.27 - 2.85 - 2.22	p = 0.2344
C	19 - 2 - 2	21 - 6 - 5	20.94 - 4.51 - 3.56	p = 0.0518
D	25 - 2 - 2	25 - 7 - 7	25.70 - 5.24 - 4.66	p = 0.0888
E	13 - 1 - 2	13 - 4 - 4	13.34 - 2.85 - 3.22	p = 0.1406
F	21 - 1 - 1	23 - 4 - 6	22.03 - 3.24 - 4.44	p = 0.1042
G	27 - 2 - 2	31 - 6 - 7	30.16 - 4.90 - 9.68	p = 0.0240
H	23 - 2 - 2	35 - 7 - 9	34.92 - 5.63 - 5.87	p = 0.0411
J	21 - 1 - 2	23 - 4 - 6	22.56 - 3.24 - 4.34	p = 0.0651

Figure 11: Total Cost of Individual Paths

From these values we can now establish the overall cost for the belief formation process for a candidate event:

Path	average case	probability	proportional cost
A	5.33 - 1.00 - 1.67	p = 0.2500	1.33 - 0.25 - 0.42
B	13.27 - 2.85 - 2.22	p = 0.2344	3.11 - 0.67 - 0.52
C	20.94 - 4.51 - 3.56	p = 0.0518	1.08 - 0.23 - 0.18
D	25.70 - 5.24 - 4.66	p = 0.0888	2.28 - 0.46 - 0.41
E	13.34 - 2.85 - 3.22	p = 0.1406	1.88 - 0.40 - 0.45
F	22.03 - 3.24 - 4.44	p = 0.1042	2.29 - 0.34 - 0.46
G	30.16 - 4.90 - 9.68	p = 0.0240	0.72 - 0.12 - 0.23
H	34.92 - 5.63 - 5.87	p = 0.0411	1.43 - 0.23 - 0.24
J	22.56 - 3.24 - 4.34	p = 0.0651	1.47 - 0.21 - 0.28
		total	15.59 - 2.91 - 3.19

Figure 12: Total Average Cost of Belief Formation

From the table we can also see that the best overall case occurs on path A where the information is not considered significant enough to be further analysed. In the converse case we can see that the worst possible case occurs when the processing incorporates all possible forms of analysis and enquiry, as in H. In addition, the weighted average can be calculated, based on the relative likelihoods indicated, by multiplying each average figure by the respective likelihood and adding the costs so derived.

The average case can thus be established and the best case is the best case of path A and the worst case is to be found in path G:

Best Case:	6 - 0 - 1
Worst Case:	35 - 7 - 9
Average Case:	15.59 - 2.91 - 3.19

6.7 The Cost of Source Reevaluation

The Source Reevaluation Process works relatively independent from the Belief Formation Process and they differ in their perspective. While belief formation works on the short term basis of assigning beliefs to individual pieces of information, source reevaluation works on a more long-term perspective to adjust the source models used by belief formation. Again, source reevaluation branches out into a number of different components which are united on the top level by an importance analysis to cut down on excessive processing.

6.7.1 The Architecture of Source Reevaluation

Similar to the architecture of the Belief Formation process the Source Reevaluation process has a number of components which are operated through a control level. The following diagram shows this control level in general terms and how the sub-components of the source reevaluation process are connected. Following the presentation of the separate components which can be found in Appendix B, and the analysis of the cost, we consider the overall cost of the source reevaluation process.

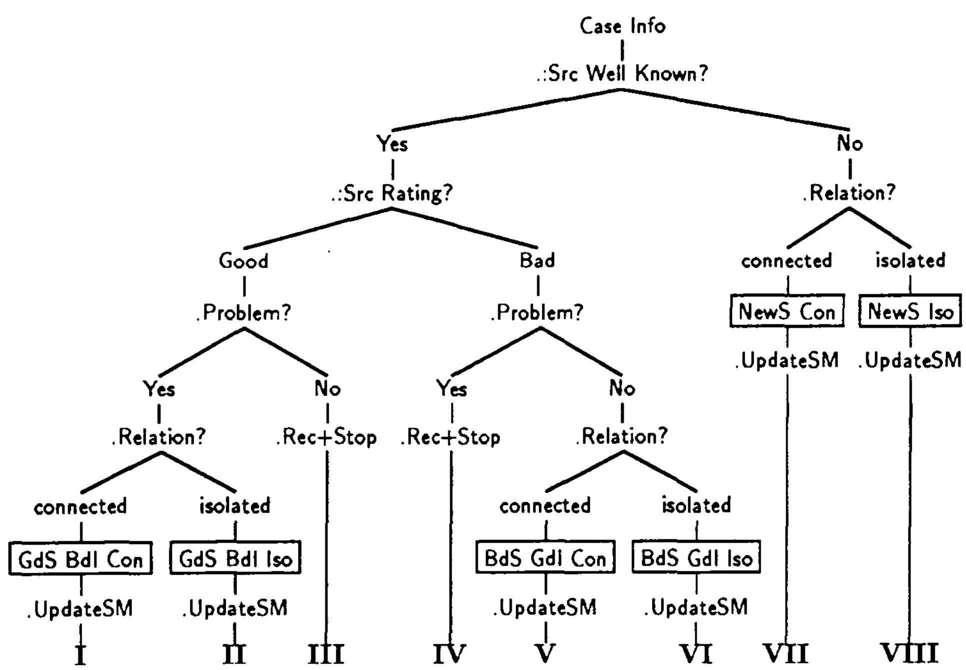


Figure 13: Overview of Source Reevaluation

6.7.2 Source Reevaluation Control

The Source Reevaluation Control needs to determine whether to perform a source model reevaluation or whether to just record the incidence as a basis for analysis sometime in the future. [Src Reev] shows the process of distinguishing those cases which will be referred to a full reevaluation from those merely registered and recorded.

As with the Control Regime in Belief Formation one needs to distinguish the different paths for the calculation of the overall cost of Source Reevaluation (labelled I to VIII). As there are no alternative branches to each path, the best-, worst- and average cases coincide, and the cost of the framework, without the cost of components, can be established as follows:

Path	cost
I	3 - 2 - 0
II	3 - 2 - 0
III	2 - 2 - 0
IV	2 - 2 - 0
V	3 - 2 - 0
VI	3 - 2 - 0
VII	2 - 1 - 0
VIII	2 - 1 - 0

Figure 14: Paths of Source Reevaluation

6.7.3 Source Reevaluation Analysis

There are three major types of Source Reevaluation Analysis each of which separates into two parts; isolated and connected. They reflect three main categories, namely good information from bad sources ([GoodS BadI Iso] and [GoodS BadI Con]), bad information from good sources ([BadS GoodI Iso] and [BadS GoodI Con]) and information from new, little known sources ([NewS Iso] and [NewS Con]). Again there are connections between different parts of the tree according to the same principle seen in Belief Formation. These six situations are distinguished by the Source Reevaluation Control and only one is executed eventually (if any) and they are therefore presented separately:

Case	best	worst	average
[GoodS BadI Con]	3 - 1 - 0	3 - 2 - 1	3.36 - 1.67 - 0.19
[GoodS BadI Iso]	2 - 1 - 0	4 - 3 - 0	3.34 - 1.88 - 0.00
[BadS GoodI Con]	2 - 1 - 0	5 - 2 - 0	4.29 - 1.67 - 0.00
[BadS GoodI Iso]	2 - 1 - 0	2 - 3 - 0	2.48 - 1.85 - 0.00
[NewS Con]	3 - 1 - 0	5 - 2 - 0	4.33 - 1.67 - 0.00
[NewS Iso]	3 - 1 - 0	4 - 2 - 0	3.58 - 1.69 - 0.00

Figure 15: Component Costs of Source Reevaluation

6.7.4 The Aggregate Cost of Source Reevaluation

From the previous two tables and in concordance with the organisation of [Src Reev] we can now establish the aggregate cost of the paths I to VIII. The following table gives the best-, worst- and average-case cost for each path:

Path	best	worst	average
I	6 - 3 - 0	6 - 5 - 1	6.36 - 3.67 - 0.18
II	5 - 3 - 0	7 - 5 - 0	6.34 - 3.88 - 0.00
III	2 - 2 - 0	2 - 2 - 0	2.00 - 2.00 - 0.00
IV	2 - 2 - 0	2 - 2 - 0	2.00 - 2.00 - 0.00
V	5 - 3 - 0	8 - 4 - 0	7.29 - 3.67 - 0.00
VI	5 - 3 - 0	5 - 5 - 0	5.48 - 3.85 - 0.00
VII	5 - 2 - 0	7 - 3 - 0	6.33 - 2.67 - 0.00
VIII	5 - 2 - 0	6 - 3 - 0	5.56 - 2.69 - 0.00

Figure 16: Cost of Paths of Source Reevaluation

Similar to the Belief Formation process we need to calculate the probabilities of the separate paths in order to calculate the average cost of source reevaluation. Since there is no bias towards particular branches the probabilities are evenly divided (i.e. to get to path I the probabilities are $p = 1/2 \cdot 1/2 \cdot 1/2 \cdot 1/2 = 1/16$ through 4 branching points). The following diagram shows the probabilities:

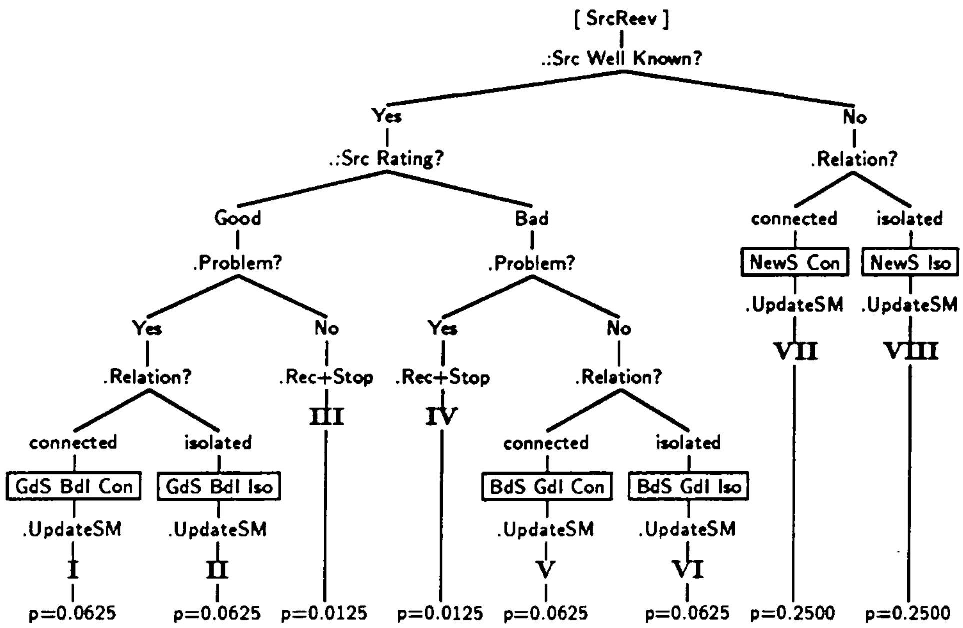


Figure 17: Probabilities of Paths

The overall best and worst cost of the Source Reevaluation Process can be seen from the previous table and the average cost can be calculated:

Path	average	prob.	proportional cost
I	6.36 - 3.67 - 0.18	p=0.0625	0.40 - 0.23 - 0.01
II	6.34 - 3.88 - 0.00	p=0.0625	0.21 - 0.12 - 0.00
III	2.00 - 2.00 - 0.00	p=0.1250	0.02 - 0.02 - 0.00
IV	2.00 - 2.00 - 0.00	p=0.1250	0.02 - 0.02 - 0.00
V	7.29 - 3.67 - 0.00	p=0.0625	0.46 - 0.23 - 0.00
VI	5.48 - 3.85 - 0.00	p=0.0625	0.34 - 0.24 - 0.00
VII	6.33 - 2.67 - 0.00	p=0.2500	1.58 - 0.67 - 0.00
VIII	5.56 - 2.69 - 0.00	p=0.2500	1.40 - 0.67 - 0.00
		total	4.43 - 2.20 - 0.01

Figure 18: Average Cost of Source Reevaluation

This yields the following results:

Best Case:	2 - 2 - 0
Worst Case:	6 - 5 - 1
Average Cost:	4.43 - 2.20 - 0.01

6.8 Likely Case Analysis

Until now we have considered the cost of the SCM in isolation. It was appropriate to proceed in this way to establish an objective measure of the cost involved, independent of any particular implementation and application. We established the best and worst case, thus defining the *bandwidth* of the SCM and, in addition, we calculated the average case to give a sense of what performance can be expected given that the SCM will deal with a variety of cases and that if cannot be determined *a priori* what the pattern of input will be.

In this section we want to consider the impact of the information pattern we expect the SCM to deal with. The average case is indiscriminate across all different possible situations and we now want to consider what the *likely cost* may be across a typical scenario. There are a number of points which may have a bearing on this:

- nature of information
- pattern of sources
- system goals

In usual, everyday situations the great majority of information is of low relevance. In the terms of the SCM, this would mean that the majority of information would not have an endorsement of being particularly important to the system. Looking at the importance analysis it is clear that that is one of the major reasons for disqualification; extensive analysis requires either a sufficient system-importance or the source being important to the SCM. Given that situation the information is not likely to be subject to enquiries, as the analysis would be stopped at that point. Whether the information will be thoroughly analysed, will be much more dependent in the SCM taking an interest in the source. A considerable amount of information is also likely to be isolated, thus removing the need for a conflict resolution. This means that the bias would be heavily towards paths A, B, C and E, thus cutting down on the more expensive paths like D, G and H.

Everyday experience will also convince us that we usually deal with a pool of about 30-50 sources we are better acquainted with, together with the occasional new source. This in turn has an effect on the Source Reevaluation process as we are not likely to spend a great deal of time dealing with new sources, which we would need to get to know, in order to build up an adequate source model. As time progresses the source models we have should settle down to a more steady state, not requiring continuous revision and adjustment. Both these aspects should therefore cut down considerably on the amount of source reevaluation being performed.

Another factor which has an influence on the performance of the SCM is the goal of the system to listen to good sources rather than sources which are persistently bad. This has a considerable effect on both the belief formation and source reevaluation. Firstly, unless the system shows an interest in the information, the SCM is unlikely to want to spend a great deal of effort on analysing the information and the source reevaluation process is unlikely to do much reevaluation as it expects bad information from bad sources. As a consequence, the system is more likely to deal with good information from good sources which means that there is less likelihood for enquiries to take place and that excludes all the most expensive paths D and F to J. It also implies that there is less demand for source reevaluation as the system expects to get good information from good sources there will be no need to constantly reevaluate these source models.

This leaves new sources, which are expensive as the system doesn't know them well and may have to sort out problems and to learn about them in order to build up a source model. But, the system should in usual circumstances not have to deal with excessive amounts of information from unknown sources and therefore not have to constantly put effort into building up source models.

In order to estimate the expected cost in a scenario like this, suppose 60% of information is divided evenly between categories A and B, and of the remaining 40% only a small proportion lead to an enquiry and, as we said before, less than half of the information has connections. As a result, we could get something like the following set of percentages with the associated proportional cost (calculated from the average costs stated previously):

Case	%	proportional cost
A	30	1.60 - 0.30 - 0.50
B	30	3.98 - 0.85 - 0.67
C	10	2.09 - 0.45 - 0.36
D	5	1.28 - 0.26 - 0.23
E	15	2.00 - 0.43 - 0.48
F	3	0.66 - 0.10 - 0.13
G	2	0.60 - 0.10 - 0.19
H	1	0.35 - 0.06 - 0.06
J	4	0.90 - 0.13 - 0.17
	total	13.46 - 2.68 - 2.79

Figure 19: Likely Cost of Belief Formation

For the source reevaluation what we said about the typical about the typical pattern of cases this means that there will be a considerable concentration of 60% on cases III and IV, as well a moderate amount of information from new sources up to 20%, leaving the other 20% to be divided between cases I, II, V and VI:

Case	%	proportional cost
I	5	0.31 - 0.18 - 0.01
II	5	0.17 - 0.09 - 0.00
III	30	0.60 - 0.60 - 0.00
IV	30	0.60 - 0.60 - 0.00
V	5	0.36 - 0.18 - 0.00
VI	5	0.27 - 0.19 - 0.00
VII	10	0.63 - 0.27 - 0.00
VIII	10	0.56 - 0.27 - 0.00
total		3.50 - 2.38 - 0.01

Figure 20: Likely Cost of Source Reevaluation

If we compare these figures with the average cost derived in the previous sections we can see that there would be an improvement on the belief formation process of about 10% and on the source reevaluation process about 10% as well.

6.9 Discussion of Results

From the results derived for the cost of Belief Formation *BF* and Source Reevaluation *SR* we can say that the costs can be established to be between the following limits:

$$(35a + 7b + 9c) \geq BF \geq (6a + 1c)$$

$$(6a + 5b + 1c) \geq SR \geq (2a + 2c)$$

where a is cat2, b is cat 2 and c is cat3. We can also say that, on average, the cost is:

$$BF_{average} = 15.59a + 2.91b + 3.19c$$

$$SR_{average} = 4.43a + 2.20b + 0.01c$$

Given that the cat3, which represent external enquiries are definitely more expensive than cat2 and cat1, it is the cat3 cost which has the decisive effect on the overall cost. From our earlier discussion of complexity theory we see that there is an assumption about the complexity being a function of the size of the input string. This is not the case with respect to the SCM and therefore the measures of asymptotic behaviour which have typically been used to describe the bounds of complexity of algorithms (o,O,~,Θ and Ω see [WIL86] pp. 9-11) are not meaningful, as they are built on that assumption. However, if we consider that o describes the upper bound then the equivalent concept is expressed by the *worst case* and likewise the *best case* is comparable to O which describes the lower bound. The *average case,* as we have seen in the case of quicksort is an acceptable measure that is independent of whether or not the function is dependent on some measure of the input and therefore applicable in the situation we are considering, Θ

expresses the concept of the cost being within a constant range; although that can be said for cat1 and cat2, which are internal and under the control of the SCM, it is more difficult to state for all possible systems the SCM might be linked to. In finding related information the cat3 request might incur a cost of searching for information which might be subject to the usual costs of searches. At the same time the system the SCM is connected to might have had to search through the database anyway in the process of trying to accommodate the new information and getting related information becomes a by-product.

It is difficult to give more precise estimates, as it is dependent on the particular application what the actual cost involved is, especially with respect to external queries represented by cat3 and to a lesser extent the help given to facilitate work done by cat2 functions. Given this uncertain situation, we have separately accounted for the different costs. In this way we show the impact of the separate classes, and once the factors which have an influence on cat2 and cat3 are known, it is easy to establish the cost for a particular application.

6.10 Discussion of Complexity

In this chapter we have described the algorithm of the SCM by way of the decision process the SCM goes through in its analysis. We have separately presented the different components of the SCM by giving the decision tree(s) for the component and calculated the cost of the best-, worst- and average cases. We then proceeded to establish the overall cost of belief formation and source reevaluation. In this section we want to discuss the implications of these results and discuss a few additional points:

- the cost as *n* grows without bound
- the SCM as an algorithm, not a family
- the cost of modification to the SCM algorithm

6.10.1 The cost as *n* grows without bound

We recall from our discussion of complexity analysis in general, at the beginning of the chapter, that the complexity of algorithms is usually considered as a function of the size of the input *n*, when *n* becomes arbitrarily large. The point of doing this is to show the nature of algorithms which deal with a variable input string; thus the time spent in using a given algorithm for inverting a matrix, or for sorting a list, increases as the matrix or the list becomes larger.

Clearly, the SCM deals with items presented one at a time and the cost of using the SCM is not a function of the number of items. Therefore the cost for each item will not increase as more and more events are being processed. From the calculations it is clear that the cost should be on average a constant multiplicative factor for each item, with the reservation that if cat3 steps involve searches which are carried out specifically for the SCM the complexity of the search applies, but otherwise the resulting complexity of the algorithm should be linear.

The cost of the SCM can under certain circumstances even be slightly lower than the average cost indicated. This situation arises if there is a reasonably steady pool of sources with respect to membership and their behaviour, paired with a relatively small amount of information from new sources. This would imply that once the SCM has generated adequate source models, their application will become fairly straightforward and there will be little need of enquiries and source reevaluation. Also, as time progresses the SCM would learn about the bad sources and gradually be less inclined to listen to them and therefore be less likely to have to sort out problems with bad information. This in turn means that the cost of the SCM will be kept to a minimum. The converse case, when sources are constantly changing and are frequently causing problems the SCM will operate at a higher cost, but as we have already discussed in the preceding chapters, the use of the SCM may not be recommended in those circumstances. However, even in those cases the cost should still be linear, though at a higher rate.

6.10.2 The SCM, an Algorithm, not a Family

In previous chapters we described the principles and design for the SCM and together with the decision trees presented in this chapter the algorithm of the SCM has been outlined. The SCM algorithm we have presented is a singular algorithm, whose purpose is to determine whether given pieces of information from human sources can be believed, and if so, to what extent. The design, provided in chapter 5 showed that the SCM can adequately deal with the examples, and the algorithm described in this chapter translates the design into an implementable decision process. The algorithm we have presented fulfils its purpose adequately and a distinction has to be made between the algorithm of the SCM we present, and variations on the algorithm, which may be desirable for the requirements of a particular application. We present a singular algorithm and not a family of algorithms. At the same time it is conceivable to produce variations of the algorithm, although one should carefully consider the motivations and implications for a departure from the SCM we presented. The SCM is capable of dealing adequately with the examples we introduced and when considering refinements one also has to consider, for example, that there is no point in producing a level of refinement of the SCM which cannot be appreciated by a natural language processing system the SCM would be attached to. Also, from a human perspective a more sophisticated reaction may be plausible in particular situations, but personal impressions and judgements are usually the product of a plethora of considerations, such as years of experience, a refined understanding of language, human nature and the nature of the world, and those aspects are not really the responsibility of the SCM. We argued in chapter 4 that the source model in its present form sufficiently covers the major aspects of source behaviour which are relevant for the purpose of assigning beliefs.

It is possible to conceive of an elaboration of the basic principles in order to produce a different SCM for a particular application. We discuss the implications of such changes on the complexity of the SCM in the following section.

6.10.3 The Cost of Modification to the SCM Algorithm

If we want to see the SCM as part of a family of algorithms, then there are four ways in which the principles of the SCM can be elaborated:

- adding new slots in the existing source model
- adding items to the existing slots
- increasing the resolution of values
- adding new rules

Adding New Slots in the Existing Source Model

It is difficult to see what kind of slot could be introduced to the SCM which would not be covered by the present Ability and Trust Indices. Thus the four main categories in the ability index of expertise, reasoning, judging and personal experience appear to cover all the major types of information and whatever further refinement would be desirable would probably fall within these categories. Supposing a separate ability index were to be found, it would have the effect that at those branching points in the belief formation process where the different categories of the ability index are distinguished (to ensure that the right index is being used in the evaluation) another branch would have to be introduced. As it is difficult to see how such a categories would differ from the others it would imply the introduction of a sub-tree in the decision process, similar to the sub-tree of the other alternatives. As a consequence the sub-tree would expand, however the cost of using it would not significantly, as the number of subsequent steps involved would be the same. The exception perhaps is the cost of the function which makes the distinction between the various cases and which has to consider an additional possibility. In the same way the elimination of a slot would imply the removal of the associated sub-tree with the effect of simplifying the function making the distinction. Again this will not significantly affect the cost of the respective component and we can conclude that the effect on the cost of the process would be minimal.

Adding Items to the Existing Slots

A change in the number of items in a slot such as specific indices for different kinds of expertise (like, for example, distinguishing between expertise in Artificial Intelligence and Traditional Computer Science as opposed to Computer Science in general) is already possible in the present design.

At the same time it should be stressed that obviously the need of such refinement will be born out of distinctions made in that area by an NLP for a particular application and which uses the SCM to determine and maintain its beliefs. Obviously, there is little purpose in introducing refinements in the SCM which the NLP is not able to effectively use. Given that the NLP was able to distinguish and use that information it would be reasonable for the SCM to deal with it provided that would have an effect on the belief formation or source reevaluation problem. For example, there may be many interests and beliefs which do not affect the veracity of the source and should not be introduced as they would unnecessarily clutter the indices. The point should be stressed that some source properties may be interesting in themselves for other reason, but only those properties which affect the belief formation or source reevaluation could have a legitimate claim to be explicitly represented in the source models and as a basis for forming beliefs. If a relevant distinction was made in the NLP, the SCM could obviously draw on the NLP's ability to distinguish and use that information in the belief formation or source reevaluation process. Consequently, the additional cost would be minimal, involving distinguishing a different case and applying a different index. As far as the decision diagrams are concerned there would be no formal change and only a marginal additional cost to distinguish the additional case in the selection function.

Increasing the Resolution of Values

Finally, the last possibility would be to change the number of possible values from the five values currently used in the SCM.

It is important at this point to make a distinction between the resolution of the strength of belief, of the values of the source model indices and the level of refinement of distinctions made in the belief formation and source reevaluation processes. As mentioned earlier, the SCM, as presented, works with five levels of belief and five values of source indices. The distinctions made in the belief formation and source reevaluation process, by contrast, are different and of two kinds; firstly, there are a number of examples of filters where we do not consider the whole spectrum of possible cases separately, but single out a particular, significant case and as a result follow one course of action rather than another. Secondly, at other points we effectively distinguish bands of values which in the end will lead to the same conclusion in the decision process.

The reader will notice that the decision on how to treat the piece of information is usually one of four possibilities, namely to reject, reduce, modify or accept the piece of information:

- to reject means to either expel the information from the database or reduce the belief to nil,
- to reduce the information means to reduce the belief to below the action point,
- to modify means to adjust the belief according to the level of ability of the source and
- to accept is to accept the information as presented by the source

As a result of this arrangement of filtering and banding, it is possible to have any refinement in the levels of belief and levels of source model indices as it affects only the modifying functions (which will adjust the levels of belief at the end of the belief formation process) or the functions calculating the new index during source reevaluation. The functions which make distinctions in the belief formation process can easily be implemented in a generic fashion to determine which band the particular level of the index falls into. Seen from this perspective, the effect of refinement would therefore cause a negligible amount of additional cost.

The situation would be different if there was a requirement for additional refinement in the SCM's response. If this is a matter of shifting the limits of the bands then that would only require a change in the distinguishing function to classify the case according to the new principle. Alternatively, if the requirement was to introduce additional, more narrow bands instead of the existing ones the question would arise as to whether the existing measures of fixing the belief (reject, reduce, modify and accept) would be sufficient or whether others would have to be introduced. In either case, the number of steps in the decision process for individual cases (depth of the tree) would still remain unchanged, although the tree would grow in width through the addition of the appropriate sub-trees.

Adding New Rules

Finally, it is conceivable that one could want to add new rules to make a further distinction which may appear relevant for certain cases. This would imply that at a certain point in the decision tree a new node would be inserted, and for each of the outcomes of the new rule similar sub-trees would be introduced. In this case the total number of steps for paths affected by the insertion of the rule will increase by one step and the whole tree will grow in width through the addition of new subtrees. If the designer decides to do something different following the outcomes of that new rule, then the effect on cost is dependent on the number of steps following on from that point and the size of sub-trees involved.

6.10.4 Conclusions

There are a number of conclusions to be drawn from these considerations which affect the complexity of modifications to the SCM. As mentioned earlier, the present algorithm is based on filtering and banding, which leaves room for modifications without affecting the present organisation, as filtering and banding does not exhaustively consider all possible cases, but only those of special interest or by uniting others into common classes. Therefore, one has to distinguish those refinements which do not affect the respective decision tree or process and those which do. The former category is negligible and requires only marginal adjustments in the operation of the process whereas the effects on the latter category are more visible, although it should be stressed, that even then the number of steps, and hence the cost of processing individual pieces of information, remains roughly unchanged.

The effect of introducing additional categories into the source models has the effect of the introduction of additional sub-trees and in a given uniform sub-tree the expansion is determined by the marginal effect of new categories on the existing number of categories.

The marginal effect of introducing refinement in cases distinguished at decision points in the belief formation process again causes an expansion of the decision tree. This will be proportional to the marginal effect of new cases on existing ones (supposing a uniform tree) and a product of the marginal effects if several refinements were introduced successively along the same decision path.

Finally, the addition of new rules has an effect on the number of steps to complete the respective tree and affects the subtrees following on from that point. It is not possible to determine the cost *a priori* unless the new rule does not disturb the sub-tree that was there before. In this case only one step is added to the cost and identical subtrees are added below the new rule, thus causing an expansion determined by the original size of the subtree and the number of branches generated by the new rule.

To conclude, there are four types of refinement with different effects on the complexity. Firstly, those refinements which observe the organisation of the present decision process and operate within it, do not significantly affect the complexity of the algorithm. Secondly, those refinements which require the extension of existing classifications, staying within the principles of filtering and banding (by considering more bands) will expand the decision trees in proportion to the increase of additional cases over existing ones on the parts of the decision tree affected by the change. Thirdly, the increase in the width of the decision trees caused by exhaustive distinctions will result in the product of successive decision steps affected (i.e. three successive distinctions of n possible values each, amounts to a tree of n^3 possible outcomes). Fourthly, the addition of new rules extends the path by a further node and causes an addition of new subtrees. It has to be stressed that although the second and third option entail an expansion of the tree in width they do not affect its depth and therefore in any of the cases so described, the number of steps (and their category type) remain unchanged and it is only through the addition of rules that the paths are potentially extended and the cost increased.

6.11 Application of the SCM to Production Management

In the previous sections we analysed the cost of the SCM in order to establish its feasibility from a computational perspective. To give a sense of a realistic application the SCM could be applied to and to give a sense of the rate of input it would be expected to deal with, we briefly look at an example of an application of the SCM in a manufacturing domain as part of a decision support system for integrated design optimization and production route engineering.

A proposal for a collaborative project (see Appendix C) to produce such a system has been motivated by problems caused by the fragmentation of the various engineering and planning functions in the manufacturing industry and the need to introduce effective IT approaches to integrate these functions. The interface between the design and production engineering functions are considered to be of particular importance since the production costs can to a considerable extent be affected by the decisions taken at this stage. The aim is to improve product quality and reduce design and manufacturing lead times and costs by parallel and synchronous product and process design by bringing together the engineers involved in the different parts of the process; this is termed the simultaneous engineering approach. The production routes on the shop floor have traditionally been generated manually by planners and the process is often based on personal preference and previous experience. Frequently, when new products are to be manufactured which have similarities with previous products, the new route is constructed by modifying the previous route. The process is very subjective and slow and does not take much notice of issues such as batch size variations, equipment utilisation and process optimisation potentials. Also the product designers do not usually participate in the route generation activity and this does not promote integration and may lead to a poor performance of the manufacturing process.

It is common for specialists to have contradictory points of view as to how to optimize a design in terms of its functionality, manufacturability and quality. However, the simultaneous engineering approach tends to be most effective at the early design stage, at which there is still a considerable amount of uncertain and incomplete information about certain aspects of the design and manufacturing process. This establishes a need for techniques to manage uncertainty and resolve conflicting viewpoints.

The aim of the project is to develop an AI tool consisting of a Route Generator and a Source Control System. The function of the tool is to assist the designers and production engineers in the process of route generation and design optimisation with the following objectives:

- To develop a Route Generator to allow the rapid evaluation of alternative designs and manufacturing strategies.
- To develop a Source Control System for the management of uncertainty and resolution of conflicts in information supplied by members of the simultaneous engineering team to aid design review in order to improve manufacturability.

Design reviews of candidate plans involves the evaluation of uncertain information from different sources, which may be contradictory on the subject of materials, functionality, quality and manufacturability. This requires an adaptation of the source control system to the application domain through the development class-based default source models for designers and process planners. Also a domain-specific set of heuristics needs to be generated to optimise the source control system for the problem domain. In addition, information about typical terminology and forms of argument typically used haveto be incorporated as well as the development of a specialised user interface.

The members of the simultaneous engineering team would be expected to input their analyses and these will be considered in conjunction with the output of the route generator. If there are conflicts between the sources (including the route generator), then a conflict resolution cycle will report problems and suggest a solution. The engineers can then enter their views and the system will go through a further cycle of conflict resolution. The intention is to make the users aware of problems and subjective elements and to mediate a solution. In addition the system will use the information from interaction with the users to revise the source models in order to improve its response for the future.

The source control approach could be particularly valuable when design reviews are conducted at the conceptual design stage, where information about product or process details is frequently contradictory or uncertain. The source control approach would deal with the human element, by managing uncertainty and resolving conflicts in information regarding the evaluation of routing solutions and the assessment of lead time implications of various design options. Together with a suitable route generator this should provide a basis for a consistent and integrated design support which should help substantially in the task of speed-up and optimisation of process planning and thereby increase productivity and profitability.

Through the complexity analysis we have shown that the SCM is computationally feasible. This example of the application to process planning shows that the source control approach is relevant to real problems faced by the commercial community and that it can make a considerable contribution in conjunction with other components in a decision support system. As part of the proposed project the SCM would initially be prototypes in a functional programming language before developing it in the same expert system environment used to implement the route generator. The AI tool is eventually intended to run on a SPARC workstation. The following diagram gives a general overview of the tool:

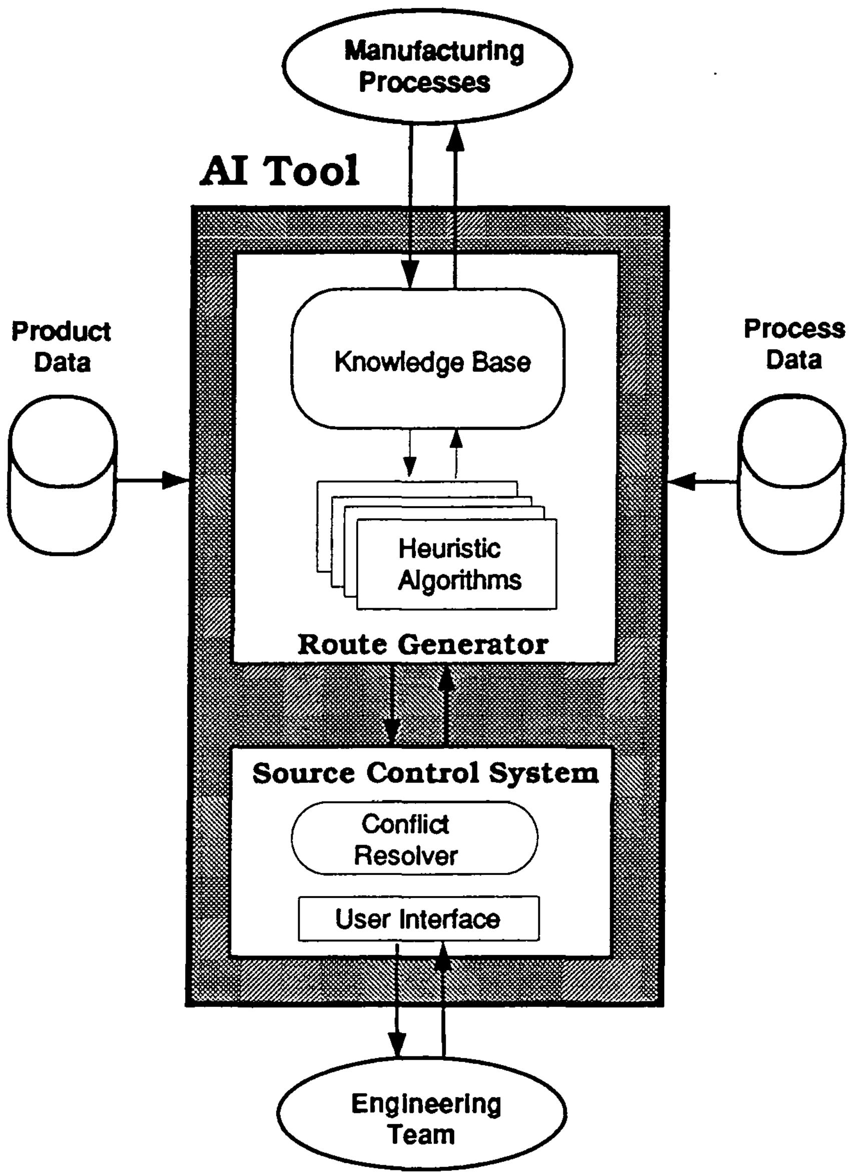


Figure 21: Overview of the AI Tool

6.12 Stability

Stability is concerned with the implications of refinements of a given system; whether these refinements are reversible and respect the existing framework, or whether they break out of it and thereby generate a fundamentally different system. A transformation can be shown to be stable if the relation between the original system and a new system is such that the result that would have been derived for a particular input in the original system can be reconstructed from the output of the new system.

To maintain stability may be of interest to someone wishing to carry out adaptive maintenance on the SCM, either retrospectively to establish whether a given transformation that was previously carried out maintains this property, or, alternatively, to establish whether anticipated changes satisfy the constraints or whether they generate a different system.

This property may not appear very significant at first, but supposing the original system has important properties which have perhaps been formally proven, it may be important to ensure that these properties are maintained in a given refinement of the system. To ensure that changes satisfy the constraints may therefore be highly desirable.

It will be obvious that if we have all the details of a given transformation, it will in most cases be a trivial task to carry out the reconstruction required to determine stability. What is more significant, is to be able to carry out the reconstruction without the specific details of the transformation.

This is a typical situation when maintenance is being carried out on a system where there is little or no documentation to say precisely what changes were introduced. The situation may be further aggravated if the system is sufficiently complex to make it very difficult to determine retrospectively the nature and location of changes. Supposing that the original system has important properties we want to preserve and the new system has additional properties we also want to preserve; therefore we do not want to start with the original system and completely rebuild the new system plus the functionality we want to add, but we want to determine whether the transformations was stable before embarking on further changes. We need to make a step forward rather than start again from the beginning. In this scenario one has to determine on the basis of limited information whether the transformation was stable or not.

In order to determine the stability of given transformations one therefore has to start with the output of the new system and with general knowledge about the nature of the transformation without necessarily knowing where exactly changes were made, and to reconstruct the result which the original system would have produced. If we do know the precise nature and location of changes then the task of determining whether the transformation is stable becomes trivially easy, but if we don't know anything about the transformation then the task is nigh impossible. Consequently, the important issue is to be able to operate with minimal information.

In this section we want to analyse the stability of systems potentially derived from the SCM. There are different ways in which the SCM could be changed. Although it is not our task to state what these potential changes are, we need to specify the conditions that modifications have to fulfil, to ensure that stability is maintained. It appears that there are four different types which are of particular interest for our purposes:

- partitioning
- deletion of rules
- addition of rules
- changing rules

By partitioning we mean that a given scale of possible values is increased by further dividing individual values to allow the system to operate at a higher level of resolution. Alternatively, one could modify a given system by removing or changing existing rules or adding new ones. These modifications are primitives and it is possible to break down more complex transformations into a series of steps of this kind.

In the following we first introduce a number of concepts and definitions. We then examine different types of transformation in general before analysing the SCM in this light.

6.12.1 Concepts and Definitions

In order to be able to analyse stability we need a number of definitions whose purpose will become clear as we progress in the analysis. For the purpose of our analysis we use the term machine to refer to a system as defined below. We also need to define stability. In the subsequent lemma we show that the property of stability is transitive. Due to the nature of the SCM, it is convenient to look at it in terms of a production system, where, in a chain of application of rules the input is transformed into the output. In the following, a number of definitions are given and are followed in the next section by an analysis of the properties of different types of partitioning.

Definition 1: - ***Machine***

A machine M can be represented as a nesting of functions

$$M = A(b_1(c_{11},\ldots,c_{1i},\ldots),\ldots,b_j(c_{j1},\ldots,c_{ji},\ldots),\ldots)$$

Given that the nesting is allowed to be of arbitrary depth, and can be recursive, any machine can be represented.

Given the definition of a machine we now can define stability as a property of a transformation. Stability as we have seen in the introduction is dependent on whether it is possible to reverse the transformation and to reproduce the result the original system would have produced by applying the inverse-transformation to the new system. More formally we define stability as:

Definition 2: - *Stability*

Given a set of machines $\mathcal{M}$: {Input $\rightarrow$ Output} with $M_1 \in \mathcal{M}$, and a transformation $T : \mathcal{M} \rightarrow \mathcal{M}$ such that $T(M_1) = M_2$, and the domain $D(M_1) = D(M_2)$ then T is *stable* **iff** there exists an $T^{-1} : \mathcal{M} \leftarrow \mathcal{M}$ such that $\forall y \in$ Input,

$$T^{-1}(M_2(y)) = M_1(y)$$

We can now establish the property of transitivity of stability for a series of transformations which are themselves stable.

Lemma 1: - *Transitivity of Stability*

Given T_1 , T_2 : $\mathcal{M} \rightarrow \mathcal{M}$, if $T_1(M_1) = M_2$ is *stable* and $T_2(M_2) = M_3$ is *stable*, then $T_2(T_1(M_1)) = M_3$ is *stable*.

Proof:

(1) Choose $y \in D(M_1)$ then there exists an T_1^{-1} such that:

$$T_1^{-1}(M_2(y)) = M_1(y)$$

(2) Choose $y' \in D(M_2)$ then there exists an T_2^{-1} such that:

$$T_2^{-1}(M_3(y')) = M_2(y')$$

By definition of T_1, $D(M_2) = D(M_1)$. Similarly $D(M_3) = D(M_2)$ and hence, without loss of generality, $y' = y$ and so (2) becomes $T_2^{-1}(M_3(y)) = M_2(y)$ substituting this in (1) gives the required result:

$$T_1^{-1}(T_2^{-1}(M_3(y))) = M_1(y)$$

Systems like the SCM deal with pieces of data which have a number of indices and where the analysis is largely dependent on the values of indices. It therefore seems appropriate to maintain the distinction between data and levels of indices in our analysis. We use the notion of labelled bodies where the bodies represent identifiable packages of data and the labels correspond to the levels of indices. These items are passed from one function of rule to another; they are the output of one function (or rules) and the input of another function of the system and are represented as internal values (*IntV*). The definitions are followed by two further technical definitions required later on.

Definition 3: - *Label*

A label is a member of a finite ordered set of tokens.

Definition 4: - *Body*

A body is a uniquely identifiable constant (e.g. o_i) belonging to a countable set.

Definition 5: - *IntV*

An internal value *IntV* is some value which is the input or output of the functions contained in a machine and consists of a label and a body, e.g. o_i^l.

Definition 6: - *Under*

Given some progression of functions in a machine, a function f under a function g, f *under* g is defined as follows:

$$f \textit{ under } g \equiv \exists\, o_i^x \in R(g)\,.\,(\, o_i^x \in D(f) \vee (\, \exists\, h\,.\, o_i^x \in D(h)\ \wedge f \textit{ under } h\,))).$$

Definition 7: - *Level*

We define *Level* to be the set of all functions immediately under a function f such that:

$$\exists o_i^x \,.\, (o_i^x \in R(f) \wedge o_i^x \in D(g) \rightarrow g \in Level(f))$$

Definition 8: - *Path Set*

The rule for inclusion in the set $PathSet(o_x^y)$ is:

$$\forall o_i^{y'} \, \forall f \, \forall g \,.\, (\, o_x^y \in R(f) \wedge g \textit{ under } f \wedge (\, o_i^{y'} \in R(g) \vee o_i^{y'} \in D(g)) \rightarrow o_i^{y'} \in PathSet(o_x^y))$$

6.12.2 Analysis of Partitioning

Given these initial definitions and the transitivity of stability we now can give definitions and analyse the properties of different types of partitioning. The SCM is essentially non-numeric; despite the fact that different values are ordered on a scale, it does not follow that they are by nature numeric. We therefore define and analyse partitioning in terms of ordered sets of labels rather than numbers.

Definition 9: - *Partition*

Given a set $\mathcal{L}$ of ordered labels $L_1,...L_n$, a partition of $\mathcal{L}$ is a set $\mathcal{L}'$ of ordered labels $L_1',...L_m'$ such that $m > n$ and if $L_i, L_j \in \mathcal{L}'$ then $i < j$.

In order for a partitioning to be stable it needs to be reversible. Therefore it is necessary that the partitioning only sub-divides existing labels and that the mapping from an old set of labels to a new set is such that no two different labels from the old set map onto the same new label. If that was the case, it would not be possible to determine which of the two old labels it originated from. In the following we give a definition of a consistent partitioning, followed by a de-partition to reverse a given partitioning.

Definition 10: - *Consistent Partitioning*

Given two machines M_1 and M_2, a partitioning is *consistent* **iff**

$$\forall o_1^y, o_2^{y'} \in IntVM_1 \,.\, (\, y \neq y' \rightarrow (\forall x \in LabelM_2 \,.\, o_1^x \in IntVM_2 \rightarrow o_2^x \notin IntVM_2)).$$

Definition 11: - *De-Partition*

Given $\mathcal{L}', \mathcal{L} \subseteq Label$ a de-partition P^{-1} is a function $\mathcal{L}' \rightarrow \mathcal{L}$ where $\mathcal{L}' = L_1',\ldots, L_m' \wedge \mathcal{L} = L_1,\ldots, L_n$ such that:

1) $m \geq n$
2) $L_i', L_j' \in \mathcal{L} \rightarrow i < j$ (order preserving)
3) $\forall x' \,.\, x' \in \mathcal{L}' \,.\, \exists! x \,.\, x \in \mathcal{L}(P^{-1}(x') = x)$ (de-partition of a consistent partition)

Given a consistent partitioning, there are different ways in which it can be applied and this has an effect on whether a transformation consisting of such a partitioning preserves stability.

Definition 12: - *Path Consistency*

A consistent partitioning is said to be *path consistent* **if** $\forall o_1^y, o_2^{y'} \in IntVM_1 \,.\, o_2^{y'} \in PathSet(o_1^y) \,.\, (\, y = y' \rightarrow (\forall x \in LabelM_2 \,.\, (\, o_1^x \in IntVM_2 \rightarrow (\, o_2^z \in PathSet(o_1^x) \rightarrow z = x)).$

From the definitions we have given so far, we can now prove that a de-partitioning can actually be generated for a path consistent partitioning. This is an important step for the re-construction process which requires that the transformation can be reversed.

Lemma 2: - *Existence of a De-Partitioning*

If T is a transformation consisting only of a path consistent partitioning P, then we can build a de-partitioning P^{-1}:

Proof:

We need to show that $\forall o^y \in IntVM_2$, if $\exists x \in LabelM_1 \wedge P(x) = y \cup \ldots, \rightarrow P^{-1}(o^y) = o^x$. If, without loss of generality, we choose an o^y then:

1) **If there is no $x \in LabelM_1$ such that $P(x) = y$ then proven.**
2) **If $\exists\, x \in LabelM_1$ such that $P(x) = y$ then both**
   a) **x is unique. Proof: Assume $x_1 \neq x$ such that $P(x_1) = y$ then the partitioning could not have been consistent. It is assumed that the partitioning was consistent and therefore we have a contradiction.**
   b) **$o^x \in IntVM_1$. Proof: Assume $o^z \in IntVM_1, z \neq x$ for $T(o^z) = o^y$ then $P(z) = y$ and so from uniqueness we have a contradiction.**

A consistent partitioning which has not been applied to the whole path, can cause a mismatch between a labelled body which is the output of one rule and the input of the next rule in the path and thereby cause processing to be stopped short, since the next rule will not recognise the different label as a legal input. Rules in a path are connected by the output of one rule being the input to the next. If the output of one rule is no longer the input to the next rule, the path is broken at that point as it is impossible to continue the path. A transformation which consists only of a path consistent partitioning cannot cause this situation to arise and in the following we show that path consistent partitioning does not result in broken paths.

Lemma 3:

A transformation T consisting only of a path consistent partitioning P does not cause a broken path.

Proof:

Let $T(M_1) = M_2$ where T is a path consistent partitioning, then $\exists o_i^y, o_j^x \in IntVM_1 . o_i^y \in PathSet(o_j^x) \wedge \exists o_i^{y'}, o_j^{x'} \in IntVM_2 . o_i^{y'} \notin PathSet(o_j^{x'})$.

We can prove that P does not cause a broken path by induction on n, where n is the number of levels in g *under* f, $o_i^x \in R(f)$ and $o_j^y \in R(g)$.

1) The base case. $n = 1$. Then $g(o_i^x) = o_j^y$ and $T(g(o_i^x)) = o_l^z$ where $l \neq j$. As this is the base case, there cannot have been a broken path before. Path consistent partitioning is a subset of partitioning, and by definition of partition $\forall x\ T(o_x^y) = o_x^z$. So $l = j$ and therefore we have a contradiction.
2) Induction Step. Assume, by the induction hypothesis, that there are no broken paths up to a level n below f, then $\exists\ h_n\ \exists\ o_l^z \in IntVM_1 . o_l^z \in R(h_n) \wedge g(o_l^z) = o_j^y$. By the induction hypothesis there is no broken path upto h_n.

   Assume a break, then $\forall o_l^{z'} \in IntVM_2 . o_l^{z'} \in R(h_n) \rightarrow g(o_l^{z'}) \neq o_j^{z''}$ for some z''. But since $g(o_l^{z'}) = o_j^{z''}$ and the only possible change in g is the label of the $IntV$ (by path consistency) by definition of T. Therefore these labels may only have been changed consistently, and it follows that $g(o_l^{z'}) \neq o_j^{z''}$ is a contradiction.

Lemma 4:

A transformation T consisting only of a path consistent partitioning P does not add a new path.

Proof:

Let $T(M_1) = M_2$ where T is a path consistent partitioning. Assume there is an added path then:

$\exists o_i^{y'}, o_j^{x'} \in IntVM_2 . o_i^{y'} \in PathSet(o_j^{x'}) \wedge$

$\exists o_i^y, o_j^x . o_j^x \in IntVM_1 \wedge o_i^y \notin IntVM_1$.

Without loss of generality let g *under* f, $o_i^x \in R(f) \wedge o_j^y \in R(g)$. Then there are two possible ways in which a path could have been added:

a) Addition of a Body: Then $g(o_i^x) = o_j^y$ and $T(g(o_i^x)) = o_l^z$ where $l \neq j$. Path consistent partitioning is a subset of partitioning, and by definition of partitioning:

$$\forall x \; T(o_x^y) = o_x^z$$

So $l = j$, and therefore we have a contradiction.

b) Addition of a New Rule: Assume an added rule, then $\exists h \in M_2 \wedge h \notin M_1 \,.\, \forall o_l^z \in IntVM_2 \,.\, o_l^z \in R(h) \rightarrow h(o_l^z) = o_m^r \wedge T^{-1}(o_m^r) \notin IntVM_1$.

By assumption T consisted only of a path consistent partitioning and therefore the path consistent partitioning must have added a new rule. But since $P : \mathcal{L} \rightarrow \mathcal{L}$ the only possible change is the label of $IntV$ and it follows that no rule could have been added.

Given our definitions and proofs that a path consistent partitioning does not cause a broken path or add a new path, and that we can build a de-partition, we can now show that path consistent partitioning preserves stability.

Theorem 1: *Stability of Path Consistent Partitioning*

If T is a transformation consisting only of a path consistent partitioning then $T(M_1)$ is stable.

Proof:

This requires that we can prove that there exists an T^{-1} such that:

$$T^{-1}(M_2(y)) = M_1(y)$$

Without loss of generality we choose $M_1(y) = o_i^x$ and $M_2(y) = o_j^z$

1) We need to prove that $i = j$. There are two conditions for this to be true.
   a) that there are no broken paths - proven by Lemma 3.
   b) that there are no added paths - proven by Lemma 4.

2) We also need to prove that one can get x from z by de-partitioning - proven by Lemma2.

Corollary 1:

If T is a transformation consisting only of a path consistent partitioning P then we can build a T^{-1} such that $T^{-1}(M_2(y)) = M_1(y)$.

Proof:

By Theorem 1 $M_2(y)$ can differ from $M_1(y)$ at most in the labels and so by Lemma 2 we can de-partition it. Hence T^{-1} is the de-partitioning P^{-1}.

In order for a transformation to be stable, it is necessary that the precise result of the original system can be re-constructed. In addition, we define another interesting property, closely associated with stability, which applies when the re-construction does not yield the precise result, but a result which falls into a specified neighbourhood. We call this *compatibility*.

Definition 13: *Compatibility*

Given $o_i^x, o_j^y \in IntV$ and a metric $N :: (IntV, IntV) \rightarrow R$, o_i^x is compatible with o_j^y ($o_i^x \; comp_n \; o_j^y$) **iff** $N(o_i^x, o_j^y) < n$ where n is the acceptable neighbourhood measure inbuilt in the system.

Definition 14: *Path*

There is a path $P(o_i^x, o_j^y)$ in M **iff**

$$\exists g_1 \dots g_n \in M \,.\, g_1 \dots g_n(o_i^x) = o_j^y$$

Definition 15: *Body Path*

There is a *body path* $Bp(o_i^x, o_j^y)$ in M **iff**

$$\exists \, g_1 \dots g_n \;\; M \,.\, \exists \, x \,, y \in LabelM \,.\, g_1 \dots g_n(o_i^x) = o_j^y.$$

From these definitions and the definition of partitioning we can make an important observation about the properties of partitioning.

Observation 1:

Path Consistent Partitioning cannot add any new body paths. Path consistent partitioning is a subset of consistent partitioning and consistent partitioning only affects the labels and not the bodies and therefore no new bodies can have been added.

When, as we shall see later on, we are dealing with a result from the new machine and there are several places where this result could have been produced (given that we do not know where exactly the transformation changed the original system), one may be able to infer this information by backtracking, by moving back in the system. This requires that one can reconstruct the path leading up to that point.

Definition 16: *Path Above*

$\forall y , x \,.\, x \in IntVM , y \in PathAbove(x)$ **iff**

$$(\exists\, g \in M \,.\, g(y) = x) \;\vee (\exists\, f \in M \wedge \exists\, z \in IntVM \,.\, F(z) = x \wedge y \in PathAbove(z))$$

Lemma 5:

$\forall x \in IntVM$ we can build $PathAbove(x)$

Proof: by cases

1) $x \in InputM$, then x is at the top of the tree and therefore $PathAbove = \{\,\}$
2) $x \notin InputM$, then x must occur somewhere in the tree representation of M. There are three steps we need to follow to build the $PathAbove(x)$:
    1) Pick all $g \,.\, x \in R(g)$, then
    2) $\forall\, y \in D(g) \,.\, g(y) = x \rightarrow y \in PathAbove(x)$.

3) Repeat steps 1) by picking $f \,.\, y \in R(f)$ and step 2) by picking $z \,.\, f(z) = y$ until reaching case 1).

Theorem 2: *Non Path-Consistent Partitioning*

Assume a transformation $T : \mathcal{M} \rightarrow \mathcal{M}$ is a consistent partitioning, but is not path consistent, then:

i) By observation 1 we have not created any new body paths.
ii) We may, or may not have broken a path. Since we are not given the details of T, we must assume the worst case i.e. that a broken path has occurred.

Let $M_2(y) = y'$ where y is unknown and y' is known:

1) There is only one possible $y' \in IntVM_2$ and $y' \in OutputM_1$. By de-partitioning the original result is obtained and T is *stable*.
2) There are several $y' \in IntVM_2 \wedge y' \in OutputM_1$. Again the original result can be obtained by de-partitioning and T is *stable*.
3) There is only one possible $y' \in IntVM_2$ and $y' \notin OutputM_1$.
   3.1) No additional information is required and the computation can be continued in M_1 and the original result be generated. T is *stable*.
   3.2) Additional information is required, which can be uniquely inferred by going backwards and the original result be produced and the transformation is *stable*.
   3.3) Additional information is required but can not be uniquely inferred. In this case one has to continue processing all possible paths in the original machine to get a set of results:
   a) all results are identical and therefore T is *stable*.
   b) all results are *compatible* and the transformation is *compatible*
   c) otherwise *unstable*.
4) There are many $y' \in IntVM_2$ and several are internal. This situation is equivalent to 3.3) above.

Proof:

1) The original result can be obtained by de-partitioning and can be checked by simple observation.
2) Again, the result only needs to be de-partitioned to immediately get the original result and can be checked by simple observation.

3.1) We apply M_1 to the result of the de-partitioning of y' i.e. $M_1(P^{-1}(y'))$. Either there are no further choice points, or if there are, we have the information to decide and get to a unique result. By definition of stability we can produce the original result and the transformation T is *stable*.

3.2) From our assumption all information is available in the *PathAbove*. From Lemma 5, it is accessible. As the information needed is unique (by assumption) it is possible to continue the computation to a unique value and therefore by the definition of stability the transformation T is *stable*.

3.3) From Lemma 5 we can access all the inferable information from the path above. As some information is either not available or not unique (by assumption) we need to follow all continuation paths until we can go no further. This will produce a set of values $y_1 ,..., y_n$ and there are three possible situations:

a) They are all equal and by the definition of stability T is *stable*.
b) They are all compatible; then T is *compatible* by the definition of compatibility.
c) otherwise, when the results are outside the compatible neighbourhood R is *unstable*.

4) If there are several internal possibilities we follow all continuation paths. This will produce a set of values $y_1 ,..., y_n$:
   1) They are all equal and by the definition of stability T is *stable*.
   2) They are all compatible; then T is *compatible* by the definition of compatibility
   3) otherwise, when the results are outside the compatible neighbourhood T is *unstable*.

6.12.2 Deletion of Rules

While partitioning only affects the labels, leaving the bodies and therefore the structure of the system intact, deletion affects the actual structure. Through the deletion of a rule, the affected path is broken at that point and cases which would have come down that path would stop. However, the effect is identical as both cause a break of the path albeit in different ways. While consistent partitioning which is not path consistent causes a break by a mismatch of labels previously connecting the output of one rule to the input of the next, deletion of a rule has the same effect by not being able to generate a particular IntV required to maintain the connection.

Theorem 3: *Deletion without Partitioning*

Assume a transformation $T : M \rightarrow M$ consists only of a deletion of a rule without a partitioning. This is equivalent to a consistent partitioning which is not path consistent (without the need to de-partition).

If a deletion of a rule without a partitioning is equivalent to a consistent partitioning which is not path consistent, then the results of Theorem 2 follow.

Proof: by observation

A partitioning which is not path consistent, causes a break of the path by partitioning the labels of one function but not the labels of the next function in the path. A deletion of a rule causes a break in the path by a missing *IntV* no longer generated by the deleted rule. After reinstatement of the deleted rule the path is reconnected and the same situation arises as in Theorem 2, except that no de-partitioning is required.

Similar to an ordinary deletion is a deletion which is combined with a consistent partitioning and we can immediately state the following theorem:

Theorem 4: *Deletion with Consistent Partitioning*

Assume a transformation $T: M \rightarrow M$ consisting of a deletion and a consistent partitioning such that $T(M_1(y)) = y'$, then the situation is equivalent to a consistent partitioning which is not path consistent.

Proof: - by observation

Once y' has been de-partitioned all that is left is a deleted rule (in the worst case).

If $P^{-1}(y))$ is put into M_1 then we have the same situation as in Theorem 2 section 3.3.

6.12.3 Addition of Rules

Transformations based only on partitioning and deletion cannot add anything new to the paths of the original system. Only the addition or change of rules has this potential. There is also the possibility that the addition of a new rule could cause non-determinism, as can be seen in the following Lemma.

Lemma 6: *Addition of Rules*

Assume a transformation $T: M \rightarrow M$ consists only of the addition of one rule, then either the rule has no effect, or the rule was added at the end of the tree (assuming the system is deterministic).

Proof: by contradiction

Suppose that a rule was added, but not at the end of the tree, then there is an $f \in M_2$ such that:

i) $f \notin M_1$ (there is a new rule)

ii) $\exists g \in M_1 \,.\, D(f) \subseteq R(g)$ (the rule is connected)

iii) $\exists h \in M_1 \,.\, D(f) \supseteq R(g)$ (the new rule is not connected at the end)

It follows that:

a) $h \notin M_2$, thus a deletion must have occurred and that is a contradiction.

b) f and h have the same input at some place in M_2 and therefore create a non-determinism. That is a contradiction.

Observation 2:

If there is more than one addition of rules and the new rules have been added in different places, the analysis has to be carried out for each addition separately.

If all the additions are at the same point (i.e. a chain of added rules), then the reconstruction can be carried out in the same process.

Theorem 4: *Addition of Rules*

Assume a transformation $T : M \to M$ which consists only of an addition of a rule, then there are two possibilities.

1) M_2 is deterministic
2) M_2 is not deterministic

If M_2 is deterministic, then there are two possibilities:

1.1) The rule has no effect
1.2) The rule was added at the end of the tree

In the worst case, given a new rule $f : x \to y$, then $\forall x \in X \,.\, \exists y \in X \,.\, T(x) = y$ and x *comp* y then T is *compatible*, otherwise T is *unstable*.

Proof:

1.1) If there is no change, then by definition of stability T is *stable*.

1.2) By Lemma 6, f was added at the end. This means that $D(f) \subseteq R(M_1)$. Hence, if $D(f)$ is compatible with $R(M_2)$ and T is therefore *compatible*, otherwise it is *unstable*.

Observation 3:

By assumption $R(M_1)$ - $D(T) = R(M_2)$ - $R(f)$. If $R(f)$ *comp* $R(T)$, then M_1 compatible with M_2.

Definition 17: *Range Compatibility*

$D(f)$ *comp* $R(f) \equiv \forall x \,.\, x \in D(f) \,.\, \exists y \in R(f) \,.\, f(x) = y \cap x$ *comp* y

6.12.4 Change of Rules

Finally, it is possible to change an existing rule. This is equivalent to a deletion and an addition at the same place. The mere addition of a new rule parallel to an existing one would either have no effect or would cause non-determinism. The way to proceed therefore is by deletion of an existing rule and addition of a new one which, for example might be used to single out a case thought of as being of special significance and requiring special treatment.

Theorem 5: *Change of Rule* (Deletion and Addition)

Given a transformation $T : \mathcal{M} \rightarrow \mathcal{M}$, such that $\exists f \,.\, f \in M_2 \wedge f \notin M_1 \wedge \exists g \,.\, g \in M_1 \wedge g \notin M_2$ then we can have one of two different cases:

1) $D(f) = D(g)$ and the new rule replaces the old.
2) $D(f) \neq D(g)$ and the new rule was added in a different place.

Proof: by cases

1) The new rule replaces the old and there are three different cases:
   a) g is last in the chain: then f is last in the chain. If $D(f)$ *comp* $R(T)$, then by the assumption of the definition of compatibility $D(g)$ *comp* $R(f)$. Since we know that g is last in M_1 then it follows that if $\forall h \in M_1 \,.\, R(h) \subseteq R(M_1) \rightarrow D(h)$ *comp* $R(h)$, then T is *compatible.*

b) g is followed by h and $R(f) \supseteq D(h)$ (g replaces h in the same slot):
If $D(f)$ *comp* $R(g) \wedge D(h)$ *comp* $R(h)$ then T is *compatible*.

c) otherwise: (like a deletion) If $R(g)$ is not compatible with $R(f)$ then T is *unstable*. If they are compatible, then the transformation may be *compatible* according to the rules from Theorem 2.

2) There are two possibilities: Either f is disconnected and never executed (by the conditions of Theorem 2), or f is added at the end (by Theorems 2 and 3). Since the only other "free" place is that left by g, T is *stable* or *compatible* under the same conditions as Theorem 2 and 3 applied successively.

6.12.6 Stability and the SCM

In the preceding sections we have analysed different types of modifications and their stability. The modifications discussed are primitives and by a combination of a number of these, more complex modifications can be achieved. The results of the analysis can be used to analyse retrospectively the stability of transformations already carried out on a given system. In addition, someone intending to make modifications can also use the results to check whether prospective changes would maintain stability.

In the following we look at how the results of the analysis of the preceding sections apply to the SCM. Before discussing the potential stability of the various types of modification, we first need to explain how the basic concepts of the stability analysis map onto the SCM.

Mapping the Concepts

At the beginning of the section we introduced a number of definitions which we have to map onto the SCM. These include the following:

- *Machine:* The decision trees in Appendices A and B describe the architecture of the SCM and collectively constitute the machine.

- *Functions:* The nodes of the decision trees correspond to functions or rules. There are a number of ways in which the SCM could be implemented. As it is fundamentally a system of rules, the functions are considered equivalent to rules.
- *IntV:* These are the internal values which are generated by one function and passed to the next. In the SCM the information that is passed from one rule to the next consists of three items, namely the event, the source model and the case.
- *Label:* The label corresponds to the level of a given index. In the SCM there are a number of indices, although only one of them may be required by a particular rule. The label in the SCM is therefore not a single but a composite item.
- *Body:* In our analysis we distinguished bodies and labels to maintain a separation between the variable part of the information and the constant by which the item can uniquely be identified. It is the information contained in the label which affects the decision process, but in order to ensure a proper sequence of processing, unique body names are required to connect the output of one rule to the input of the next.
- *PathSet:* The *PathSet* describes a path as the set of all IntV which would be generated along that path by the functions of the machine. In the decision trees of the SCM a path is a direct connection between a given leaf of a tree and its root node and beyond the component the decision tree is part of the path continues through the whole SCM, thus representing one of the possible outcomes of the system.
- *BodyPath:* While the PathSet describes one of the possible paths, where the functions of the system are able to generate all the *IntV* in the *PathSet*, the *BodyPath* is less specific. Thus, when we have a partitioning which has not been consistently applied, thereby causing a broken path, the BodyPath still exists as the functions of the machine are still taking the same bodies as their input, although through a mismatch of labels processing along a particular path would stop at the first point where one of the functions was unable to process the new labels. The decision trees

do not show the details of input and output of the various rules, but only indicate by a connecting line that the output of one function is connected to the next function along the path. This implies that there is a connection and in essence that there is a *PathSet*. Consequently, it would be necessary to indicate a break in the diagrams to show a *BodyPath* which does not constitute a *PathSet*. (Note: There is an equivalent *BodyPath* for each *PathSet*, but not every *BodyPath* has an equivalent *PathSet*.)

- *PathAbove:* Given a particular *IntV* the *PathAbove* is the *PathSet* from the *IntV* to the *Input*. In terms of the SCM, given the output of a particular rule the *PathAbove* is the path that leads from that output back to the root node.

In a rule-based system there are no hard connections between rules as in a nested function call for example, but the connections are between rules by virtue of the output of one rule being the unique input of only one other rule in the system. Each *IntV* therefore has to be a uniquely identifiable instance of a schema (or abstract data type) containing the data of a labelled body. If the SCM was implemented as a rule-based system, then unique body names are required, so that rules which are designed to operate one after the other, each can recognise the *IntV* which is the output of the previous rule as uniquely being their domain, and processing can be carried out in an organised fashion.

Given that a rule has a domain and a range and maps from *IntV* to *IntV*, and that rules are connected to each other by virtue of the connection between the range of one rule being the domain of the next, a chain of these connections constitute a path. In the decision trees of the SCM the nodes are the rules and the connection above signifies the domain and the named branches below are the range of the rule.

Stability requires that the original result is reproducible. In terms of the SCM this means that the path the information would have originally have taken through the system can be reconstructed. The requirement for *compatibility* is that the result is within a specific neighbourhood of the original result. The SCM uses four different forms of modification, namely to *reject* the information, to *reduce* belief to a lower level, to *modify* belief dependent on the ability or to *accept* the information as given. In addition there are a number of modifications in the conflict resolution component which are equivalent but deal with the formation of a single belief based on two pieces of information rather than a single one. Thus, *ignoring* the new information implies *rejection*, *merging* means *accepting*, *adding* the new source means *modification* and *levelling* means *reducing*. *Rejecting*, *reducing*, *modifying* and *accepting* are in decreasing order of the severity with which they recommend to reduce the belief in the information as indicated by the source. In terms of the SCM this means that compatibility would constrain the result to the neighbouring type of modification. For example, if the original result was *modifying*, then *accepting* or *reducing* would be compatible but *rejection* would not. Given that the forms of modification were refined in conjunction with a partitioning one could restrain the limits of compatibility to a simple neighbourhood within the new order of modifications.

Partitioning

In contrast to the analysis of stability, where we dealt with simple labels, the SCM has a number of indices and therefore there are a number of labels, only one or two of which may be relevant to a given rule. A given partitioning will obviously apply only to one index amongst the number of indices in the SCM. For a rule to be able to handle the new partitioning it will be necessary to change the legal set of labels. As not every rule will use the newly partitioned index, it will depend on how the system is implemented where an partitioning which is not path consistent causes the break of a path. If the data type of the index was something like an enumerated type, then any rule which was not changed to accept the new enumerated type would stop the analysis at that point. If the data types were less discriminatory, then the break would happen only at the next rule which uses the index, but is not able to work with the new labels.

Partitioning therefore involves a change to the legal set of labels. Path Consistent Partitioning means that this has been done for all rules along a path, and for all rules in the system, if partitioning is globally applied. Supposing we changed the ability index from 5 labels (*N,L,M,H,T*) to 10 by subdividing each label into a upper and lower division (*N+,N-,L+,L-,M+,M-,H+,H-,T+,T-*), then there would not be any visible change to the structure of the decision trees concerned. This would however mean, that if the modification was globally applied, each rule would now recognise the new labels, even though they would still be in the same bands as before. Thus the rule Abil? would now put *N+,N-,L+* and *L-* into the Low band as opposed to *N* and *L* before and similarly for the other bands. Given that the partitioning was not applied path consistently, given the case is affected by the change, we would get an *IntV* which is the output of the last rule, before the point where the path is broken, as the next rule is not able to recognize this *IntV* as being in its domain.

If the partitioning is consistent, then this implies a refinement which respects the old divisions, by refining within them, rather than across the boundaries. This means that processing will still take the same path, albeit with more refined labels.

If we are dealing with a transformation consisting only of a path consistent partitioning then we have shown in Theorem 1 that the transformation is stable. This is dependent on the transformation not having caused a break in the path and the transformation not having added new paths. In Lemma 3 we have shown that a path consistent partitioning cannot have caused broken paths and in Lemma 4 we demonstrate that no new paths can have been added either. The reconstruction is however dependent on being able to successfully reverse the partitioning and in Lemma 2 we show that such a de-partitioning can be generated. In addition, through Lemma 1 we are also able to establish that if a series of transformations are carried out, each one of which is stable individually, then they are stable together. This implies, for example, that a series of path consistent partitionings could be carried out on the SCM (whether globally applied or only on selected paths), without jeopardising stability.

If we are dealing with a partitioning in the SCM, which is not path consistent, then it depends on whether the case can uniquely be identified and whether it is possible to successfully reconstruct the original result. If a particular case, or a number of candidate cases were outputs of the original system, then they were not affected by a broken path. Therefore the original result can immediately be reconstructed by de-partitioning and the transformation is stable. If the result is an internal *IntV*, then stability is dependent on whether there is only a unique situation where this *IntV* could have occurred. If the SCM was implemented as a rule based system, the *IntV* would be unique for each component and could therefore immediately be identified. As shown in the decision trees of the Control Regime, there are different paths the analysis can take and a given component will be used in a number of paths, thus creating a number of situations in which such an *IntV* could have occurred. From the information contained in the case, it is however possible to reconstruct the path the analysis took and thereby correctly identify the situation. Supposing a partitioning, which was not path consistent, caused a broken path in the Conflict Resolution Module, then there are only two paths upto that point (path C and D or G and H which are not distinguished at that point yet). From the information contained in the case, it can be reconstructed whether there was an enquiry, or alternatively what the result of the Importance Analysis was, in order to determine whether we are dealing with case C and D or case G and H. The SCM would therefore also be stable for cases of non-path consistent partitioning.

Deletion

Deletions of rules, like non-path-consistent partitioning, cause a break of a path. The difference is that in the case of a deletion of a rule the break is caused by a missing body, while in the case of non-path-consistent partitioning the break is due to a mismatch of labels. In either case the occurrence of this situation is likely to be due to a mistake by the maintenance programmer rather than deliberate. It results in the generation of an *IntV* which has only partially been processed by the system and is not designed to be one of the usual outputs.

A deletion may only affect some of all possible cases. An example of this would be the deletion of one or more of the rules ChEnv? in the Conflict Resolution Module. Not all cases include a conflict resolution and not all those that do, necessarily get to the point where the rule ChEnv? would have been used.

Theorem 3 shows that a transformation which consists only of a deletion is equivalent to a consistent partitioning which is not path-consistent. Both cases cause a break in the affected path and the stability of the transformation is subject to the same conditions. As Theorem 2 shows, the success of the reconstruction is dependent on the uniqueness of the result. As we have seen, the path the analysis took in terms of the Control Regime (without going into the details of the components involved) can be reconstructed from the case. As cases affected by a deletion result in an IntV which is not an output of the system, one can determine in which component and where the IntV came from and resume the analysis with the original SCM. From Observation 1 we know that no new paths could have been added to the SCM through deletion and the IntV in question is therefore one which would have been produced in the original SCM and is the domain of one of the rules of the original system. By continuation of the analysis the original result can be reproduced.

The case is similar if the deletion was combined with a partitioning. Again, once de-partitioning as been accomplished, we are left with one of a number of possible situation (as described in section 3 of Theorem 2). This means that the transformation is stable if the case can be uniquely identified, or if the different cases have the same result. Otherwise the result may still be compatible if it falls within the specified neighbourhood. The only feature which is different from a simple deletion, is the added partitioning, and therefore it requires a de-partitioning as well as to reinstate the missing rule to carry out the reconstruction.

Addition

From Theorem 4 on Addition we can see that an addition on its own would have to occur at the end of the tree. Otherwise the added rule would either be disconnected or cause a non-determinism. It is difficult to imagine how the latter two cases could have been an intended modification. To introduce a new rule in the middle of a tree requires changes to other rules and will be discussed in the next section. Supposing we are not dealing with these cases, the only possibility is that the rule was added at the end of the tree.

In the SCM such an addition would have to occur in the Control Regime as the IntV which are at the bottom of the trees of the various components like Information Evaluation, Conflict Resolution and so on are signals for the Control Regime that processing in the respective component has finished and that the Control Regime has to decide what to do next with the case. In order to introduce a rule at the end of one of the components therefore requires changing the last rule to produce an *IntV* with a different identifier so that the new rule can safely be introduced. This actually requires a change as well as the addition but we shall ignore this technical point for the moment (this will be discussed in the following section). Supposing a new rule was introduced at the end, then it depends on whether the result is different from the one that would have been obtained from the original system. If the result is the same, then the transformation is stable, but otherwise it will depend on whether the result is close enough to make the transformation at least compatible. With deletion processing will stop at the point where the deletion took place thereby indicating the location of the deleted rule. An added rule is more difficult to locate as processing is not stopped but may carry on to completion. As indicated earlier, we may not have the details of the location of the changes, due to a lack of documentation about the transformations. To successfully carry out a reconstruction one has to retrace to see whether the case was affected by this change and to find the appropriate point to continue the computation in the original system.

If, for example, the new rule was added only at a few points in Enquiry 1, the change may not affect the majority of cases, but only a small number of them. In this case we are able to say that the transformation is stable for all cases which do not go through Enquiry 1. The success fo the reconstruction depends on how much information is still available at that point. If the information about the case is still accessible one can determine whether the analysis included Enquiry 1, and if so whether the result of Enquiry 1 might subsequently have been superseded by Conflict Resolution. However, in order for the transformation to be stable for all cases one would have to be able to restart the analysis, and that requires that the missing information can at least be inferred. If the missing information can uniquely be inferred, then the transformation can be shown to be stable. Otherwise, we are in the same situation as in Theorem 2 point 3.3, and it is dependent on the result of pursuing all candidate cases, whether the transformation turns out to be stable, compatible or unstable.

Another approach is to consider the range of the entire tree. If the addition of the new rule caused a change which is within the limits of compatibility of the range of the original tree, then the transformation is compatible. Supposing that the addition was a done in conjunction with partitionings on the strength of belief and belief modification functions, to create 10 values instead of the existing 5 and a lenient and severe modification function for each of the existing ones e.g. modify-l and modify-s) to take advantage of the further refinement. The transformation is stable if it keeps within the limits of the original divisions or compatible if it conforms to the conditions above. As the SCM is composed of a number of components, compatibility and stability only hold for the whole SCM if they do not have a knock-on effect and cause a different behaviour later on. In the SCM, the actions of the Important Analysis and other modules are not based on the particular modification recommended, but on the endorsements of whether there are problems with the information. Thus additions which cause small changes (like those mentioned above) in one module would not have a knock-on effect if they do not affect the problem-endorsements.

Change of Rules

From the previous section it can be seen that it is indeed difficult to add a new rule without actually changing any of the other rules involved. The effect on stability of changing rules is very similar to that of addition of rules. A change of rule is essentially a deletion and an addition combined. As can be seen from Theorem 5, there are those changes of rules where the addition was in a different place than the deletion. This case has already been covered in previous sections.

We therefore consider the situation where deletion and addition are in the same place. Theorem 5 distinguishes three cases, namely a change of a rule at the bottom of the tree, change of a rule in the middle of a tree and a change where the new rule remains disconnected. The latter is identical to a deletion as discussed in previous sections.

Given that the change was at the bottom of the tree, then this is similar to addition of a new rule as discussed above. If, for example, one wanted to change the Abil? rules which have two branches, at the left side of the decision tree for Information Evaluation so as to make them uniform with the other incidences which have three branches by making a division between medium and high, then it depends on whether the new medium and high branches both have the same result as in the original system, in which case the change was stable. Alternatively, if the range of the new rules are compatible with the old rules then the transformation would at least be compatible, but unstable otherwise.

Changing rules in the middle of a tree can be more difficult as there are specific sub-trees connected to the range of the old rule. The old rule has a specific number of branches and if the change introduces more branches, then whole new sub-trees may be required. Supposing, in the Information Evaluation module, one wanted to change the Trst? rule to split the medium-high into medium and high, then if a second, identical sub-trees was introduced, the result would be the same and the transformation would be stable. Alternatively, if the ranges of the sub-trees are within the limits of compatibility, the transformation would at least be compatible.

6.13 Conclusion

In previous chapters we have introduced the principles of the SCM, as well as providing a specification and design. Two issues which are particularly of interest to potential implementers of the SCM are its complexity and stability.

There are a number of ways in which the complexity of algorithms can be assessed, depending on the type and properties of the algorithm. We found that given the multi-threadedness of the SCM and the variation of its behaviour dependent on a number of factors a best-, worst- and average case analyses over the range of all possible cases appeared to be the most appropriate approach. Subsequently, we carried out the analysis based on the decision trees provided in the appendices. As the precise cost of the SCM will be partially dependent on the particular implementation and application environment we established the cost of the SCM based on three categories to highlight the potential impact of these factors. In the subsequent we also looked at the cost of making changes.

To give a sense as to how the SCM can be used in practice in commercial applications, we also briefly introduced a current proposal for the application of the SCM as part of a decision support system for production management. This serves to highlight both its potential to provide solutions to real-life problems.

There are in fact two different types of cost involved in making changes to the SCM we presented. Firstly, there is the cost in terms of complexity and the computational cost involved in these changes and secondly there is the potential cost of stability, which was the subject of the last part of the chapter. There are in fact few systems which will remain unchanged over their entire life-span; most systems undergo periodic changes to adapt them to new requirements. It is also not uncommon to find that changes are often badly documented, and that can cause considerable problems for maintenance programmers. Particularly when a systems in its original specification had a number of important (perhaps formally established) properties which need to be preserved in the process of refining the system, an analysis of stability will be relevant to ensure that the link between the original and the new system was maintained. We gave a formal analysis of stability by looking at a number of different forms of modification. As the SCM is essentially rule-based and processing is guided by the levels of various indices, we carried out an analysis geared towards rule-based systems and also taking into account the emphasis on indices by an analysis based on labels. We analysed partitioning which is particularly relevant to the refinement of indices as well as various types of change to the rules, and applied it to the SCM to demonstrate the capacity of the SCM to cope with changes within the limits of stability.

In the complexity analysis we showed the SCM to be computationally feasible and relevant and in the stability analysis we showed its capacity for change. These are important criteria for the selection of candidate algorithms in the process of serious system development and our analysis showed the SCM to be well placed to face these demands. In addition, the discussion of application of the SCM showed one way in which the SCM can be put to practical use in AI systems to provide solutions to current problems.

Chapter 7

Conclusion and Further Work

7.1 The Problem

Human agents are continually faced with the need to take decisions in response to information they receive from their environment. There is a considerable amount of uncertainty involved in that information, and in order to respond adequately to the demand for decisions, it is essential to analyse and assess this information to decide whether that information can be believed. Human agents are very successful in dealing with this problem, and much can be gained from modelling this behaviour with a view to enabling AI systems to deal with a situation of this kind. The thesis addresses this problem with a model for uncertainty management.

Different kinds of uncertainty were presented and which show that the problem of uncertainty, in general, is extensive. The analysis also showed that there are considerable differences between the various types of uncertainty and that the problem is amorphous, and in many places not clearly defined or definable. There are also a number of different types of uncertainty which require special treatment.

This thesis deals with uncertain information from human sources, which is a class of uncertain information which features prominently in our everyday experience. This choice is significant for a number of reasons:

Key Position: The class is important, because a substantial proportion of information we have to deal with in everyday situations is from human sources and one cannot avoid having to deal with them.

Distinctive Nature: The problems of uncertainty represented in this class are significantly different from other types of uncertain information and require special treatment.

Measuring Success: In order to be able to assess the success of a model it is important to have a clear class of problems, to determine whether the model can produce the expected results.

7.2 Main Criticisms of Other Approaches

The problem of management of uncertainty has been addressed in a number of areas, but the issues which have been addressed are very different from the issues we are considering. Apart from the notable exceptions of Rescher, Cohen and Garigliano, none of the approaches address the specific problem of uncertain information from human sources.

Probabilistic representations of uncertainty have been widely used to associate precise, numeric values with statements, or sets of statements to express their likelihood or express belief.

Rescher argues that probabilistic approaches have two shortcomings. Firstly, while providing a machinery to propagate probabilities they lack tools to assess initial probabilities. Secondly, the deterioration-effect of propagating probabilities over chains of reasoning produces information which is useless, because of its extremely low likelihood. He proposes to assesses initial plausibilities with static source models and he uses a different propagation algorithm, which does not deteriorate plausibility levels in the same way. While we agree with the basic principle of source control, we argue that non-numeric approaches are more adequate, and that the source control needs to be more sophisticated, more dynamic and more adaptive, to produce a realistic behaviour.

Although a probabilistic strategy will be appropriate for certain problems, in our experience, the majority of everyday situations do not seem particularly suited to this approach. This view is shared by Cohen, who proposes to replace the traditional quantitative analysis with a qualitative evaluation of the arguments for and against a particular proposition. Unfortunately, as he admits himself, his approach can be very cumbersome and difficult to use for complex situations, and that a complementary approach is needed, which is easier to operate. We agree with this position, and in addition we argue that there are a considerable number of situations which cannot be adequately modelled by an analysis of arguments alone, but which require a source control approach to take into account source's abilities and trustworthiness.

7.3 The Fundamental Idea

An adequate analysis of uncertainty requires a qualitative analysis of the information, but an exclusive reliance on a semantic analysis is problematic. As Cohen admits, a pure analysis of the arguments for and against can be unwieldy. In addition, we argue that it also neglects important considerations about human nature. The basic position of this thesis is that the uncertainty of information from human sources has to be seen through a model of the source. The idea of static source models was first put forward by Rescher in a very rudimentary way, but not taken to its full conclusion.

Considerable uncertainty and complexity make it difficult for human agents to completely understand the world they live in. In the attempt of dealing with this situation, human agents tend to generate a world model and try to keep a grip on reality by continuously updating and revising this world model with data from sensations and perceptions. A considerable proportion of that information is from other human agents and these agents differ in the quality of information they provide. Therefore it is necessary to maintain models about their strengths and weaknesses to be able to assess information from them. The thesis adopts this successful source modelling approach in order to enable systems to deal with this problem.

7.4 The Solution

We have shown that there are general principles and heuristics which govern a source modelling approach, and that they can be translated into an operational model which can deal with realistic situations and generate the expected behaviour.

The SCM generates and maintains a source model for each source, recording the source's abilities in a number of areas, as well as the source's trustworthiness, interests and fundamental beliefs. The system learns about the varying abilities and trustworthiness of its sources and adapts the source models accordingly. This model is used in the assessment of information from that source.

The SCM is able to evaluate information and suggest beliefs, based on the analysis of the information and the properties of the source. In addition, it deals with resolving conflicts between pieces of information from the same source or different sources and is able to use enquiries to resolve difficulties. Together with an understanding of the basic situation in which the information was given, it is able to form beliefs in a reasoned manner.

In order to produce a solution to the task a number of important building-blocks were provided:

- A systematic set of principles for this task was generated, as well as a phenomenological explanation of their operation.
- A high-level design has been given, by which such a mechanism can be reconstructed, and, with the help of two examples, we have also demonstrated that the mechanism works satisfactorily.
- An analysis of the complexity and stability of the mechanism has been carried out and which establishes the cost of the mechanism as well as showing the conditions which have to be obeyed in refining it.

7.5 The Significance of the Solution

The Source Control Mechanism provides a qualitative technique for belief formation and conflict resolution for information from human sources with the help of a sophisticated source control approach. This combination makes a significant contribution for a number of reasons.

While we agree with the principles of a qualitative analysis of uncertainty as proposed by Cohen, one cannot afford to ignore basic human nature when dealing with information from them. His approach also requires a great deal of domain specific knowledge. The source control approach, by contrast, is specifically designed to deal with the problem of assessing information from human sources independent of specific domains, and manages to address important issues connected with source's ability and trustworthiness which have not been dealt with by Cohen.

The source control mechanism also makes a significant contribution by adding moving to a qualitative rather than quantitative analysis, thereby adding a sophisticated machinery to the basic concept proposed by Rescher, and further elaborated by Garigliano, who introduced adjustable source models with two indices. The source control mechanism substantially improved this model by adding a whole range of indices together with a more sophisticated machinery to use and adjust them, as well as extending conflict resolution to deal with multiple sources.

Finally, in the complexity and stability analysis the cost of the mechanism and the preconditions of a further refinement were established. Because of the heuristic nature of the approach there is no mathematical function by which results are obtained in the SCM, but a mathematical treatment of the complexity was achieved through an analysis of the decision process the SCM goes through. A quantitative measure of the cost was established, showing the SCM to be feasible. The stability analysis also showed the requirements which have to be fulfilled in refining the system to produce a more sophisticated response within the limits of the present design.

7.6 The Completion of the Task

We have demonstrated that the model is capable of dealing adequately with complex, realistic situations and that we have completed the task we set ourselves. Although it is possible to optimize the model for particular applications, these optimisations are likely to be domain specific and are therefore will not tend to add to the generality of the model. Provided there is a natural language processor which translates the information from human sources into the appropriate form, and which can provide the input required by the source control mechanism, the source control mechanism is capable of dealing with the task of belief formation independently. Given also that there is a mechanism by which queries can be made and reasoning can be checked, and definitions of words and concepts be supplied, in a closer integration between source control mechanism and natural language processor, the full capabilities of the source control mechanism can be realised. The model is well-contained and appears to perform satisfactorily in the problem space. The complexity analysis looked at the various components of the SCM and their connections in order to establish the expected overall cost of using the SCM. In addition we looked at the stability of the SCM to examine the potential for refinement in the present framework.

7.7 Further Work

As far as future developments are concerned, there are a number of possibilities which can be put under the following headings:

- implementation and connection to a NLP
- optimization for particular applications
- integration with other models of reasoning

Having specified the source control mechanism the next step is to implement it and to connect it to a natural language processor. A specification of the SCM to industrial standards is currently being developed and will be implemented as part of a M.Sc. project. Upon completion it will be connected to the LOLITA system.

The model could be applied in a number of different areas, such as in decision support systems which have to deal with uncertain information from human sources. It was part of a proposal for a project to develop a workstation for decision support for crisis management in a centre of operations to deal with uncertainty involved in reports from human sources. The SCM could also prove useful for a problem faced by the Italian judiciary of needing automated tools for sifting through large amounts of data in search for evidence. This may also require the source control model to be optimised for particular applications, given a more restricted type of situation and the requirements specific to the problem. The source control approach is also currently part of a project proposal for a decision support tool to help in the organisation and automatisation of shop floor layout and production management, as briefly described in the previous chapter. There the requirements and personal preferences of the engineers involved have to be balanced with efficiency and optimisation requirements and source control would help to deal with the management of uncertain information from human sources and the need to reconcile conflicting viewpoints.

As already mentioned, there is scope for the integration of the source control model with other reasoning modules. Thus, another area of future development would be to integrate the source control mechanism with modules for plausible reasoning techniques or with an endorsement model in order to make progress towards an integrated reasoning system. Especially as far as the model of endorsement is concerned, that would be very profitable, as they appear to be complementary. Considering, for example, an application to the problem faced by the Italian judiciary, an integration with other techniques for uncertainty management would appear advantageous.

Irrespective of what one may wish to apply the model to, the research has shown that uncertainty management through source control is both successful and feasible and it can be added as another building block on the road towards the construction of integrated and autonomous general reasoning systems. Response from industry in the cases of decision support systems for crisis management and production management have also shown that there is considerable scope for the source control approach making an important contribution to a solution to problems that are currently being addressed.

Glossary

action point - level of belief at which one may be prepared to act on the basis of a piece of information

analogy - form of reasoning by which inferences are made about properties of two classes based on their resemblance

assumption based truth maintenance - truth maintenance system which allows to consider multiple solutions in a belief system to be considered simultaneously.

background theory - theory through which data are interpreted, such as data from a telescope being dependent on a theory of optics.

backtracking - process by which inconsistencies in a truth maintenance system are traced back to their origin; used to eliminate contradictions

Bayes' Theorem - common form of probability propagation technique commonly used in AI

behaviour - phenomenological response of a system or model

belief-space - describing a collection of beliefs entertained by a system

cases - collection of events and corresponding source models about a single incidence in the source control mechanism; used for organisational purposes

class - collection of items which have one or more common properties

classification mechanism - feature of the source control mechanism by which default source models are produced for sources once the source can be attributed to a particular class of sources

closed-world assumption - strategy to avoid dealing with uncertainty, whereby it is assumed that there is no more information to be had than is existent in the model or system; i.e. if you don't know about something then that thing does not exist.

conflict resolution - strategy by which competing entities are resolved; in the source control mechanism used to describe process by which conflicting beliefs are unified into a single one, or none at all

conflicts of information - when used in the context of the source control mechanism it denotes that there are problems within a single piece of information.

conviction - in the source control mechanism it is an indication of a source about its level or intensity of belief in a given piece of information

corroboration - denotes the inertia of a belief which is supported by a number of sources and which therefore is less likely to change through antagonism of a single source

culprit - in truth maintenance used to denote that formula which is chosen to be excluded in the attempt to restore consistency in the set of beliefs

degree of determination - in the source control mechanism used to describe the element of doubt remaining of a given source in a given piece of information

Dempster/Shafer theory of evidence - a variant of Bayes' theorem which uses a probability bounding notion whereby a probability can be ascribed to a set without having to determine the exact probability of each constituent member

diversification - technique used to limit the impact of uncertainty by spreading the risk

enquiry - process in the source control model where active strategies are engaged to find explanations and solutions to problems; often in context with asking sources to provide further information to clarify the problem

events - representation of information by the natural language environment

evidential reasoning - reasoning method involving probabilistic methods in inferentially connected belief spaces

experience - index in the source model of the SCM to record source's ability to handle its own experiences

expertise - term used in the source control mechanism to denote proficiency or competence in a particular subject area

formal approaches - as opposed to heuristic approaches, approaches which have been formalised and proven

frequency theory of probability - statistical variant of probability theory, built on the concept of rate of occurrence

general reasoning system - reasoning systems which are not domain specific and which use reasoning techniques with a wide area of application to deal with real-world environments

heuristics -

importance analysis - in the context of the source control mechanism a process to determine potentially how significant the information or situation is to the system to determine whether and if so how much to investigate.

indication of belief - in the SCM the source's indication of its conviction in the information as represented in the event.

inductive argument - argument from a few examples and a strong assumption of homogeneity of reality to a general conclusion.

interests - in the SCM denoting an index in the source model to record particularly strong interests i.e. making profit for a business-man

judging sources - index in the SCM to describe the source's ability to deal with, and correctly assess information from human sources

justification based truth maintenance - variant of truth maintenance system which only allows a single context and where conclusions can be calculated in an inferentially connected belief space by the propagation of justifications from

literal meaning input - input from natural language based on the assumption that the input can be taken literally and that there is no use of metaphors, figurative or other elements of that kind.

logical theory of probability - variant of probability theory where the probability of a given statement can be logically determined if it can be determined at all.

management of uncertainty - strategies for the assessment of uncertain information and for determining the impact of uncertain information on system's of belief

message - in the SCM that part of the communication which carries the meaning.

qualification - in the SCM that part of the communication which puts the interpretation by the source on the message

model of endorsement - strategy for assessing the uncertainty of information through an analysis of the arguments for and against and relative to the purpose the information is to be used for

monotonicity - assumption in a system of logic that propositions do not change their truth value once it has been assigned

natural language environment - system which is capable of taking natural language input and transforming it into a machine representation

noise-free data - assumption that all the data given are relevant and genuine

pars deterior principle - strategy whereby the probability or plausibility or strength of belief of a conclusion of an inference is at least as strong as the weakest premiss

personalistic theory of probability - variant of probability theory where probabilities are determined by personal preference

plausible reasoning - term used to describe human forms of reasoning either for uncertainty management or the management of uncertain reasoning techniques

plausible reasoning techniques - uncertain reasoning techniques modelled on human forms of reasoning like analogy, induction, abduction etc. which are not guaranteed to be truth preserving

possible no-good - type of justification in a truth maintenance system which introduces a representation of uncertainty by relegating it to the realm of hypotheses

problem solver - a system to designed to find solutions to problems

real-world - denoting the real environment as opposed to toy worlds

reasoning - an index in the SCM to record the source's ability to follow and construct complex pieces of reasoning

robustness - when referring to uncertainty it denotes that the information is not likely to change dramatically and unexpectedly

solidity - in the SCM denoting that the information is of good quality and backed by a good source

source control - process by which information is evaluated with the help of considerations about the known abilities and trustworthiness of the source

source model - model recording various aspects about the abilities and trustworthiness as well as interests and beliefs

source model reevaluation - process by which source models are readjusted following new evidence.

sources - human agents known to the system and communicating with it

system of belief - set of pieces of information believed by the system and which are expected to be mutually consistent

theorem prover - type of system designed to prove theorems

truth maintenance - the process of ordering sets of beliefs and to eradicate inconsistencies either by expulsion of formulae or by producing multiple

world model - representation of the real world as a set of beliefs

References

[ADE88a] Adey S, High Level Control of Simulated Ships and Aircraft, Alvey Workshop on Multiple Agent Systems, Redhill, UK 14-15 April 1988

[ALE80] McAllester D., An Outlook on Truth Maintenance, MIT AI Memo 551, 1983

[ALV88] Alvey Workshop on Multiple Agents, Proceedings, Redhill, UK, 14-15 April 1988

[AMB87] D'Ambrosio B, Truth Maintenance with Numeric Certainty Estimates, Proc of the 3rd Conference on Artificial Intelligence Applications, Kissimmee FL USA 23-27 Feb 1987 pp 244-9

[AMB88] D'Ambrosio B, A Hybrid Approach to Resoning under Uncertainty, International Journal of Approximate Reasoning (USA), vol 2 no 1 Jan 1988 pp 29-45

[BAL87] Baldwin J F, Evidential Support Logic Programming, Fuzzy Sets Systems (Netherlands), vol 24 no 1 Oct 1987 pp 1-26

[BAM86] Balakrishnan N & Mahapatra PR & Nayak K R & Poulose M M & Krishna M G, An Expert System for Air Traffic Control, 7th European Conference on Electrotechnics : Advanced Technologies and Processes in Communication and Power Systems EUROCON 86, Paris France, 21-23 April 1986, pp 97-104

[BAV88] Backer E , Van der Lubbe J C A & Krijgsman W, On Modelling of Uncertainty and Inexactness in Expert Systems, Proc. of the 9th Symposium on Information Theory in the Benelux, Mierolo Netherlands 26-27 May 1988 pp 105-11

[BEA89] Beardon c, Artificial Intelligence Terminology; A Reference Guide, Ellis Horwood, Chichester UK, 1989

[BEL88] Bellamy R K E, Agents and Actions: A Framework for Describing the Interaction Between User and Intelligent Systems, Alvey Workshop on Multiple Agents, Redhill, UK, 14-15 April 1988

[BER87] Berenji H & Lum H Jr, Applications of Plausible Reasoning to AI-based control systems, Proc. of the 1987 American Control Conference, Minneapolis MN USA 10-12 June 1987 vol 3 pp 1655-61

[BES87] Bestougeff H & Ligozat G, Plausible Reasoning in Expert Systems, Expert Systems, Theory and Applications; IASTED Int Conference, Geneva, Switzerland 16-18 June 1987 pp 216-19

[BGA90] Bokma A & Garigliano R, Towards a Conflict Resolution Mechanism for Multiple Sources, Proceedings, International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Paris France, 2-6 July 1990

[BIR82] Birman K P, Rule Based Learning for more accurate ECG Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence (USA), vol PAMI-4 no 4 July 1982 pp 369-80

[BRA31] Braithwaite R B, Jeffrey's Theory of Probability, Mind 40, 1930

[CAR51] Carnap R, The Logical Foundations of Probability, Chicago 1951

[COH85] Cohen P R, Heuristic Reasoning About Uncertainty: An Artificial Intelligence Approach, Pitman, London, 1985

[COL87a] Cohen P & Levesque H, Persistence, Intention and Committment, Technical Report No 415 SRI International CA USA

[COL87A] Cohen P R & Levesque H J, Rational Interaction as the Basis for Communication, Proc of the Symp on Intentions and Plans in Communication and Discourse, California USA 1987

[COM89] Collins A & Michalsky R, The Logic of Plausible Reasoning: a Core Theory, Cognitive Science (USA) vol 13 no 1 Jan 1989 pp 1-49

[COS88] Connah D, Shiels M & Wavish P, The Architecture of an Agent, Alvey Workshop on Multiple Agents, Redhill, UK, 14-15 April 1988

[CRA46] Cramer H, Mathematical Methods of Statistics, Princeton UP 1946 p.149

[DEC86] Dechter R, Learning while Searching in Constraint-Satisfaction-Problems, Proc AAAI-86; 5th National Conference on Artificial Intelligence, Philadelphia PA USA 11-15 Aug 1986 vol 1 pp 178-83

[DEC87] Dechter R, Constraint-Network Approach to Truth-Maintenance, UCLA CA USA Report No. CSD-870009

[DED80] McDermott D & Doyle J, Nonmonotonic Logic I, Artificial Intelligence No.13 pp. 41-72, 1980

[DOR88] Doran J, The Structure and Emergence of Hierarchical Organisations, Alvey Workshop on Multiple Agents, Redhill, UK, 14-15 April 1988

[DOY78] Doyle J., Truth Maintenance Systems for Problem Solving, AI-TR-419, AI Labs, MIT, 1978

[DOY79] Doyle J., A Truth Maintenance System, Artificial Intelligence No.12, pp. 232-272, 1979

[DOZ88] Dontas K & Zemankova M, APPLAUSE: An Experimental Plausible Reasoning System, 3rd Proceedings of the 3rd Int. Symposium of Methodologies for Intelligent Systems, Turin Italy 12-15 Oct 1988 pp 29-39

[DRE87] Dressler O, Extensions of the Basic ATMS, GWAI-87: 11th German Workshop on Artificial Intelligence, Proceedings, Geseke FRG 28 Sept - 2 Oct 1987, pp 185-94

[DRE88] Dressler O, Extending the Basic ATMS, ECAI 88, Proceedings of the 8th European Conference on Artificial Intelligence, Munich FRG 1-5 Aug. 1988, pp 535-40

[DRS87] Drummond M & Steel S & Kelleher G, Truth Maintenance Systems: A Major-topic Group-report from the 6th Alvey Planning SIG, Alvey Expert Systems Research Theme; Report of the 6th Workshop of the Planning Special Interest Group, Cambridge UK 2-3 April 1989 pp 1-21

[FAL87] Falkenhainer B, Towards a General-purpose Belief Maintenance System, Report No UIUCDCS-R87-1329, Uni of Illinois, Urbana-Champaign, USA April 1987

[FAN88] Fangqing Dong & Nakagawa H, Extended ATMS for Knowledge with Uncertainty, Journal of Japanese Society for Artificial Intelligence, vol 3 no 1 Jan 1988 pp 86-93

[FAP85] Farreny H & Prade H, Mechanising Default Reasoning in Terms of Possibility, Hardware and Software Components and Architectures for the 5th Generation. Congress and Exhibition, 5-7 March 1985, AFCET Paris France pp 353-64

[FAY88] Fayyad U M & Van Voorhis K E & Wiesmeyer M D, Learning Control Information in Rule Based Systems: A Weak Method, Proc. Of the 4th Conference on Art. Int. Applications, San Diego CA USA 14-18 March 1988 pp 188-93

[FIN37] de Finetti, La prevision: ses lois logiques, ses sources subjectives, Annales de l'Institut Henri Poincare 7 1937

[GAB88a] Garigliano R, Bokma A & Long D, A Model for Learning by Source Control, in Bouchon B. and Yager R.R. (eds.) Uncertainty in Knowledge- Based Systems, LNCS, Springer-Verlag, 1988

[GAB88b] Garigliano R & Bokma A, Problem Independent Heuristics for Conflict Resolution, Alvey Wokshop on Multiple Systems, Phillips Research Laboratories, Redhill, UK April 1988

[GAL89] Garigliano R. and Long D.: A Formal Model through Homogeneity Theory of Adaptive Reasoning, PRG Monograph, Oxford University, 1989

[GAR86] Garigliano R.: A Consistency-Recovering System for Inference Engines, in Bouchon B. and Yager R.R. (eds.) Uncertainty in Knowledge- Based Systems, LNCS, Springer-Verlag, 1986

[GIN88] Ginsberg M L, Multivalued Logics: A Uniform Approach to Reasoning in Artificial Intelligence, Computational Intelligence (Canada), vol 4 no 3, August 1988, pp 265-316

[GIS85a] Gilmore J F & Semeco A C, Route Generation and Navigation in Robotic Vehicles, Computer Vision for Robots, Proc. SPIE Int Soc Opt Eng (USA), Cannes France 2-6 Dec 1985, pp 313-17

[GIS85b] Gilmore J F , Semeco A C & Eamsherangkoon P, Knowledge-Based Route Planning Through Natural Terrain, INSPEC Conference 1985 : Applications of Artificial Intelligence II, Arlington VA USA, 9-11 April 1985, Proceedings of the SPIE Int Soc of Opt Eng (USA), vol 548 pp 128-36

[GLL87] Galliers J R, A Definition of Cooperation (but not Benevolence) for Multi-Agent Planning, Alvey Exp. Sys. Research Theme: Report of the 7th Workshop of The Planning Special Interest Group, Ipswich UK 26-27 Nov 1987

[GLL88] Galliers J R, A Strategic Framework for Multi-Agent Cooperative Dialogue, Proceedings of the 8th European Conference on Artificial Intelligence, Munich FRG 1-5 August 1988, pp 415-20

[GOO84] Goodwin J W, WATSOM: a Dependency directed inference system, Proceedings, Non-Monotonic Reasoning Workshop, pp. 103-114, 1984

[HER88] Hern L E C, Describing DAI Models: A Framework and Examples, Alvey Workshop on Multiple Agents, Redhill, UK, 14-15 April 1988

[HOH86] Horvitz E J & Heckerman DE & Langlotz C P, A Framework for Comparing alternative formalisms for plausible reasoning, Proc AAAI-86 5th National Conference on Artificial Intelligence, Philadelphia PA USA 11-15 Aug 1986, pp 210-214 vol 1

[HOS88] Hori S & Sheu P C, Plausible Reasoning in Expert Systems for Troubleshooting, Applications of Artificial Intelligence VI, Orlando FL USA 4-6 April 1988, Proceedings SPIE - The International Society for Optical Engineering, vol 937 pp 202-9

[IKM89] Ikeda M & Mizoguchi R & Kakusho O, Student Model Description Language SMDL and Student Model Inference System SMIS, Transactions of the Institute of Electronics, Information and Communication Engineers D-II (Japan), vol J72D-II no 1 Jan. 1989 pp 112-20

[ION89] Ionnidis Y E & Sellis T K, Conflict Resolution of Rules Assigning Values to Virtual Attributes, ACM SIGMOD Int. Conf. on Management of Data, Portland Oregon USA 31 May - 2 June 1989 vol 18 no. 2 pp 205-14

[KAI89] Katai O & Iwai S, Pluralistic Evaluation of Belief Plausibility and its Application to Nonmonotonic Reasoning, International Journal of Approximate Reasoning, vol 3 no 2 March 1989 pp 219-37

[KEY21] Keynes J M, A Treatise on Probability, New York 1921

[KLE84] de Kleer J, Choices without Backtracking, Proceedings AAAI-84

[KLE85] de Kleer J, Assumption Based Truth Maintenance, GWAI-85: 9th German Workshop on Artificial Intelligence, Proceedings, Dassel/Solling FRG 23-27 Sept 1985, p 258ff

[KLE86a] De Kleer J, An Assumption Based TMS, Artificial Intelligence, No. 28, 1986, pp. 127-162

[KLE86b] De Kleer J, Extending the ATMS, Artificial Intelligence, No. 28, 1986, pp. 163-196

[KLE86c] De Kleer J, Problem Solving with the ATMS, Artificial Intelligence, No. 28, 1986, pp. 197-224

[KYB61] Kyburg H E, Probability and the Logic of Rational Belief, Wesleyan University Press, Middleton Connecticut USA 1961

[LAU84] Laurent J P, Operation of a Knowledge Base: Choices and Strategies, 4eme Congres. Reconnaissance des Formes et Intelligence Artificielle, Roquencourt, France, 23 - 27 January 1984, pp 219-33

[LOL92] Garigliano R, Morgan R & Smith M: LOLITA: Progress Report 1, Technical Report No 12/92, Department of Computer Science, University of Durham, 1992

[LON87] Long D, A Formal Model for Reasoning by Analogy, D.Phil. Thesis, PRG, University of Oxford, 1987

[MAR83] Martins J, Reasoning in Multiple Belief Spaces, Proceedings IJCAI No.12 pp. 370-373, 1983

[MAS84] Martins J & Shapiro S, A Model for Belief Revision, Proceedings Non-Monotonic Reasoning Workshop, pp. 241-294, 1984

[MAS86] Martins J. & Shapiro S, Hypothetical Reasoning in Applications of AI in Engineering Problems vol2, ed Sriram D & Adey R, Springer Verlag 1986

[MAS88] Martins J. & Shapiro S, A Model for Belief Revision, Art. Int. vol 35 pp 25-79

[MIS28] von Mises R., Wahrscheinlichkeit, Statistik und Wahrheit, Wien 1928

[MON86] Morris P H & Nado R A, Representing Actions with an Assumption-based Truth Maintenance Suystem, Proc AAAI-86; 5th National Conference on Artificial Intelligence, Philadelphia PA USA 11-15 Aug 1986 vol 1 pp 13-17

[MOR86] Morris M E, Meta-Level Reasoning for Conflict Resolution in Backward Chaining, Proceedings WESTEX-86 IEEE Western Conference on Knowledge-Based, Engineering and Expert Systems, Anaheim CA USA 24-26 June 86 pp 13-18

[NEI87] O'Neill J L, Plausible Reasoning, Australian Computer Journal (Australia), vol 29 no 1 Feb 1987 pp 2-15

[NEY50] Neyman J, First Course in Probability and Statistics, New York 1950

[OGN81] Ogami K & Nishiyama S & Kakusho O, Hierarchical Production System, System Organisation and Implementation, Technical Report Kensai University Japan, no 22 March 1981 pp 45-55

[ORG88] Orgill C H, A Knowledge Bus: A Framework for Multiple Knowledge Models, Alvey Workshop on Multiple Agents, Redhill, UK, 14-15 April 1988

[PAA86] Paass G, Consistent Evaluation of Uncertain Reasoning Systems, 6th International Workshop on Expert Systems and their Applications, Avignon, France, 28-30 Aplril 1986 pp 73-94

[PAU85] Pau L F, Artificial Intelligence System for Failure Detection and Monitoring in Electronics and Communications, Expert System in Government Symposium, McLean VA USA 24-25 Oct 1985 pp 4341-8

[PEA88a] Pearl J, Evidential Reasoning Under Uncertainty, UCLA, Computer Science Department CA USA Rep. No. CSD-880055

[PEA88b] Pearl J, Non-Bayesian Formalisms for Managing Uncertainty, UCLA, Computer Science Department CA USA Rep. No. CSD-880054

[PET86] Petrie C J Jr, A Diffusing Computation for Truth Maintenance, Proc of the 1986 Int Conf on Parrallel Processing, St Charles IL USA 19-22 Aug 1986, pp 691-5

[POZ87] Popchev I P & Zlatareva N P, Intelligent Problem-independent Tool for Plausible Inference, Automatic Control - World Congress 1987, Selected papers from the 10th Triennial World Congres of the Int. Fed. of Automatic Control, Munich FRG 27-31 July 1987 vol VII pp. 341-6

[PRA85] Prade H, A Computational Approach to Approximate and Plausible Reasoning with Applications to Expert Systems, IEEE Transactions, Pattern Analysis & Machine Intelligence (USA), vol PAMI-7 no 3 May 1985 pp 260-83

[PRO88] Provan G M, Solving Diagnostic Problems using Extended Truth Maintenance Systems, ECAI 88, Proceedings of the 8th European Conference on Artificial, Intelligence, Munich FRG 1-5 Aug. 1988, pp 547-52

[QUI85] Quinlan J R, Internal Consistency in Plausible Reasoning Systems, New Generation Computers (Japan), vol 3 no 2 1985 pp 157-80

[RAM31] Ramsey F P, The Foundations of Mathematics, New York 1931

[RAS88] Rake S T & Smith L D R, Use of an Assumption-based Truth Maintenance System to Record and Resolve Ambiguity in Cardiac Angiograms, Proc. of the 4th Int Conf. on Patterm Recognition, Cambridge UK 28-30 March 1988, pp 656-65

[RES76] Rescher N.: Plausible Reasoning, Van Gorcum, Assen/Amsterdam 1976

[ROD89] Rodi W L, A New Algorithm for Truth Maintenance, Proceedings of the Annual AI Systems in Government Conference, Washington DC USA 27-31 March 1989, pp 14-21

[RUS48] Russell B, Human Knowledge, New York 1948

[SAV54] Savage L J, The Foundations of Statistics, New York 1954

[SHH86] Shao-Hung Liu G, Causal and Plausible Reasoning in Expert Systems, Proc AAAI-86 5th National Conference on Artificial Intelligence, Philadelphia PA USA 11-15 Aug 1986, pp 220-225 vol 1

[SHA86] Shaw M J, A Pattern-Directed Approach to FMS Scheduling, Second ORSA/TIMS Conference on Flexible Manufacturing Systems, Ann Arbor Michigan USA Aug 1986, Annals of Operations Research vol 15 no 1-4 Oct 1988 pp 353-76

[SMK88] Barbara Smith & Gerald Kelleher, Reason Maintenance Systems and Their Applications, Ellis Horwood, Chichester, UK, 1988

[STE81] Stefik M, Planning and Meta-Planning (MOLGEN: II), Artificial Intelligence (Netherlands), vol 16 no 2 May 1981 pp 141-69

[THO79] Thompson A, Network Truth-maintenance for deduction and modelling, Proc. IJCAI-79, pp 877-879

[VEA88] Venkat Rangan P & Ashany R, The Use of AI to Secure Communication in distributed Systems for Industrial Applications, Proc of the International Workshop on Artificial Intelligence for Industrial Applications IEEE AI '88, Hitachi City Japan 23-27 May 1988 pp 565-9

[VEN86] Venn J, The Logic of Chance, London and New York, 1886

[WAL37] Wald A, Die Widerspruchsfreiheit des Kollektivbegriffs der Wahrscheinlichkeitsrechnung, Ergebnisse eines mathematischen Kolloquiums 8, 1937

[WHS86] White C C & Sykes E A, A User Preference Guided Approach to Conflict Resolution in Rule-Based Expert Systems, IEEE Transactions Systems , Man & Cybernetics (USA), vol SMC-16 no 2 March-April 1986 pp 276-8

[WID89] Widman L E, Expert System Reasoning about Dynamic Systems by Semi-Quantitative Simulation (medicine), Computer Methods and Programs in Biomedicine, vol 29 no 2 June 1989 pp 95-113

[WIL86] Wilf H S: Algorithms and Complexity, Prentice-Hall, Englewood Cliffs, New Jersey, USA, 1986

Bibliography

[ADE88b] Adey S J, Two Consistency Systems with their Similarities and Differences, UK IT 88 Conference Publication, Swansea UK 4-7 July 1988

[ALE82] McAllester D., Reasoning Utility Users Manual, MIT AI Labs, 1983

[ALS88] D'Aloisi D & Stock O & Tuozzi A, An Implementation of the Propositional Part of KRAPFEN, a Hybrid Knowledge Representation System, Methodologies for Intelligent Systems; Proc. of the 3rd Int. Symp., Turin Italy 12-15 Oct 1988 pp 200-9

[ANB75] Anderson A R & Belnap N D, Entailment, Princeton University Press, 1975

[ARJ88] Arlaboss F & Jean-Bart B & Porte N & de Ravinel B, An Efficient Problem Solving Architecture Using ATMS Tested on a Non-toy Case Study, AI Communications, vol 1 no 4 December 1988 pp 6-15

[BAF81] Barr A & Feigenbaum E A : The Handbook of Artificial Intelligence: William Kaufmann Inc. ; Los Altos CA, USA, 1981

[BAP85] Bazakos M E & Panda D P, Stereopsis and Scene Partitionning for Terrain Interpretation, Applications of Artificial Intelligence II, Arlington VA USA, 9-11 April 1985, Proceedings of the SPIE Int Soc of Opt Eng (USA), vol 548 pp 84-91

[BET59] Beth, E, The Foundations of Mathematics, Amsterdam, North Holland, 1959

[BIG88] Bigham J, Truth Maintenance in Inference, IEE Colloquium on 'Inference", London UK March 1988, Digest No 37 pp 1-2

[BOM86] Bowen J & Mayhew J, Consistency Maintenance in the REV graph Environment, Technical Report AIVRU 020, University of Sheffield 1986

[BRO86] Brown F M, A Comparison of the Commonsense and Fixed Point Theories of Non-Monotonicity, Proceedings of the AAAI 1986

[BRO88] Brown A L Jr, Logics of Justified Belief, ECAI 88, Proceedings of the 8th European Conference on Artificial Intelligence, Munich FRG 1-5 Aug. 1988, pp 507-12

[BRK89] Brookshaer J G, Theory of Computation, Benjamin Cummings Publishing, Redwood City, California 1989

[BUC87] Buckley J J, The Utility of Information and Risk Taking Fuzzy Expert Systems, Proceedings of the 26th IEEE Conf. on Decision and Control, Los Angeles CA USA 9-11 Dec 1987

[CAR50] Carnap R.: Logical Foundations of Probability (1950), 2nd ed. 1962 by University of Chicago Press:Chicago

[CAR80] McCarthy J, Circumscription - A form of Nonmonotonic Reasoning, Artificial Intelligence, No. 13, 1986, pp. 27-39

[CEC88] Cebulka K D & Carberry S & Chester D L, Solving Dynamic-input Interpretation Problems Using the Hypothesize-test-revise paradigm, Proceedings of the 4th Conference on Artificial Intelligence Applications, San Digo Ca USA 14-18 March 1988

[CHE86] Chehire W, SYPRUC: A Knowledge Representation and Manipulation System, 6th International Workshop on Expert Systems and Their Applications, Avignon France 28-30 April 1986 pp 933-46 vol 2

[CLL78] Collins A M, Fragments of a Theory of Human Plausible Reasoning, TINLAP-2, pp 194-201

[DEC88] Dechter R, A Distributed Algorithm for the ATMS, Technical Report R-109, UCLA Dept of Comp. Sci. 1988

[DEL87] Delgrande J P, A First-Order Conditional Logic for Prototypical Properties, Artificial Intelligence No.33, pp. 105-130, 1987

[DER83] McDermott D, Contexts and data dependencies: A synthesis, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-5, pp. 237-246

[DOY81] Doyle J., A Truth Maintenance System, in Readings in Artificial Intelligence, eds. N J Nilson, Morgan Kaufmann 1981

[DOY83] Doyle J., The Ins and Outs of Reason Maintenance, Proceedings, IJCAI pp. 349-351, 1983

[DUP84] Dubois D & Prade H, On Distances between Fuzzy Set Points and their use for Plausible Reasoning, Proc. od the International Conference on Systems, Man and Cybernetics, Bombay & New Delhi India 29 Dec 1983 - 7 Jan 1984, pp 300-3 vol 1

[EEL82] Eells E.: Rational decision and causality, Cambridge University Press, Cambridge 1982

[ESH87] Eshghi K, Abductive Planning with Event Calculus, Report of the 7th Workshop of the Planning Special Interest Group, Alvey Conference, Ipswich UK 26-27 Nov 1987 p 22ff. (also in: Logic Programming: Proceedings of the 5th Int. Conference and Symposium, Seattle WA USA 15-19 Aug 1988, pp 562-79 vol 1)

[FOR86] Forbus K D, Qualitative Process Engine, Report No UIUCDCS-R-86-1288, Uni of Illinois, Urbana-Champaign, USA Dec 1986

[FRE64] Frege G.: Begriffsschrift, ed. I.Angelelli, Hildesheim 1964

[GAB82] Gabbay D M, Intuitionistic Basis for Non-Monotonic Logic, 6th Conference on Natural Deduction, Springer Verlag, pp. 260-273, 1982

[GLT84] Glymour C & Thomason R, Default Reasoning and the Logic of Theory Perturbation, Procedings of the AAAI Workshop on Non-monotonic Reasoning, New York, 1984

[GOO87] Goodwin J W, A Theory and system for non-monotonic reasoning, Linkoping Studies in Science and Technology, Dissertations, no. 165, Dept of Comp. and Inf. Sci., Linkoping Uni , Sweden, 1987

[GRA87] Gray M A, Implementing an Intelligent Design Machine in a TMS-based Inference System, Proc of the 1987 Int Conf on Systems, Man, and Cybernetics, Alexandria VA USA 20-23 Oct 1987 vol 1 pp 163-72

[GUT84] Guida G & Tasso C, A New Approach to the Design of Expert System architectures, Artificial Intelligience and Information-Control Systems of Robots, Smolenice, Czechoslovakia 11-15 June 1984, pp 405-13

[HAV88] Havlicsek B L, Integrating Diagnostic Knowledge, AUTOTESTCON 88 ; Symposium Proceedings: Futuretest. IEEE Int. Automatic Testing Conference, Minneapolis MN USA 4-6 Oct 1988

[HAY75] Hayes P., A Representation for Robot Plans, Proceedings IJCAI-4, 1975

[HAY83] Hayes-Roth B, The Blackboard architecture: a general framework for problem solving, Stanford Comp Sci Dept Report no. HPP-83-30

[HEK86] Herman M & Kanade T, Artificial Intelligence, No. 30, 1986, pp. 289-314, 1986

[HRY88] Hrycej T, Intelligent Backtracking with Structured Contexts, ECAI 88, Proceedings of the 8th European Conference on Artificial Intelligence, Munich FRG 1-5 Aug. 1988, pp 589-98

[HUZ83] Hummel R & Zucker S, On the Foundations of Relaxation Labelling process, IEEE Trans. PAMI, 5, pp. 267-304, 1983

[KIM88] Kimura F, Computer Graphics and Artificial Intelligence, Information Processing Society of Japan, vol 29 no 10 1988 pp 1223-30

[KLB80] De Kleer J & Brown, A Qualitative Physics Based on Confluences, Artificial Intelligence, No. 24, 1980, pp. 7-83, 1980

[KLF87] De Kleer J & Forbus K & Williams B, Tutorial No: TA 4. Truth Maintenance Systems, Proc 6th Nat Conf on AI - AAAI-87

[KLK89] De Kleer J & Konolige K, Eliminating the Fixed Predicates from a Circumscription, Artificial Intelligence, vol 39 no 3 July 1989, pp 391-8

[KLS80] De Kleer J & Sussman G, Propagation of constraints applied to circuit synthesis, Circuit Theory and Applications, pp. 127-144

[KLW86] de Kleer J & Williams B C, Back to Backtracking: Controlling the ATMS, Proc AAAI-86; 5th National Conference on Artificial Intelligence, Philadelphia PA USA 11-15 Aug 1986 vol 2 pp 910-17

[KLW87] De Kleer J & Williams B, Diagnosing Multiple Faults, Artificial Intelligence, No. 32, 1987, pp. 97-130

[KOR81] Kornfeld W A & Hewitt C E, The Scientific Community Metaphor, IEEE Transactions on Systems, Man and Cybernetics, 1981 SMC-11:1 pp 24-33

[KUH70] Kuhn Th. S.: The Structure of Scientific Revolution, University of Chicago Press 1970

[LEN88] Lenzerini M & Nardi D, Belief Revision as Meta-Reasoning, ECAI 88, Proceedings of the 8th European Conference on Artificial Intelligence, Munich FRG 1-5 Aug. 1988, pp 577-9

[MAR86] Van Marcke K, A Parallel Algorithm for Consistency Maintenance in Knowledge Representation, ECAI '86: 7th European Conference on Artificial Intelligence, Proceedings, Brighton UK 21-25 July 1986 vol 1 pp 278-90

[MAS89] Matsumoto Y & Satoh K, Nonmonotonic Logic and Commonsense Reasoning, Information Processing Society of Japan, vol 30, no. 6 1989 pp 674-83

[MIN74] Minsky M, A Framework for Representing Knowledge, MIT AI Labs, AIM-306, Cambridge Mass.

[MOO85] Moore R C, Semantical Considerations on Nonmonotonic Logic, Artificial Intelligence, No. 25, 1985, pp. 75-94

[MOR88] Morris P H, The Anomalous Extension problem in Default Reasoning, Artificial Intelligence, vol 35 no 3 July 1988, pp 383-99

[NAP86] Nardi B A & Paulson E A, Multiple Worlds with Truth Maintenance in AI Applications, ECAI '86: 7th European Conference on Artificial Intelligence, Proceedings, Brighton UK 21-25 July 1986 vol 1 pp 437-44

[NIL71] Nilsson N J Problem Solving Methods in Artificial Intelligence, McGraw-Hill, New York, 1971

[NIL80] Nilsson N J, Principles of Artificial Intelligence, Palo Alto CA Tioga, 1980

[NOZ69] Nozick R.: Newcomb's problem and two principles of choice in Essays in Honour of Carl G. Hempel, ed. Rescher N. by Reidel:Dordrecht 1969

[NUD83] Nudel B, Constraint-Labelling Problems and Their Algorithms: Expected-Complexities and Thory-Based Heuristics, Artificial Intelligence, no.21, pp 135-178, 1983

[NUT86] Nute D, LDR: A Logic for Defeasible Reasoning, ACMC Research Report 01-0013, Advanced Computational Methods Center, University of Georgia

[REB79] Rescher N and Brandom R, The Logic of Inconsistency, Basil Blackwell UK, 1979

[REI80] Reiter R, A Logic For Default Reasoning, Artificial Intelligence, 1980 vol 13 pp 81-132

[REI80] Reiter R.: A Logic for Default Reasoning, Artificial Intelligence, vol 13, pp.81-132, 1980

[RES64] Hypothetical Reasoning, North Holland, 1964

[RIN86] Rine D C, Some Relationships between Logic Programming and Multiple-valued Logic, Proc. of the 16th International Symposium on Multiple-valued Logic, Blacksburg VA USA 27-29 May 1986 pp 160-3

[SCH84] Schor M I, Using Declarative Knowledge Representation Techniques: Implementing, Truth Maintenance in OPS5, 1st Conference on Artificial Intelligence Applications, Denver, CO USA 5-7 Dec 1984 pp 261-6

[SHA87] Shadbolt N, Report of the Assumption Based Reasoning Theme, Proc of the Alvey Knowledge Based Systems Club Explanation Special Interest Group; 2nd Workshop, Guildford, UK 8-9 Jan 1987 pp 171-86

[SMU68] Smullyan R, First Order Logic, Springer, 1968

[TAS87] Tanabe K Suzuki A, A Formalization of Assumption-based Truth Maintenance System in Relevance Logic, Report of the Graduate School of Electronics, Science and Technology, Shizuoka University Hamamatsu Japan, no 8 March 1987 pp 65-71

[TUR53] Turing A M et. al. : Digital Computers applied to games. in B V Bowden (ed.) Faster than thought : London Pitman pp. 286-310, 1953

[WAH87] Wah B W, Design Methodologies of Computers for Artificial Intelligence Processing, Proceedings of the 2nd International Conference on Supercomputing, San Francisco, CA USA 4-7 May 87, vol 1 pp 494-5

[WAL81] Walker A & Leakey R E F, The Hominids of Eastern Turkana, Scientific American, August 1981, pp 54-66

[WAP81] Warren D & Pereira F, An efficient easily adaptable system for interpreting natural language queries, Edinburgh DAI Res Labs Paper No.155

[WHR25] Whitehead A N & Russell B :Principia Mathematica; CUP England vol. 1, 1925

[WIL86] Williams C, Expert Systems, Knowledge Engineering and AI Tools - An Overview, IEEE Expert (USA) vol 1 no 4 Winter 1986 pp 66-70

[WIL86] Williams B, Doing Time: Putting Qualitative Reasoning on Firmer Ground, Proc. AAAI-86, pp 105-112

[WIN84] Winston P.: Artificial Intelligence (2nd ed.), Addison-Wesley Reading, Massachusetts, 1984

Appendix A

Belief Formation Diagrams

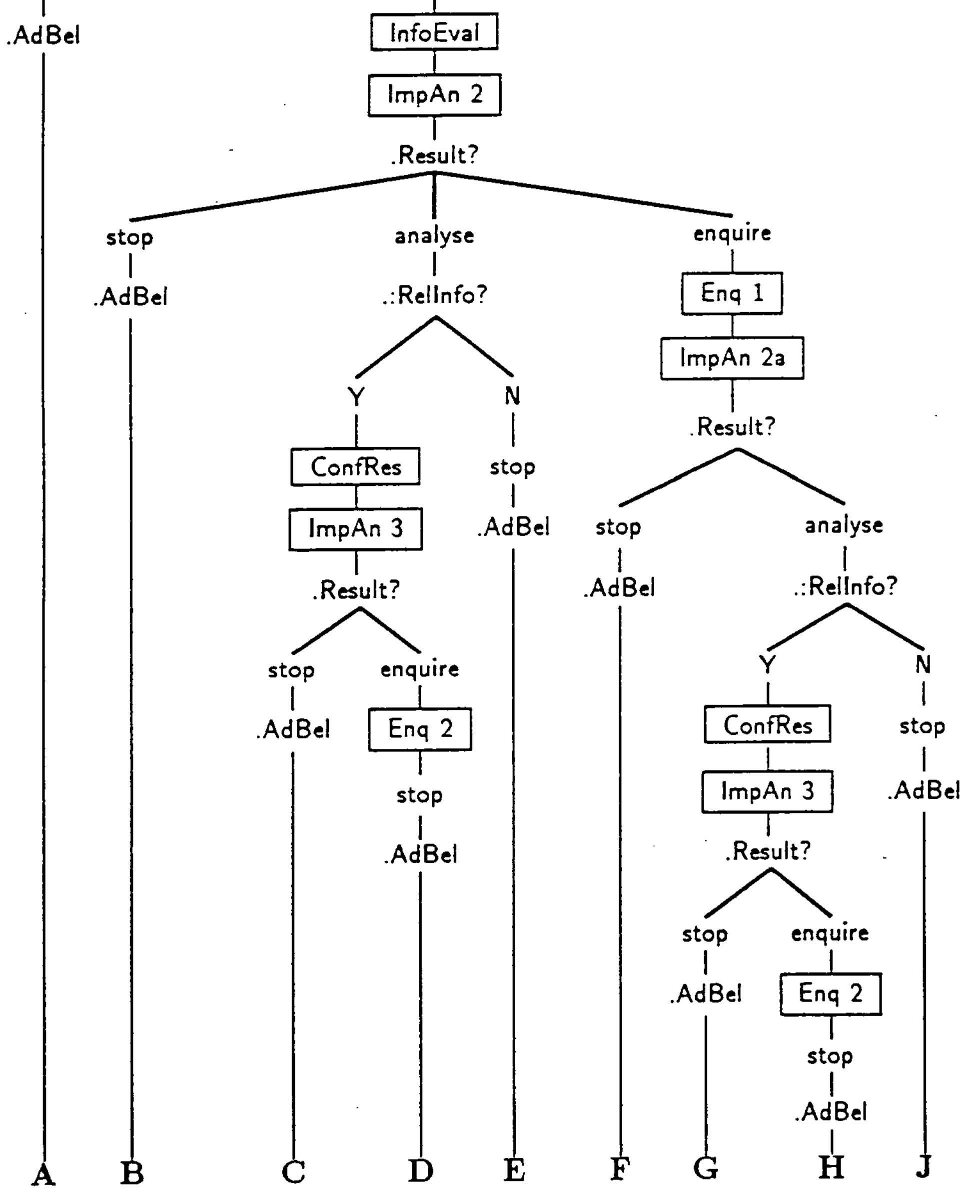
This appendix contains the decision trees corresponding to the belief formation component of the Source Control Mechanism. The decision trees are divided in separate sections, corresponding to the various subcomponents as detailed in the principles and design chapters and preceeded by a legend, explaining the items on the decision trees.

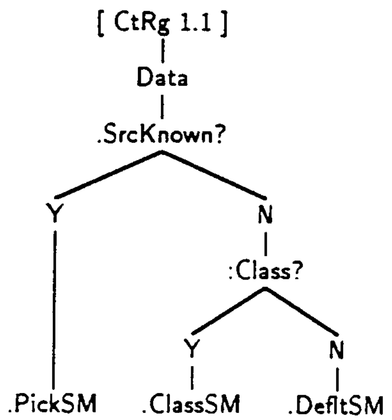
Control Regime

Cont Reg	the label of the diagram
Data Input	an event: individual piece of information
Data+SM	an event and the source model
Result?	decision of importance analysis
stop	recommendation not to do further analysis
AdBel	action of adopting the belief recommended by the SCM
analyse	recommendation to do further analysis
enquire	recommendation to do an enquiry
RelInfo?	is there a connection to other information
Y	yes
N	no
A - J	label to identify different paths
CtRg 1.1	the label of the diagram
SrcKnown?	whether a source model for the source exists
PickSM	get source model from database
Class?	does source belong to a known class
ClassSM	get default source model for class
DefltSM	get default source model

Note that boxed items refer to other parts of the SCM which are executed at this point and upon completion processing continues on the original diagram:

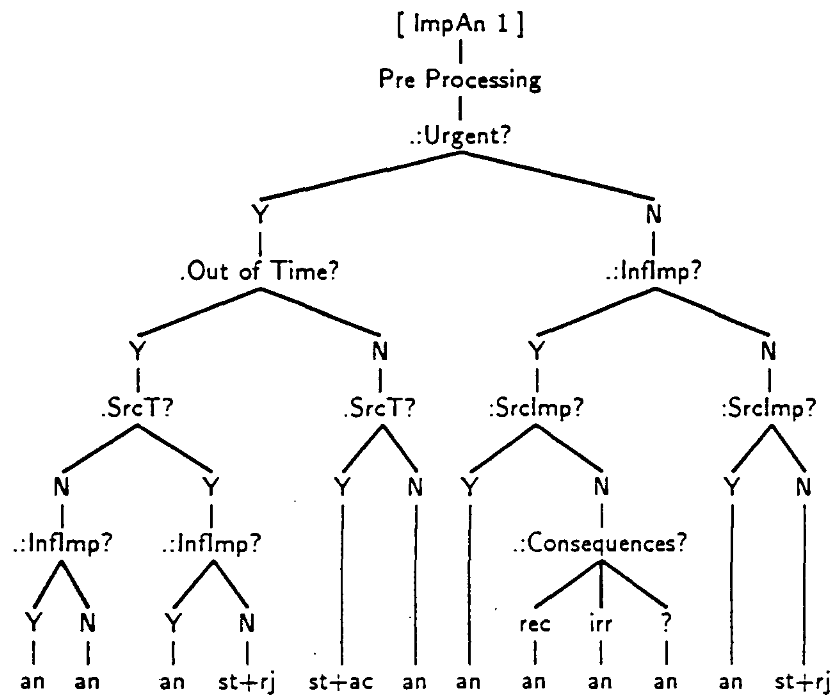
ImpAn 1	Importance Analysis (see [ImpAn 1])
InfoEval	Information Evaluation (see [InfoEval])
ImpAn 2	Importance Analysis (see [ImpAn 2])
ConfRes	Conflict Resolution (see [ConfRes])
Enquiry 1	Enquiry for Information Evaluation (see [Enquiry 1])
ImpAn 2a	Importance Analysis (see [ImpAn 2a])
ImpAn 3	Importance Analysis (see [ImpAn 3])
Enquiry 2	Enquiry for Conflict Resolution (see [Enquiry 2])

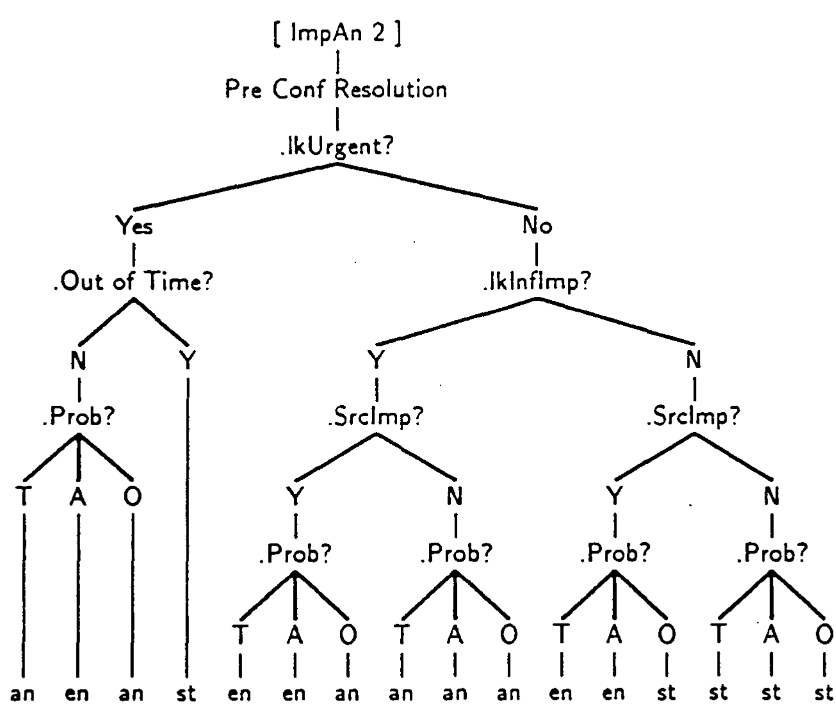

[Cont Reg]
Data Input
CtRg 1.1
Data+SM
ImpAn 1
.Result?
stop
analyse
.AdBel
InfoEval
ImpAn 2
.Result?
stop
analyse
enquire
.AdBel
.:RelInfo?
Enq 1
ImpAn 2a
Y
N
.Result?
ConfRes
stop
ImpAn 3
.AdBel
stop
analyse
.Result?
.AdBel
.:RelInfo?
stop
enquire
Y
N
.AdBel
Enq 2
ConfRes
stop
stop
ImpAn 3
.AdBel
.AdBel
.Result?
stop
enquire
.AdBel
Enq 2
stop
.AdBel
A
B
C
D
E
F
G
H
J

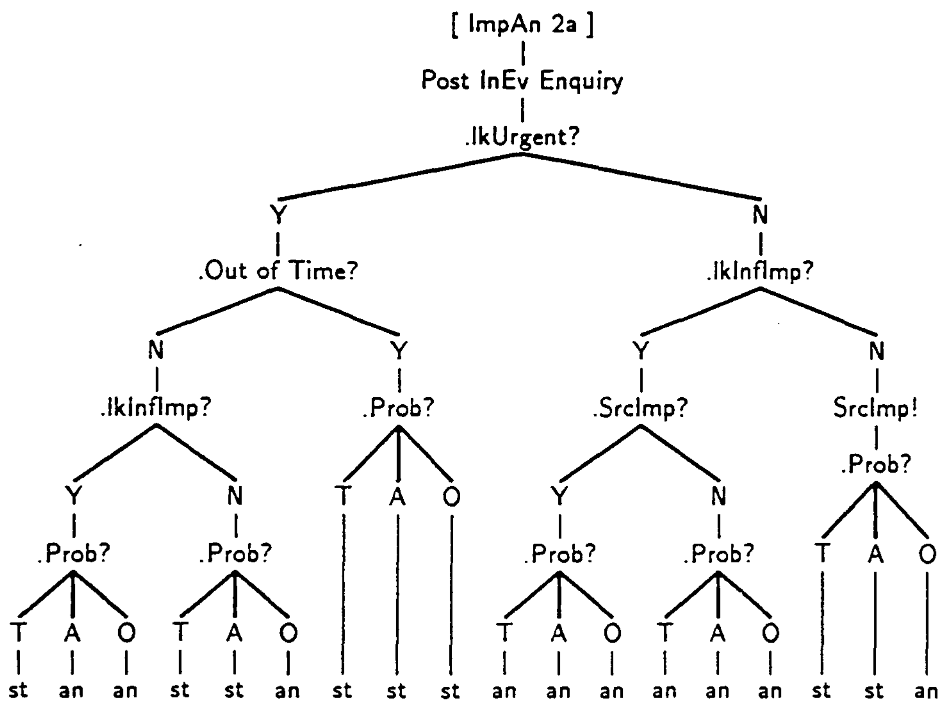

[CtRg 1.1]
Data
.SrcKnown?
Y
N
:Class?
Y
N
.PickSM
.ClassSM
.DefltSM

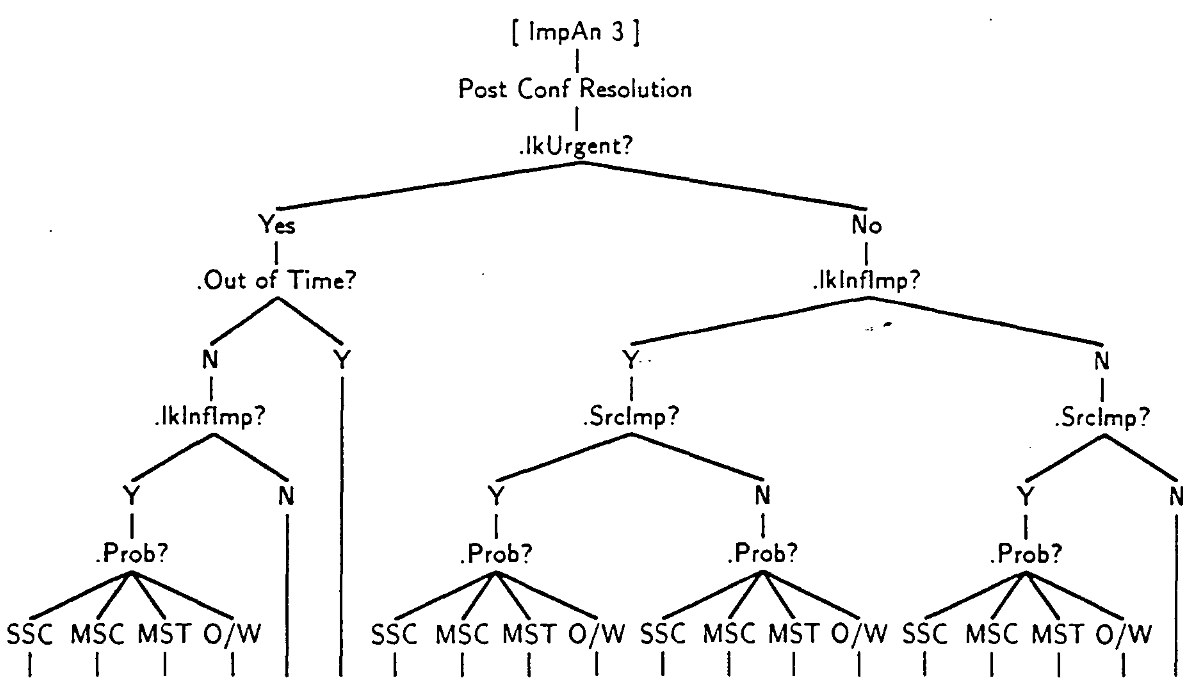
Importance Analysis

[ImpAn 1]	label of the diagram
Pre Processing	a comment that ImpAn 1 applies before InfoEval
Urgent?	determine whether the case is urgent
Y	yes
N	no
Out of Time?	determine whether time left for further analysis
SrcT?	determine whether source is trustworthy
InfImp?	ask the NLP whether information is important
SrcImp?	determine whether source is important to SCM
Consequences?	determine how serious the consequences are
rec	consequences recoverable
irr	consequences irrecoverable
?	indeterminate (don't know)
an	analyse
st+rj	stop and reject information
st+ac	stop and accept the information
[ImpAn 2]	Importance Analysis [ImpAn 2]
Pre Conf Resolution	comment to say that diagram applies before ConfRes
lkUrgent?	look up whether the case is urgent
lkInfImp?	look up whether information is important
Prob?	determine type of problem with information
T	record trust problem
A	record ability problem
O	record o.k. (i.e. no problem)
en	enquire
st	stop
[ImpAn 2a]	Importance Analysis [ImpAn 2a]
Post InEv Enquiry	comment to say that diagram applies after Enquiry 1
SrcImp!	comment; source is important
[ImpAn 3]	
SSC	single source contradiction
MSC	multiple source contradiction
MST	multiple source tie
O/W	otherwise


[ImpAn 1]
Pre Processing
.:Urgent?
Y
N
.Out of Time?
.:InfImp?
Y
N
Y
N
.SrcT?
.SrcT?
:SrcImp?
:SrcImp?
N
Y
Y
N
Y
N
Y
N
.:InfImp?
.:InfImp?
.:Consequences?
Y
N
Y
N
rec
irr
?
an
an
an
st+rj
st+ac
an
an
an
an
an
an
st+rj

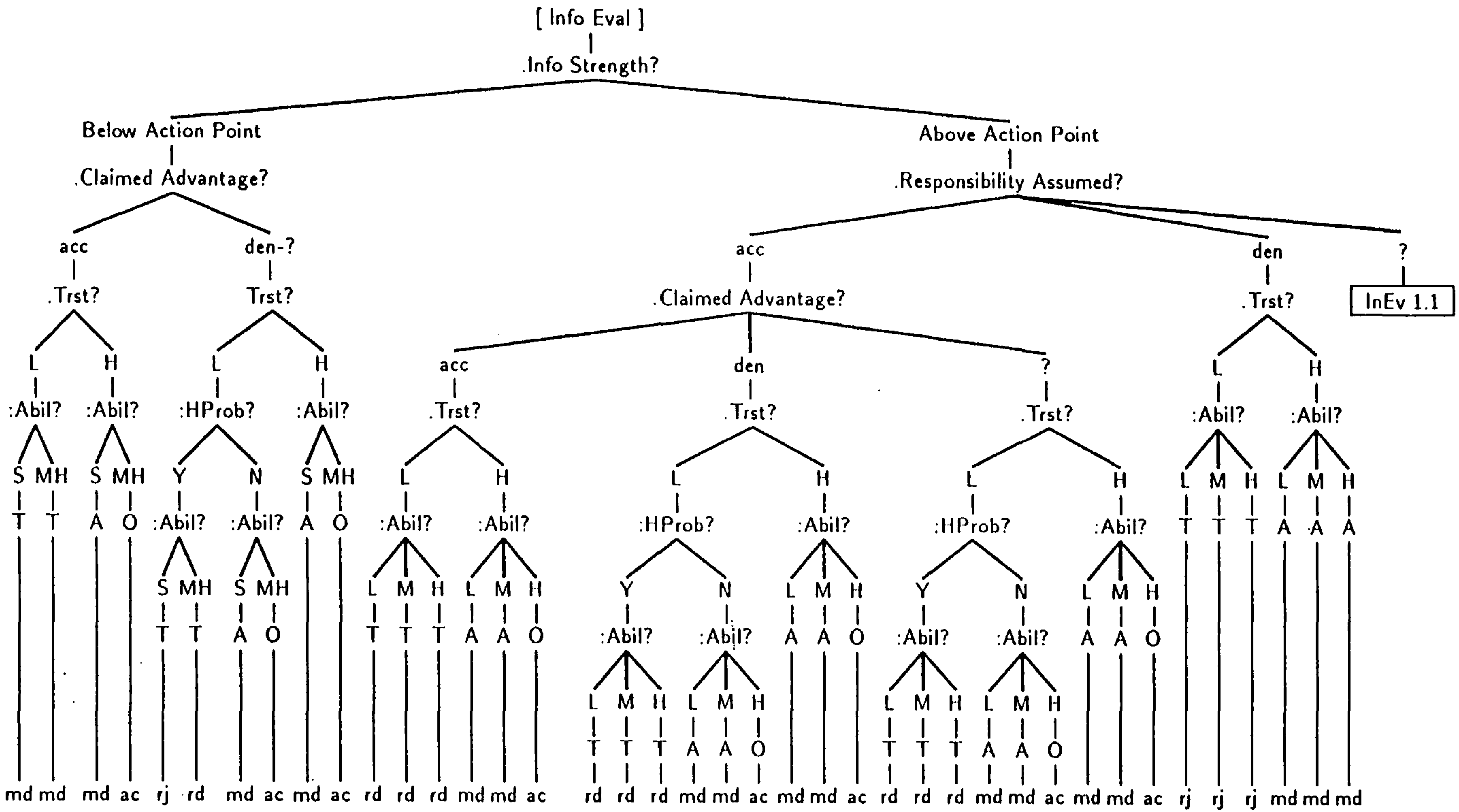

[ImpAn 2]
Pre Conf Resolution
.IkUrgent?
Yes
No
.Out of Time?
.IkInfImp?
N
Y
Y
N
.Prob?
.SrcImp?
.SrcImp?
T
A
O
Y
N
Y
N
.Prob?
.Prob?
.Prob?
.Prob?
T
A
O
T
A
O
T
A
O
T
A
O
an
en
an
st
en
en
an
an
an
an
en
en
st
st
st
st

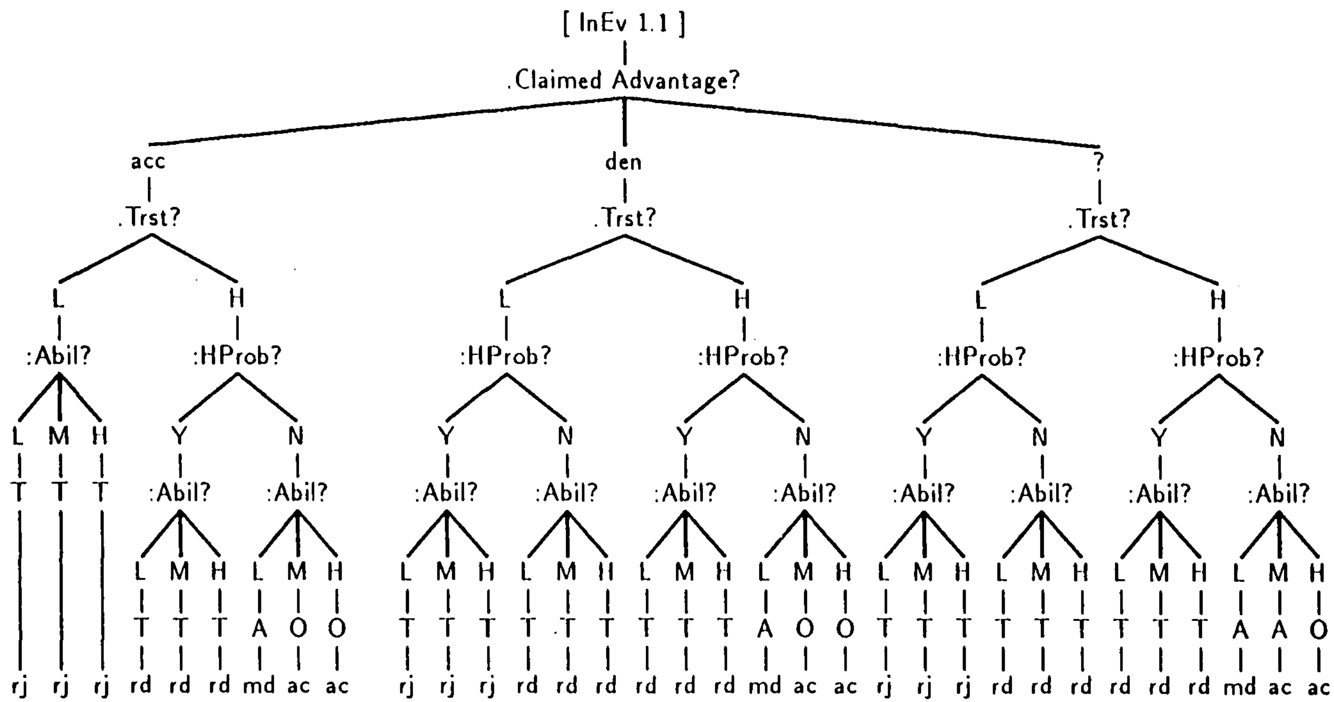

[ImpAn 2a]
Post InEv Enquiry
.IkUrgent?
Y
.Out of Time?
N
.IkInfImp?
Y
.Prob?
T A O
st an an
N
.Prob?
T A O
st st an
Y
.Prob?
T A O
st st st
N
.IkInfImp?
Y
.SrcImp?
Y
.Prob?
T A O
an an an
N
.Prob?
T A O
an an an
N
SrcImp!
.Prob?
T A O
st st an


[ImpAn 3]
Post Conf Resolution
.IkUrgent?
Yes
.Out of Time?
N
.IkInfImp?
Y
.Prob?
SSC MSC MST O/W
en en en st
N
st
Y
st
No
.IkInfImp?
Y
.SrcImp?
Y
.Prob?
SSC MSC MST O/W
en en en st
N
.Prob?
SSC MSC MST O/W
en en en st
N
.SrcImp?
Y
.Prob?
SSC MSC MST O/W
en en en st
N
st

Information Evaluation

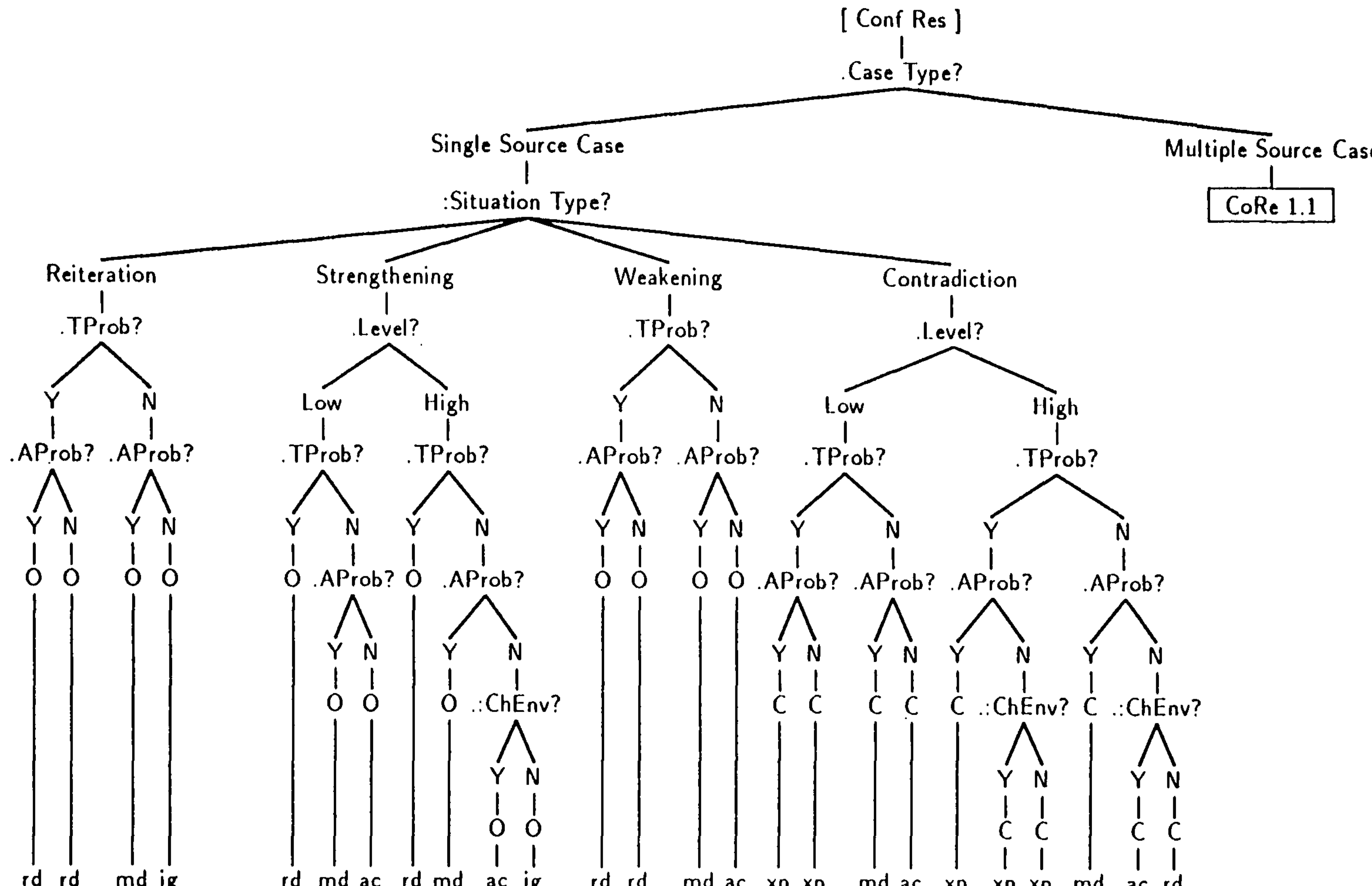
[Info Eval]	label of the diagram
Info Strength?	determine strength of belief
Below Action Point	certainty less than medium
Above Action Point	certainty medium or higher
Claimed Advantage	whether source acknowledges vested interests
acc	accepted
den-?	denied or questionnable
Trst?	level of trustworthiness
Abil?	level of ability
HProb?	is there a problem with helpfulness
L	level is low
M	level is medium or higher
H	level is high
T	record trust problem
A	record ability problem
O	record o.k. (i.e. no problem)
Responsib. Assumed?	whether source assumes responsibility
den	denied
?	questionnable (i.e. don't know)
md	modify belief as a function of source's ability
ac	accept information as given by source
rj	reject information (i.e. belief is nil)
rd	reduce belief to below the action point

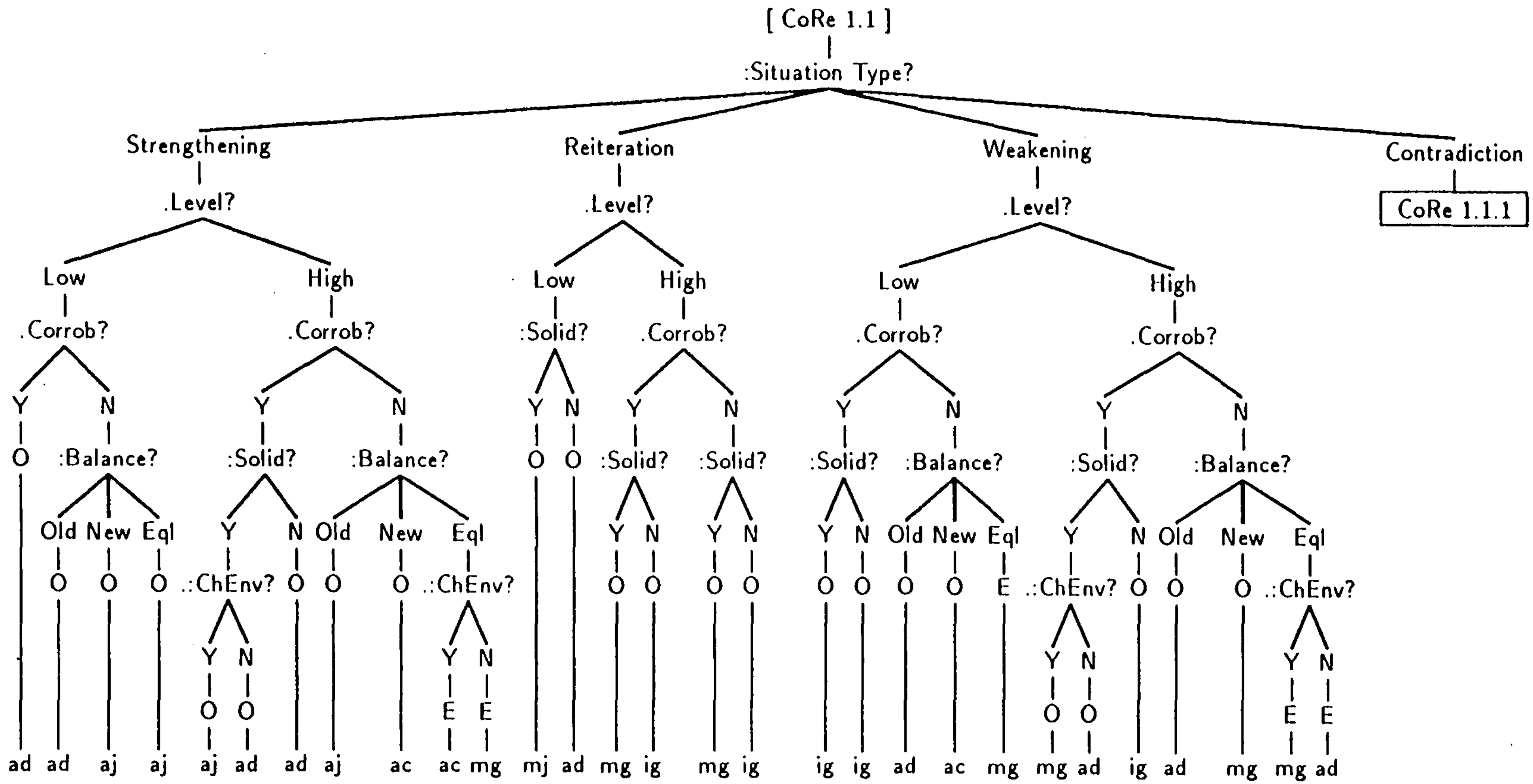

[Info Eval]
.Info Strength?
Below Action Point
Above Action Point
.Claimed Advantage?
.Responsibility Assumed?
acc
den-?
den
?
.Trst?
Trst?
InEv 1.1
:Abil?
:HProb?
L
M
H
S
MH
Y
N
T
A
O
md
ac
rj
rd

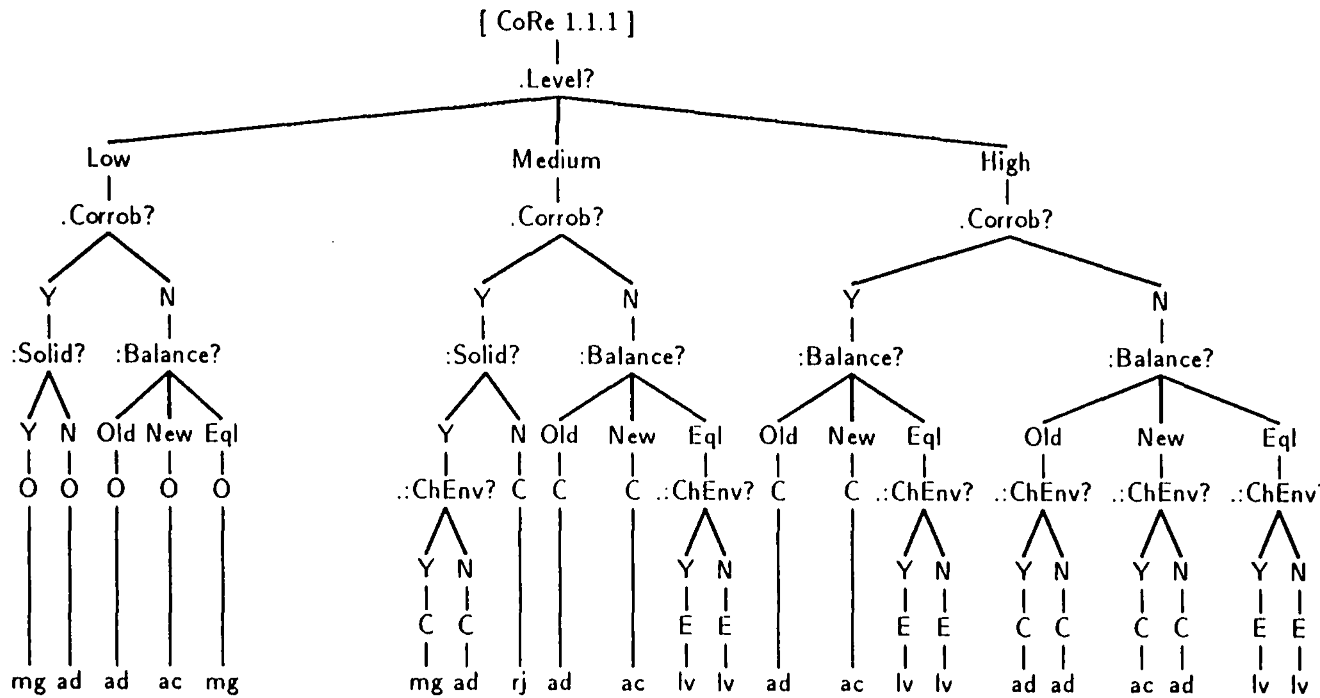

[InEv 1.1]
.Claimed Advantage?
acc
den
?
.Trst?
.Trst?
.Trst?
L H L H L H
:Abil? :HProb? :HProb? :HProb? :HProb? :HProb?
L M H Y N Y N Y N Y N Y N
T T T :Abil? :Abil? :Abil? :Abil? :Abil? :Abil? :Abil? :Abil? :Abil? :Abil?
L M H L M H L M H L M H L M H L M H L M H L M H L M H L M H
T T T A O O T T T T T T T T T A O O T T T T T T T T T A A O
rj rj rj rd rd rd md ac ac rj rj rj rd rd rd rd rd rd md ac ac rj rj rj rd rd rd rd rd rd md ac ac

Conflict Resolution

[Conf Res]	label of the diagram
Case Type?	whether it is a single-source or multiple-source case
Single Source Case	both pieces of information are from same source
Multiple Source Case	the pieces of information are from different sources
Situation Type?	relation between the pieces of information
Reiteration	both are virtually the same
Strengthening	new information stronger (higher certainty) than the old
Weakening	new information weaker than the old
Contradiction	new onformation contradicts the old
Y	yes
N	no
TProb?	is there a problem with trust?
AProb?	is there a problem with ability?
O	record o.k. (i.e. no problem)
C	record contradiction problem
Level?	strength of the stronger piece of information
Low	both pieces of information below medium
High	at least one piece of information at medium or higher
ChEnv?	is a change in the environment plausible
rd	reduce belief to below the action point
md	modify belief as a function of source's ability
ig	ignore information (i.e. retain old piece of information as is)
ac	accept information as given by source
xp	expel both pieces information (i.e. belief is nil - or erase)
[CoRe 1.1]	label of diagram
Corrob?	whether old information is supported by more than one source
Balance?	determine which of the two positions is stronger
Old	old information stronger than new
New	new information stronger than old
Eql	both pieces of information are equally strong
Solid?	whether information is without problem and supported by a strong source
ChEnv?	is a change in the environment plausible
O	record o.k. (i.e. no problem)
C	record contradiction problem
E	record tie problem
ad	add source to list of sources of information
aj	adjust belief of information according to relative strength of sources
mg	merge both pieces of information
[CoRe 1.1.1]	label of diagram
Low	both pieces of information below medium
Medium	at least one piece of information at
High	at least one piece of information at higher level
lv	set belief at even level (i.e. low/low or medium/medium)


[Conf Res]
.Case Type?
Single Source Case
Multiple Source Case
:Situation Type?
CoRe 1.1
Reiteration
Strengthening
Weakening
Contradiction
.TProb?
.Level?
.TProb?
.Level?
Y
N
Low
High
.AProb?
.TProb?
O
C
.ChEnv?
rd
md
ig
ac
xp


[CoRe 1.1]
:Situation Type?
Strengthening
Reiteration
Weakening
Contradiction
CoRe 1.1.1
.Level?
Low
High
.Corrob?
:Solid?
:Balance?
Old
New
Eql
.:ChEnv?
ad ad aj aj aj ad ad aj ac ac mg mj ad mg ig mg ig ig ig ad ac mg mg ad ig ad mg mg ad

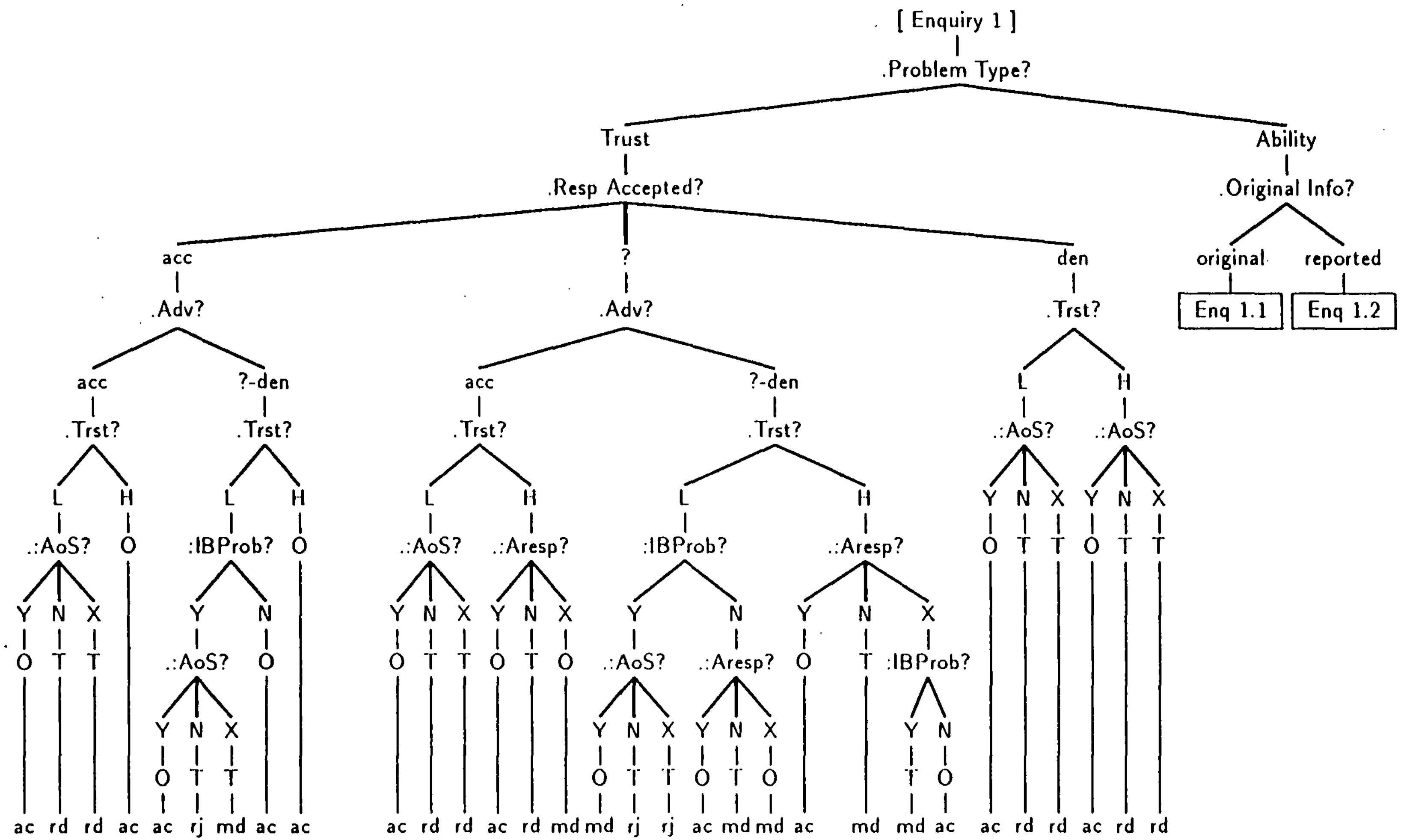

[CoRe 1.1.1]
.Level?
Low
Medium
High
.Corrob?
.Corrob?
.Corrob?
Y N
Y N
Y N
:Solid? :Balance?
:Solid? :Balance?
:Balance? :Balance?
Y N Old New Eql
O O O O O
mg ad ad ac mg
Y N Old New Eql
.:ChEnv? C C C .:ChEnv?
Y N Y N
C C E E
mg ad rj ad ac lv lv
Old New Eql
C C .:ChEnv?
Y N
E E
ad ac lv lv
Old New Eql
.:ChEnv? .:ChEnv? .:ChEnv?
Y N Y N Y N
C C C C E E
ad ad ac ad lv lv

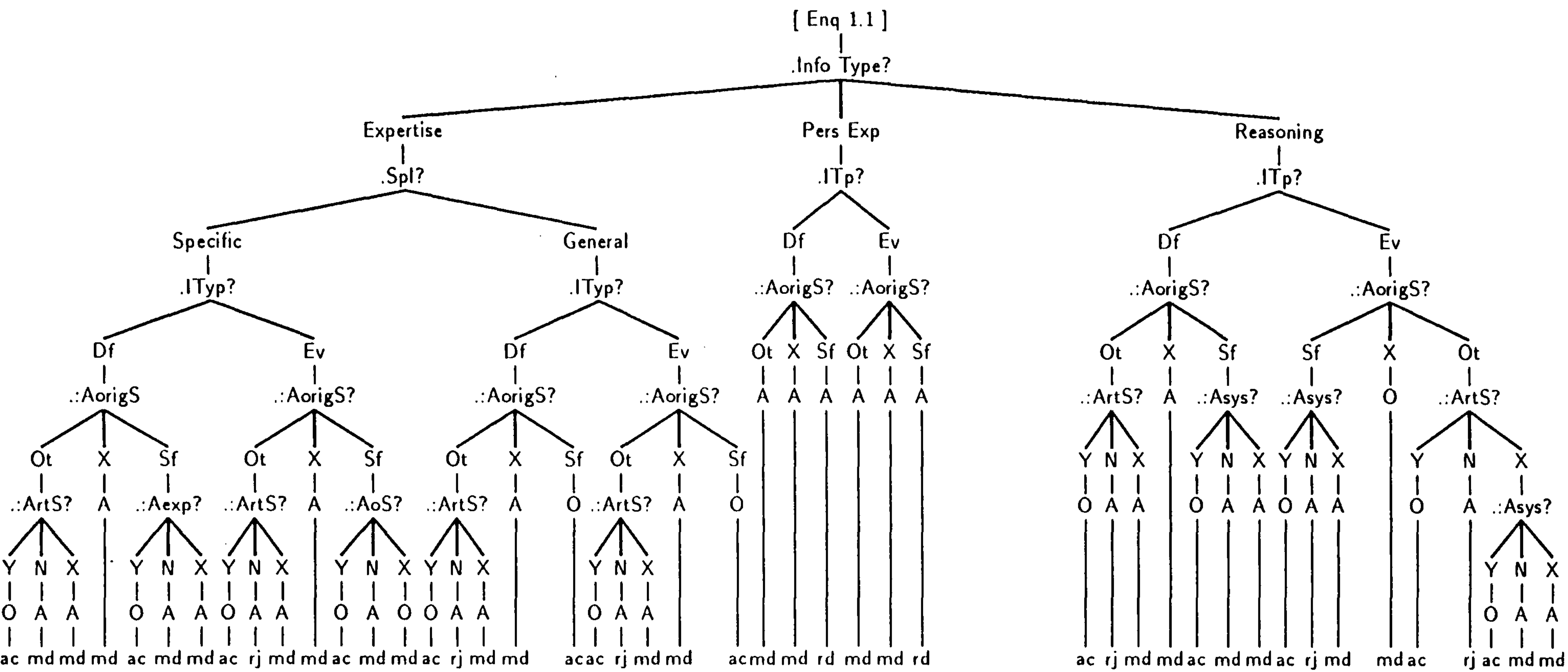
Enquiry 1

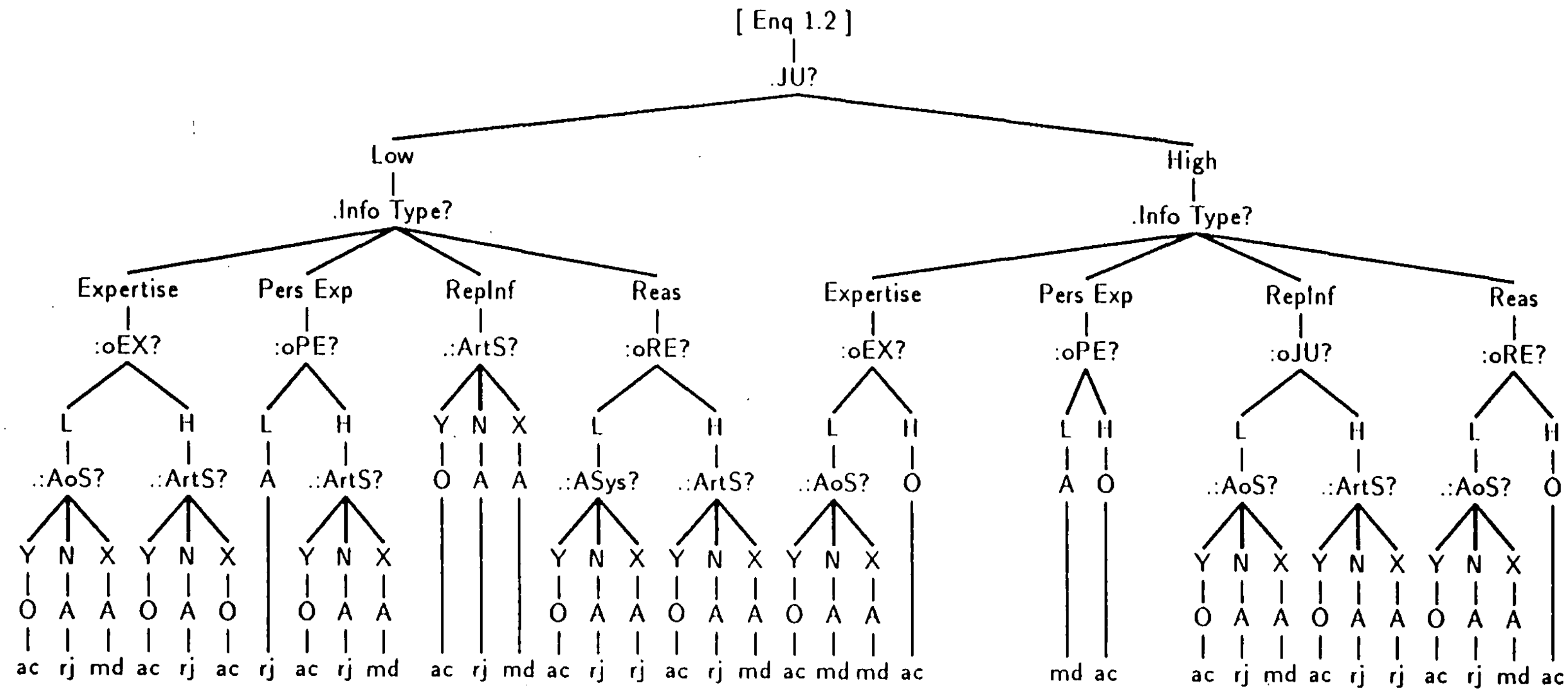
[Enquiry 1]	label of the diagram
Problem Type?	determine the type of problem
Trust	problem of trust
Ability	problem of ability
Responsib. Accepted?	whether source acceptss responsibility
acc	accepted
?	questionnable (i.e. don't know)
den	denied
Adv?	whether the source admits advantage
?-den	denied or questionnable
Trst?	level of trustworthiness of source
L	level of below medium
H	level of medium and above
AoS?	ask another, independent, reliable source for confirmation
Y	yes (i.e. confirmed)
N	no (i.e. denied)
X	request refused by the NLP
O	record o.k. (i.e. no problem)
T	record trust problem
IBProb?	whether there is a problem with interests or beliefs
Aresp?	ask source to accept responsibility
Original Info?	is the information originally from the source
original	information appears to be from the source
reported	information reported from another source
ac	accept information as given by source
rd	reduce belief to below the action point
rj	reject information (i.e. belief is nil)
md	modify belief as a function of source's ability

[Enq 1.1]	label of the diagram
Info Type?	ability category of the information
Expertise	ability type - expertise
Pers Exp	ability type - personal experience
Reasoning	ability type - reasoning
Spl?	whether the source model contains a special index for the subject
Specific	special index exists
General	general index used
ITyp?	index type
Df	index based on default
Ev	index is built on concrete evidence
AorigS?	ask source whether information is original
Ot	information is from someone else
X	request refused by the NLP

Sf	information is from source itself
ArtS?	ask the original source for confirmation
Aexp?	ask source whether it has expertise
A	problem of ability
Asys?	ask system (NLP) to check soundness of reasoning
[Enq 1.2]	label of the diagram
JU?	level of ability to judge sources
Low	level of index is lower than medium
High	level of index is medium or higher
Info Type?	type of information and ability required
RepInf	reported information from another source
Reas	information based on reasoning
oEX?	level of expertise index of original source
oPE?	level of personal experience index of original source
oRE?	level of reasoning index of original source
oJU?	level of source judging index of original source


[Enquiry 1]
.Problem Type?
Trust
Ability
.Resp Accepted?
.Original Info?
acc
?
den
original
reported
Enq 1.1
Enq 1.2
.Adv?
.Adv?
.Trst?
acc
?-den
acc
?-den
L
H
.Trst?
.Trst?
.Trst?
.Trst?
.:AoS?
.:AoS?
L
H
L
H
L
H
L
H
Y
N
X
Y
N
X
.:AoS?
O
:IBProb?
O
.:AoS?
.:Aresp?
:IBProb?
.:Aresp?
O
T
T
O
T
T
Y
N
X
Y
N
Y
N
X
Y
N
X
Y
N
Y
N
X
O
T
T
.:AoS?
O
O
T
T
O
T
O
.:AoS?
.:Aresp?
O
T
:IBProb?
Y
N
X
Y
N
X
Y
N
X
Y
N
O
T
T
O
T
T
O
T
O
T
O
ac rd rd ac ac rj md ac ac
ac rd rd ac rd md md rj rj ac md md ac md md ac
ac rd rd ac rd rd

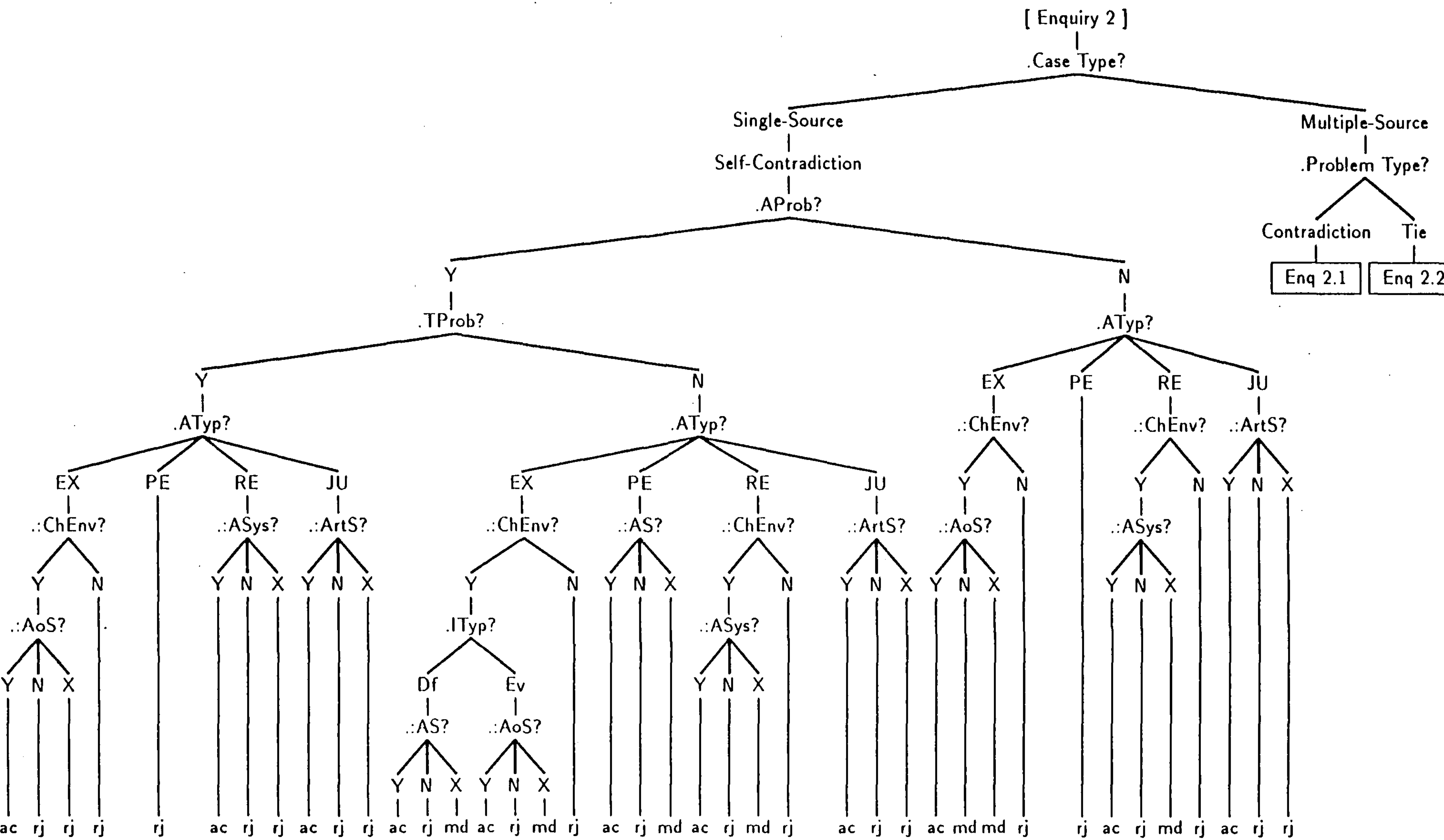

[Enq 1.1]
.Info Type?
Expertise
Pers Exp
Reasoning
.Spl?
.ITp?
.ITp?
Specific
General
Df
Ev
Df
Ev
.ITyp?
.ITyp?
.:AorigS?
.:AorigS?
.:AorigS?
.:AorigS?
Df
Ev
Df
Ev
Ot X Sf Ot X Sf
Ot X Sf Sf X Ot
A A A A A A
.:AorigS
.:AorigS?
.:AorigS?
.:AorigS?
.:ArtS? A .:Asys? .:Asys? O .:ArtS?
Ot X Sf Ot X Sf Ot X Sf Ot X Sf
Y N X Y N X Y N X Y N X
.:ArtS? A .:Aexp? .:ArtS? A .:AoS? .:ArtS? A O .:ArtS? A O
O A A O A A O A A O A .:Asys?
Y N X Y N X Y N X Y N X Y N X
Y N X
O A A O A A O A A O A O O A A O A A
O A A
ac md md md ac md md ac rj md md ac md md ac rj md md acac rj md md acmd md rd md md rd
ac rj md md ac md md ac rj md md ac rj ac md md

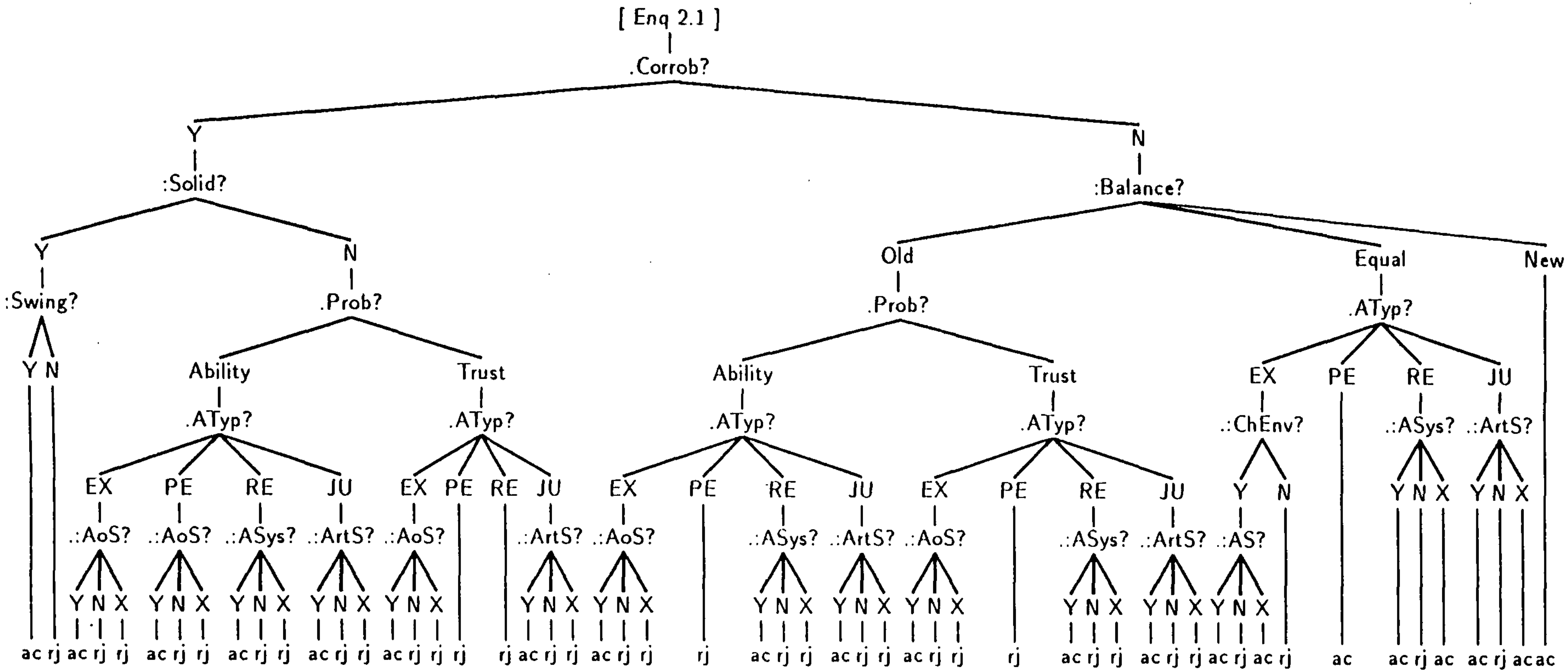

[Enq 1.2]
.JU?
Low
.Info Type?
Expertise
:oEX?
Pers Exp
:oPE?
RepInf
.:ArtS?
Reas
:oRE?
.:AoS?
.:ArtS?
.:ASys?
High
.Info Type?
Expertise
:oEX?
Pers Exp
:oPE?
RepInf
:oJU?
Reas
:oRE?
ac rj md ac rj ac rj ac rj md ac rj md ac rj rj ac rj md ac md md ac
md ac ac rj md ac rj rj ac rj md ac

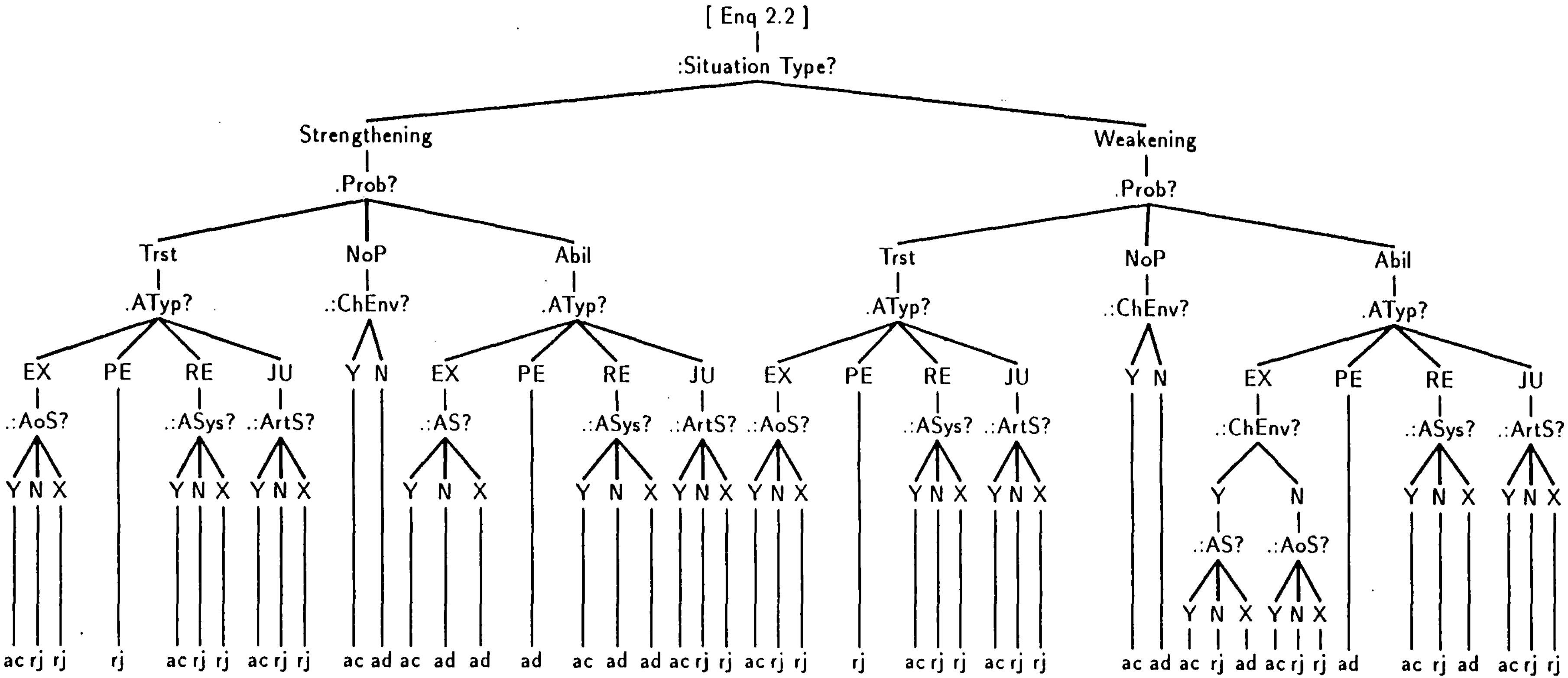
Enquiry 2

[Enquiry 2]	label of the diagram
Case Type?	type of relation between pieces of information
Single-Source	both pieces of information from the same source
Multiple-Source	the pieces of information are from different sources
Self-Contradiction	comment that it is a self-contradiction problem
AProb?	whether there is a problem of ability
Y	yes
N	no
TProb?	whether there is a problem of trust
ATyp?	type of ability
EX	expertise
PE	personal experience
RE	reasoning ability
JU	judging sources
ChEnv?	whether it is likely that there is a change in the environment
AoS?	ask another, independent reliable source for confirmation
X	request denied by NLP
ASys?	ask the NLP to check reasoning
ITyp?	index type; whether default of based on evidence
Df	default index
Ev	index built on evidence
ArtS?	ask the original source for confirmation
AS?	ask source for confirmation
Problem Type?	whether problem is a tie or contradiction
Contradiction	problem of contradiction
Tie	problem of tie between pieces of information
ac	accept information as given by source
rj	reject information (i.e. belief is nil)
md	modify belief as a function of source's ability
[Enq 2.1]	label of the diagram
Corrob?	whether old information is corroborated
Y	yes
N	no
Solid?	whether the new information is without problem and supported by strong source
Swing?	whether the new information swings the balance of the corroborated information
Prob?	whether problem is ability or trust
Ability	problem of ability
Trust	problem of trust
Balance?	which of the two pieces of information wins on balance
Old	the old information
Equal	both pieces of information are equally strong
New	the new information

[Enq 2.2]	label of the diagram
Situation Type?	type of relation between pieces of information
Strengthening	new information reinforces the old
Weakening	new information weakens the old
Prob?	what type of problem there is with the new information
Trst	problem of trust
NoP	no problem
Abil	problem of ability
ad	add source to list of supporting sources of information


[Enquiry 2]
.Case Type?
Single-Source
Multiple-Source
Self-Contradiction
.Problem Type?
.AProb?
Contradiction
Tie
Enq 2.1
Enq 2.2
Y
N
.TProb?
.ATyp?
Y
N
EX
PE
RE
JU
.ATyp?
.ATyp?
.:ChEnv?
.:ChEnv?
.:ArtS?
EX
PE
RE
JU
EX
PE
RE
JU
Y
N
Y
N
Y
N
X
.:ChEnv?
.:ASys?
.:ArtS?
.:ChEnv?
.:AS?
.:ChEnv?
.:ArtS?
.:AoS?
.:ASys?
Y
N
Y
N
X
Y
N
X
Y
N
Y
N
X
Y
N
Y
N
X
Y
N
X
Y
N
X
.:AoS?
.ITyp?
.:ASys?
Y
N
X
Df
Ev
Y
N
X
.:AS?
.:AoS?
Y
N
X
Y
N
X
ac
rj
rj
rj
rj
ac
rj
rj
ac
rj
rj
ac
rj
md
ac
rj
md
rj
ac
rj
md
ac
rj
md
rj
ac
rj
rj
ac
md
md
rj
rj
ac
rj
md
rj
ac
rj
rj


[Enq 2.1]
.Corrob?
Y
N
:Solid?
:Balance?
Y
N
Old
Equal
New
:Swing?
.Prob?
.Prob?
.ATyp?
Y N
Ability
Trust
Ability
Trust
EX
PE
RE
JU
.ATyp?
.ATyp?
.ATyp?
.ATyp?
.:ChEnv?
.:ASys?
.:ArtS?
EX PE RE JU
EX PE RE JU
EX PE RE JU
EX PE RE JU
Y N
Y N X
Y N X
.:AoS?
.:AoS?
.:ASys?
.:ArtS?
.:AoS?
.:ArtS?
.:AoS?
.:ASys?
.:ArtS?
.:AoS?
.:ASys?
.:ArtS?
.:AS?
Y N X
ac rj ac rj rj ac rj rj ac rj rj ac rj rj ac rj rj rj rj ac rj rj ac rj rj rj ac rj rj ac rj rj ac rj rj rj ac rj rj ac rj rj ac rj ac rj ac ac rj ac ac rj ac ac


[Enq 2.2]
:Situation Type?
Strengthening
.Prob?
Trst
.ATyp?
EX
.:AoS?
Y N X
ac rj rj
PE
rj
RE
.:ASys?
Y N X
ac rj rj
JU
.:ArtS?
Y N X
ac rj rj
NoP
.:ChEnv?
Y N
ac ad
Abil
.ATyp?
EX
.:AS?
Y N X
ac ad ad
PE
ad
RE
.:ASys?
Y N X
ac ad ad
JU
.:ArtS?
Y N X
ac rj rj
Weakening
.Prob?
Trst
.ATyp?
EX
.:AoS?
Y N X
ac rj rj
PE
rj
RE
.:ASys?
Y N X
ac rj rj
JU
.:ArtS?
Y N X
ac rj rj
NoP
.:ChEnv?
Y N
ac ad
Abil
.ATyp?
EX
.:ChEnv?
Y
.:AS?
Y N X
ac rj ad
N
.:AoS?
Y N X
ac rj rj
PE
ad
RE
.:ASys?
Y N X
ac rj ad
JU
.:ArtS?
Y N X
ac rj rj

Appendix B

Source Reevaluation Diagrams

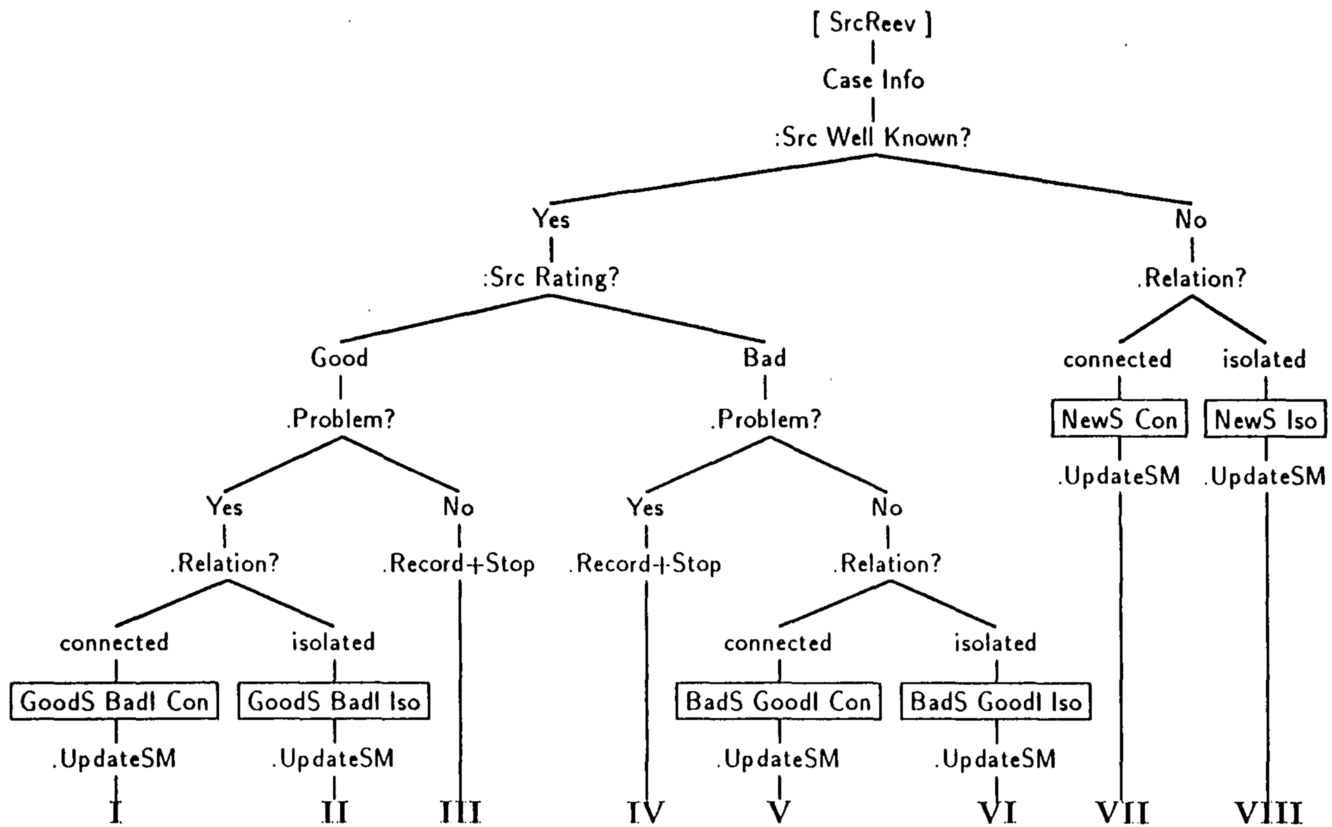
This appendix contains the decision trees corresponding to the source reevaluation component of the Source Control Mechanism. The decision trees are again divided in separate sections, corresponding to the various subcomponents as detailed in the principles and design chapters and preceeded by a legend, explaining the items on the decision trees.

Source Reevaluation Control

[SrcReev]	label of the diagram; control level of the source reevaluation
Case Info	comment to say that input to the process is the case information
Src Well Known?	whether the source is well known to the SCM
Src Rating?	whether the source is considered a good or a bad source
Good	source considered to be good
Bad	source considered to be bad
Problem?	whether there is a problem with the information
Relation?	whether the information is related to other information
connected	information is connected to other information
isolated	information is isolated; no connection
UpdateSM	update source model according to recommendations
Record+Stop	record incident in evidence but no full reevaluation
I-VIII	label to identify the different paths

Note that boxed items refer to other parts of the Source Reevaluation which are executed at this point and upon completion processing continues on the original diagram:

GoodS BadI Con	good source, bad information, connected
GoodS BadI Iso	good source, bad information, isolated
BadS GoodI Con	bad source, good information, connected
BadS GoodI Iso	bad source, good information, isolated
NewS Con	new source, connected
NewS Iso	new source, isolated


[SrcReev]
Case Info
:Src Well Known?
Yes
No
:Src Rating?
.Relation?
Good
Bad
connected
isolated
.Problem?
.Problem?
NewS Con
NewS Iso
.UpdateSM
.UpdateSM
Yes
No
Yes
No
.Relation?
.Record+Stop
.Record+Stop
.Relation?
connected
isolated
connected
isolated
GoodS BadI Con
GoodS BadI Iso
BadS GoodI Con
BadS GoodI Iso
.UpdateSM
.UpdateSM
.UpdateSM
.UpdateSM
I
II
III
IV
V
VI
VII
VIII

Source Reevaluation Analysis

[GoodS BadI Con]	label of the diagram
Type?	whether it is a single or multiple source situation
Case Type?	whether single source or multiple source
SSMIP	single source multiple information package
MSMIP	multiplesource multiple information package
Situation Type?	relation between pieces of information
Weak	new information is weakening the old
Cntr	new information is contradicting the old
Prob?	type of problem with the information
Trust	problem of trust
Abil	problem of ability
Pat?	whether there is a pattern in source's past performance to suggest that current index is maladjusted
Y	yes
N	no
Frc?	whether there is a fracture in the source's past performance to suggest that there is a need for an additional index
ChEnv?	whether a change in the environment is plausible
Corrob?	whether the old information is corroborated
Solid?	whether the new information is without problems and supported by a strong source
ci	change index
si	split index
rs	record and stop

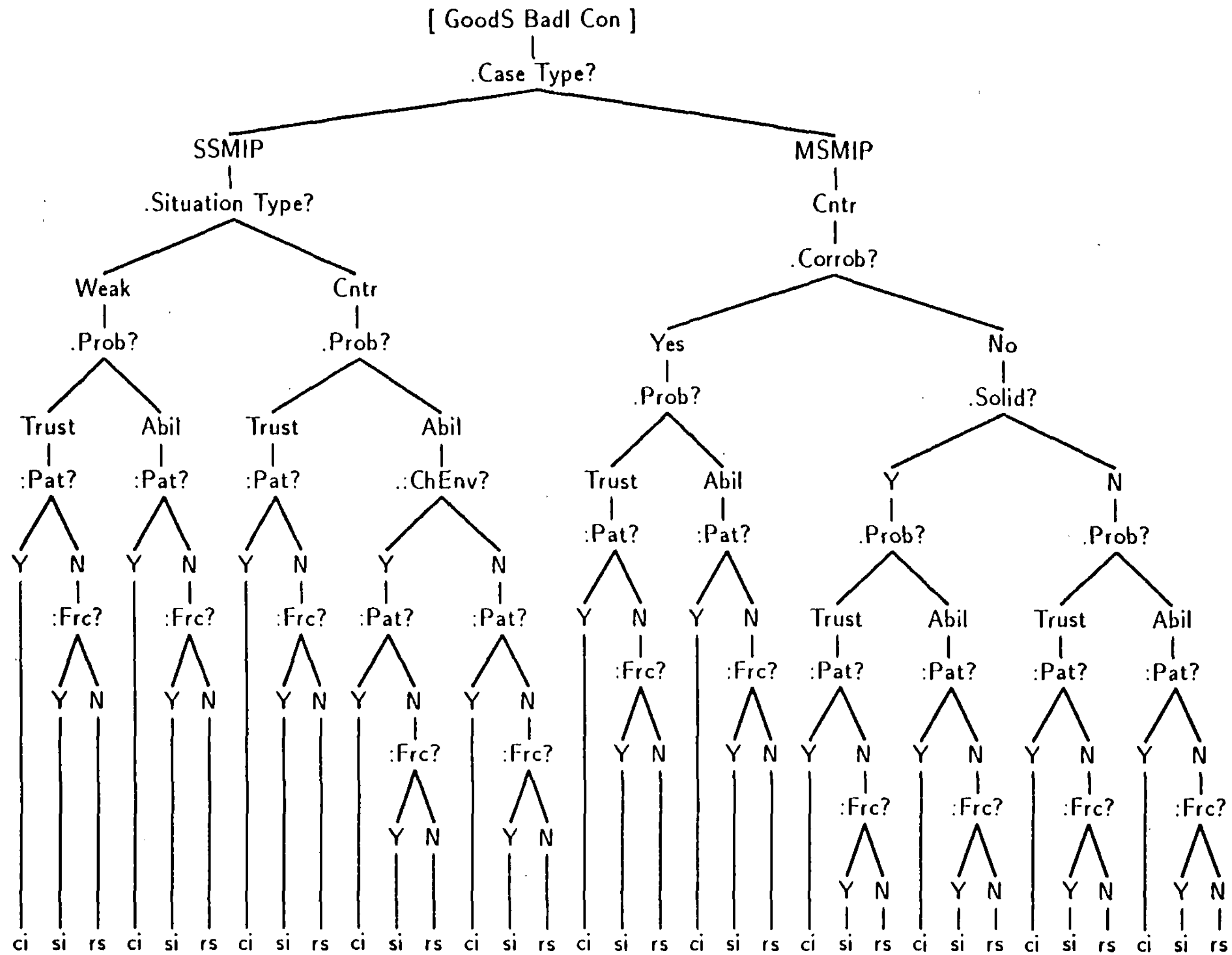
[GoodS BadI Iso]	label of the diagram
Prob Type?	what type of problem there is
Resp	problem with responsibility
Adv	problem with advantage
HlpProb?	whether there is a problem of helpfulness
RelTyp?	whether trust relation type is general or specific
Cls?	whether the pattern is connected with a known class
IntBel?	whether interests or beliefs are involved
AType?	what ability type is required for the information
Expertise	expertise
Reasoning	ability to reason
Judg Srces	ability to judge sources
Pers Exp	ability to handle personal experience
IndTyp?	type of index; whether general or specific
gen	general index
spec	specific index
nc	introduction of a new index for the class

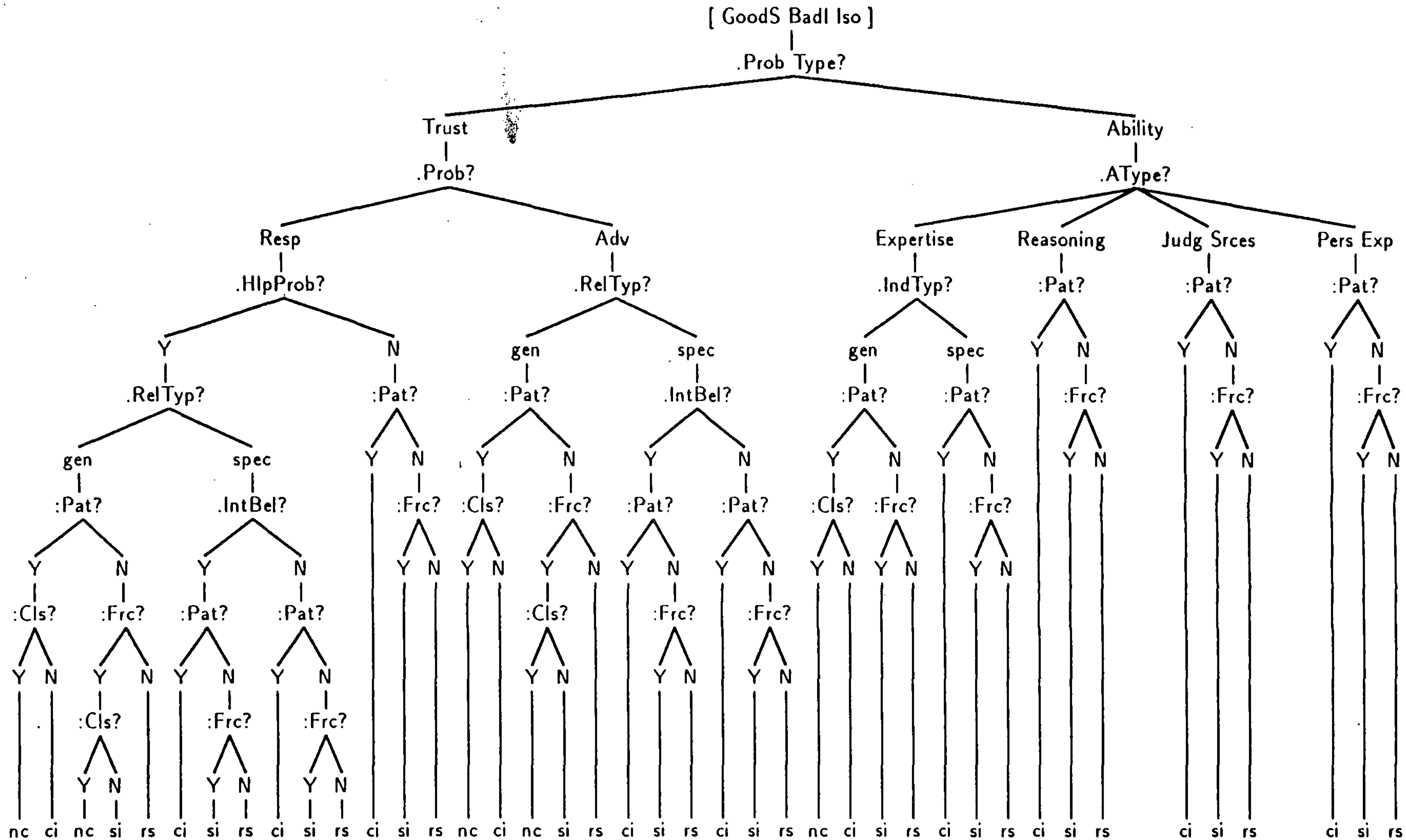
[BadS GoodI Con]	label of the diagram
Typ?	whether belief was based on trust or ability
Reit	new information reiterates the old
Strng	new information strengthens the old

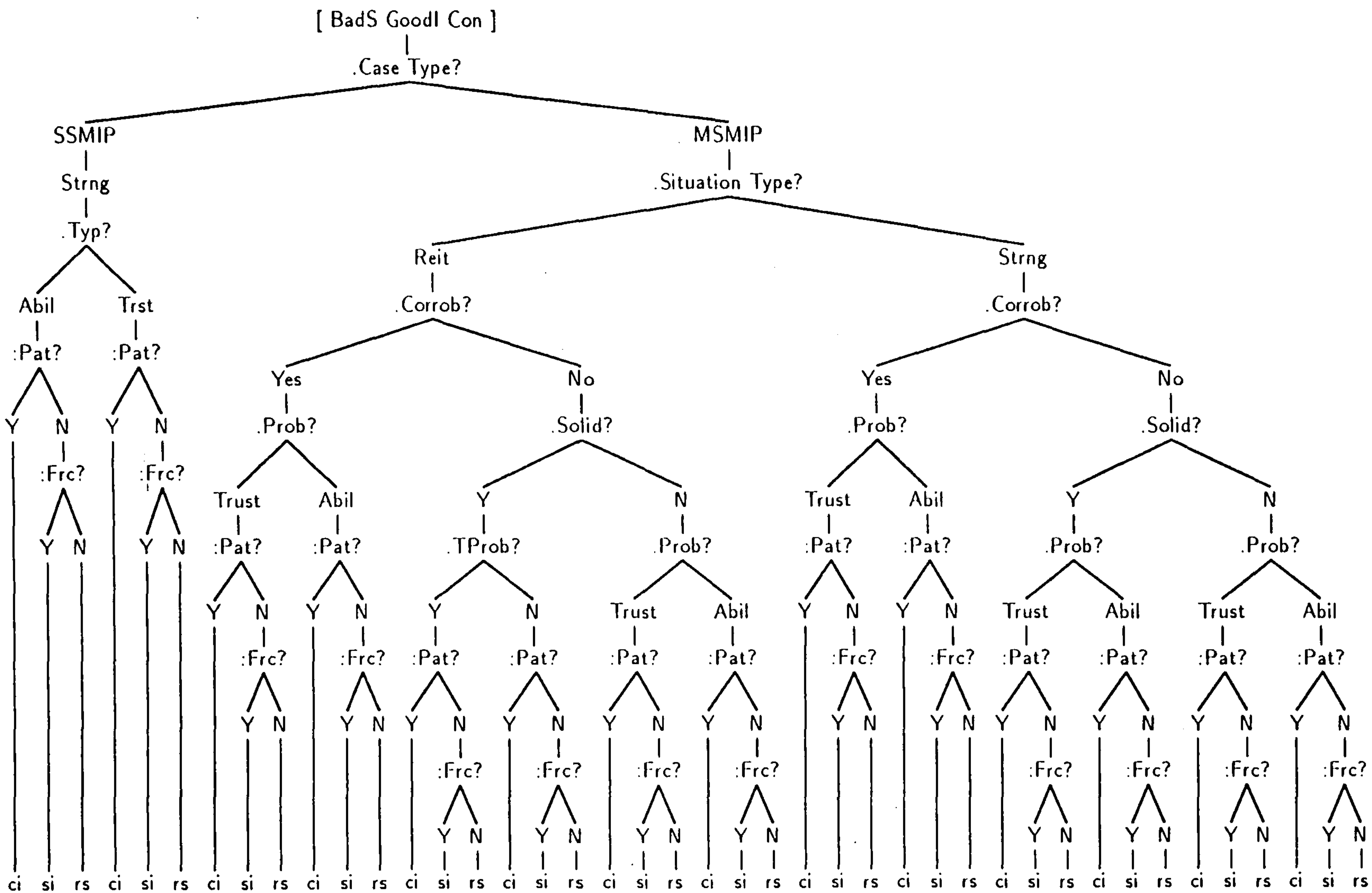
[BadS GoodI Iso]	label of the diagram
Type?	whether belief was based on trust or ability
IndTyp?	whether the index employed was a general or specific index

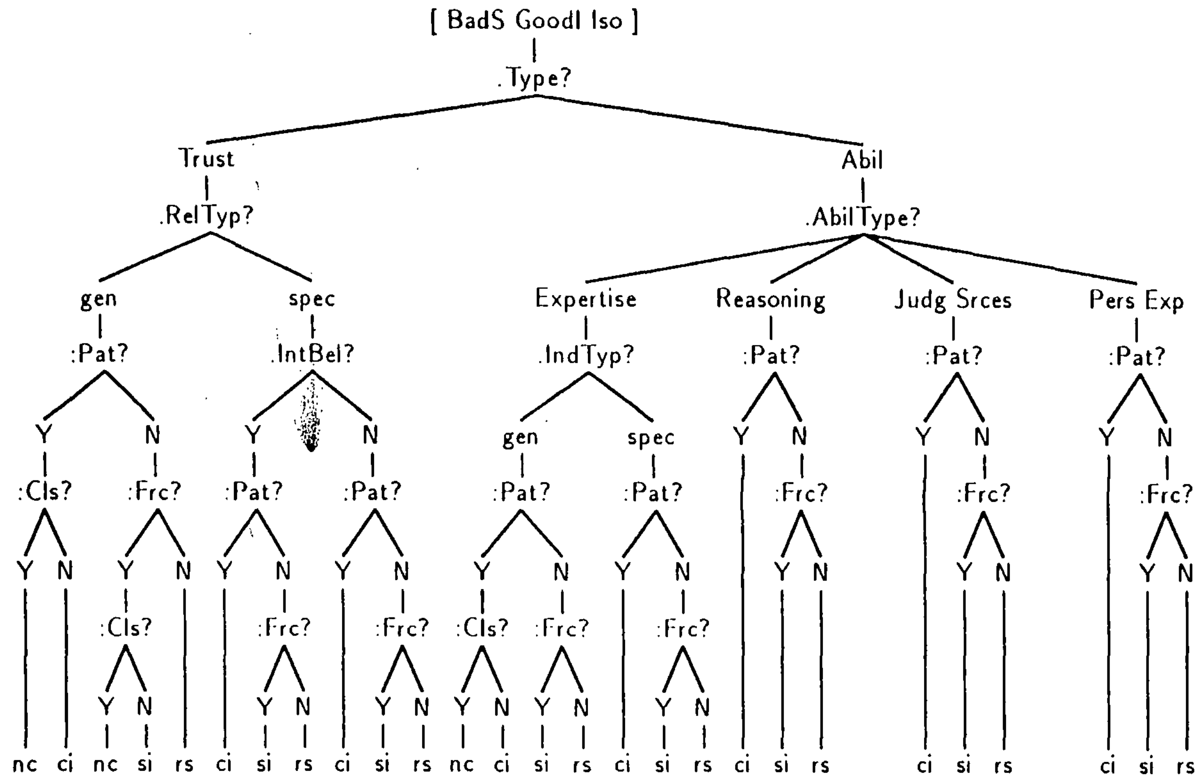
[NewS Con]	label of the diagram
Evd?	whether the case is strong enough evidence to replace a default index
ri	replace default index with index based on evidence
[SoRe 1.1]	label of the diagram

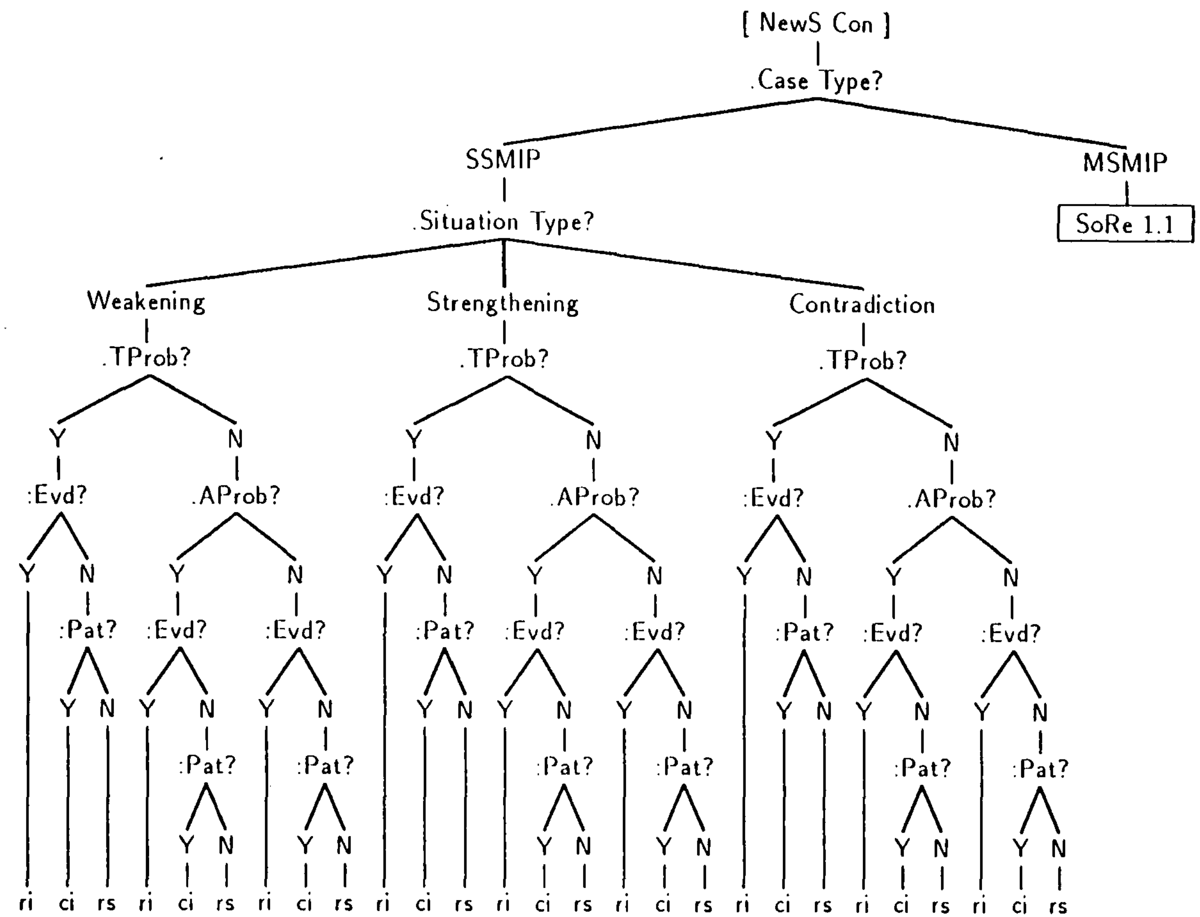
[NewS Iso]	label of the diagram
EX	expertise
RE	reasoning
JU	judging sources
PE	personal experience
Resp	problem of lack of responsibility
LowI	low index; not enough ability

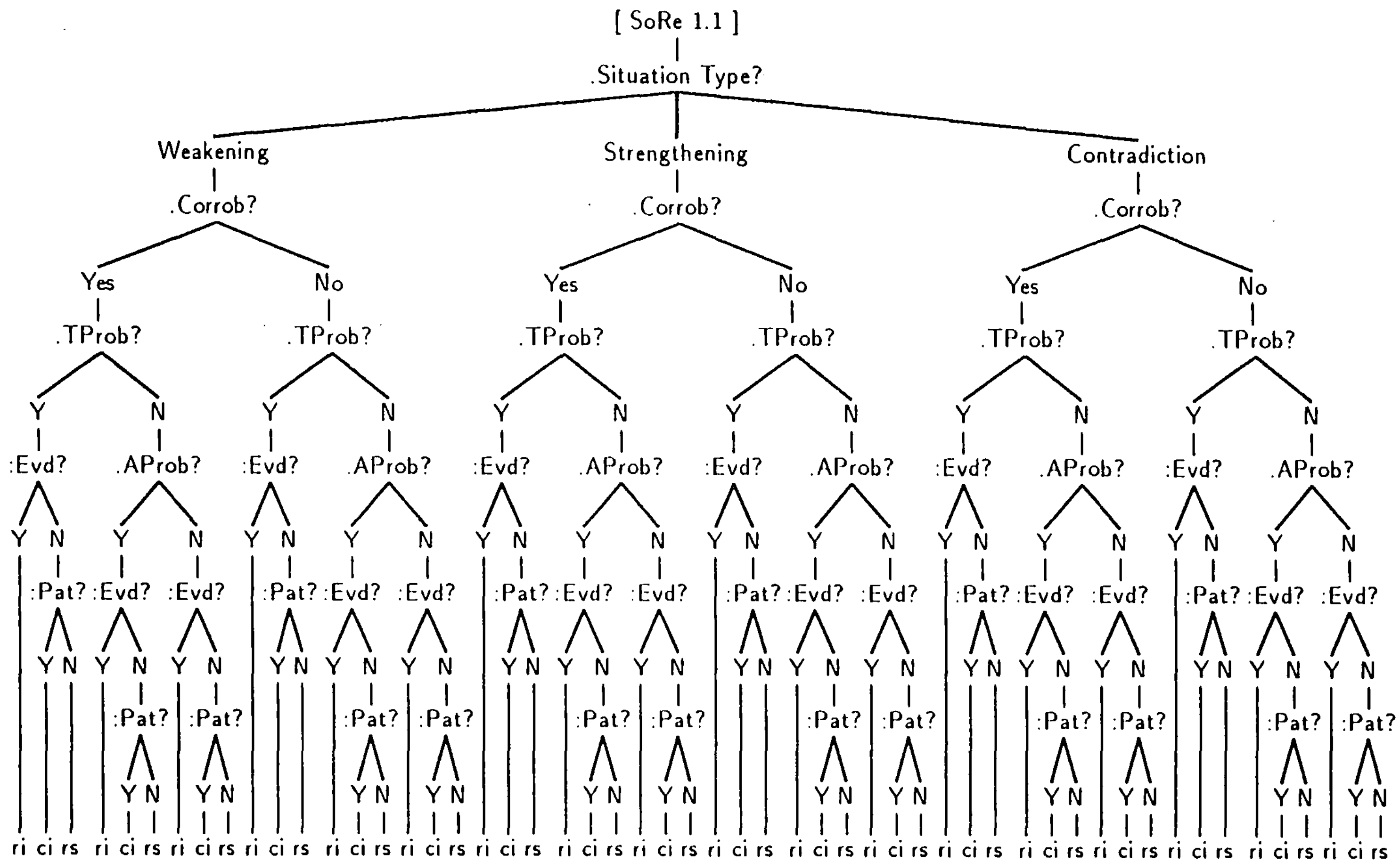

[GoodS BadI Con]
.Case Type?
SSMIP
.Situation Type?
Weak
.Prob?
Trust
:Pat?
Abil
Cntr
:ChEnv?
:Frc?
MSMIP
Cntr
.Corrob?
Yes
No
.Solid?
Y
N
ci
si
rs

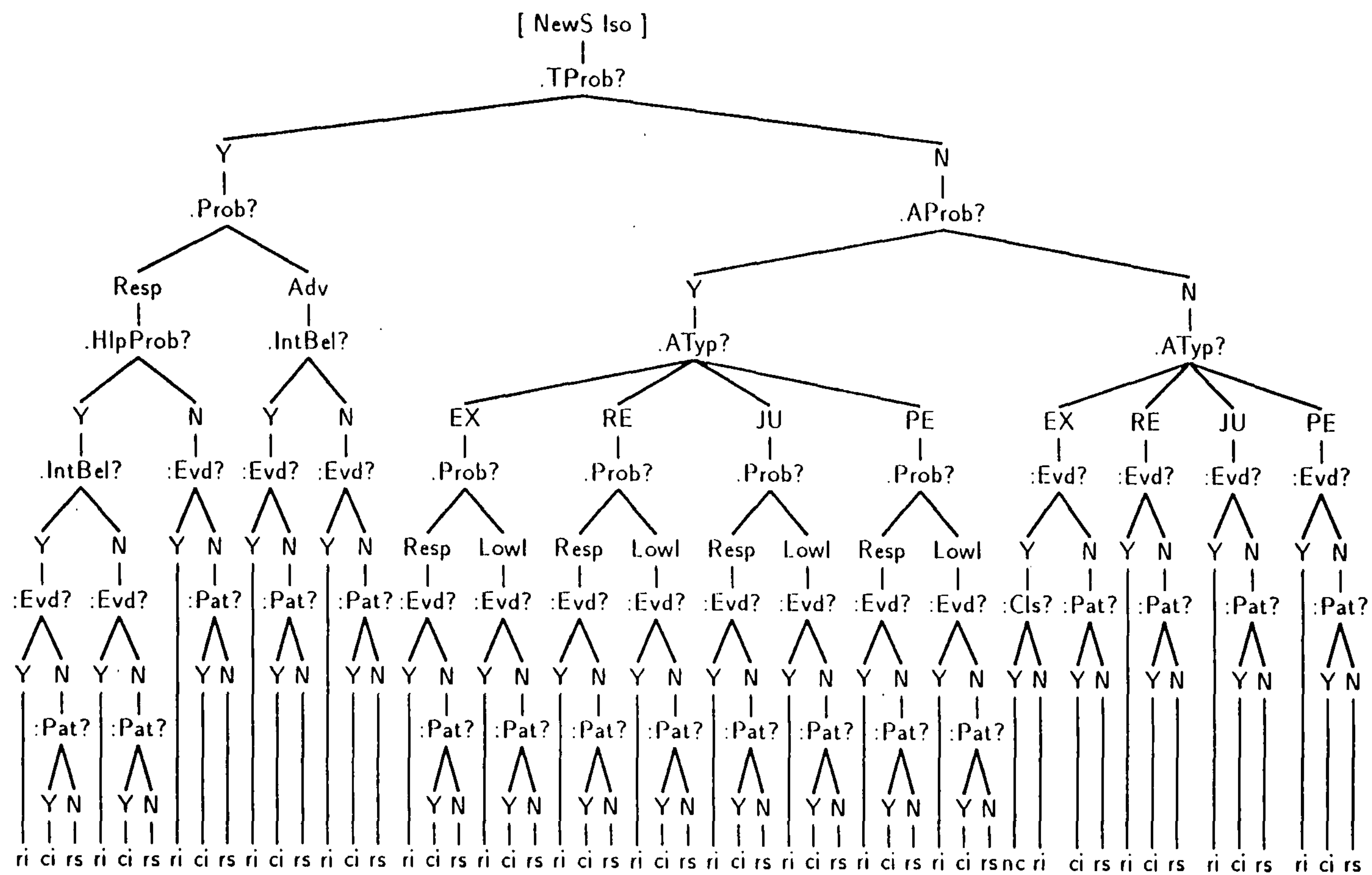

[GoodS BadI Iso]
.Prob Type?
Trust
Ability
.Prob?
.AType?
Resp
Adv
Expertise
Reasoning
Judg Srces
Pers Exp
.HlpProb?
.RelTyp?
.IndTyp?
:Pat?
:Pat?
:Pat?
Y N
gen spec
gen spec
Y N Y N Y N
.RelTyp?
:Pat?
:Pat?
.IntBel?
:Pat?
:Pat?
:Frc?
:Frc?
:Frc?
gen spec
Y N Y N Y N Y N Y N Y N Y N Y N
:Pat?
.IntBel?
:Frc?
:Cls?
:Frc?
:Pat?
:Pat?
:Cls?
:Frc?
:Frc?
Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N
:Cls?
:Frc?
:Pat?
:Pat?
:Cls?
:Frc?
:Frc?
Y N Y N Y N Y N Y N Y N Y N
:Cls?
:Frc?
:Frc?
Y N Y N Y N
nc ci nc si rs ci si rs ci si rs ci si rs nc ci nc si rs ci si rs ci si rs nc ci si rs ci si rs ci si rs ci si rs ci si rs


[BadS GoodI Con]
.Case Type?
SSMIP
MSMIP
Strng
.Situation Type?
.Typ?
Reit
Strng
Abil
Trst
.Corrob?
.Corrob?
:Pat?
:Pat?
Yes
No
Yes
No
Y N Y N
.Prob?
.Solid?
.Prob?
.Solid?
:Frc?
:Frc?
Trust
Abil
Y
N
Trust
Abil
Y
N
Y N Y N
:Pat?
:Pat?
.TProb?
.Prob?
:Pat?
:Pat?
.Prob?
.Prob?
Y N Y N Y N Trust Abil Y N Y N Trust Abil Trust Abil
:Frc? :Frc? :Pat? :Pat? :Pat? :Pat? :Frc? :Frc? :Pat? :Pat? :Pat? :Pat?
Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N
:Frc? :Frc? :Frc? :Frc? :Frc? :Frc? :Frc? :Frc?
Y N Y N Y N Y N Y N Y N Y N Y N
ci si rs ci si rs ci si rs ci si rs ci si rs ci si rs ci si rs ci si rs ci si rs ci si rs ci si rs ci si rs ci si rs ci si rs


[BadS GoodI Iso]
.Type?
Trust
Abil
.RelTyp?
.AbilType?
gen
spec
Expertise
Reasoning
Judg Srces
Pers Exp
:Pat?
.IntBel?
.IndTyp?
:Pat?
:Pat?
:Pat?
Y N Y N gen spec Y N Y N Y N
:Cls? :Frc? :Pat? :Pat? :Pat? :Pat? :Frc? :Frc? :Frc?
Y N Y N Y N Y N Y N Y N Y N Y N Y N
:Cls? :Frc? :Frc? :Cls? :Frc? :Frc?
Y N Y N Y N Y N Y N Y N
nc ci nc si rs ci si rs ci si rs nc ci si rs ci si rs ci si rs ci si rs ci si rs


[NewS Con]
.Case Type?
SSMIP
MSMIP
SoRe 1.1
.Situation Type?
Weakening
Strengthening
Contradiction
.TProb?
Y
N
:Evd?
.AProb?
:Pat?
ri
ci
rs


[SoRe 1.1]
.Situation Type?
Weakening
Strengthening
Contradiction
.Corrob?
Yes
No
.TProb?
Y
N
:Evd?
.AProb?
:Pat?
ri ci rs


[NewS Iso]
.TProb?
Y N
.Prob? .AProb?
Resp Adv Y N
.HlpProb? .IntBel? .ATyp? .ATyp?
Y N Y N EX RE JU PE EX RE JU PE
.IntBel? :Evd? :Evd? :Evd? .Prob? .Prob? .Prob? .Prob? :Evd? :Evd? :Evd? :Evd?
Y N Y N Y N Y N Resp Lowl Resp Lowl Resp Lowl Resp Lowl Y N Y N Y N Y N
:Evd? :Evd? :Pat? :Pat? :Pat? :Evd? :Evd? :Evd? :Evd? :Evd? :Evd? :Evd? :Evd? :Cls? :Pat? :Pat? :Pat? :Pat?
Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N
:Pat? :Pat? :Pat? :Pat? :Pat? :Pat? :Pat? :Pat? :Pat? :Pat?
Y N Y N Y N Y N Y N Y N Y N Y N Y N Y N
ri ci rs ri ci rs ri ci rs ri ci rs ri ci rs ri ci rs ri ci rs ri ci rs ri ci rs ri ci rs ri ci rs ri ci rs ri ci rs nc ri ci rs ri ci rs ri ci rs ri ci rs

Appendix C

SCM Application Proposal

This appendix contains the proposal for the application of the Source Control Mechanism as part of a decision support system for design optimisation and production route generation.

JOINT FRAMEWORK FOR INFORMATION TECHNOLOGY

FIRST CALL OF JOINT DTI / SERC
INTELLIGENT SYSTEMS INTEGRATION PROGRAMME

PROJECT PROPOSAL

A.I. support for integrated design optimization and production route generation

CONSORTIUM PARTNERS

WARNER ELECTRIC LIMITED, Bishop Auckland
(Lead Partner)

NEI REYROLLE SWITCHGEAR Ltd., Hebburn on Tyne,
Tyneside

UNIVERSITY OF DURHAM, School of Engineering and Computer Science,
Durham

JANUARY 1993

Contents

AI support for integrated design optimization and production route generation

Executive Summary

1. The Consortium

-Warner Electric Limited, St. Helen Auckland, Bishop Auckland, Co. Durham DL14 9AA (lead partner)
Contact: Mr J.Summerbell
-NEI Reyrolle Switchgear, Hebburn on Tyne, Tyneside NE31 1UP
Contact: Mr C.Jones
-Durham University, School of Engineering and Computer Science, Science Laboratories, South Road, Durham DH1 3LE
Contact: Dr P.G.Maropoulos

2. The industrial problem

A key issue in today's manufacturing industry is the fragmentation of the various engineering and planning functions. The interface between design and production engineering functions is particularly important since the majority of production costs are decided at the early design stages. The concept of Simultaneous Engineering is aiming at improving product quality and reducing design and manufacturing lead times and costs by the parallel and synchronous product and process design. The generation of routes for a discrete component or sub-assembly is the most direct link between design and production engineering. Currently, the route generation in industry is subjective, slow and does not consider batch size variations, equipment utilization, reported quality problems and process optimization capabilities. Furthermore, the designers do not participate in the route generation activity and this does not promote integration. In a simultaneous engineering environment, it is common to review a design in terms of its manufacturability after an initial routing and lead time have been worked out. Thus, the process of design modification is closely related to the route generation activity. The two activities are performed independently and sequentially using separate systems and this results in increased design lead times and poor market response. For example, at NEI Reyrolle routing and cost estimation for a typical product takes between 2 to 6 weeks and during this time there is no effective design feedback. Also, there is no systematic procedure for the evaluation of incomplete, conflicting or contradictory information available during the process of design modification and this frequently results in suboptimal design solutions.

3. Research track record

A major consideration during the process of route generation is the identification of all generic processes required for producing a component or sub-assembly and the subsequent selection of suitable resources available on the shop floor. At Durham, a considerable amount of research has been undertaken on process capability identification and resource selection methodologies for discrete manufacturing environments. Dr Maropoulos is the principal investigator on this research. Theoretical AI research at Durham has resulted in the specification of new methods for the management of uncertainty. In particular, a unique 'Source Control System' has been developed by Dr Garigliano's research team that reasons about a specific problem using a model of sources. These projects have formed an AI/manufacturing technology framework which will be valuable for the development of the proposed AI tool.

4. Research aims and objectives

The overall aim of the proposed research is to develop an A.I. tool that will consist of two sub-systems namely, a route generator and a source control system. The function of the AI tool will be to assist the designers and production engineers during the process of route generation and design optimization. The research has the following objectives.

- To investigate the application of a knowledge base and heuristic algorithms for generating production routes for discrete components and simple sub-assemblies at an early design stage.
- To specify and develop a route generator that will enable rapid evaluation of alternative design configurations and alternative manufacturing strategies for existing designs by allowing lead time comparisons of the corresponding production methods.
- To investigate the application of source control techniques in a design environment and to evaluate the problems posed by the new domain.
- To specify and develop a source control system for the management of uncertainty and the resolution of conflicts in information supplied by members of a simultaneous engineering team during the process of reviewing a design in order to improve its routing and manufacturability.

- To test the viability and effectiveness of the AI tool and evaluate the scalability constraints of the tool in a CAD/CAPP environment.
- To promote integration of design and production engineering that will be based on a single function (routing) that crosses these traditional departmental boundaries.

5. Benefits and industrial applications

It is believed that the route generator will;
- result in more efficient routes and reduced manufacturing lead times
- guide the designer in making product improvements by displaying the link between product attributes and routes
- improve designer awareness on potential quality problems due to process limitations
- aid designer learning about the production system and its capabilities
- improve production consistency by reducing the lead time variance between different routing options.

The source control system will result in reduced design lead times by managing uncertainty and resolving conflicts in information regarding the evaluation of routing solutions and the assessment of lead time implications of various design options.

The AI tool will act as an integrating factor by providing objective and consistent decision support with regard to manufacturing routes both at the design and the process planning level. This will eventually reduce the number of iterations between designers and planners because both will be using the same decision support tool. The AI tool will also become the focusing point of a simultaneous engineering team and will be the facilitator of common action.

The overall philosophy of the proposed research will be applicable to any manufacturing company that designs and manufactures discrete products. The generic aspects of the source control system will be suitable to most simultaneous engineering environments. In terms of specific industrial sectors, the route generator will be particularly relevant to batch manufacturers where there is considerable production variety and significant process variability. The route generator will be less relevant in high volume/low variety environments where specialized production equipment is used. However, even in this sector, the route generator will be useful when assessing the cost and viability of state-of-the-art production methods and deciding on subcontracting bids.

It is envisaged that, if successful, the proposed AI tool will form an advanced technology demonstrator the future commercial exploitation of which will provide UK manufacturers with a critical competitive advantage in the area of integrated product and process development and improvement.

6. Resources required and costs

The staff involvements of Warner Electric and NEI Reyrolle will be 4.56 and 4.67 man years respectively. The University requests support for two Research Assistants (1A, point 6), each for three years. Computing equipment, consumables as well as travel and subsistence allowances are also required by all partners. The costs per partner are shown below.

	Warner Electric	NEI Reyrolle	Durham University	Total
Project Total	254.900	246.900	158.886	660.686
Grant Requested	87.096	84.361	158.886	330.343
(%)	approx. 34.168	approx. 34.168	100	50

7. Summary description of the AI tool

The tool will comprise a route generator and a source control system. The proposed route generator will have a tandem architecture and will comprise a knowledge base and a number of heuristic algorithms. The algorithms will deal with quantitative components of the route generation problem (such as lead time calculation), guaranteeing the rigorous generation of routes. At the same time the knowledge based system will deal mainly with qualitative and probabilistic elements of the routing problem such as the quality implications of different routes. The route generator will be embedded in a suitable expert system shell which will be used to develop an integrated, intelligent routing system.

The source control system of the proposed research will be built by adapting the existing, general framework to the requirements of the new application domain. The source control system will comprise source models, importance analyser, information consistency analyser, heuristics set, conflict resolution unit, enquiry unit and source re-evaluation unit. The output of the route generator will be considered by the source control system together with the analyses of members of a simultaneous engineering team. A number of enquiry and conflict resolution cycles will be executed and if a mediated solution cannot be found, several alternatives will be output together with the evidence and corresponding arguments. Finally the system will update its internal records.

A loose coupling of the sub-systems is the preferred arrangement for generating the AI tool and will involve producing a package with the same front-end and appropriate data interfacing. A working prototype of the tool will be installed and tested in both companies and will share company-specific design and planning data such as generic description of operations, machine details and standard operation times and cost rates.

8. Quality
The School of Engineering and Computer Science, through the Centre of Software Maintenance, has substantial research and practical/industrial experience in software quality issues. Professor Bennett leads the CSM, and is currently managing a collaborative IEATP project on process assessment, maturity and improvement. Prof Bennett will be responsible for the quality plans and methods in relation to the software produced in this project.

9. Project overview
The proposed project will be over three years and its deliverables include the software, major reports that will be written after each main part of the work and quarterly reports for the meetings with the DTI project monitoring officer. The industrial partners will play a leading role in the specification of the industrial problem, operational requirements and performance assessment criteria. The industrial environments will be used as test bed for the AI prototype tool which will be developed at the University. The AI tool will address the industrial problem in a generic manner and will be tested using company-specific data with regard to a limited range of products and processes. The University team will have the leading role on technical issues. In particular, Dr Maropoulos will be responsible for the engineering and routing aspects of the project. Dr Garigliano will be responsible for the computing aspects and the source control system. Prof Spooner will be responsible for the design elements and Prof Bennett will be responsible for the software process modelling and quality related issues.

The research has four main phases and during the first phase, the Research Assistants will perform a general literature review and collect information. The partners will exchange visits in order to introduce key staff to the main themes of the research and promote the partnership idea. The second phase is aiming at the specification of detailed system requirements and the development of the first routing and source control prototypes. The route generator will be the responsibility of one RA whilst the other one will develop the source control system. The prototypes will be tested in the collaborating companies using case studies. The industrial staff will identify the products needed for the case studies and will prepare the information required. The third phase will result in the definition of the final system architecture and the integration of the route generator and the source control system into an AI tool. During this period the industrial contribution will be mainly in relation to the generation of the product-specific databases and of interfaces with existing company data. Finally, during the fourth phase the tool will be tested using additional case studies.

10. Project management and co-ordination
Mr J.Summerbell of Warner Electric will be the project manager and Dr P.G.Mapopoulos of Durham University will be the technical co-ordinator. A project co-ordination committee will be formed comprising one member from each of the collaborators plus the DTI Monitoring Officer.

11. IPR and exploitation
The partners have identified what constitutes background IPR and have agreed to split all foreground IRP equally. The industrial partners are keen to test and fully evaluate the prototype tool and the integration methodology resulting from the proposed investigation. Academically, dissemination of the generated knowledge will be achieved by publishing papers in learned journals and by presenting the work in conferences. The University has recently set up the Mountjoy Research Centre (MRC) in order to strengthen its interface with industry and the members of the research team maintain strong links with software houses in MRC. Also, the School is the lead partner in the TEAMwork project (STRIDE programme of the DTI) and this is another good platform for disseminating the results of the project to North East companies. Finally, the partners will explore the possibility of future exploitation and will discuss the research with suitable IT vendors such as CAD, CAPP and AI suppliers.

1. The industrial problem and current state-of-the-art

This section includes a description of the industrial problem and outlines current industrial practice. It also refers to state-of-the-art research work in this area.

1.1 The industrial problem

A key issue in today's manufacturing industry is the fragmentation of the various engineering and planning functions. The interface between design and production engineering functions is particularly important since the majority of production costs are decided at the early design stages. The concept of Simultaneous Engineering is aiming at improving product quality and reducing design and manufacturing lead times and costs by the parallel and synchronous product and process design. The characteristics of a design define the manufacturing processes required for its manufacture. Subsequently, the route for the design is generated by selecting and sequencing production machines which can perform the required manufacturing processes. The generation of routes for a discrete component or sub-assembly is the most direct link between design and production engineering and frequently results in alterations of the original design. The designer deals predominantly with functionality issues and subsequently the production routes are generated manually by planners and this process is based on personal preference and previous experience. When similar products can be found (i.e., families of products), the new route is constructed by modifying previous routes and this is the normal way variant process planning systems operate. Hence, the route generation in manufacturing industry is subjective, slow and does not consider batch size variations, equipment utilization, reported quality problems and process optimization capabilities. Furthermore, the designers do not participate in the route generation activity and this does not promote integration.

In a simultaneous engineering environment a design may be revised a number of times in order to improve its functionality and manufacturability. It is common to review a design in terms of its manufacturability after an initial routing and lead time have been worked out. Thus, the process of design modification is closely related to the route generation activity. Currently, the two processes are performed independently and sequentially using separate systems and procedures and this results in increased design lead times and poor market response. For example, at NEI Reyrolle routing and cost estimation for a typical product takes between 2 to 6 weeks and during this time there is no effective design feedback. It is also common for specialists to have contradictory points of view as to how to optimize a design in terms of its functionality, manufacturability and quality. An additional problem is that the simultaneous engineering process is most effective during the conceptual design stage at which there may be incomplete information about certain aspects of the design or manufacturing process. Currently, there is no systematic procedure for the evaluation of incomplete, conflicting or contradictory information available during the process of design modification and optimization and this frequently results in suboptimal design solutions.

1.2 State-of-the-art research work

Recent research in Design and Simultaneous Engineering has resulted in conceptual methods that can be used to infuse knowledge of down stream production activities into the design process (1, 2). Significant progress has been achieved in equipment selection systems for machining operations (2), cost estimation tools to support product design (3, 4) and the standardization of product representation data (5). An area that received little attention is the generation of production routes (1, 2) and the application of AI tools which will assist a simultaneous engineering team during the process of design optimization (1).

A major consideration during the process of route generation is the identification of all generic processes required for producing a component or sub-assembly and the subsequent selection of suitable resources available on the shop floor. At Durham, a considerable amount of research has been undertaken on process capability identification and resource selection methodologies for discrete manufacturing environments. In particular, a knowledge based technology demonstrator has been developed for machine tool selection (2). The system is using process and equipment models for establishing the main correlation parameters and selection rules. A constraint network is also applied during the selection process in order to take into account layout, quality and planning considerations. This system is currently extended to incorporate a work-handling model which will allow the full evaluation of the associated layout characteristics and space limitations. At a more detailed level, process capability studies have resulted in the development of knowledge based systems for turning and milling tool selection (6, 7). Dr Maropoulos is the Principal Investigator of the research described above which is funded by SERC and industry and has been tested in industrial environments with very encouraging results. These projects have formed a technology framework which will be valuable for the development of the proposed route generator. A recent literature review identified that linear programming techniques, such as the assignment, distribution and Vogel's approximation method, have limited suitability for route generation since they cannot incorporate constraints due to design characteristics and process capability. Also, AI techniques have been developed for scheduling and sequencing (8) and those have some relevance to the route generation process.

Theoretical AI research at Durham has resulted in the specification of new methods for the management of uncertainty. In particular, a unique 'Source Control System' has been developed that reasons about a specific problem using a model of sources (9, 10). The system was initially designed to be used as part of an advanced natural language processing unit and it incorporates a generic conflict resolution mechanism for processing information that may be incomplete and/or contradictory and is obtained from several different sources (11). Dr Garigliano is the principal investigator on this AI research. Other methods for managing uncertainty is by using numerical (probabilistic, fuzzy) or endorsement based systems. When compared with the source control system, the numerical systems are easier to use but less efficient in assessing the likelihood and richness of information. On the other hand, the source control technique is better than the endorsement approach in terms of ease of use and for assessing the likelihood of information.

This AI work provides an sound theoretical background for managing uncertainty in new domains such as in a design environment. It is believed that the source control system can be successfully adapted and employed in a design environment, in order to aid in the interpretation of information from different experts, resolve conflicts and handle qualitative design representations. A basic problem with traditional quantitative design representations is that the underlying relationships between design parameters and constraints are often lost or hidden. Also, the formulation of most design problems is highly non-linear and the number of design variables and constraints can be large. Numerical methods for nonlinear optimization have been developed but they are only applicable for relatively simple unconstrained problems (12). Monotonicity analysis (13) has been used for the simplification of design constraint networks. This technique is of limited use since it requires global monotonicity that cannot be found in most engineering designs.

The School of Engineering and Computer Science of Durham University has also a proven track record in integration methodologies in the area of process planning (14) and in quality considerations during software development (15, 16).

2. Aims and objectives

The overall aim of the proposed research is to develop an A.I. tool that will consist of two sub-systems namely, a route generator and a source control system. The function of the AI tool will be to assist the designers and production engineers during the process of route generation and design optimization. The research has the following objectives.

- To investigate the application of a knowledge base and heuristic algorithms for generating production routes for discrete components and simple sub-assemblies at an early design stage.
- To specify and develop a route generator that will enable rapid evaluation of alternative design configurations and alternative manufacturing strategies for existing designs by allowing lead time comparisons of the corresponding production methods.
- To investigate the application of source control techniques in a design environment and to evaluate the problems posed by the new domain.
- To specify and develop a source control system for the management of uncertainty and the resolution of conflicts in information supplied by members of a simultaneous engineering team during the process of reviewing a design in order to improve its manufacturability.
- To test the viability and effectiveness of the AI tool and evaluate the scalability constraints of the tool in a CAD/CAPP environment.
- To promote integration of design and production engineering that will be based on a single function (routing) that crosses these traditional departmental boundaries.

3. Method of research

The research proposed herein will be conducted by the partners of the consortium and will be applied to the specific industrial environments. The research method will comprise the following steps.

1. A particular range of products and the associated manufacturing processes will be specified in each collaborating company and will be used as a test bed for the project. The specification will also include detailed criteria for performance evaluation.
2. A prototype AI tool will be built based on AI/Manufacturing expertize at Durham and guided by the specific product and process knowledge of the industrial partners.

3. The prototype will be tested using industrial case studies and its performance will be assessed against the criteria defined at the specification phase.
4. Following the first case studies, the tool will be appropriately revised and further developed.
5. A second testing phase will take place in industry using additional case studies.
6. The tool will be interfaced with existing company systems to produce a realistic test bed for assessment of effectiveness, robustness and scalability.

4. Research techniques and novel approaches

The aim of the proposed research is the specification and development of an AI tool for solving integration problems between design and production engineering. Figure 1 shows the overall structure of the tool that consists of a route generator and a source control system. The research techniques are described below together with their industrial relevance.

4.1 A tandem architecture route generator

The research will seek to develop a route generator that will employ a novel, tandem architecture. Figure 1 shows the aggregate structure of the proposed route generator which will consist of a knowledge base and a number of algorithms. The algorithms will deal with quantitative components of the route generation problem (such as lead time calculation), guaranteeing the rigorous generation of routes. Various operations research based formulations will be assessed before deciding on the specific algorithms for route generation and the application of various constraints such as work handling and space limitations, layout details and batch size variations. As part of a previous project, heuristic algorithms have been successfully developed for solving the machine tool selection problem (2). It is believed that heuristic algorithms will prove to be the most suitable method for solving routing problems at an industrial scale.

The number of algorithms will depend on the range of processes required for manufacturing the limited range of discrete components and simple sub-assemblies which will be considered by the project. Previous experience indicates that in most cases a single algorithm will be used for each process category such as machining, forming, welding and assembly. However, when there is a considerable variation within one process, special versions of the main algorithm may be required. At the same time the knowledge based system will deal mainly with qualitative and probabilistic elements of the routing problem such as the quality implications of different routes. The tandem architecture will reduce the reliance on accurate routing algorithms since the knowledge base will receive and use feedback from the shop floor. A knowledge acquisition procedure has been developed for collecting, processing and storing tooling information from the shop floor in a systematic manner (6, 7). It is believed that this technique can be modified for collecting quality feedback from processes and for monitoring actual lead times of a limited number of routes.

The feedback information is essential for investigating the validity of routing assumptions in relation to lead times and for studying the quality implications (tolerances, surface finish, scrap rates, functionality) of different resource selection options. It is believed that, for a specific range of products, it will be possible to develop a history of quality problems and lead times associated with various design options and processes. The operation of the system will be of increased accuracy as more historical data becomes available and the knowledge base is enlarged and refined with routing data for many components. Finally, the tandem architecture will be embedded in a suitable expert system shell which will be used to develop an integrated, intelligent routing system.

4.1.1 Industrial relevance of the route generator

The effectiveness of the route generator prototype will be assessed by testing it using a limited range of products/processes in each company. If successful, the technique can be adapted and expanded in future to cover most aspects of discrete component design and manufacture. The aim is to provide rapid feedback on production routes and lead times corresponding to a particular set of design attributes and to make explicit the relationship between the routing and the design attributes. Displaying the link between product attributes and routes will guide the designer in making product improvements and will aid designer learning about the production system and its capabilities. The proposed system will also be used for exploring alternative manufacturing possibilities for existing designs. Existing routings that are manually generated are likely to be sub-optimal and may show wide variance in lead times. The route generator is expected to improve production consistency by reducing the lead time variance between different routing options. Also, the comparison of possible routings with previously generated and tried routings for similar designs will result in more efficient manufacturing processes since the tried routings are based on processes that satisfy all quality considerations. At a later stage, the system could also be used for comparing the actual cost of designs (based on existing processes), with the cost when using state-of-the-art processes and equipment. This function will be particularly useful for deciding whether to make in house or to subcontract as well as for assessing the value-for-money of various subcontracting bids.

4.2 A source control system for design optimization

Design reviews are normally conducted after the generation of routes and the calculation of the corresponding lead times. Reviewing an electro-mechanical design, particularly at conceptual level, involves the assessment and evaluation of information from different sources, which may be contradictory or uncertain, concerning materials, mechanical and electrical functionality, safety, quality and manufacturability. The present project will address issues in the management of uncertain infronmation at the design stage by using a source conrol system. The system will be composed of source models, importance analyser, information consistency analyser, heuristics set, conflict resolution unit, enquiry unit and source re-evaluation unit.

The source control system of the proposed research will be built by adapting the existing general framework to the requirements of the new application domain. It is thought that the adaptation procedure will be extensive and will involve the specification and development of the following domain specific components. Firstly, detailed, a-priori models for classes of specific sources are required and these models will include designers and process planners. Track records of individual participants will also be formed. A knowledge elicitation exercise will be performed to gather specialist knowledge about terminology, forms of argument etc., used in the domain. A specialist set of heuristics will be developed to deal with such extra information. The importance and enquiry units will also be adapted to the task at hand. Problem specific knowledge will also be acquired to be used during the conflict resolution phase and a special front-end will be designed. Finally, an important research task will be to investigate the various problems posed by the new domain and to assess its effectiveness, scalability and robusmess by testing it in real design environments.

The members of a simultaneous engineering team will input, through the specialised interface, their analyses (routing options, machines selected, sequences etc.) and these will be considered together with the output of the route generator and any quality problem reported. An enquiry or conflict resolution cycle may occur at this stage. Some generally acceptable solutions will be proposed and the participants will then enter their solutions. Again, an enquiry or conflict resolution cycle may occur. If the conflicts cannot be solved through disambiguation or source models, the endorsement sub-unit will be activated. Participants will be made aware of problems or clear subjective elements. If a mediated solution cannot be found, several alternative suggestions will be output, together with the evidence, arguments and authority pro and cons. Finally, the system will take stock of the session by updating its internal records.

4.2.1 Industrial relevance of the source control system

The source control system will be particularly valuable to companies where design reviews are conducted at the conceptual design stage where there is frequently contradictory or uncertain information about product or process details. It will also be useful when a product is designed in order to meet customer specifications or when there is pressure to reduce manufacturing lead times and/or satisfy new standards concerning its operation. The source control system will result in reduced design lead times by managing uncertainty and resolving conflicts in information regarding the evaluation of routing solutions and the assessment of lead time implications of various design options. It is also believed that this consistent and integrated design support will increase profit margins.

4.3 An AI tool for integrated route generation and design optimization

A prototype AI tool will be formed by integrating the route generator and the source control system. An interesting research question will be to define the most appropriate level of integration between the two sub-systems. At the outset there are two answers to this question namely, a tight or a loose coupling of the sub-systems. The loose coupling will involve producing a package with the same front-end and limited data interfacing. In particular, the source control system will receive the ouput of the route generator and will include essential routing knowledge and a set of heuristics. Proprietary windowing systems will be used for this integration. The advantage of a loose coupling is that the source control system will be more flexible and could be used for other considerations during design review such as material selection and general process capability issues. The tight coupling will involve building into the source control system a considerable amount of routing knowledge and an enlarged set of heuristics for dealing with this knowledge. It is expected that a tight coupling will improve the accuracy of predictions of the source control system in relation to routings but it will limit its applicability to deal with information from other domains. A loose coupling will be the initial arrangement and after the testing and evaluation period the research team will produce design rules for the tight coupling of the sub-systems. A working prototype of the tool will be installed in both collaborating companies covering a limited range of manufacturing processes. The prototype will share company-specific design and planning data such as generic description of operations, machine details and standard operation times and cost rates.

4.3.1 Industrial relevance of the AI tool

The integration objective of the project will be achieved by employing a common system at the design and process planning interface. The tool will be used either independently by designers/planners or by a simultaneous engineering team.

(i) When the tool is used independently, the integration objective is achieved by providing objective and consistent decision support with regard to manufacturing routes both at the design and the process planning level. This will eventually reduce the number of iterations between designers and planners because both will be using the same decision support tool. The result of this approach is the specification of an integration process that is based on a single function (routing) that crosses traditional departmental boundaries (design-process planning).

(ii) The route generator and the source control system will also be used by a simultaneous engineering team during the process of design modification/optimization. The feedback from the shop floor and the source control system are considered particularly important in this context since they will provide the capability to assess the effect of previous routing decisions and evaluate potentially conflicting or incomplete information. Thus, the unified system will become the focusing point of a simultaneous engineering team and will be the facilitator of common action.

4.4 Quality

The School of Engineering and Computer Science, through the Centre of Software Maintenance, has substantial research and practical/industrial experience (15, 16) in software quality issues. Professor Bennett leads the CSM, and is currently managing a collaborative IEATP project on process assessment, maturity and improvement. This has led to a novel approach to assessment, which indicates also how improvement should subsequently be carried out. We are familiar with Tickit, BS5750, TQM, SEI CMM, Quantum etc. and appropriate quality plans and methods will be established for the software produced in this project.

5. Description of the research plan

5.1 Project overview

a) Overall project description

The proposed project will be over three years and the aggregate research plan is shown in Appendix 1. There will be four main phases and major reports will be written after each major part of the work. Quarterly reports will be written for the meetings with the project monitoring officer. The project deliverables are grouped into internal and external and are listed in Appendix 2. This research is characterized by the need to solve the industrial problem, as outlined in Section 1.1, using AI computing techniques. The industrial partners will play a leading role in the definition of the industrial problem, the specification of realistic operational requirements and performance assessment criteria. The industrial environments will be used as test bed for the AI prototype tool which will be developed at the University. Information in relation to products and processes will also be provided by the companies and their staff will design the interface of the computer based system with company data. The University team will have the leading role on technical issues and will be responsible for the specification of advanced research techniques and the development of an AI prototype tool which will provide the required functionality and will address the industrial problem in a generic manner.

There is a clear need to introduce the University staff to the industrial activities and to perform a detailed scoping of the industrial problem in order to ensure that the generated methods and systems are realistic and can be tested. Equally, there is a need for company personnel to improve their awareness of what AI, and in particular the proposed tool, can do in their specific environments. The industrial evaluation and testing of the tool largely depends on the ability of company staff to operate and support it in their environments. Thus, regular meetings between the partners and the formation of interdisciplinary teams which will include University and company staff are considered to be essential pre-requisites for the success of the project.

b) Scope of the research

In order to achieve the objectives of the research it is essential to focus both the computing and the manufacturing aspects of the research. The AI tool will be developed as an experimental prototype which provides the required functionality for a limited range of products and will be tested using company-specific data. It is also important to limit the manufacturing aspects by considering parts that require a limited and strictly defined range of manufacturing processes. A preliminary investigation in the collaborating companies has resulted in identifying a number of simple sub-assemblies that comprise machined, sheet metal and fabricated parts. Also, a limited range of tools and jigs are used for the corresponding assembly operations and it will be possible to consider the capability of these simple assembly areas.

5.2 Plan of research work

The research has four main phases and the detailed workplans for the University, Warner Electric and NEI Reyrolle are shown in Appendices 3,4 and 5 respectively.

a) First Phase.
During the first phase, the Research Assistants will perform a general literature review and collect information. The partners will exchange visits to introduce key staff to the main themes of the research and to promote the partnership idea.

b) Second Phase.
The second phase is aiming at the specification of detailed system requirements and the development of the first routing and source control prototypes. These prototypes will be tested in the collaborating companies using case studies. Appendix 3 shows that the route generator will be the responsibility of one RA whilst the other one will develop the source control system. The industrial staff will prepare the information required for the case studies as shown in Appendix 4 and 5. Company staff will analyze the type of considerations (quality issues, lead times etc.) and the type of available information (qualitative or quantitative) during design for manufacture. This will help the University team identify the requirements for the Source Control System. In relation to routing, Reyrolle engineers will identify an electro-mechanical sub-assembly of a low-voltage switchgear and Warner engineers will identify a family of solenoids. A detailed examination of these sub-assemblies will follow in order to identify all individual components and generate a list of manufacturing operations required for each component together with standard times, cost rates and resource description. The prototypes will be tested using appropriate case studies.

c) Third Phase
The third phase will result in the definition of the final system architecture and the integration of the route generator and the source control system into an AI tool as detailed in Appendix 3. Both RAs will be responsible for developing the AI tool and for testing it at the companies. During this period the industrial contribution will be mainly in relation to the generation of the product-specific databases and preparation of local interfaces for the prototype with existing company data.

d) Fourth Phase
Finally, during the fourth phase the tool will be tested using real discrete components and sub-assemblies. This phase is the most labour intensive period with regard to the industrial partners. Computer support staff will identify computer files which contain the information identified during phase 3 and will assist in the interfacing of the AI tool with existing datafiles. The case studies will involve the generation of routes, their implementation on the shop floor and the collection of lead time and quality feedback using the knowledge acquisition process. The source control system will be assessed by holding a number of simultaneous engineering trials for specific components.

6. Project management plan

The main management aims will be to satisfy the project objectives as detailed in Section 2 and to produce the required deliverables.

6.1 Project management and co-ordination

a) Central project management and co-ordination
Mr J.Summerbell of Warner Electric will be the project manager and Dr Maropoulos of Durham University will be the technical co-ordinator. A project co-ordination committee will be formed comprising one member from each of the collaborators plus the DTI Monitoring Officer. The committee will meet formally at least once per quarter to review progress against the research plan and take decisions on technical streategy. It is essential that the University work is closely interfaced with the industrial work and for this purpose a wider project review group will be established. The project review group will consist of Mr J.Summerbell and Dr S.Devgan of Warner Electric, Mr C.Jones and Mr N. Allonby of NEI Reyrolle Switchgear and Dr P.G.Maropoulos, Dr R.Gargliano, Professor E.Spooner and Professor K.H. Bennett of the University of Durham. The review group will have monthly meetings to assess progress and plan the details of the next

phase of the work. The Research Assistants and other company staff who are involved in the project will be invited to participate in the review meetings.

b) University work management
The University research team will produce a short monthly progress report, a quarterly report and a more substantial report after each major part of the work. Dr Maropoulos will be responsible for the engineering and routing aspects of the project as well as for the management of the University work. Dr Garigliano will be responsible for the computing aspects and in particular the source control system and its integration with the route generator. Dr Maropoulos' and Dr Garigliano's involvements will be at 35% an 30% respectively. Prof Spooner will be responsible for the design elements of the project and his involvement will be at 10%. Prof Bennett will be responsible for the software process modelling and quality related issues in the development process and will maintain links with current trends and standards. Prof Bennett will contribute 15% of his time to the project.

c) Management of Warner Electric work
The Project Manager will supervise the work at Warner Electric. Dr Devgan will be responsible for the co-ordination of all work conducted at Warner Electric. His duties will include liaising with University staff and ensuring easy access to information and personnel as required. He will produce short monthly reports, reflecting the project status at Warner. On the whole it is envisaged that he will contribute 20% of his time to the project.

d) Management of NEI Reyrolle Switchgear work
Two supervisors (representing senior design and production management) would participate at 5% involvement each. They would attend key meetings and take ultimate responsibility for resources and quality within Reyrolle. Mr N. Allonby will co-ordinate project activities and resources within NEI Reyrolle. He will be responsible for monthly reporting and liaison between Research Assistants and Reyrolle staff. Participation will be at 30%.

6.2 Allocation of resources

a) University resources
The resources requested can be found in Appendix 6. Support is requested for two Research Assistants (1A), each for three years. This is a challenging research project that requires high calibre graduates hence a scale point 6 is requested. A travel and subsistence allowance is required since the Research Assistants will spend significant periods of time at the premises of the collaborating companies collecting data and studying existing procedures and systems. Support is also requested for attending five key conferences namely the IPMU '94, CAPE '95 as well as JFIT '94, '95 and '96. The proposed knowledge base, source control system and the routing algorithms will require fast processing and large computer memory. Two SPARCstations are required for developing the computer based prototype. A limited amount of proprietary software will also be required including an expert shell (Nexpert Object), a prototyping language (Miranda) and a data management system.

b) Warner Electric Ltd resources
The resources required by Warner Electric are shown in Appendix 7. Five different levels of staff will take part in the project and the total involvement will be 4.56 man years. The computing hardware and software required is identical to that of the University and Reyrolle so that the various versions of the AI tool and its sub-systems can be tested at Warner over extended periods of time and interfaced with specific product and process data. The only consumable cost is material for testing the routing and design configuration solutions with regard to certain types of solenoids and clutches. Finally a travel amount is required for visiting the other partners.

c) NEI Reyrolle Switchgear resources
The resources required by NEI Reyrolle are shown in Appendix 8. Support is requested for 4.67 man years with most of the labour required during the third year. The other items required are identical with those of Warner Electric with the exception of consumables and the distribution of engineering support. NEI Reyrolle has more manufacturing processes than Warner Electric and the individual components are larger and more expensive thus, a larger amount is required for consumables and considerable engineering support will be required during the second phase of the project.

7. Financial budget

The financial budget of the project is shown in Appendix 9. The total cost of the project is £660,686 and the total grant requested is £330,343. Each industrial partner requests a DTI grant at 34% of the corresponding costs and the University requests 100% funding from SERC.

8. Benefits and industrial applications

8.1 Industrial benefits

The research will allow a Simultaneous Engineering team to review the effects of alternative manufacturing routes and processes at the design stage, before design finalization. The following questions will be asked:
- Is lead time more important?
- Is cost more important?
- Is quality the prime consideration?

The above considerations will allow the team, which includes designers, industrial engineers and sales engineers to optimize the particular design and route to meet any specific customer expectation. The benefits resulting from this capability include:
- Improved quality, stemming from early awareness of potential problems.
- Improved customer service, stemming from realistic lead times.
- Improved profit margins, stemming from the above and a more accurate costing.

All of the above can be measured and compared with present levels. Additionally, the companies will benefit from:
- Improved inter-departmental co-operation including design engineering, production engineering and sales.
- Improved efficiency and utilization of engineering personnel.
- Improved training of engineering personnel producing better all round engineers.
- Reduced risk of human errors of judgement.
- Continuous updating and improvement of the knowledge base.

8.2 Industrial applications

The overall philosophy of the proposed research will be applicable to any manufacturing company that designs and manufactures discrete products. In terms of specific industrial sectors, the route generator will be particularly relevant to batch manufacturers where there is considerable production variety and significant process variability. It will also be useful in make-to-order environments where the product is designed according to customer requirements and a good market response is required. The route generator will be less relevant in high volume/low variety environments where specialized production equipment is used. However, even in this sector, the route generator will be useful when assessing the cost and viability of state-of-the-art production methods and deciding on subcontracting bids.

The generic aspects of the source control system will be suitable to any simultaneous engineering environment. The system could also be extended to deal with information from other domains such as specific functionality issues and conformance to standards.

It is envisaged that, if successful, the proposed AI tool will form an advanced technology demonstrator the future commercial exploitation of which will provide UK manufacturers with a critical competitive advantage in the area of integrated product and process development and improvement.

9. Consortium details

It is essential that the proposed AI research work is put in the context of discrete manufacturing industry that covers the whole range of activities from design to manufacture. The proposed consortium comprises two North East manufacturing companies, NEI Reyrolle Switchgear and Warner Electric Limited and Durham University. Appendix 10 shows the summary profiles of the industrial partners. Company brochures are also included in the documentation of the submission.

9.1 Durham University, School of Engineering and Computer Science
Science Laboratories, South Road, Durham DH1 3LE

The School of Engineering and Computer Science has 43 academic staff and its research income is over £4m. Currently the School participates in several collaborative projects that involve the application of advanced computing techniques in manufacturing (planning and technical functions), engineering design, telecommunications, software maintenance, natural language processing and geotechnical engineering. Most projects have high industrial relevance and funding is obtained from a variety of sources that include the Research Councils (SERC, ESRC), DTI, European Community Programmes (Esprit, Brite) Charitable Foundations and industry.

The School has set up the Centre for Industrial Automation and Manufacture for providing interdisciplinary expertize and ensuring high quality project management when dealing with large industrial projects like the one proposed herein. There are also large research groups in Manufacturing, AI and Software Modelling/Quality. Engineering has got a top research rating in the last research assessment exercise.

9.2 Warner Electric Limited
St. Helen Auckland, Bishop Auckland, Co. Durham DL14 9AA
Warner Electric Limited, a subsidiary of the DANA Corporation, has been situated in the North East of England, County Durham, since 1946. The company is based on two sites, a main plant of 150,000 square feet at Bishop Auckland and a smaller one of 20,000 square feet at Spennymoor. Warner Electric's principle line of business is in the development and high volume manufacture of electromagnetic products associated with motion and motion control. The portfolio includes products for controlling rotary and linear motion such as clutches, brakes, solenoids, linear actuators and tension control systems. Some of the products such as the solenoids, clutches and brakes are custom designed to suit each particular application.

The types of industry Warner Electric supply to are;

1) Textile equipment manufacturers
2) Vending manufacturers
3) Information Technology (hardware).

Warner Electric export approximately 25% of all manufactured product. The company is keen to develop links with Durham University and this project represents the first example of a research collaboration between the two parties.

9.3 NEI Reyrolle Switchgear
Hebburn on Tyne, Tyneside NE31 1UP
A.Reyrolle & Co. was established at Hebburn on Tyne (its present site) in 1901 to manufacture electrical switchgear and gained reputation for world leading products. Northern Engineering Industries (NEI) was formed by the merger of A.Reyrolle & Co. with turbine manufacturer C.A.Parsons in 1968, and with power engineers Clarke-Chapman in 1977. In 1989 NEI became part of the Rolls Royce Industrial Power Group. NEI Reyrolle currently has 1,815 staff and a turnover of £50m/year.

NEI Reyrolle Switchgear is the main business unit within NEI Reyrolle and manufactures switchgear and associated equipment for the construction of complete substations in the range 3.3 kV to 550 kV. The product range also includes a limited range of other high technology products such as nuclear reactor instrumentation penetrations. The company is Quality Assured to BS 5750 Part I. There is a strong relationship between Reyrolle and the University of Durham that has been built over a number of years through research contracts, teaching company schemes and student sponsorship. This relationship provides a firm basis for the proposed work.

10. Experience and background of the individuals

The team of the proposed investigators is interdisciplinary since the project requires expertize in production engineering, design and computer science.

Mr Jim Summerbell
Mr Summerbell has been employed by Warner Electric Ltd for some 32 years. Commencing as a Technical Apprentice he was trained as an Electro-Mechanical Engineer. He spent 10 years as a Design and Development Engineer on various products including transformers, clutches, brakes and solenoids. The following 10 years he was a Project Leader in R&D Engineering specifically guiding a design team involved with solenoids ans associated products. For the last 6 years he has been in charge of the Electro-Mechanical Design Engineering (13 personnel) as Engineering Manager, the last 2 years of which he has been a Company Director.

Dr Sanjeev Devgan
Dr Devgan is the Principal Engineer at Warner Electric heading a team of experienced and dedicated engineers. In total he has spent 5 years at Warner. Although responsible for the introduction of advanced

computational analysis for solving electromagnetic problems, his main interests lie in the field of manufacture and associated high volume manufacturing techniques and standards. Dr Devgan obtained his PhD from UMIST, the research being the design and manufacture of a new type of electrical rotating machine. Other than Warner, Dr Devgan has also been employed by Brush Electrical Machines for research and development purposes.

Mr Chris Jones
After gaining a degree in Electrical Engineering from the University of Liverpool, Chris Jones joined NEI Reyrolle as a graduate trainee in 1978. After a number of responsibilities for product development he is now Chief Engineer and responsible for all product development and product engineering activities. An active member of various international switchgear technical committees and working groups, he also supervises and participates in research collaboration with Universities including Liverpool, Newcastle and Strathclyde.

Mr Nathan Allonby
As a Senior Design Engineer he is currently involved in product development for light voltage switchgear, and has co-ordinated research collaboration with Newcastle University. Nathan Allonby graduated from Newcastle University in Mechanical Engineering in 1984 and after experience at British Aerospace and Newcastle University, joined NEI Reyrolle in 1991. During 1987-90 he was employed as a Research Assistant at Newcastle University working on an industrial collaborative project with APV Baker as part of the SERC programme on High Speed Machinery.

Dr Paul Maropoulos
Dr Maropoulos is Lecturer in Engineering at the School of Engineering and Computer Science of Durham University. He was appointed in 1989 and participated in several research projects funded mainly by SERC, DTI and industry. His main research interests are the specification of new simultaneous engineering methodologies and innovative applications of AI techniques. A current project funded by ACME resulted in the development of a sucessful knowledge based system for cutting tool selection. Other current projects include a project in intelligent machining funded by SERC and industry and a teaching company scheme with Thorn Lighting Ltd..

Dr Roberto Garigliano
Dr Garigliano is Lecturer in Artificial Intelligence at the School of Engineering and Computer Science of University of Durham. He participates in several research projects funded by SERC and industry and his main research interests are reasoning under uncertainty, natural language processing and genetic algorithms. He is the principal investigator of the project that developed the system 'Large-scale, Object-based, Linguistics Interactor, Translator and Analyser', which has been selected for the Royal Society Soiree 1993.

Professor Ed Spooner
Prof Spooner was appointed in 1991 to the Chair in Electrical Engineering at the School of Engineering and Computer Science, University of Durham. Prof Spooner's main research interests are the design of advanced electrical machines and wind turbine units. Currently, he participates in several research projects funded by SERC, UK and European industry.

Professor Keith Bennett
Prof Bennett was appointed in 1986 to the Chair in Computer Science at the University of Durham, School of Engineering and Computer Science. He was a founder member of the Centre for Software Maintenance, the world's first centre addressing the topic of key industrial and commercial maintenance. Prof Bennett is currently Principal Investigator on 2 SERC grants, and two IEATP grants and he has demonstrated that collaborative projects can lead to new engineering results and exploitable products.

11. Intellectual property and exploitation strategy

11.1 IPR issues
The partners have agreed that the generic specification of the existing source control system and of the machine tool selection system constitute background IPR that belongs to the University. They have also agreed that all product design, engineering and process knowledge constitute background IPR which belongs to the companies.

The partners have agreed to split the results of the project equally i.e., all foreground IRP including methodologies and computer systems.

11.2 Exploitation

The industrial partners are keen to test and fully evaluate the prototype tool and the integration methodology resulting from the proposed investigation and they also want to ensure that the ideas generated are relevant, realistic and can be tested. The membership of Reyrolle in the NEI group, which in turn is the industrial power group of Rolls Royce and of Warner Electric in the DANA group provides an additional way of exploiting the results of the research in a number of other companies.

Academically, dissemination of the generated knowledge will be achieved by publishing papers in learned journals and by presenting the work in conferences. Subject to the usual constraints concerning matters of company confidentiality, NEI Reyrolle and Warner Electric will not seek to restrict the publications arising from the research. The generic aspects of the research will also be discussed with other manufacturing companies and software houses. The University has recently set up the Mountjoy Research Centre (MRC) in order to strengthen its interface with industry and the members of the research team maintain strong links with software houses in MRC. Also, the School is the lead partner in the TEAMwork project that has just been funded from the STRIDE programme of the DTI. The STRIDE programme is another good platform for disseminating the technology resulting from the proposed project to North East companies. Finally, the partners will explore the possibility of future exploitation and will discuss the research with suitable IT vendors such as CAD, CAPP and AI suppliers.

References

1. Finger, S. and Dixon, J.R., 'A review of research in Mechanical Engineering Design, Part II,' *Research in Engineering Design*, Vol. 1, No. 2, 1989, pp. 121-137.
2. Maropoulos, P.G., 'Integration of design and production engineering using a KBS for route generation and design cycle optimization', European *Simulation Symposium*, November 5-8, 1992, Dresden, Germany.
3. Allen, A.J. and Swift, K.G., 'Manufacturing process selection and costing', *Proc Instn Mech Engrs*, Vol. 204, 1990, pp. 143-148.
4. Wong, J.P., Parsaei, H.R. and Imam, I.N., 'An integrated cost estimating system for concurrent engineering environments', *Computers ind. Engng*, Vol. 21, No. 1, 1991, pp. 589-594.
5. PDES/STEP, 'Product Data Exchange Specification', First working draft, NISTIR 884004, National Institute of Standards and Technology, National Engineering Laboratory, 1989, Gaithersburg, MD 20899, USA.
6. Maropoulos, P.G., 'Cutting tool selection: An intelligent methodology and its interfaces with technical and planning functions', *Proc Instn Mech Engrs*, Part B, Vol 206, 1992, pp 47-60.
7. Maropoulos, P.G., 'EXTEP : Intelligent work planning for milling'. Proceedings of the Seventh National Conference on *Production Research*, Hatfield, September 1991, pp178-182.
8. Kusiak, A. and Ahn, J., 'Intelligent scheduling of automated machining systems', *Computer-Integrated Manufacturing Systems*, Vol. 5, No. 1, February 1992, pp. 3-14.
9. Garigliano, R., Long, D. and Bokma, A., 'A model for learning by source control', in Bouchon, Saitta and Yager (eds.) Uncertainty and Intelligent Systems, *LNCS 313*, Springer-Verlag, 1988.
10. Garigliano, R. and Bokma, A., 'Heuristics for knowledge extraction in source control', Proceedings of International Conference on *Information Processing and Management of Uncertainty*, IPMU-92, 6-9 July 1992, Majorca, Spain.
11. Gariglíano, R. and Bokma, A., 'Towards a conflict resolution mechanism for multiple sources', *LNCS 521*, Springer-Verlag 1990.
12. Arora, J.S., *Introduction to Optimum Design*, McGraw Hill International, 1989.
13. Papalambros, P.J. and Wilde, D.J., *Principles of Optimal Design*, Cambridge University Press, New York, 1988.
14. Maropoulos, P.G., 'Integration of Process Planning and Tool Management Functions in Machining Environments'. Proceedings of the Eighth International Conference on *Computer Aided Production Engineering*, Edinburgh, 16-19 August 1992.
15. Bennett, K.H. and Hinley, D., 'Developing a model to manage the software maintenance problem', Proc. IEEE, Conference on Software Maintenance, Orlando, Nov 1992, pp. 174-182.
16. Bennett, K.H. and Hinley, D., 'A process modelling approach to managing software process improvement', Proc. Software Quality Management, British Computer Society, 30 March to 1 April 1993 (to appear).

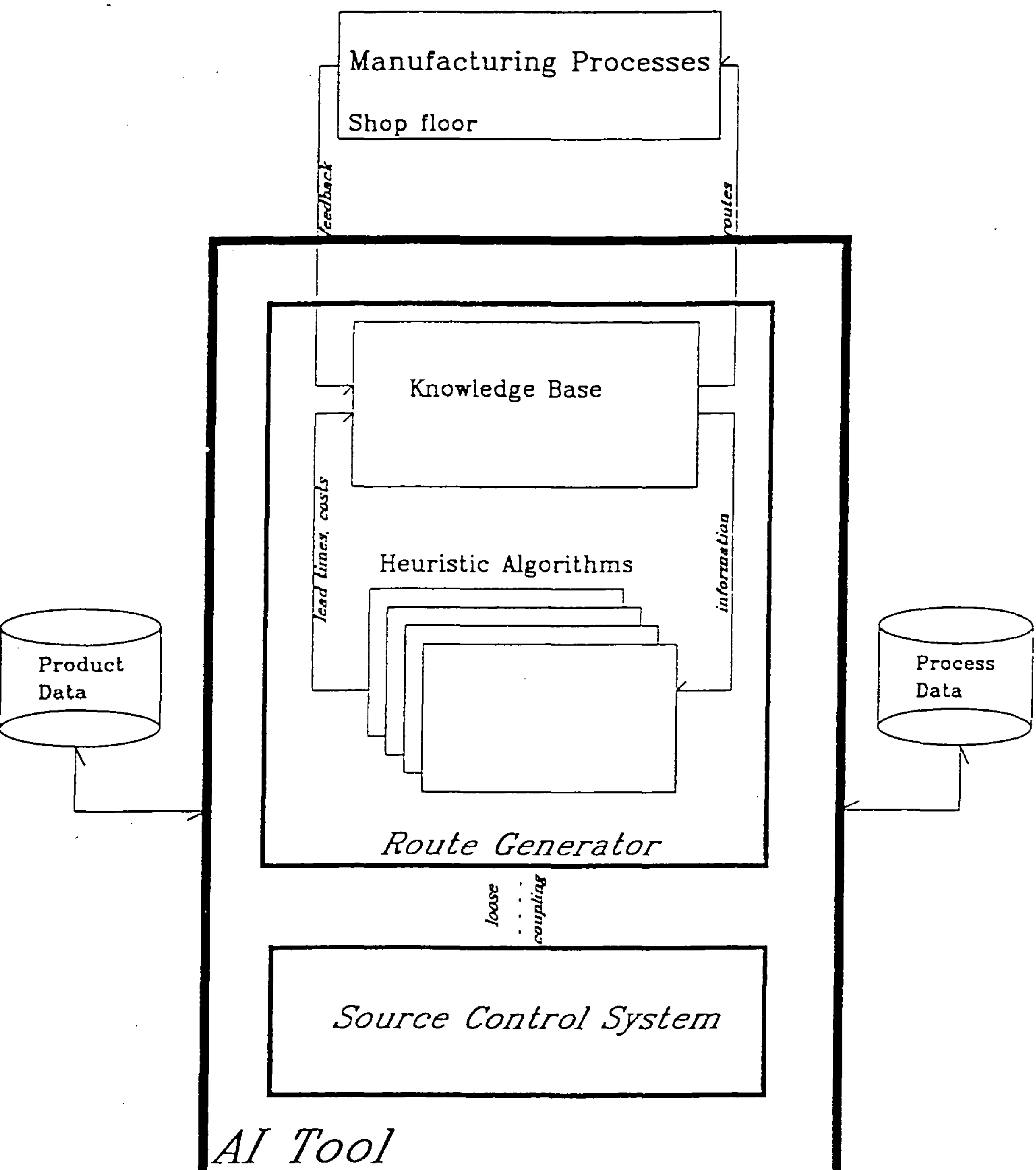


Figure 1. Overall structure of the proposed AI tool

APPENDIX 1

Overall research plan

APPENDIX 2

Project deliverables

Description	Originator	Type	Month
First Phase			
D1, Literature Review	RA1, RA2	E	3
D2, Company visits reports	RA1, RA2	I	3
Second Phase			
D3, Interim system specification report	RA1, RA2	E	7
D4, System specification report	RA1, RA2	E	11
D5, First prototype of RGT	RA1	E	15
D6, Report on first RGT protype	RA1	E	15
D7, First prototype of SCS	RA2	E	15
D8, Report on first SCS protype	RA2	E	15
D9, Report on initial testing of RGT	RA1	I	17
D10, Report on initial testing of SCS	RA2	I	17
Third Phase			
D11, RGT prototype	RA1	E	25
D12, Major report on RGT architecture	RA1	E	25
D13, SCS prototype	RA2	E	25
D14, Major report on SCS architecture	RA2	E	25
D15, AI tool prototype	RA1,RA2	E	28
D16, Report on AI tool prototype	RA1,RA2	E	28
Fourth Phase			
D17, Report on second testing phase	RA1, RA2,	E	32
D18, Warner additional testing period report	SD	I	35
D19, Reyrolle additional testing period report	NA	I	35
D20, List of presentations/papers	RA1, RA2	E	35
D21, Quarterly project management reports	JS, PGM	E	35
D22, Minutes of project co-ordination committee meetings	NA	E	35
D23, Minutes of project review committee meetings	SD	I	35
D24, Final project report	PGM, RG	E	35

D1 to D24 Deliverables
SCS=Source Control System, RGT=Route Generator
RA1, RA2=Reasearch Assistants, JS=Jim Summerbell, SD=Sanjeev Devgan, NA=Nathan Allonby
PGM=Paul Maropoulos, RG=Roberto Garigliano

APPENDIX 3

Plan of University research work

Months	Staff	Activities
First Phase - General Introduction, Literature Review (Months 0-3)		
0-2	UC, Both RAs:	Familiarisation, background reading, literature review.
2-3	CC, Both RAs:	Introduction to company staff, team building. Familiarization with company practice and systems, collection of information.
Second Phase - Specification of System Requirements Model, Rapid Prototyping (Months 3-17)		
3-5	UB, Both RAs:	Initial requirements specification. Functionality and overall interfacing requirements.
5-7	CB, Both RAs:	Knowledge elicitation and data gathering.
7-11	UC, RA1:	Specification of resource selection criteria, routing algorithms and constraints. Definition of the initial knowledge base structure and input data requirements.
	UC, RA2:	Adaptation of source control mechanism and definition of interfacing requirements. Specification of decision support system and of source control sub-system (SCS).
11-15	UB, RA1:	Development of a rapid prototype of the route generator (algorithms and the first version of the knowledge base).
	UB, RA2:	Development of a rapid prototype of SCS.
15-17	CB, RA1:	Initial testing of the first route generator prototype.
	CB, RA2:	Initial testing of the basic SCS prototype.
Third Phase - Development of System Architecture Model, Integration of Systems (Months 17-28)		
17-25	UC, RA1:	Development of the tandem route generation architecture. Definition of feedback information and of collection procedure. Development of the final version of the knowledge base.
	UC, RA2:	Development of decision support system and SCS architecture.
25-28	UC, Both RAs:	Integration of the route generator and SCS into an AI tool. Interfacing methodology with company systems.
Fourth Phase - Industrial Testing, Benefits Measurement, Documentation (Months 28-36)		
28-32	CB, Both RAs:	Industrial testing of the AI tool (route generator and the SCS). Measure benefits and deliverables based on agreed performance measures.
32-35	UB, Both RAs:	Specify the basic generic framework for applying the results of the research to the collaborating companies and other manufacturers. Preparation of final reports and material for dissemination.

Key: UB = University based, UC = University centred (75% of time)
CB = Company based, CC = Company centred (75% of time)

APPENDIX 4

Plan of Warner Electric Ltd. work

Months Activities

First Phase - General Introduction, Literature Review (Months 0-3)

0-2 Familiarisation with research theme, internal meetings, visits to University.
Management meetings.
2-3 Introduce RAs to company practice and systems, provide information on processes and products.
Management meetings.

Man years of Phase one: 0.015 supervision, 0.05 management, 0.25 eng/prod, 0.32 technical support

Second Phase - Specification of System Requirements Model, Rapid Prototyping (Months 3-17)

3-5 Provide user requirements specification.
Management meetings.
5-7 Participate in the knowledge elicitation and data gathering process.
Management meetings.
7-11 Provide information on resource and product characteristics, standard times, costs.
Provide technical support and data for initial trials of subsystems with company data.
Management meetings.
11-15 Management meetings.
15-17 Initial testing of the first route generator prototype and of the basic SCS prototype. Shop floor trials.
Management meetings.

Man years of Phase two: 0.07 supervision, 0.23 management, 0.20 eng/prod, 0.23 technical support,
0.3 shop floor

Third Phase - Development of System Architecture Model, Integration of Systems (Months 17-28)

17-28 User support and guidance during the development of the AI tool.
Work on interfacing with company-specific data concerning product and process details.
Management meetings.

Man years of Phase three: 0.055 supervision, 0.18 management, 0.09 eng/prod

Fourth Phase - Industrial Testing, Benefits Measurement, Documentation (Months 28-36)

28-32 On site testing of the AI tool. Shop floor and simultaneous engineering trials.
Measure benefits and deliverables based on agreed performance measures.
Management meetings.
32-35 Further testing with additional company-specific data, simultaneous engineering trials.
Consolidation of prototype and final industrial assessment.
Management meetings.

Man years of Phase four: 0.06 supervision, 0.14 management, 0.50 eng/prod, 0.67 tech. support,
0.90 shop floor

Project Total: 4.26 man years

Project Management

0-35 Management and co-ordination of project. 0.3 man years

Warner Electric Total: 4.56 man years

APPENDIX 5

Plan of NEI Reyrolle work

Months Activities

First Phase - General Introduction, Literature Review (Months 0-3)

0-2 Familiarisation with research theme, internal meetings, visits to University.
Management meetings.
2-3 Introduce RAs to company practice and systems, provide information on processes and products.
Management meetings.

Man years of Phase one: 0.03 supervision, 0.075 management, 0.17 eng/prod, 0.225 technical support

Second Phase - Specification of System Requirements Model, Rapid Prototyping (Months 3-17)

3-5 Provide user requirements specification.
Management meetings.
5-7 Participate in the knowledge elicitation and data gathering process.
Management meetings.
7-11 Provide information on resource and product characteristics, standard times, costs.
Provide technical support and data for initial trials of subsystems with company data.
Management meetings.
11-15 Management meetings.
15-17 Initial testing of the first route generator prototype and of the basic SCS prototype. Shop floor trials.
Management meetings.

Man years of Phase two: 0.12 supervision, 0.35 management, 0.555 eng/prod, 0.19 technical support, 0.60 shop floor

Third Phase - Development of System Architecture Model, Integration of Systems (Months 17-28)

17-28 User support and guidance during the development of the AI tool.
Work on interfacing with company-specific data concerning product and process details.
Management meetings.

Man years of Phase three: 0.07 supervision, 0.275 management, 0.22 eng/prod

Fourth Phase - Industrial Testing, Benefits Measurement, Documentation (Months 28-36)

28-32 On site testing of the AI tool. Shop floor and simultaneous engineering trials.
Measure benefits and deliverables based on agreed performance measures.
Management meetings.
32-35 Further testing with additional company-specific data, simultaneous engineering trials.
Consolidation and final industrial assessment of systems.
Management meetings.

Man years of Phase four: 0.08 supervision, 0.20 management, 0.165 eng/prod, 0.745 tech. support, 0.60 shop floor

Project Total: 4.67 man years

APPENDIX 6

Resources required
University of Durham, School of Engineering and Computer Science

Summary costs

Description	Year 1	Year 2	Year 3	Total
Staff	37,664	39,168	41,130	117,962
Capital	25,398	1,763	1,763	28,924
Consumables	500	500	500	1,500
Travel	3,000	3,750	3,750	10,500
Total	66,562	45,181	47,143	158,886

Staff costs

Description	Year 1	Year 2	Year 3	Total
2 RAs, 1A, pt 6	37,664	39,168	41,130	117,962
Total	37,664	39,168	41,130	117,962*

Capital equipment costs

Description	Year 1	Year 2	Year 3	Total
2 SPARCstations	17,635	0	0	17,635
Software	6,000	0	0	6,000
Maintenance	1,763	1,763	1,763	5,289
Total	25,398	1,763	1,763	28,924

Consumable costs

Description	Year 1	Year 2	Year 3	Total
Stationery items	200	200	200	600
Computer consumables	300	300	300	900
Total	500	500	500	1,500

Travel costs

Description	Year 1	Year 2	Year 3	Total
Travel to companies	2,500	2,500	2,500	7,500
Conferences, JFIT, IPMU'94 and CAPE'95	500	1,250	1,250	3,000
Total	3,000	3,750	3,750	10,500

* This figure does not include indirect costs which are shown in the RG2 form.

APPENDIX 7

Warner Electric Resources

Summary

Description	Year 1	Year 2	Year 3	Total
Labour	52,550	59,550	115,500	227,600
Capital	18,100	1,100	1,100	20,300
Consumables	0	1,000	3,000	4,000
Travel	1,000	1,000	1,000	3,000
Total	71,650	62,650	120,600	254,900

Labour (assume 5% inflation)

Description	Base	O/H	Man Years	Mn Yr1	Year 1	Mn Yr2	Year 2	Mn Yr3	Year 3	Total
Project Management	30,000	0.98	0.30	0.10	8,250	0.10	8,650	0.10	9,100	26,000
Supervision	30,000	0.98	0.20	0.06	4,950	0.06	5,200	0.08	7,200	17,450
Management	25,000	0.98	0.60	0.20	13,750	0.20	14,450	0.20	15,150	43,350
Engineering / Production	18,000	0.98	1.04	0.25	12,400	0.26	13,500	0.53	28,900	54,800
Eng/production Support	15,000	0.98	1.22	0.32	13,200	0.23	9,950	0.67	30,500	53,650
Shop Floor	9,000	1.34	1.20	0.00	0	0.30	7,800	0.90	24,550	32,350
Total			4.56	0.93	52,550	1.15	59,550	2.48	115,500	227,600

Capital equipment

Description	Year 1	Year 2	Year 3	Total
1 SPARCstation	11,000	0	0	11,000
Software	6,000	0	0	6,000
Maintenance	1,100	1,100	1,100	3,300
Total	18,100	1,100	1,100	20,300

Consumables

Description	Year 1	Year 2	Year 3	Total
Material for routing tests	0	1,000	3,000	4,000
Total	0	1,000	3,000	4,000

Travel and subsistence

Description	Year 1	Year 2	Year 3	Total
Travel to NEI and Durham	1,000	1,000	1,000	3,000
Total	1,000	1,000	1,000	3,000

APPENDIX 8

NEI Reyrolle Switchgear Resources

Summary

Description	Year 1	Year 2	Year 3	Total
Labour	75,100	55,900	91,100	222,100
Capital	18,100	1,100	1,100	20,300
Consumables	0	2,000	6,000	8,000
Travel	1,000	1,000	1,000	3,000
Total	94,200	60,000	99,200	253,400

Labour (assume 5% inflation)

Description	Base	O/H	Man Years	Mn Yr1	Year 1	Mn Yr2	Year 2	Mn Yr3	Year 3	Total
Supervision	25,000	1.75	0.30	0.10	6,900	0.10	7,200	0.10	7,600	21,700
Management	18,000	1.75	0.90	0.30	14,900	0.30	15,600	0.30	16,400	46,900
Engineering / Production	18,000	1.75	1.11	0.53	26,200	0.41	21,300	0.17	9,300	56,800
Eng/production Support	15,000	1.75	1.16	0.31	12,800	0.10	4,300	0.75	34,100	51,200
Shop Floor	13,000	1.75	1.20	0.40	14,300	0.20	7,500	0.60	23,700	45,500
Total			4.67	1.64	75,100	1.11	55,900	1.92	91,100	222,100

Capital equipment

Description	Year 1	Year 2	Year 3	Total
1 SPARCstation	11,000	0	0	11,000
Software	6,000	0	0	6,000
Maintenance	1,100	1,100	1,100	3,300
Total	18,100	1,100	1,100	20,300

Consumables

Description	Year 1	Year 2	Year 3	Total
Material for routing tests	0	2,000	6,000	8,000
Total	0	2,000	6,000	8,000

Travel and subsistence

Description	Year 1	Year 2	Year 3	Total
Travel to Warner and Durham	1,000	1,000	1,000	3,000
Total	1,000	1,000	1,000	3,000

APPENDIX 9

Financial budget

	Warner Electric	NEI Revrolle	Durham University	Total
Year 1 Labour	52.550	75.100	37.664	165.314
Capital	18.100	18.100	25.398	61.598
Consumables	0	0	500	500
Travel	1.000	1.000	3.000	5.000
Total	71.650	94.200	81.627	247.477
Year 2 Labour	59.550	55.900	39.168	147.536
Capital	1.100	1.100	1.763	3.963
Consumables	1.000	2.000	500	3.500
Travel	1.000	1.000	3.750	5.750
Total	62.650	60.000	60.848	183.498
Year 3 Labour	115.500	91.100	41.130	278.930
Capital	1.100	1.100	1.763	3.963
Consumables	3.000	6.000	500	9.500
Travel	1.000	1.000	3.750	5.750
Total	120.600	99.200	63.595	283.395
Total	254.900	253.400	158.886	667.186
Grant Requested			158.886	333.593
(%)			100	50

APPENDIX 10

Company profiles

	Warner Electric Ltd	NEI Reyrolle Switchgear
Main Business Activity	Electro-Mechanical Engineering	Electro-Mechanical Engineering
Turnover	£14.7m	£40.0m
Employees	466	785
Type of Production	Make-To-Order, Make-To-Stock	mainly Make-To-Order
Volume/Variety	High/High	Low/High
Number of Manufacturing Processes	10	17
Typical Batch Size	250-500	1-20 (and reducing)
Designers/Production Personnel	7/307	19/440
Process Planners/Production Personnel	5/307	10/440
Customer driven design	Yes	Very little
Design for Manufacture (Assembly)	Yes	Yes