A Knowledge Level Model of Context and Context use in Diagnostic Domains

by

Pınar Öztürk

Submitted to the Department of Computer and Information Science, Norwegian University of Science and Technology, NTNU in partial fulfillment of the requirements for the degree of Doctor Scientiarum

December 1999
Abstract

The motivation behind this research is to improve artificial intelligence methods for diagnostic problem solving by developing a knowledge level account of contextual knowledge and of its use in diagnosis. This thesis focuses mainly on the conceptual, implementation-independent aspects of contextual knowledge and, on this basis, constitutes a step toward a generic model of contextual knowledge and a methodology featuring its use for diagnostic knowledge-based systems.

Diagnostic problems often involve weak and dynamic domains, where the domain knowledge can not be expressed in the form of strong theories, and where information becomes available gradually. Therefore, abductive reasoning characterizes diagnostic problem solving. The computational implementations of abductive inference are known to suffer mainly from two important problems: the low quality of abductive conclusions and intractability of the abductive process. This thesis addresses these problems through investigating the connections between the notions of context and abductive inference, and works out a context-sensitive abductive inference account. The nature of the thesis is analytical and therefore its main contributions in this respect is at the conceptual level.

The literature lacks a holistic look into the context notion within problem solving. Even though context has been studied in several disciplines, it has been given many different meanings and roles. Our strategical goal is to combine theoretical findings and ideas from various disciplines and to take a step towards a unified theory of context. Thus, we are taking an interdisciplinary look at context, emphasizing the interconnections between psychology, philosophy, and AI.

This thesis attempts to answer the question of ‘what the notion of context means to diagnostic problem solving’, from two distinct but closely related perspectives. According to these perspectives, context is viewed (i) as a knowledge type capturing portions of the focus of attention and (ii) as a means for triggering a shift in the focus of attention.

In order to achieve these goals we have adopted a research strategy which puts a structure on our work. This strategy involves the following steps: (i) construction of a model of diagnostic ‘process’ knowledge. This serves to explicate the relevant application subtasks (i.e., diagnostic subprocesses), (ii) identification of the loci of contextual effects (i.e., the subtasks where context may have an influence), (iii) identification of the role that contextual elements plays at each locus (i.e., the way it effects each subtask), (iv) identification of the source of contextual elements that can play this roles.

Parts of the theory has been implemented for demonstrative purpose, main method being a knowledge-intensive case-based reasoning method.

Main contributions of this thesis are in (i) analysing and modeling contextual knowledge, and (ii) incorporating contextual elements into both general and specific domain knowledge, (iii) identifying the role of contextual elements in selective processing of knowledge, in particular, its role in focused abductive reasoning. These constitute a step toward context-sensitive models of diagnosis.
Acknowledgements

I am deeply indebted to several persons who have provided invaluable help, support and encouragement throughout my work with this thesis.

First of all, I want to thank my supervisor Agnar Aamodt at IDI for extensive comments, vivid discussions, and lasting co-operation. He gave me freedom to cross the borders between disciplines of artificial intelligence, cognitive psychology, and philosophy.

Agnar's other students, Morten Grimnes, Arnt Lockert, Ole Martin Winnem, Ståle Heitman, Jan Ståle Berg, and Heidi Fjorden have been a continuous inspiration source during my IDI years.

Colleagues at IDI have given me a nice atmosphere to work in. In particular, I want to thank Arvid Holme for discussions, practical advice, and his sense of humour that made IDI a rather exciting place for me. The comments from Keith Downing on the manuscript, and his English proofreading have been valuable for me.

I want to thank the administrative staff at IDI, in particular Turid Reinfjord, Kirsten Sandersen, Eli Løvø and Anne-Berit Dahl for their friendships, supports and help. Without you, days would not be the same.

I would also like to express my gratitude to people in my new work place Sintef, Telecommunication and Informatics, who always showed interest in my thesis work, and encouraged me to finish it.

I wish to direct my sincere thanks to Serdar Uckun (Md, and PhD in AI) for his creative ideas, discussions and comments, to Dr. med. Osman Senturk and Dr. med. Ole Rosvoll for their nice cooperation.

I also want to thank to my family. My parents who managed to make me feel their warm love from thousand kilometres. Special thanks to my lovely sister and her husband who has been like a brother to me as well. I want also thank to every single member of the family of my husband for their continuous support. In particular, Nahide who came to Norway and stayed with us for some time and supported me in so many ways that I even could not dream about. It was her nice friendship with my daughter Sirin that help her to hold up with a mother who had hurry most of the time.

I don't write anything to or about my husband as he wanted me not to do so.

Once again, thank you all, without your invaluable supports, this thesis could never be finished the way it is now.

This work was financed through a governmental research grant governed by NTNU, department of Computer and Information Science.
# Table of Contents

1  **Introduction and Overview**  
   1.1 Motivation ........................................ 11  
   1.2 Research goals ................................... 12  
   1.3 The research strategy ............................ 13  
   1.4 The context ontology .............................. 15  
   1.5 The generic mechanism for attention focusing. 16  
   1.6 The methodology for analysing and modeling context 16  
   1.7 The abductive inference pattern .................. 17  
   1.8 Diagnostic knowledge-based systems ............... 18  
   1.9 The context-sensitive, iterative diagnosis model .. 19  
   1.10 The knowledge-intensive case-based method ...... 20  
   1.11 Overview of the thesis ........................... 21

## PART I: Problem solving and Inference

2  **Medical problem solving**  
   2.1 Introduction ....................................... 25  
   2.2 Main considerations of a clinician ................ 25  
   2.3 Overview of Clinical practice ..................... 27  
   2.4 The correspondence between cognitive and physical activities .... 33  
   2.5 Accounts of the differences between novice and expert diagnosticians 34  
   2.5.1 The "hypothetico-deductive" model ............ 36  
   2.5.2 The "pattern recognition" model ............... 37  
   2.5.3 The relationship between the hypothetico-deductive and the pattern-recognition models ........... 38  
   2.6 Accounts integrating process and content .......... 39  
   2.6.1 Three-stage model of developing expertise ....... 40  
   2.6.2 The "holistic" model and metacognition .......... 43  
   2.7 Teaching Medical expertise ....................... 44
3 Abductive inference- an evidential approach

3.1 Abduction as plausible reasoning ........................................ 49
3.2 History of Abduction ....................................................... 49
3.3 The link between philosophical and computational accounts of abductive inference ......................................................... 50
  3.3.1 The computational point of view ...................................... 50
  3.3.2 The philosophical point of view ...................................... 52
3.4 Inference as an evidential process ....................................... 52
  3.4.1 Classifying inferences according to validity of conclusions they draw ................................................................. 53
  3.4.2 Classifying inferences according to productiveness of conclusions ................................................................. 53
  3.4.3 Classifying inferences according to the explaining-ability of their conclusions .................................................. 54
    3.4.3.1 Harman’s view ..................................................... 55
    3.4.3.2 Josephson’s view .................................................. 56
  3.4.4 What does an abductive conclusion explain? ....................... 57
    3.4.4.1 What does “inference to the best explanation” explain? ................................................................................. 57
  3.4.5 Ambiguity of the term ‘explanation’ .................................... 58
    3.4.5.1 The “relationship” ................................................... 58
  3.4.6 The relationship between abduction and induction ............... 59
3.5 Summary ............................................................................. 60

4 Abductive Inference- a methodological approach

4.1 Peirce’s method of scientific inquiry .................................... 61
4.2 Josephson’s method of diagnosis .......................................... 62
4.3 How contradictory are Peirce’s two classifications? ............... 63
4.4 The justification problem .................................................... 66
  4.4.1 The notion of self-correctiveness ..................................... 66
    4.4.1.1 Peirce and self-correctiveness .................................. 66
    4.4.1.2 Other views on self-correctiveness ............................. 68
  4.4.2 Self-correctiveness of only induction? .............................. 68
  4.4.3 Revision of Peirce’s self correctness claim: A circular, progressive method .................................................... 69
4.5 Pragmatism and continuity of the scientific method .............. 71
  4.5.1 A connection between pragmatism and justification of the method ................................................................. 71
  4.5.2 The active role of the reasoner and judgment of the inquiry method ................................................................. 72
4.6 Summary ............................................................................. 72

5 A knowledge level account of inference

5.1 Introduction ........................................................................ 75
5.2 Knowledge level in practice ................................................. 76
5.2.1 Some practical perspectives ............................................. 78
5.3 The connection between various abduction accounts ................. 79
5.3.1 The role of logic in analysing knowledge .............................. 80
5.4 The analysis of abductive inference patterns ............................. 81
5.5 The ‘why’ question .......................................................... 82
5.6 Abductive inference revisited from a knowledge level view .......... 83
5.7 The abstraction level of abductive inference ............................ 85
5.7.1 Abductive inference covers only the first stage ....................... 86
5.8 Summary ................................................................. 87

PART II: Relevance of Context

6 A methodology for modeling context ...................................... 91
6.1 Introduction .................................................................... 91
6.2 Distinction between the elements and the roles of context .......... 91
6.2.1 The role of context - relevance and focus ........................... 91
6.2.2 Elements of context ...................................................... 92
6.3 Research goals .................................................................. 92
6.4 A Task-centered methodology for studying context .................... 93
6.5 Categorizing contextual knowledge types ............................... 95
6.6 Categorizing context effects ............................................... 96
6.6.1 Context effects vary ....................................................... 97
6.7 The map between context elements and effects ......................... 98
6.8 Summary ....................................................................... 99

7 Context Studies in Cognitive Psychology .................................. 101
7.1 Cognitive science and AI .................................................... 101
7.2 Empirical research on memory ............................................. 102
7.3 Theoretical Research on Memory .......................................... 104
7.4 Recall Models .................................................................. 106
7.5 Only interactive context affects recognition ............................. 108
7.5.1 Context influence and ambiguous words ............................ 108
7.6 Summary ....................................................................... 110

8 Towards a content theory of context ....................................... 111
8.1 Introduction .................................................................... 111
8.2 Two views on context ...................................................... 111
8.3 Context as a knowledge type .............................................. 112
PART III: An Architecture for context-sensitive diagnostic problem solving

9 A context-sensitive iterative diagnostic model 123
9.1 Introduction .............................................. 123
9.2 A context-sensitive iterative approach to diagnostic problem solving 124
9.3 The view of a diagnostic label as an explanation 125
9.4 Iterative application of abduction-prediction-induction cycle 125
9.5 Internal context utilization ................................. 126
9.5.1 The shift of focus is goal-driven ........................ 126
9.5.2 Generated hypotheses provide further focus .......... 127
9.6 External context utilization ............................... 127
9.7 Loci of context effects in diagnosis ....................... 130
9.7.1 Context assistance in hypotheses generation ......... 130
9.7.2 Context in testing a hypothesis ........................ 134
9.7.2.1 Selection of the strategy for gathering information 134
9.7.2.2 Contextual role in prediction ....................... 135
9.7.2.3 Context and information gathering - planning ...... 135
9.7.3 Evaluative power of contextual knowledge .......... 136
9.8 Why does agreement with context imply plausibility ..... 137
9.8.1 Distributed evaluation ................................. 139
9.9 Summary .................................................. 140

10 ConSID-Creek - domain and task model 141
10.1 Introduction .............................................. 141
10.2 Methodology for Studying Context ....................... 142
10.3 Components of knowledge level; Domain, task and method 143
10.4 The ‘disease process’ and the ‘content’ .................. 144
10.5 The content knowledge in ConSID-Creek ............... 145
10.5.1 General domain knowledge .......................... 145
10.5.2 The domain ontology .................................. 147
10.5.3 Specific knowledge - case base ...................... 149
10.5.4 Integration of general and specific knowledge ...... 151
10.6 The control knowledge in ConSID-Creek ............... 152
10.7 The map between the control and the content knowledge .... 152
10.8 The loci, role, and source of contextual knowledge in the ConSID-Creek system ................................................. 153
10.9 The Task model in ConSID-Creek ........................................ 154
10.9.1 The task model of diagnosis .......................... 155
10.10 Summary ...................................................... 158

11 ConSID-Creek Methods and Mechanisms 159

11.1 Introduction .................................................. 159
11.2 Creek’s subtasks and methods .................................. 160
11.3 Mechanism for focus of attention .......................... 161
11.4 ConSID-Creek methods ..................................... 164
11.5 Method for T-generate-hypotheses ............................ 165
   11.5.1 Method that realizes T-formulate-hypotheses .......... 167
   11.5.2 Method for T-select-to-be-tested-hypotheses ........ 171
   11.5.2.1 Method for T-elaborate-formulated-hypotheses .... 171
   11.5.3 Method for T-select-hyp. ................................ 176
11.6 Methods for T-test-hypotheses .............................. 176
   11.6.1 Method for T-determine-hyp-testing-strategy ......... 177
   11.6.2 Method for T-determine-info-needs .................. 177
   11.6.3 Method for T-generate-expectations .......... 178
   11.6.4 Method for T-determine-tobe-gathered ............ 178
11.7 Method for gathering information ............................ 179
   11.7.1 Method for T-establish-info-gathering-context ...... 179
11.8 Method for T-make-info-gathering-plan ...................... 181
   11.8.1 Methods for T-generate-plan .......................... 181
   11.8.1.1 Method for T-formulate-plan .................... 182
   11.8.1.2 M-evaluate-and-select-plan-by-deep-knowledge .... 182
   11.8.2 Methods for T-test-plan ............................... 183
11.9 Method for T-assess-gathered-info .......................... 184
11.10 Summary ..................................................... 185

PART IV: Evaluation, Conclusion, and Future Work

12 Example 189

12.1 Introduction .................................................. 189
12.2 The patient reminds the system ............................. 190
12.3 Generation of diagnostic hypotheses ...................... 190
   12.3.1 Formulation of explanatory hypotheses .............. 190
   12.3.2 M-Activate-hypotheses: Case-based recall .......... 191
   12.3.3 M-Explain: Assess dEEper similarity ............. 196
   12.3.4 Selection of to-be-tested cases: M-Focus .......... 200

7
12.4  Test of the selected hypotheses ...................................................... 200
   12.4.1  Strategy selection ............................................................. 200
   12.4.2  Determination of information needs ..................................... 201
   12.4.3  Establishing the information-gathering context ....................... 202
   12.4.4  T-generate-plan ............................................................. 204
   12.4.4.1  T-formulate-plans - Activate ....................................... 205
   12.4.4.2  Selection of the plan to be performed: ............................ 206
   12.4.4.3  Selection of the most plausible plan candidate - focus .......... 208
   12.4.4.4  Reuse of past plans: T-test-plan .................................... 208

12.5  Summary ................................................................. 210

13  Evaluation ........................................................................... 211
   13.1  Introduction ................................................................. 211
   13.2  Research goal and aimed contribution ................................. 212
   13.3  What implies to consider context as vital in diagnostic knowledge bases 213
   13.4  Ontology ........................................................................ 214
   13.5  How good is the methodology? .............................................. 215
   13.6  How well, general and detailed is the abductive pattern? Does it cohere with and use the proposed context ontology? .................. 216
   13.7  Generality of the attention mechanism, its effects on relevance and efficiency .............................................................. 217
   13.8  How well the ConSID approach integrates the methodology, ontology, the inference patterns and the focusing mechanism .......... 218
   13.9  ConSID-Creek .................................................................. 218
   13.10  Related work ............................................................... 222
   13.11  Summary ........................................................................ 224

14  Conclusion ........................................................................ 225
   14.1  Contributions ............................................................... 225
   14.1.1  Epistemological adequacy ................................................. 226
   14.1.2  Adequacy of information processing .................................... 226
   14.1.3  Basis of the context framework ......................................... 227
   14.1.4  Expert/novice differences ............................................... 227
   14.2  CBR and External Context .................................................. 228
   14.3  Multi-directionality of inference to the diagnosis. .................. 229
   14.4  Characteristics of diagnostic problems .................................. 229
   14.4.1  Impacts of context in a diagnostic system .......................... 230
   14.5  A high-level context ontology .............................................. 230
   14.6  The methodology for analysing and modeling contextual knowledge .............................................................. 230
   14.7  (Purpose-oriented) logical aspect of abductive reasoning .......... 231
   14.8  Psychological aspects of abductive reasoning .......................... 232

225
14.9 Synergistic and additive effects of contextual features 233
14.10 Holistic case representation 234
14.11 A context-sensitive, iterative approach to diagnosis 234
14.12 A context-sensitive, CBR-based method for diagnosis 235
14.13 Future work 235

15 References 237
Chapter 1

Introduction and Overview

1.1 Motivation

Real world problems often involve weak and open domains, where the domain knowledge can not be expressed in the form of strong theories, and where information becomes available gradually. Consequently, problem solving in weak and open domains is characterized by scarce information and uncertainty, and relies on plausible reasoning such as abduction. Plausibility identifies what is rational, what reasonable people would agree on. Abduction allows inference to the most plausible hypothesis, from a set of observations and has been commonly agreed as to play a key role in human reasoning activities involved in, for example, text comprehension and diagnosis ([Pople 82], [Reggia 83], [Ramoni 92], [Josephson 94]).

The computational implementations of abductive inference are known to suffer mainly from two important problems: the low quality of abductive conclusions and intractability of the abductive process [Reggia 83]. This thesis addresses these problems through investigating the connections between the notions of context and abductive inference, and works out a context-sensitive abductive inference account. The nature of the thesis is analytical and therefore its main contributions in this respect is at the conceptual level.

The motivation behind this research is to improve artificial intelligence methods for diagnostic problem solving by developing a knowledge level account of contextual knowledge and of its use in diagnosis. This thesis focuses mainly on the conceptual, implementation-independent aspects of contextual knowledge and, on this basis, constitutes a step toward a generic model of contextual knowledge and a methodology featuring its use for diagnostic knowledge-based systems. Since Newell put forward the knowledge level hypothesis [Newell 82], a variety of researchers has viewed this level as the right level of abstraction for knowledge acquisition and engineering ([Bylander 88], [Musen 88], [Steels 90], [Wielinga 93]). Newell distinguishes between notions of knowledge and representation where the latter refers to the actual data structures and the algorithms in an AI program. He stated that AI research focused, to that date, more on the representation notion and less on studying the nature of knowledge. Since then, significant effort is put into the knowledge level research. The knowledge engineering community has continuously provided abstract descriptions of the expertise in a problem domain by modelling the tasks and the involved domain knowledge. The results of this kind of research are reusable ontologies and methodologies for acquiring and structuring knowledge. However, though the significance of context has, in general, been acknowledged in the AI field and the knowledge engineering community, contextual knowledge has been ignored in knowledge level accounts.
The thesis argues that a research involving a knowledge-level model of context and of context-sensitive abductive inference should precede the implementation efforts for improving flexibility and efficiency of knowledge-based diagnostic systems. The work presented here concentrates mainly on this ‘knowledge-level’ research. It explores at this level

(i) the notion of context, its elements, its role with respect to ‘knowledge’ and ‘knowledge use’ issues, (ii) the issues of justification of abductive process and the abductive conclusions, and (iii) the connection between context and abductive inference with respect to experts’ diagnostic reasoning

The proposed knowledge level account of context is implementation-independent and can be implemented in various ways. In order to exemplify the developed conceptual framework, a case-based reasoning method is proposed that reflects a ‘memory-based abductive inference method’. This is only one way of implementing the proposed context model. In this implementation medical diagnosis problem has been chosen as the example domain, and a medical problem solver has been partially implemented as a testbed and for demonstrative purpose.

1.2 Research goals

The overall goal of the thesis is to better understand the context notion with respect to (diagnostic) expert’s abductive reasoning, and to develop a holistic model of context bringing to bear the related research from diverse disciplines. In order to achieve this goal, the thesis sets up the following subgoals:

1. development of a high-level, reusable contextual knowledge ontology for diagnosis
2. development of a generic, context-driven mechanism for attention focusing, which provides a basis for evaluation of the efficiency of the diagnostic process
3. development of a methodology for analysing and modeling contextual knowledge for diagnostic tasks
4. development of a context-sensitive formulation of abductive inference that provides a basis for evaluation of diagnostic conclusions
5. exemplifying the implementation of the proposed context-sensitive diagnosis model through one particular method, the knowledge-intensive case-based reasoning method

These are all (except the last one) theoretical issues and the results obtained from these research tasks can be applied across various diagnostic problems. The questions addressed are ‘what is context’, ‘how its elements can be identified’, ‘how its role can be determined’, and ‘how context can guide selective processing of knowledge/information’.

To couple the proposed theoretical framework with the research into knowledge-based systems, a system architecture is developed and a test system implementing the core methods of the suggested architecture has been developed. This demonstrative system concentrates only on some parts of the theory.
The research strategy adopted for achieving these research goals are described in the next section. Following this, in the rest of the chapter we elaborate on each of the subgoals presented in this section.

### 1.3 The research strategy

Researchers in artificial intelligence have recently been concerned with the notion of context in connection with developing decision-support systems that help to solve problems in open and weak theory domains. Though it has been recognized (e.g., [Pauker 76], [Console 93], [Turner 94]) that problem solving needs to be a context sensitive process in domains where knowledge is not represented by strong theories and the problems are to be solved with incomplete information, the notion of context has not yet been thoroughly investigated at the knowledge level. We intend for this thesis to be a significant step toward recognizing, acquiring, structuring, and using contextual knowledge in knowledge based diagnostic systems.

Context has been investigated in various disciplines such as linguistics, psychology and artificial intelligence (AI). However the literature lacks a holistic look into the context notion within problem solving because context has been given different meanings and roles across different disciplines. This may be a sign that different disciplines study only a particular aspect of this notion. We argue that a unified and more holistic account of context is needed. Therefore, in our attempt to give a particular meaning to the notion of context in AI, we adopted the strategical goal of combining theoretical findings and ideas from various disciplines.

In this work we first explore the context research in other disciplines than AI in order to figure out which diverse disciplines may be studying the same problem or parts of it, how they study it, and what they have found. Figure The figure shows the disciplines and some of the concepts which we found relevant within these disciplines that we benefitted from for an AI account of context. illustrates the disciplines and the research fields within each discipline that have helped to shape our context account in diagnostic problem solving.

Researchers in linguistics and communication have studied the effects of context on the interpretation of an utterance. Their aim has been to understand how the meaning of an utterance changes when the context of the utterance changes.

Researchers in cognitive psychology have studied the context effects on various cognitive processes such as perception, learning and remembering.

In the medical research literature, a great majority of medical researchers and people practising medical problem solving agree that diagnosis involves a stage of generating hypotheses and one of testing these hypotheses. However, disagreements remain as to the way these two tasks are carried out, and as to whether novices and experts do carry out these tasks in the same way. It has been argued, in instructional research (or educational psychology) that there is a difference between novices and experts regarding the methods to generate hypotheses and regarding the type of memory used to accomplish this task. A prominent approach that has emerged in instructional research is that developing of expertise goes hand in hand with developing an episodic memory [Schmidt 93]. This agrees
with the argument posed by medical research that experts, unlike novices, apply a pattern recognition technique when reasoning towards a diagnosis.

![Diagram](image)

**FIGURE 1.1** The figure shows the disciplines and some of the concepts which we found relevant within these disciplines that we benefited from for an AI account of context.

Medical research does not provide much material that explains the nature of contextual knowledge in the medical reasoning process, except referring to a list of contextual elements they note in a patient history. This is not sufficient for modelling context. However, cognitive psychology provides useful research results regarding various types of context and their effects, though the tasks for which context has been studied are different from
medical problem solving. The way context is studied in cognitive psychology has
provided us insight into the study of context in the diagnostic domains.

We have devoted separate chapters to empirical study of context in cognitive psychology,
thoretical accounts of context in cognitive psychology, and abduction accounts in philos-
ophy, in particular, in connection with Peirce’s [Peirce 58] inference accounts and the sci-
entific inquiry model. The underlying reason for this particular interest and consequently
devoting special chapters to these issues is that complex real-world domains typically
involve abductive inference, since uncertainty is high in such domains, and context is typ-
ically utilized in case of uncertainty and ambiguity.

From this first stage, i.e., exploring the context study in other disciplines that may be rele-
vant for our context account in AI, we picked up a set of theoretical and experimental
results on which we base parts of our context account.

The methodology for analysing and modelling context, proposed in chapter 6, reflects
also our research strategy. Because, after deciding that the notion of context needs be
broken down into its elements and its various types of roles, we put a structure on what
we have found useful in the literature. The elements of context lead us towards a context
ontology. The study of medical reasoning and instructional research as well as inference
study and study on scientific inquiry has given insight into the our development of the task
structure pertinent to the diagnostic process. Having an ontology and a task structure
reflecting diagnostic thinking, we could locate context effects in the task structure. This
implicated the research subtask of identifying and explicating the various types of context
roles. Based on the idea of the dependence of context roles on tasks characteristics, we
started developed a task-oriented ‘attention-shift’ and an ‘attention-focusing’ mecha-
nism.

1.4 The context ontology

In recent approaches ([Wielinga 93], [Van Heijst 97]) knowledge engineering is viewed as
a modeling process not as a mining process as viewed earlier. Ontologies are important in
this modeling approach; there is agreement on this. However, there is no standard defini-
tion of ontology. Moreover, the developed ontologies vary in several respects. An impor-
tant distinction is between domain ontologies, and task and inference ontologies. In
addition the existing ontologies varies enormously with respect to the degree of details.
However, one thing common to almost all the existing ontologies developed in relation
with diagnostic knowledge engineering is that context has not been taken into particular
consideration. In this thesis we present a high-level context ontology. Since there exist no
agreed definition of the term, we will first give a definition reflecting what the context
ontology in the thesis means. In this thesis, the term ontology is used in a similar way as in
[Schreiber 95]. So, the context ontology described in this thesis is an

...explicit, partial specification of a conceptualization that is expressible as a meta-
level viewpoint on a set of possible domain theories ... [Schreiber 95].

The main purpose of developing such a context ontology is to triggering explication of a
shared conceptualization. ‘Shared conceptualization include conceptual frameworks for
modelling domain knowledge’ (Tom Gruber - in a message to the ‘Shared Reusable
Knowledge Bases mailing list, 1994 -, referred to in [Guarino 96]). As we develop a high-level context ontology reflecting the commonalities between various diagnostic domains, the developed ontology is reusable by others interested in contextual knowledge for different diagnostic domains.

Thus, an important role of the context ontology is to put a structure on the contextual knowledge pertinent to diagnostic domains, and to identify the basic context-sources which may guide how the knowledge engineer can elicitate this type of knowledge in a systematic and deliberate way.

At the same time, since the methodology we develop for analysing and modeling contextual knowledge and its use relies on a context ontology, the ontology also guides modeling the role of contextual knowledge and use by diagnostic tasks.

1.5 The generic mechanism for attention focusing

This thesis considers a diagnostic problem solving session as a sequence of task accomplishments. The order of the task execution determines the line of the problem solving. During the accomplishment of each task, the attention of the problem solver is focused on a small set of information and knowledge in the domain. Part of this focus consists of contextual elements, and the other part of core domain elements. The thesis presents in chapter 10 a context-sensitive iterative diagnosis (ConSID) model that describes, for each task where the reasoner needs to focus her attention, through a mechanism that we refer to as “perspective”. The task perspective specifies the type of information to be processed at that moment. When a new task is invoked, the focus of attention will be moved (if needed) toward the knowledge portion determined by the new task perspective. A task is invoked by a specific type of internal context element, namely the goal, and the focus is captured by other contextual elements in combination with core domain elements.

1.6 The methodology for analysing and modeling context

After Newell’s distinction between the implementation independent knowledge level and the level of implementation and representational constructs which he referred to as the symbol level [Newell 82], research in the area of knowledge modeling and engineering has produced several methodologies for analysing and modeling knowledge and information at the knowledge level. Influential examples of methodologies that support knowledge analysis and modeling at the knowledge level are the Generic Tasks approach [Chandrasekaran 92], the COMMET methodology [Steels 93], and CommonKADS [Breuker 94]. Inspired form these methodologies we developed a methodology that guides analysis and modeling of contextual knowledge.

The methodology is grounded on a distinction between two aspects of context: its role and its elements. We argue that a systematic investigation of the notion of context needs to be organised along these two aspects. Regarding the roles that context plays we distinguish between two basic issues. These are related to its role in capturing the focus of attention, and in imposing a shift in the focus of attention [Öztürk 98a; Öztürk98b].
The process of modeling the contextual knowledge then can be considered around two research activities: (i) typing/categorizing contextual knowledge, and (ii) identifying what kind of role each type plays, and where/when this happens during problem solving. Implicitly suggested in these two activities are the two perspectives that guide us toward an adequate account of context at the knowledge level. These perspectives are not mutually exclusive, but closely interrelated. They consider context, respectively as

- a knowledge type, and
- a trigger for shifting the focus of attention.

The first perspective leads to the study of how contextual knowledge can be categorized, and aims at constructing a context ontology that guides the modelling of contextual domain knowledge. The thesis presents a high level diagnostic context ontology that draws from the research on recall in particular, and memory in general, in cognitive psychology [Öztürk 97]. This constitutes the basis for acquisition and modelling of the contextual knowledge in a particular application domain. In addition to the context ontology a methodological framework is required in order to analyse and model the relevant aspects of context. The second perspective deals with the issue of ‘focus of attention’ and lead to the study of the relationships between a task and the type of the contextual elements it relies on. This perspective takes, as a starting point, an application task (in our case diagnosis) and identifies the subtasks of the overall task process on which some contextual influences are anticipated, identifies the role of contextual knowledge in each subtask, and determines which type of contextual knowledge plays this role for the particular subtask.

We may also say that the first perspective adopts a static view, while the second one is more concerned with dynamic aspects of context, i.e., its use.

1.7 The abductive inference pattern

Explanatory reasoning, which is tightly connected to abduction constitutes an important part of a medical expert’s reasoning toward a diagnosis.

Philosophy has been the pioneer in investigating ‘explanatory reasoning’, usually under the heading of abductive inference. According to recent interpretations abductive inference has been studied in philosophy from two perspectives, referred to as the evidential and methodological views. Our approach to abductive inference is from a third view, the computational point of view.

We have questioned in this research whether and how we could take advantage of philosophical accounts in order to develop a powerful knowledge level model of abductive reasoning. In this model we wish to reflect the contextual aspects of that type of inference - which lacks, or in the best case, is insufficient in the existing accounts- through the notion of evaluation. We have discovered that a computational approach should subsume and integrate two dimensions of abductive inference which are respectively studied by evidential and methodological aspects. In a sense, our computational approach may be seen as putting these philosophical research results into a unified framework. This is possible because the evidential and methodological approaches to abductive inference are not mutually exclusive, as has long been considered in philosophy, but orthogonal. Each of
these approaches investigate only one aspect of abductive inference while a computational account of abductive inference, on the other hand, must consider both aspects.

The computational view should first describe abduction at the knowledge level. We model abductive inference as a type of task at that level. Knowledge level model of a task is described in terms of elements such as goal, input, output, and the method that carries it out. We can conceptualize such a model as consisting of two main components. These components are related to

- the content of abductive inference, described in terms of its goal, input and output (i.e., captures the logic of abductive argument). The evidential view investigates this aspect.

- the process of abductive inference, which captures the logic of the strategy underlying abductive inference.) The methodological view investigates this aspect.

We need to clarify the connections between the computational and the philosophical account of abductive inference. The evidential account of abductive inference, which is expressed in terms of premises and conclusions, provides insight into the input and output of the model of abduction task at the knowledge level. We augment the evidential considerations by taking into account the notion of context. We claim that the addition of contextual considerations into computational models of abductive inference will improve the efficiency of the computation as well as the plausibility of the results.

To summarize, we have utilized evidential accounts of abduction in order to understand the logic of the argument of an abductive inference which provided insight into describing tasks at the knowledge level, in terms of their input, output and epistemological requirements. On the other hand, the methodological approach guided us when constructing a task model where the complex task of diagnosis is decomposed into its subtasks, as well as the order in which these subtasks are accomplished.

Regarding the method of abductive inference, we experiment with a knowledge-intensive-case-based reasoning method carrying a psychological soul as well as a logical one.

1.8 Diagnostic knowledge-based systems

A diagnostic problem is a problem in which one is given a set of manifestations and is expected to find an explanation to why they are present. Diagnostic problems can be found in various areas, for example, in diagnosing a patient's complaints, determining why a car will not start, or in localizing a fault in an electronic circuit.

Because of ubiquity of diagnostic problem solving and its being an integral part of every day life, developing knowledge based systems that support the decision making of human diagnosticians, as well as systems that facilitate teaching and learning of diagnostic competence has been an important issue in AI since 70's.

Medical diagnosis has been one of the earliest domains to which AI applications were suggested. Some early examples are MYCIN [Shortliffe 76], PIP [Szolovits 78], CASNET [Weiss 78], and INTERNIST-1 [Miller 82].

Though reasonable methods have been proposed for solving each of these problems, the slow advancement in the development of knowledge based systems capable of supporting
competent and efficient complex diagnosis (e.g., medical diagnosis) strongly indicates that a more holistic approach to these problems taking into consideration the notion of context would be worth to investigate.

A vital drawback of most existing knowledge-based systems is that they are not capable of justifying their behaviour. In connection with the study of abductive inference, the recent research started to explore the notion of justification ([Ng 90], [Thagard 89], [Josephson 94]). If a system is able to justify its behaviour it may evaluate its own behaviour by integrating an evaluatory character into its reasoning. The notion of context has only arbitrarily been considered in a few of these without explicitly formulating the relations of this notion to the notion of justification in abductive inference. In fact some recent approaches accepted the importance of additional factors other than manifestations for justification. For example, the age and the sex of the patient, and the interdependence of diseases has been mentioned to be important for diagnosis. However, a thorough study of this kind of ‘additional’ knowledge has not been made. In this thesis we focus on the prominence of justification - self-evaluation may be a better term - in diagnostic systems, and explore the possible connections between the notions of context and evaluation in diagnostic abduction.

The thesis puts a distinction between two kinds of evaluation of the generated diagnostic hypotheses. The first one is justification of a hypothesis’ being formulated at the start, based on the (initially) available evidence. This happens during formulation of hypotheses. The second is evaluation of a formulated hypothesis (thus after a hypothesis is adopted for further scrutiny) on the basis of information gathered after a set of hypotheses are selected for testing against new evidences.

The first of these include both evaluation of the abductive process underlying the activation of hypotheses, as well as evaluation of the conclusion (i.e., hypotheses) drawn by the process. The second one involves evaluation of formulated hypotheses by gathering and checking against new evidence (i.e hypothesis testing). In both kinds of evaluation the notion of context plays a central but different role.

The basic idea for justification of plausibility of formulating a hypothesis at the first place is due to Peirce’s abduction study [Peirce 58] and can be described as that

- the hypothesis must be likely in itself
- the hypothesis must render the manifestations likely

In most existing systems the second point has been taken as the only concern underlying the justification issue. The problem with these systems, in fact, is that even such a distinction has not been explicitly formulated. In this respect, the developed context-sensitive abduction model takes into consideration both the likelihood of a formulated hypothesis in itself and the hypothesis’ rendering the manifestations likely.

1.9 The context-sensitive, iterative diagnosis model

We introduce a diagnostic problem solving model, named ‘Context-Sensitive Iterative Diagnosis’ (ConSID). ConSID conceptualizes diagnosis as a context sensitive process
integrating reasoning and action. The main principles on which the ConSID approach is based can be listed as follow:

1. diagnosis is a combination of explanation and planning

2. diagnosis is realized by an iterative application of an abduction-prediction-induction cycle

3. the problem solving process is goal-driven, with a shift in the focus of attention triggered by a change in the active goal.

4. the focus of attention is captured in terms of core and contextual entities

5. pattern recognition is an important component of diagnostic reasoning at the expert level

This approach to diagnosis is context-sensitive in two ways. First, the reasoning line is goal-driven. Second, the focus of attention is described in terms of a combination of contextual and core domain elements.

1.10 The knowledge-intensive case-based method

Studies in cognitive psychology (chapter 7) emphasize the role of "experience" in learning and remembering. An experience is something more than, for example, a word to be learned and includes the context in which the word is learned. We find a close parallel between studies in cognitive psychology on episodic memory and the case-based reasoning paradigm in AI. Tulving's distinction between episodic and semantic memories (Tulving 73) matches perfectly with a distinction between case specific and general domain knowledge. This thesis conceptualizes the representation of what is learned as an experience which embeds a combination of contextual and core domain elements in the form of a case.

Other disciplines that provided much support and insight are medical research and instructional research. Prominent researchers (e.g., [Patel 91]) argue that medical experts use a 'pattern recognition' method in diagnosis agrees with the idea behind the case-based reasoning paradigm; experienced reasoners match the new problem with the patterns they have in their minds, and use these patterns as short-cuts. On the other hand, the disease model, referred to as illness scripts by Schmidt [Schmidt 93] and his colleagues has been the basis for our case content. The link between process and content, emphasized by Barrows, related to teaching medical expertise guided us toward the way we study the contextual knowledge, a continuous attempt to maintain the link between the reasoning and the disease processes.

In the medical research literature, we often encounter statements revealing intuition in medical diagnosis, in particular, in connection with 'guessing' a diagnostic hypothesis. The connection between the notions of intuition and experience has been recognized [Bastic 82]. We suggest that the process of making guesses can be realized by case-based reasoning. This we refer to as a memory-based abduction method.

The Creek [Aamodt 91; Aamodt 94] framework offers a case-based approach to diagnosis. The demonstrative system (ConSID-Creek) developed within this work is an imple-
mentation of the ConSID problem solving approach that takes Creek as the basis and refines and extends it by incorporating the context notion into it.

ConSID-Creek consists of a case-based explainer and a case-based planner. In this system new explanations are formed by retrieving and modifying previous explanations which have been useful in similar contexts. Similarly, new plans are made on the basis of a "retrieve and modify" principle.

The similarity assessment in ConSID-Creek differs from most of the case based explanation and planning applications in that it explicitly considers contextual aspects as a parameter included in the similarity measure.

Different contextual elements are relevant in case-based explanation and planning components of the system. In addition, different contextual elements are relevant in various subprocesses of case-based explanation and of case-based planning processes.

1.11 Overview of the thesis

Our aim, in general, is to understand and formulate the role of context in problem solving. For this purpose, we adopted the strategy of focusing on complex tasks which are frequently encountered in the real world either in everyday life or in scientific research. These are also the domains where abduction is typically the keystone in the problem solving process. For this reason, the relations between the notions of abductive inference and context has a vital importance in our research. Thus, our strategy has been to do basic research in explicating and modeling the elements and the role of context in solving complex real-world problems. For this purpose we chose diagnostic problem solving, in particular, medical problem solving, as the example domain.

The rest of the thesis is organised into four parts. In each part a different aspect of context has been focused on. In part I, the abductive pattern is augmented by contextual elements in order to improve the quality of abductive conclusions. In part II context is dealt with from an epistemological dimension and investigated as a special type of knowledge. It has also been viewed as a mechanism that features the focusing of attention on certain portions of knowledge. In part III context sensitive diagnostic approach (referred to as CONSID approach that stands for 'context sensitive iterative diagnosis') is presented integrating the developed ontology, the abductive pattern and the other inferences underlying the diagnostic process, and the focus mechanism. The case-based reasoning paradigm is used in the computational realisation of the abductive inference in order to alleviate the inflexibility problem. Part IV presents a detailed example from Consid-Creek and close the thesis with evaluation and discussion chapters.

Part I consists of chapters 2, 3, 4, and 5 and deals with problem solving and the role of various types of inferences in diagnostic problem solving. In chapter 2 we take a look into medical domain as an example to diagnostic problem solving and introduce how the medical professionals view their own knowledge and thinking. The medical domain is introduced from the perspective of an everyday event- an everyday event both from the medical expert and from the patient dimension. The chapter summarizes a review of literature illustrating how the medical experts talk intuitively about their task. We highlight the tasks, skills and type of information that the expert themselves present as important for
their job. Presentation of such an intuitive perspective reflects also the way we started to study the medical domain. Then we move towards a more systematic analysis of the characteristics and the components of the content and the process related aspects of the domain. We emphasize the plausible character of the reasoning underlying clinical diagnostic problem solving. In chapters 3, 4, and 5 we deal with the abductive inference which is of prominent importance for diagnostic problem solving. The openness of the clinical problem descriptions and the weakness of the medical domain theory are some central characteristics of that domain. These lead to ambiguity which relies on clarification and focus by context.

We start with studying two dominant accounts of abductive inference in philosophy, elaborated respectively in chapters 3 and 4. We propose a third, a knowledge level account which integrates philosophical ones into a more holistic approach in chapter 5. The abductive pattern we propose in this chapter is context sensitive and gives a special emphasis to the issue of quality of the abduced conclusions.

Part II focuses on the notion of context, and spans chapters 6, 7, and 8. In this part context is dealt with as a special type of knowledge. Through these chapters we address several questions that have often been asked. For example, 'Does context represent the state of the mind or the state of the world?'. Our answer is 'Both, it has both internal and external components'. Another issue is related to differences between context that has semantical relationship with the 'to be remembered events', and the one that has no such relationship. We clearly distinguish between these two types (i.e., independently and interactively encoded contextual elements) and formulate how they effect different reasoning tasks. We also model the various kinds of roles that context plays in the course of the event under consideration. This we exemplify in the context-sensitive model of the disease process.

Experiments conducted in cognitive psychology in order to investigate context have been instrumental. Various ways that the experiments are arranged give us a possibility to interpret the differences between various types of contexts. Even though the choice as to the arrangement of experiments seems to be more or less arbitrary, or at most intuitive, findings are tangled, and the results are inconsistent and unsystematically presented, we found these empirical studies rather fruitful for our construction of a general context ontology, illustrated in figure 8.2 in chapter 8.

Part III includes chapters 9 to 11 and presents an architecture for context sensitive diagnostic problem solving. Here, we combine our theoretical contributions into an AI model of diagnostic problem solving and demonstrate how such a model can be realized on a partially implemented demonstration system, ConSID-Creek, an extension of the Creek framework.

Part IV consists of chapters 12, 13, and 14. Chapter 12 goes through an example. Chapter 13 presents an evaluation of the theory and the demonstration of the theory in medical domain, and the future work that seems complementary. Lastly, chapter 14 closes the thesis with conclusions reached.
PART I
Problem Solving and Inference
Chapter 2

Medical problem solving

2.1 Introduction

The medical domain is a typical and complex example of diagnostic problem solving and abductive inference. This chapter will illustrate a real life example to human reasoning involving abductive inference and will allow us to identify the main aspects to take into consideration in a context-sensitive model of abductive inference and diagnostic problem solving.

In our attempt to constructing a knowledge level model of contextual knowledge and its use in diagnosis, we benefit from a bottom-up approach. Though we intend to construct a generic, reusable model of context we use medical domain in order to abstract from the reusable characteristics pertinent to the diagnostic contextual knowledge.

In this chapter we will identify the principal components of medical problem solving on the basis of (i) what the medical experts thinks these components are (ii) how research in educational psychology characterizes expertise.

The first part of the chapter (sections 2.2-2.4) can be seen as a survey of medical reasoning research literature. Our goal, in fact, is to be able to transfer these components from the human expertise plane to a ‘knowledge level’ expertise plane. That is, we use what we filter from medical experts’ introspection related to how they solve medical problems. At the most abstract level, medical professionals mention three elements as covering the abilities of a medical expert [Wright 79]. These are reasoning skills, physical skills, and knowledge. We aim at a model that captures reasoning skills, physical skills and the memory of the medical professional, in terms of task, method and domain knowledge. In this part, we take a look at the human model of expertise as reported by medical diagnosticians themselves.

The second part of the chapter (starting from section 2.5) discusses various accounts of expertise and its development. A particular emphasis is put on the impact of the structure and content of knowledge on the reasoning performance. The conclusion we draw in this part highlights the importance of the illness scripts/cases and the contextual knowledge for expertise.

2.2 Main considerations of a clinician

Time pressure is a characteristic of medical problem solving. The physicians are responsible for quickly deciding whether the case is urgent, whether they should initiate some treatment, or they have time to make a correct diagnosis before they start a therapy. For example, in the case of traumas and sometimes in cases of infectious diseases, the doctors
may not have enough time to find out exactly what has happened, but they nevertheless need to initiate some form of treatment in order to avoid an unrecoverable development of the illness.

From the view of the patient, the clinician listens to the complaints, asks questions, examines the patients, orders laboratory tests, and writes a recipe or proposes a clinical intervention such as surgery. This is a mental picture of the clinician's problem solving in the patient's head. For our purpose, what is more important is the mental picture in the clinician's head. Such a picture can help us to identify the types of knowledge and information that assists the clinician in various stages of the process, as well as to identify these processes themselves. We will first identify the reasoning skills, domain knowledge and the physical skills that an expert possesses. This is the static aspect of our modelling problem. A more dynamic aspect is related to linking these three components, that is explicating which reasoning skill needs which type of domain knowledge and which type of physical skill has relation with which type of information, etc. We start with determining, from the medical experts statements in the literature, and our talks with medical experts, the static elements of the expertise. We will then recognize the links between these three static elements which will guide us towards a holistic picture of a medical expert's problem solving that can be used for the purpose of modelling a computational expert.

Wright and Macadam [Wright 79] summarizes a clinician's overall concerns under four questions:

What is wrong?,
Why has it happened?,
What is going to happen?, and
What should be done?

The attempts to collectively answer these questions are embedded in the diagnosis and treatment process. The first question is related to the features of the case which have been rendered abnormal. That is, the expert notices the abnormal findings, in addition to what the patient complaints about. The words used for conveying these features are important for our purpose, because a domain model should include these. The second question seeks an answer to 'what disease has caused these abnormal features'. The diseases are also important entities in the domain model. The third question asks what more can be expected to be observed of a patient over time. These "to be expected's" are also findings (or manifestations) and will be included in our domain model. The last question concerns also which action to take in order to establish a diagnose. For example what questions to ask to the patient (referred to as history taking or interviewing), what laboratory tests to order (referred to as laboratory investigations), what physical examination to perform, etc. Further, after establishing a diagnosis, a proper treatment needs to be suggested. For our purpose this means that we should include in our domain model the physical actions that support diagnostic explorations or therapy.
2.3 Overview of Clinical practice

Problem solving in a medical setting may be likened to solving other problems in the real world. For example, the reasoning of a detective solving a criminal case, or the method used by scientists conducting empirical research.

All are ill-structured problems. The main characteristics of ill-structured problems are their having inadequate information at the outset, and the lack of definite guidelines for solving the problem. In both domains the information is insufficient at the outset in order to jump to the result directly.

Instead, the problem solver generates a set of possible, tentative hypotheses suggested by the available information. These hypotheses then guide further exploration of the patient’s problem. So, selection of the strategy for gathering information is guided by these hypotheses ([Wright 79], [Elstein 78], [Feltovich 84b], [Barrows 91], [Elstein 95]).

Needed information can be gathered in clinical problem solving via various invasive and non-invasive methods such as interview, examination, X-ray, etc. The possible questions that a doctor can ask to the patient, or possible examinations she can perform are infinite. The process of information gathering is guided by the disease hypotheses that are tentatively generated early in the patient encounter. The process of hypothesis generation will be referred to as abduction in later chapters. The hypotheses indicate the manifestations that can be expected to be observed if the hypotheses are correct. The process where hypotheses guide the expectations as to what will be observed is referred to as prediction in later chapters.

The gathered information is analysed to judge whether it leads to re-definition of the problem- this process is referred to as induction in later chapters. The collection of information that becomes available as time goes by may change the clinician’s understanding of the patient’s problem. This procedure may entail that none of the hypotheses can adequately explain the symptoms and signs observed so far. Then, new hypotheses are generated which guide further clinical investigation until either they are shown to be inadequate, or one of them is judged to be the true cause of the patient’s problem.

Identifying the following stages of the overall process can help to conceptualize the cognitive behaviour of the professionals when handling patient problems, as adopted from [Barrows 91]:

1. Perception of the initial cues - forming initial concept
2. Rapid generation of hypotheses
3. Formulating an inquiry strategy
4. Applying appropriate clinical skills in order to apply this strategy
5. Interpretation of the new data and reformulation of problem if necessary
6. Diagnostic and therapeutic decision making.

The literature indicates that medical problem solving does not, in fact, constitute a sequential processing involving these aspects. For example, the doctor perceives the initial cues that she assesses to be important at any point in time. She may, as the process proceeds,
realize that the cues she was not aware of earlier, turn out to be rather important. On the other hand, the interpretation of new data may reveal that the initial hypotheses are not adequate, and she may need to generate new ones. That is, from stage 5, she turns back to stage 2. Our computational model of the diagnostic process reflecting this cyclic process is presented in chapter 9.

These are the main tasks of a medical professional. However, different professionals may accomplish these tasks by using different reasoning skills and physical skills. These skills are captured in our model as "methods". In the following sections we go through these six stages and try to look into the characteristics of each stage (e.g., what kind of input is needed to accomplish each stage, what is the output of each stage). This will guide us in determining what needs be modelled in order to capture the medical expertise. It is, at the same time easier to analyse the knowledge involved in medical problem solving when the whole process is partitioned into smaller subprocesses. Another advantage of partitioning the whole process is that we can readily relate the particular type of knowledge to the subprocesses where it is used. This simplifies the dynamic portion of our modelling job.

The strategy of decomposing a process into its subprocesses (i.e., reasoning tasks and physical tasks accomplished by the medical expert), has been of vital importance for our study of context in a diagnostic framework. By partitioning the diagnostic process, we were able to identify the various types of contextual elements, as well as the points where they are used (see section 6.4).

In the rest of the chapter we follow Barrow's six steps. However, the content of each subsection is not necessarily due only to him, but has been gleaned from the medical research literature ([Wright 79], [Elstein 78]). For the purpose of our research, the most important steps are the first three. Therefore they have been focused on in more detail than the last three.

**Initial Conception; Perception of initial cues.** At any time there is usually more available information than what a medical expert can handle. The initial conception is shaped by what the physician perceives, and reflects her synthesis of what she perceives. Thus, the initial conception is a transformation of the real world situation into a mental image in the physician's head. It is this conception that further guides the diagnostic process. This conception may change as new information becomes available. It is important at this stage that the physician includes as many relevant cues as possible and excludes irrelevant ones. The ability of doing so is dependent of the experience of the expert. Experienced physicians gather far more relevant cues than that the patients volunteer ([Barrows 91], [Harvey 79], [Taylor 88])

Since the interview with the patient is of crucial importance, it would be fair to examine the possible obstructions that may impact the communication. Some of these are related to the patient and may be a result of emotion, anxiety, fatigue, drugs, or depression. Others are related to the expert, which again includes emotions, impatience (e.g., time pressure) and preoccupation (e.g., recent quarrel in the family) [Wright 79].

The environmental conditions may also hinder a pleasant communication. For example, the room may be too warm or smell bad, or the light may be inappropriate. An extreme example will be a patient-doctor encounter occurring in a war zone.
We consider these obstructions as portions of contextual aspects of an interview. In fact, we have found a significant correspondence between the results from empirical studies in cognitive psychology and the medical experts' statements on the three types of obstructions that may arise in an interview session. As a result, we highlighted three dimensions of context; the problem solver (the medical doctor) related one, the target-related (the patient) ones, and the environment-related ones. These correspond to the grouping which the medical experts does when explaining possible obstructions that may arise in interview.

The cues important to one clinician may be unimportant to another one. So, the value of a particular fact may be subjective. This is stated to largely depend on the clinician's past experiences and on the objectives determined by the health-care setting, in which the interview takes place, for example hospital ward, ambulance or war-zone [Barrows 80].

The components of a medical interview have been divided into six groups [Greenberger 93]: chief complaints, history of present illness, past medical history, family history, social history, review of systems.

Included in a past medical history are all the past events of type pediatric, medical, surgical, psychiatric and obstetric. In a family history, the causes of death, age of death, living environments, and family disorders are included. Social history includes work conditions, stressful situations, home life, etc.

We distinguish between these various types of information gathered in interviews. Some are included in our core-domain entities while others in contextual domain entities, as will be seen in section 8.3. The following example is presented in order to illustrate how important a role the medical expert's knowledge can play in peripheral domains when solving a patient problem (see figure 2.1). The following example illustrates the importance of knowledge from peripheral domains for a faster and more accurate interpretation of the patient's problem (from [Barrows 91]):

A 54-year old Saudi Arabian Bedouin immigrant in USA complains of bilateral ankle pain associated with activity. The patient does not speak English. The physician is American. The patient's wife translates. The wife is in traditional cloths. The patient himself is quite shy and walks clumsily. He works in a lumber warehouse. In the examination the physician notes that the patient has lax ankle joints. There is no sign of inflammation or deformity. He has significant loss of light-touch, pin-prick and vibration sense throughout his feet. The muscles of his feet and legs are normal. The physician suspects a peripheral neuropathy. She based this hypothesis on the following possibilities emerging from the background of the patient:

- infectious diseases of the middle east such as bejel or leprosy.
- exposure to toxin, which may be caused by chemicals used in the lumber industry.
- a deficiency in nutrients such as B12, which may be due to change in food culture.
- side effects of traditional home-made herbal remedies

If the patient is not an Arab, the physician would tend to ask his alcoholic habits, but as muslims do not drink and he seems rather traditional the physician takes it as given that he does not drink. This reasoning failed, since he was a Bedouin, Christian, and, in fact an alcoholic. Her ignorance in other disciplines than basic medical science played a detri-
mental role in her diagnostic performance. Figure 2.1 illustrates the fact that common sense knowledge of a medical expert may interact with medical problem solving. The case of the 54-year old Saudi Arabian Bedouin is an example of how medical experts use their knowledge from a number of domains on the periphery of medicine, which is of type common sense knowledge.

![Diagram](image)

**FIGURE 2.1** Medical reasoning relies on knowledge from peripheral domains in addition to medical domain knowledge. Medical experts use a diversity of other (contextual) cues in addition to pure medical cues. So, physicians knowledge is combination of pieces of knowledge from anthropology, social life, religion etc., and basic medical science. This is because of the close intercation between basic medical science ant the knowledge pertinent to the disciplines in its periphery.

The cues picked out in an interview are interpreted so as to synthesize, possibly tentatively, an ‘initial conception’. This interpretation to an initial conception is based on the past experience and beliefs of the physician. Hence, past experiences, determine both what cues to pick out and how to interpret these cues. As will be seen in the next chapter, possession of experience and its use has been associated with the degree of expertise.

**Hypotheses generation.** A hypothesis, according to Webster’s dictionary (Random House, 1989), is “a proposition, or set of propositions, set forth as an explanation for the occurrence of some specified group of phenomena, either asserted merely as a provisional conjecture to guide investigations (working hypothesis) or accepted as highly probable in the light of established facts”.

It has been observed that experts are able to generate one or more working hypotheses rather early in the first encounter with the patient. Even from only knowledge of age, sex and the main complaints of the patient experienced physicians can make a ‘guess’ related to the cause of the complaints. What activates these hypothesis in the mind of the clinician? Is this an art? Peirce who studied the role of abductive inference in scientific inquiry method states that hypothesis generation is a guess, but a ‘skilled guess’. The term ‘skill’ as he uses is closely related to the experiences and beliefs of the reasoner. According to [Barrow 80, p. 23] who studies the learning of medical problem solving, the hypotheses are “usually a product of the clinician’s past experiences with patient problems. Their
retrieval from the physician’s memory bank is largely an unconscious act of memory association”. Patel and Groen [Patel 91] also emphasize the phenomena of "enhanced recall" in connection with the expert-novice comparison in the medical domain. Enhanced recall refers to the fact that “experts have superior memory skills in recognizing patterns in their domains of expertise”.

A new hypothesis is activated, according to Kassirer and et al.[Kassirer 85] whenever:

- there are some manifestations to be explained. A cluster of these manifestations may activate a hypothesis
- the new information contradicts an already existing hypothesis which has been considered.

Generation of hypotheses is of crucial importance as they guide the gathering of further information to help decide which of the hypotheses in the pool is the correct diagnosis. Hypotheses may be of diverse sorts, depending on the goal of the medical expert. In a hospital-ward, the goal is to generate an explanation of patient complaints. Then the hypothesis may be etiological. However, if the encounter with patient occurs in ambulance, diagnosing will hardly be the goal. A typical goal would be managing the bleeding. Accordingly, a hypothesis can be an adequate method of managing bleeding.

The information collected during an interview contributes to hypothesis generation. How this exactly happens is still a research subject. One important point that has commonly been emphasized in the literature is that past experiences play a crucial role in hypothesis generation. So, hypothesis generation is possibly a product of a process linking the cues available in interview, past experiences, and the expert’s professional knowledge. What the term “professional knowledge” includes is also a relevant question here. For example, the knowledge that ‘psychological stress may cause people to act carelessly’ is included in the physician’s “professional knowledge”.

Hypothesis generation is the progress from concrete and specific information to more abstract knowledge that explains this information. The initial conception may be a step toward linking those most specifics to most abstracts. Computational approaches to diagnosis have utilized the concept of abstraction. Clancey’s heuristic classification is an example based on abstraction from data. In this approach, a move from data to a ‘data abstraction’ allows the reasoner to make heuristic matches between the data abstractions and patterns of problem solutions [Clancey 85].

A physician may generate a hypothesis either directly from cues, or more usually first proceeding in steps upward to initial conceptions and then to hypotheses. It depends on the existing connections between concepts in the memory of the physician. The cues which are the observables are available earliest in the patient encounter. Initial conceptions are rapidly constructed by combining various subsets of cues. The combination of initial conceptions enables the physician to reach a diagnostic hypothesis. However, sometimes, “the clinician recognizes - identifies - a patient’s illness as a case of a disease because the physician has memorized the elements that make up that pattern when elements of the right kind are present in combination” [Albert 88, p 179]. In this way the physician may reach the most abstract level (i.e., hypothesis) without consciously moving through the whole chain connecting cues to hypotheses.
The endeavours of both initial conception and hypothesis generation can be considered as attempts to transform an open-ended problem to a more closed, that is, to a framed one [Barrows 84].

A clinician cannot generate hypotheses until she starts to understand the patient's situation. If the clinical findings are scarce and of a general type that can be associated to very many hypotheses, the clinician would not be able to generate adequate hypotheses. The expertise reveals itself in the ability to quickly find additional cues that contribute to making a meaningful cluster of manifestations. An example of using such an additional clue is inferring from the perceptible deformities in the patient's physiology that the patient is abusing alcohol.

Once a set of hypotheses are formed in the mind of the physician, she needs to acquire more information to shape or refine her temporary hypotheses. The medical experts use a variety of information gathering techniques such as interview, physical examination, and laboratory tests. This process, again according to Barrows, involves a "deductive, problem-oriented" reasoning.

**Formulation of inquiry strategies.** The generated hypotheses guide the physician's inquiry in pursuit of gathering more information and data from the patient. More information is needed so as to support or weaken the hypotheses temporarily generated. By Barrows statements the hypotheses guides the physician in selecting "a strategical sequence of inquiries designed to ferret out data that will deduce, verify, deny, focus, or rank the hypotheses in his head in the most efficient and effective manner possible" [Barrows 80, p 27]. That is, further inquiry following hypotheses generation is necessary for evaluating the generated hypotheses, and recognizing the most probable one.

Nevertheless, there is no straightforward way of acquiring new information. The generated hypotheses plays a particular role in bringing a structure to information gathering.

Barrows [Barrows 80] characterizes the physician's inquiry as a combination of "search and scan" processes. The physician performs a 'search' when she knows what she is going to look for. She has some expectations and checks whether these are true. The expectations are imposed by hypotheses. Quite often she attempts to confirm or rule out a hypothesis. However, when she does not know what to do to differentiate between hypotheses, she may use 'scan' type of questions.

So, hypotheses are used to generate expectations related to what would be true (i.e., which manifestations) if the hypotheses were correct. The expert predicts the manifestations. The manifestation that are not yet known should be gathered.

As already mentioned we will not elaborate the following three steps of diagnosis in the same detail level as the preceding three steps.

**Applying appropriate clinical skills.** The physician, after deciding which further information she needs, makes a plan for acquiring that information. Then she starts to apply the plan which involves taking some actions. This is where the physical skills may be needed. The choice of the way to acquire a needed piece of information depends also the setting (i.e., location), since different settings provide different opportunities. For example in a war zone, physicians have restricted resources. We include the physical actions that an
expert takes in our core domain model, while the factors such as availability of X-ray machine is included in contextual factors.

**Interpretation of new data and reformulation of the problem.** The physician reviews her understanding of the patient as new information or data becomes available. This may entail changing her hypotheses if the expectations and observations are contradictory. Not all such data triggers the generation of new hypotheses. For example, a new fact which is rather general, such as fever, would not lead to activation of new hypotheses. More specific data may be of significant value, while fever is of low value.

**Diagnostic and therapeutic decision making.** On the basis of the degree that the manifestations and the hypotheses fit together, the expert decides upon the ultimate diagnosis. This process may vary from expert to expert. Some may make final diagnoses earlier than other and by using less information.

For each diagnosis, there may be a number of therapeutic alternatives. The choice among this set depends on various factors, many being contextual ones. For example, the age of the patient or the allergies, as well as the environmental factors such as availability of required instruments for the selected therapeutic method are important for making this decision.

### 2.4 The correspondence between cognitive and physical activities

In the preceding section of this chapter we tried to allow the reader to see our perspective, needs and interpretations with respect to clinical problem solving. Based on various approaches in the literature that attempt to explain the clinician’s problem solving, we configured figure 2.2 sketching out our understanding of how a medical expert does her job. We did this without explicitly stating what the mentioned perspective, needs and interests were. This would lead to a more natural way to getting involved in the issues at the core of this work; first the reality, then its model. As any model is a simplification of a real world phenomenon, our model does not reflect the whole reality. We modelled medical reasoning from a view shaped by the motivation of investigating context use in medical diagnosis. So, essentially the parts of medical problem solving that can be related to context are our primary focus.

The medical problem solving imposes the problem solver to continuously switch between mental and physical activities. Figure 2.2 illustrates our view of the correspondence between cognitive and physical activities of the reasoner in a diagnostic domain. Cognitive activities require epistemological resources, while physical ones need physical resources. The analysis of these resources for each activity in either the ‘mental world’ or the ‘physical world’ provides us a basis for explicating and typifying various contextual elements used in medical intelligence. The figure illustrates some of the physical activities (e.g., ‘listening’, ‘observing’) while cognitive methods (e.g., abductive inference by case-based reasoning method which realizes ‘hypothesis generation’) are not shown in the figure.
Medical diagnosis is an intelligent process as well as a practical art or skill. "As an intellectual discipline it engages in inquiry, formulates hypotheses and tests hypotheses, offers explanation, confirms hypotheses, and direct practice in accordance with scientific theories" [Albert 88]. So far in this chapter, we took an informal look into the logic of medical reasoning, as well as its intuitional aspects. In the rest of the chapter we investigate these aspects of diagnosis from a more computational view where we look into a model diagnosis in a more formal way.

![Diagram of reasoning and action in diagnosis]

**FIGURE 2.2** Reasoning and action are inter-weaved in a diagnostic process. Squares illustrate the tasks and the ellipses show the actions/methods. A task is realized by a method/action. The mental actions are not shown in the figure.

### 2.5 Accounts of the differences between novice and expert diagnosticians

*Instructional research* dealing with the development of new medical curricula explored new ways of teaching medical practice via studying the differences between experts and novices. Remarkable differences have been observed between experts and novices with respect to efficiency and quality concerns.

What is prerequisite to expertise? This question has been investigated via both a process and content orientation [Higgs 95], emphasizing respectively the cognition, and the clinical knowledge. The distinction between content and process has been summarized by Barrows and Pickell as follows: "There are two components of an expert clinical problem-solving that need to be considered separately even though they cannot be separated in practice. One is content, the rich extensive knowledge base about medicine that resides in the long term memory of the expert. The other is the process, the method of knowledge manipulation the expert uses to apply that knowledge to the patient's problem. In expert performance these components are inextricably intertwined. Both are required; a well developed reasoning process appropriately bringing accurate knowledge to bear on a problem in the most effective manner." ([Barrows 91], p. xii).

The very first attempts to analyse clinical expertise took as basis the *performance* of the professionals. This was at the same time as behaviourism was the leading paradigm in
psychology. Later, with the rise of cognitive psychology, the emphasis from *behaviour* shifted to the general cognitive *processes* underlying clinical problem solving. Before 1980, the essence of medical expertise had been explained in terms of variables such as quality of problem solving processes, efficiency, and number of hypotheses considered ([Custers 96]; [Bordage 91]. The seminal work of Elstein [Elstein 78] presented the first model of medical reasoning suggesting no difference between the strategies applied by novices and experts in medical domain; both use general strategies. This was at the same time as that Newell and Simon investigated general problem solving methods. Their aim was to invent weak methods which could be applicable across many domains irrespective of the domain specific knowledge. This presupposed a distinction between disease knowledge and the reasoning process, as also is exemplified in expert system implementations [Clancey 84].

The distinction between the process and knowledge used in clinical problem solving has had an important impact on instructional research. Since the early process-oriented research claimed that both novices and experts apply general strategies, instructional research stressed the role of teaching general problem solving strategies.

The later research on clinical reasoning anticipated that novices and experts use different problem solving strategies. Experts use strategies that rely on the knowledge base in order to limit the search. According to [Patel 86] experts apply a pattern recognition strategy relying on the rich and structured content knowledge. This approach triggered a content-oriented approach to medical reasoning.

Instead of investigating general strategies, investigations have been initiated in order to find out whether the organization and availability of domain specific knowledge could be responsible for ‘good’ medical reasoning [Feltovich 84; Bordage 84; Patel 86; Barrows 87]. The idea was to show that efficient problem solving is conditioned on exclusive and well organized knowledge about the problem domain. The findings have rendered knowledge organization1 as an important underlying reason for an expert’s diagnostic superiority. The content-oriented approach originated from the idea that “A very intelligent person might be that way because of specific local features of his knowledge organizing rather than because of global qualities of his thinking”2.

More recent trends, however stress an interdependence between clinical reasoning process and clinical knowledge [Norman 90; Schmidt 90; Boshuizen 91; Patel 90; Arocha 93] which is gaining increased acceptance. In this view, the use of different processes by experts and novices was explained in key terms of ‘experience’ of the clinician, and the ‘difficulty degree’ of the problem [Patel 90; Schmidt 90; Boshuizen 91]. As the interpretations of what the differences are between experts and novices changed, the approaches to the development of expertise also has changed. The reflection of this approach on instructional research suggests a focus on the integration of content and process. The ‘problem based learning’ (PBL) curricula presuppose that learning can be facilitated if the learning happens in the real world context in which the students need to apply what they have learned. This argument agrees perfectly with Tulving’s ‘encoding specificity’ theory since

---

1. By knowledge organisation we will understand knowledge content including its structure.
2. Due to Minsky and Papert (1974), referred to in [Glaser 88]
it also stresses the importance of the correspondence in the context at learning and remembering time. We elaborate encoding specificity theory in chapter 7.

We will, following Higgs [Higgs 95], identify four main approaches to medical problem solving in the literature; the hypothetico-deductive model, the pattern recognition model, the knowledge-reasoning integration model, and the holistic model. The first two investigate the process of the clinical reasoning. The pattern recognition model recognizes the role of knowledge content, while the hypothetico-deductive model ignores it. The last two models emphasize the importance of both process and the content. They focus on the integration of the content and the process. The holistic approach, in addition, stresses the role of context in clinical reasoning. These four approaches are presented in the following sections.

2.5.1 The “hypothetico-deductive” model

The seminal work of Elstein et. al. [Elstein 78] contrasted the medical problem solving of experts and novices, in terms of the cognitive processes employed by these two groups. Elstein et. al. attempted to identify the strategies and processes that distinguished experts from less experts. They postulated that a crucial difference between experts and novices was that experts employed different processes than novices. However, Elstein et. al., contrary to their expectations, did not find significant differences between the processes used by experts and novices; both used, in their terms, a hypothetico-deductive method. Other medical researchers also portrayed hypothetico-deductive reasoning as a model of clinical reasoning [Kassirer 78; Feltovich 84b]. This approach involves generation of a limited number of hypotheses based on clinical data and information, and testing of the hypotheses through further on-going inquiry. The generation stage is called an induction<sup>1</sup> process while the testing part is called a deductive process. [Barrows 91] interprets this ‘inductive’ stage as a way of structuring or understanding the problem. It is this understanding that governs further reasoning. Hypothesis generation occurs rather early, and is a rapid and automatic process. That is, clinicians from all levels of expertise generate, early in the workup, a set of hypotheses. In the testing stage, these hypotheses are used to predict what additional findings should be present if the hypothesis is true.

The early findings of Elstein et. al. were, in fact, calling for another research agenda. This is because even though they did not detect any difference between experts and novices with respect to ‘process’, they found other differences. These are;

1) Difference in the quality of hypotheses: An important finding of Elstein was the observation of differences in the content of the generated hypotheses set. The experts were superior to generate more correct hypotheses even in the first generation round. So, there is a prominent difference between novices and experts; even though they both used a hypothetico-deductive method, they generated different sets of hypothesis for the same patient case.

---

1. Along these lines, we will use the term ‘abduction’ instead of ‘induction’ in later chapters. The reasons for preferring this term is explained in chapters 3, 4, and 5.
2) **Difference across cases:** Elstein [Elstein 78] also found that problem solving expertise varies greatly across cases and is highly dependent of a clinician's mastery in the particular area. However, these findings have not been further investigated neither by Elstein himself nor by cognitive psychologists, as cognitive psychology focused on the role of process rather than that of knowledge at that time. As we mentioned above, the dominating research subject was the general strategies and skills. As such, the role of experience had been underrated. According to Higgs, the hypothetico-deductive approach ignored the fact that knowledge of the problem solver could provide a significant source with respect to differences between experts and novices. This implies a lack of consideration whether “effective problem solving requires a large store of relevant knowledge” [Higgs 95], p. 9). Consequently, the role of experience, its influence on organization of knowledge, and its effects, in turn, on the diagnostic process had to wait for being investigated until a shift occurred in the focus of cognitive psychology.

After this shift, however, the generality of the hypothetico-deductive method has been subject to prominent critique by cognitive psychologists who suggest that clinical reasoning may sometimes be a pattern recognition process [Patel 86], rather than a purely hypothetico-deductive process. This approach is elaborated in the next section.

### 2.5.2 The “pattern recognition” model

Patel (Patel 86; Patel 90), challenging Elstein’s hypothetico-deductive model, suggested that experts and novices may not be using the same reasoning process. Experts use a “pattern recognition” model, rather than a hypothetico-deductive method.

In this view, the process used by experts is primarily an inductive process, a forward reasoning, and is essentially a kind of pattern recognition. Forward reasoning is data driven and occurs when cues available about the patient are utilized in order to evoke diagnostic hypotheses while backward reasoning is hypothesis driven. According to Patel, medical experts do not display explicit hypothesis testing in familiar situations.

The relationship between the level of expertise and the directionality of reasoning has been investigated through a series of experiments. According to the early findings suggested by Patel’s work [Patel 86], expert cardiologists who established accurate diagnoses always used pure forward reasoning, while novices uses most backward reasoning. However, later investigations [Patel 90] showed that experts also use backward reasoning but only when the available cues used for pattern recognition are not sufficient to arrive at a solution. Hence, experts fall back on backward reasoning when confronted with “loose ends”, that is, unrelated facts. The expert ability to use forward reasoning has been attributed to the possession of highly organized knowledge of the underlying cardiac disorder.

Higgs explains pattern recognition as including categorization and prototypes. “Categorization involves grouping of objects or events. It can be related to the process of recognizing the similarity between a set of signs and symptoms and a previously experienced clinical pattern or case. The new case is placed in the same category as the past case(s) and is given the same label (diagnosis)” ([Higgs 95], p.13).

The “pattern recognition” model of clinical reasoning is based on the presupposition of existence of some knowledge patterns in the memory of the clinician which could be recalled by using the cues available in the new patient situation. Thus, this account empha-
sizes the role of content in the choice of strategy to be employed in clinical problem solving.

The whole theory originates from the idea that experts have a knowledge base with a richer content than that of novices. Further, they have an adequately organised and structured memory, a characteristic which novices lack. This is what makes experts able to use a faster and more efficient inductive process. This happens when the case is a familiar, routine case. Other results from Patel's group revealed that clinicians use a combination of forward and backward reasoning when encountering unfamiliar, difficult problems with incomplete or inaccurate data and information. This supports also Chi's [Chi 87] argument that whether a child uses a particular strategy is highly dependent of her knowledge base.

2.5.3 The relationship between the hypothetico-deductive and the pattern-recognition models

Though hypothetico-deductive and pattern recognition models have been considered mutually exclusive, we see that they are, in fact, not conflicting. In this section we present our interpretation of the relationship between hypothetico-deductive and pattern recognition models.

An aspect which has been less explicated in both the hypothetico-deductive and pattern recognition models is that both require a "hypotheses generation" process. However, depending on the richness of the experience and the structure of the knowledge base, the methods to evoke hypotheses may be different in experts and novices. Experts accomplish 'hypothesis generation' by employing a pattern recognition method while novices a hypothetico-deductive method.

This interpretation conforms with the idea adopted in AI that a task can be realized by more than one method (see [Benjamins 93]). The most adequate method is selected on the basis of various factors such as the match between a reasoner's experience in that domain and the knowledge required by the method. Consequently, although both hypothetico-deductive and pattern-recognition models include the same "task" of evoking hypotheses, they use different type of processes and different type of knowledge when doing that. Patel et.al., being more concerned with showing the absence of hypothesis testing in expert reasoning, pays less attention to the fact that experts and novices also differ in hypothesis evoking strategies.

We are inclined to contend that the general process of diagnosis consists of hypothesis-generation and hypothesis-testing components. Hypothesis testing in the pattern-recognition model is embedded in the pattern recognition method which is applied exclusively by experts, as only experts own such patterns. As Elstein [95] recently maintains, experts' hypothesis testing is usually "rapid, automatic, and often nonverbal". In this model, hypothesis testing is not explicit. However, this may not always be so, and usually is not. The low level of expertise is not the only condition for the need to collect extensive data. Even the most outstanding experts may need to acquire more data when the data at hand is scarce and thus is not adequate for establishing a correct diagnosis. Patel et. al. assume a sufficient degree of match when referring to pattern recognition. In general, experts use their past experiences, but nonetheless, because of a possible poor match, they may need explicit hypothesis testing by information gathering. In such cases, pattern recognition is
followed by hypothesis testing similar to that in the hypothetico-deductive model. The match of patterns is not absolute. There may be various degrees of a match between two patterns. In the pattern recognition stage, the clinician “recognizes” whether she has sufficient knowledge and data or should gather more. It is her confidence in the hypotheses she has generated which determines what to do further.

Another important point is that the hypotheses should be evaluated regardless of what specific method is used for generating them. This is because hypothesis generation is an abductive process, and by nature generates fallible results. Irrespective of the method used, the conclusions (i.e., hypotheses) must be evaluated; either by testing, or “pattern recognizing”. However there is an intuitive difference between these two kinds of evaluation, and we analyze this difference in the chapters where we elaborate abduction. We can briefly mention here that pattern recognition may involve the reasoner’s justification of her own reasoning. If she becomes able to perfectly justify her selection of the pattern from her memory and that this pattern matches sufficiently with the current case, then she may not see the reason for further testing her diagnostic hypothesis. However, if she is not confident of her reasoning at the end of the pattern recognition step, she may try to verify the hypotheses by gathering further evidence. So, for experts, the pattern recognition process is decisive for determining whether testing a hypothesis in the same way as in hypothetico-deductive model is necessary or not.

The need for justification of one’s own reasoning behaviour is, we believe, related to “metacognition”, an important factor which has not been given necessary prominence in these two models of clinical reasoning, but becomes a central factor in another model which we will soon present.

It seems that a clear cut distinction between the hypothetico-deductive and the pattern-recognition models of medical reasoning is artificial as they are not necessarily mutually exclusive, and on the contrary, are usually practiced in combination. They can be considered rather as speciations or variations of the same process based on the generation and evaluation of hypotheses.

2.6 Accounts integrating process and content

As the process oriented approach to understanding medical reasoning attempted to explain development of expertise in terms of procedures and skills, a similar mentality attributed the deficiency in young children’s performance to the absence of mature strategies. That is, children’s inability to solve a problem is attributed to their lack of reasoning skills. However, Chi [Chi 88] provided evidence on that the lack of knowledge or inability to access the available knowledge could also be responsible for such deficiencies. Even though the young children have acquired mature strategies, they may be unable to solve the given problem because of either lack of knowledge or inadequate representation of knowledge leading to difficulties in accessing it [Chi 87]. Chi analyses and examines how the knowledge structure in children changes with age. She defines memory development as the change in performance with age, in all kinds of memory tasks, such as recalling a sequence of digits, reconstructing experienced events, etc. Chi argues that in order to be able to use a given strategy, the young children have been observed to rely on the knowledge in their semantic memory. That is, even if the cognitive strategy necessary for solv-
ing a problem is available in the child’s stage of maturation, the child may be unable to use it. This happens if the amount and structure of content knowledge required by the strategy is not already acquired.

A recent approach investigates the characteristics of medical expertise via studying the process of development of expertise ([Schmidt 90]; [Schmidt 93]; [Bosshuizen 95]; [Custers 96]). Content is the keystone in this approach. For example, Schmidt et. al. concentrated on the domain specific knowledge and medical experts’ organisations of knowledge bases. They have hypothesized that development of expertise in the medical domain leads to structural changes in a physician’s domain knowledge. Schmidt’s team examined the changes in the knowledge structure reflecting change in the level of medical expertise. Based on the findings from their experiments Schmidt and Bosshuizen provide a theory of memory development of experts, as we will see in the following section.

Chi’s argument agrees with Schmidt and colleague’s which may be considered to take this idea a big step further towards a theory of developing the expert’s memory. There is a parallel between Chi’s examining age-related differences in memory and Schmidt and colleague’s (in Limburg) analysing the development of clinicians’ memory by experience. The Limburg group argues, similar to Chi, that in order for a clinician to make a fast and efficient hypothesis activation, she must already have acquired the required knowledge structured as illness scripts and patient cases.

2.6.1 Three-stage model of developing expertise

Educational psychology is concerned with (a) arranging the curriculum in order for the students to quickly become experts in their field, and (b) enhancing performance. The medical field has been one of the faculties that most desperately seeks the answers to these questions.

The traditional learning/teaching method is based on a preclinical and a clinical period. The preclinical period is typically devoted to learning knowledge and acquisition of some skills, while the clinical period is for learning how to use those knowledge and skills. So, the acquisition of knowledge and its use are conceptualized as two distinct tasks.

Schmidt and his colleagues [Schmidt 90; 93] denied this view and proposed that development of expertise can not only be associated to acquiring more and more knowledge and skill, but that structuring knowledge is an essential part of improving expertise [Bosshuizen 94]. This research group developed a “theory of expertise development in medicine” which describes expertise as

“...the progression through a series of consecutive phases, each of which is characterized by functionally different knowledge structures underlying performance. The first phase is characterized by accumulation of causal knowledge about disease and its consequences. Through experience with real cases, this knowledge transforms into narrative structures called illness scripts. The cognitive mechanism responsible for this transition are: Encapsulation of elaborated knowledge into high level but simplified causal models or even diagnostic categories and tuning through the inclusion of contextual information. The third phase is characterized by the use of episodic memories of actual patients in the
diagnosis of new cases. It is assumed that knowledge acquired in different phases form layers in memory through a sedimentation process" [Schmidt 93].

So, according to this theory, knowledge is structured in the form of one of the following three types:

- principled causal knowledge (rather deep)
- illness scripts (less deep)
- cases (shallow, most specific)

The last two include a large amount of knowledge which is activated as a whole. Schmidt emphasizes the role of context in the construction of such structures [Schmidt 90]. Context serves as a glue that connects clinical and biomedical knowledge so as to constitute a meaningful whole, either as an illness script or an episodic case.

The term “illness script” was first used by Feltochich and Barrows [Feltovich 84b]. In Schmidt and his colleague’s view, illness scripts describe features of prototypical patients. These structures include little deep knowledge related to pathological causes of manifestations, but rather clinically relevant information about diseases. An illness script is a cognitive structure consisting of three components and links together knowledge from these three groups of knowledge/information. These are:

1. enabling conditions which comprise the information about “circumstances in patients and their environment that may lead to illness” [Boshuizen 94, p 315]. This is the information describing the “context under which illness develops” [Schmidt 90, p 611]
2. the fault which is the “actual pathological process that is taking place” [Boshuizen 94], and
3. consequences, the signs and symptoms.

This research emphasizes the changes in the knowledge structure, contrary to approaches emphasizing the development of general problem solving and thinking skills. In this view, development of expertise in medicine is associated with a “transition from a network-type organisation to another type of structure”, the illness scripts. Thus, the approach implies that development of expertise can not be analysed independently from development of a particular type of memory, one consisting of what they call “encapsulated structures”. These structures “enable experts to grasp the structure of the problem and then proceed with solutions that bypass the novice’s lengthy search” [Glaser 90, p 33].

As to accessing knowledge in such a memory, “contrary to knowledge networks, illness scripts are activated as a whole” [Boshuizen 95] instead of as individual concepts. The expert’s hypothesis activation and testing is accounted as for by “epiphenomenon illness script activation and instantiation” [Boshuizen 95, p27]. In their account, episodic memory consists of instantiated scripts. And, these “instantiated scripts, in turn, do not become decontextualized after use, but remain available in memory as episodic traces of previously diagnosed patients, and will be used in the diagnosis of future similar problems” [Schmidt 93, p.208]. They emphasize essentially that all three types of knowledge (deep causal, illness scripts, and episodic cases) do not decay and do not become inaccessible.
All are accessible at any time. However, the most specific and appropriate one is always preferred. What is most specific and appropriate depends on the problem at hand and the richness of the memory of the reasoner.

According to Schmidt and et. al. "problem solving in routine cases is a process of script search, script selection, and script verification" [Schmidt 90, p 615]. Cases are "instantiations" of illness scripts with concrete patient data and information. As experience increases, many such cases are accumulated. Notice that the word 'script' does not imply an emphasis on distinguishing between individual episodes and more general scripts. Hence, in our interpretation, this view suggests a coherence with the case-based reasoning paradigm.

So, in most familiar patient problems cases have been used, while in less typical problems, the illness scripts are utilized. As the degree of difficulty and atypicality increases, finding appropriate scripts or instances becomes difficult, and consequently more and more principled biomedical knowledge gets utilized.

We would like to draw attention to the agreement between the 'theory of developing expertise' due to Schmidt et. al., and Patel and colleagues association of expertise with application of a pattern recognition strategy. Criticizing the hypothetico-deductive account of the medical problem solving process, they also note that experts must have some patterns and match these with the new case rather than relying on principled knowledge.

Even though they call their theory "development of expertise in medicine", as mentioned in [Schmidt 93], medicine is not a unique example of domains where problems are ill-defined; "Like medicine, many problems in the real world require some kind of diagnosis based on conceptual knowledge to characterize or understand a situation and act upon it in an appropriate way".

This memory theory can explain the findings of Elstein et. al. related to the superiority of outstanding experts in activating high quality of hypotheses, as we mentioned above. It has been shown that if the correct hypothesis is found in the hypothesis pool generated initially, experts recognize it as the ultimate solution. If it is not generated initially, on the other hand, it is not uncommon that it is not generated later either. Schmidt et. al. draw a connection between the ability of generating the correct hypothesis initially and the way diagnostic knowledge is organised. In particular, they explain generating the correct hypothesis early in the workup by the expert's ability to activate appropriate illness scripts.

They relate this ability to experts utilizing enabling conditions much better than novices do. Especially when the data is scarce, the use of enabling conditions as additional information becomes significant with regard to performance. "Enabling conditions consist partly of nonmedical background knowledge.....patient contextual factors that influence the probability that someone has a specific disease (e.g., age, sex, previous medical history, current medication, occupation, hereditary and environmental influences, risk behaviour)" ([Custers 96], p 385). So, experts are better able to utilize contextual and

---

1. The terms biomedical knowledge and core domain knowledge are used interchangeably throughout the thesis.
background information when generating hypotheses; a feature which makes them superior to novices.

An illness script represents a "whole" thing. It represents a contiguity, from start to end, similar to Schank's scripts [Schank 77]. The relationship between the enabling conditions and the fault "is of a psychological and probabilistic, rather than of a medical or causal nature: For example, risky behaviour (e.g., alcoholism) may in general increase the possibility that a particular disease (e.g., pancreatitis) is present, even though in a particular case, these events may be unrelated" [Custers 96, p385]. An important aspect with respect to enabling conditions is that they are usually available in an early stage of diagnosis.

2.6.2 The "holistic" model and metacognition

A more recent trend emphasizes the role of metacognition, which is ignored in other models. Chi recognizes three components of expertise as knowledge, cognition, and metacognition [Chi 85]. She argues that although knowledge, cognition and metacognition are all important in the development of expertise, none of them is sufficient for understanding expertise. She proposes that a holistic approach integrating domain specific knowledge, strategies and metacognition is necessary to account adequately for memory development. Considering any of these in isolation will not provide an understanding of expertise.

It has been suggested also that both medical reasoning and expertise can only be understood collectively in terms of knowledge, cognition and metacognition notions [Barrows 91]. Among these metacognition is probably the least investigated so far.

Higgs and Jones particularly emphasizes the prominence of metacognition in clinical reasoning. For them, "the ability to reason knowingly, and to justify articulately our decisions and interventions is essential for effective clinical practice and for the development of the knowledge bases of our professions" ([Higgs 95], p.3). They base their understanding of clinical reasoning on the complex interactions of many factors. Many of these factors are encapsulated in the term "context". They point out that clinical reasoning occurs in a specific clinical context. In their account, clinical reasoning occurs in several contexts:

"the immediate personal context of the individual patient/client, the unique multifaceted context of the client's clinical problem within the actual clinical setting in question, the personal and professional framework of the clinician, the broad context of health care delivery and the complex context of professional decision making. In order to understand and address the reasoning behind clinical decisions the various contextual factors that influence reasoning need to be appreciated ([Higgs 95], p 4).

Metacognition "refers to being aware of one's own cognitive process and exerting control over these processes" ([Higgs 95b], p 141). Metacognitive skill is required in order to manage knowledge and other cognitive skills.

According to Biggs "High quality human performance inevitably requires metacognitive as well as cognitive components. To perform well, one needs to be aware not only of the knowledge and algorithms required for the task, but of one's own motives and resources, the contextual constraints, and to plan strategically on that knowledge" ([Biggs 86], p 143, referred to in [Higgs 95b]).
Metacognition is a high level cognitive skill which is important for dealing with uncertainties, cognitive limitations, and ambiguities in clinical reasoning. It represents a self-monitoring ability of the clinician, which is needed to control and evaluate the knowledge and strategies (i.e., cognition) involved in clinical reasoning. It “provides an interface between general problem solving skills and domain specific knowledge” [Higgs 95b]. The following processes involved in metacognition are recognized: “realizing that important problem solving (task) information is missing or ambiguous, recognizing that the problem will be difficult..., being aware that reasoning errors have been committed, evaluating the effectiveness of reasoning strategies, modifying reasoning strategies and allocating cognitive resources” [Higgs 95b p 18].

In AI, the research community which concerns the reusability of elements of a system seems to deal with an equivalent of the notion metacognition. As we will see in details later, the community attempts to explicitly represent all elements of reasoning at the knowledge level [Newell 82]. To mention briefly, the three “components of expertise”, according to [Steels 90] are domain, task, and method, which, in our interpretation, respectively correspond to knowledge, metacognition and cognition. Figure 2.3 shows this correspondence. Task knowledge links domain knowledge and cognitive methods to each other. It also mediates ‘what is my goal now?’, and ‘what should I do next?’ type of questions.

\[\text{Cognitive psychology} \quad \text{Artificial Intelligence}\]

\[\text{cognition} \quad \text{knowledge} \quad \text{method} \quad \text{domain}\]

\[\text{metacognition} \quad \text{task}\]

**FIGURE 2.3** The links imply the correspondence between the components of medical reasoning model and AI model of expertise.

Our knowledge level model of context-sensitive diagnosis, elaborated in later chapters, suggests a combination of pattern recognition and hypothetico-deductive approaches, where the knowledge content of the diagnostic domain determines the details of this combination. So, the knowledge pertinent to ‘diagnosing’ is a key issue for the model of which contextual knowledge constitutes an important portion.

### 2.7 Teaching Medical expertise

“The objective of medical schools is to turn relative novices into knowledgeable and skilled professionals” [Boshuizen 95, p24]. The question is how to do this. Educational psychology investigates this subjects: what to teach and how to teach.

The first step is to analyse the type of knowledge necessary to learn for developing expertise. Knowledge to be taught has been classified in two main groups: basic science or biomedical knowledge, also called textbook knowledge, and clinical or heuristic knowledge.
Basic science knowledge\(^1\) in medicine includes anatomy, biochemistry, physiology, microbiology, histology and pathology. Heated discussions are centered around questions of which type is more important for problem solving performance, and which sequence these should be taught, or should they be taught simultaneously.

### 2.7.3 Distinction between “textbook” knowledge and “heuristic” knowledge

Research in medical education states that in training of medical students the primary emphasis has been on complaint exploration and physical examination [Hobus 87]. Consequently, the anatomical and pathological type of knowledge has been given the most importance, since this sort of knowledge is used for understanding the relations between the symptoms and signs, and diseases. An example of a typical textbook subject is “Myocardial ischemia means reduced blood supply to myocardium which in turn reduces myocardial compliance. This implies diastolic dysfunction (decreased ability to expand) which causes the ventricular diastolic pressure to rise, which leads to abnormal S-T segment changes. The patient may experience angina pectoris. The rise in the diastolic filling pressure may cause some patient’s experiencing acute dyspnea. The reduction in the ability of left ventricle to contract causes reduced cardiac output which cause some patients to feel fatigue...”. This example is on the pathophysiology of myocardial ischemia.

It has also been stated, however, that other source of knowledge are needed in critical points of a diagnostic process. An example is how a patient’s age, sex and race may favour certain disease. When the muscular system is regarded, juvenile rheumatoid arthritis or rheumatic fever are more common in children, while reiter’s syndrome or systemic lupus in young-adults, fibrosis in middle age, and osteoarthritis in old age. Regarding race, SLE and sarcoidosis are more usual for negroes while polymyalgia rheumatica is more often observed in caucasiains [Greenberger 93; Wright 79].

In one meeting, our medical expert\(^2\) was interpreting the signs and symptoms of a patient with cardiological problems. The patient was a black male. The doctor was suspecting an infectious disease. At one point he said “...he is black. He may have tuberculosis...”. When we asked him later why he associated the patient’s being black and the possibility that he had tuberculosis, he looked at us for some time, because he did not remember having said that. Then he remembered, but needed some time in order to find out how he may have thought. He said “black people usually have lower economical... and therefore malnutrition...tuberculosis...”

### 2.7.4 The roles of biomedical and clinical knowledge

We can briefly present the question as such: whether the roles of biomedical and clinical knowledge are rather distinct and these can be considered as two different worlds [Patel 91], or whether biomedical knowledge is fully integrated in clinical knowledge. Patel et. al. advocates the first alternative suggesting that basic science plays its role when generating a coherent explanation of the patient problem connecting various components of the

---

1. We call ‘basic science knowledge’ as ‘core domain knowledge’ in later chapters.
2. Md Ole Rosvol, cardiology, RIT, Trondheim
clinical problem. Its role is not facilitating clinical reasoning itself. The “explanation” mentioned here must be the explanation offered by the physician to third parts (the experimenters, for example, investigating the role of biomedical knowledge), after establishing the diagnosis. The proponents of the second account, Boshuizen and Schmidt and Talmon, argues that both biomedical and clinical knowledge play a role in clinical reasoning itself. However, their roles are different and occur in different phases of clinical reasoning. Biomedical knowledge” plays its role in a tacit way, leaving no trace in the think-aloud protocols” [Boshuizen 91]. In their account, the expert resorts to biomedical knowledge, in atypical cases, and when the illness script partially fits the current case. “Biomedical knowledge was supposed to be used to explain why atypicalities occur in a specific case” [Boshuizen 91]. According to this account, clinical knowledge is used to interpret and order/structure available information, while “biomedical knowledge seemed to be applied for a justification or explanation after the interpretation had been made” [Boshuizen 91]. So, the information is first interpreted, and then the interpretation is justified. This is a rather important proposition for us, as the “justification” here corresponds to what we call “justification of one’s reasoning” in connection to abductive reasoning. This is a natural process, especially when encountering a problem where information and data at hand is not sufficient for reasoning with certainty.

So, in the “Limburg account”, clinical knowledge is used for generating useful hypotheses while biomedical information is used to elaborate on the information that does not seem to agree with these hypotheses. This account recognizes the need for links between clinical and relevant parts of biomedical knowledge.

Findings from experiments conducted by both approaches show that

1) Experts use less basic science knowledge than novices or interns.

2) Atypical cases require more basic science knowledge than typical ones.

The reason why interns or novices use more basic science knowledge is that problems are more difficult or atypical for them [Barrow 91; Boshuizen 91].

2.7.5 Problem Based Curriculum

The problem based learning paradigm is a result of the search for a new curriculum. This search may be seen as a search for a new learning context which may be closer to the context of using the learned knowledge. Thus, the aim of changing the learning context was to reorganize the structure of the knowledge [Neame 84]. In a problem based curriculum, the students are given a clinical problem to solve either in a simulated or a real situation. Learning happens around the problem, and thus is problem-oriented.

It is proposed that problem based learning can be characterized by the following aspects [Schmidt 93b]:

1. Activation of prior knowledge- the initial analysis of a problem stimulates the retrieval of knowledge acquired earlier.

2. Elaboration of prior knowledge through small-group discussion, both before or after new knowledge has been acquired; active processing of new information.
3. Restructuring of knowledge in order to fit the problem presented. Construction of an appropriate semantic network.

4. Learning in context. The problem serves as a scaffold for storing cues that may support retrieval of relevant knowledge when needed for similar problems.

Adoption of a problem based approach to learning seems to be a “contextual reform of medical education. Context is used to describe the style of the teaching, the structure of materials and the framework within which the content is presented” [Neame 84]. This approach is based on the idea that the context in which the material is learned influences the development of knowledge and understanding, and subsequently the way in which the learner can use this knowledge. Barrows states that problem based learning is adequate for acquiring retrievable and reusable knowledge. “Retrieval and use of information in the task context, in medicine the clinical context, requires that the information be learned in clinical context, so that cues that appear while working in the task situation will stimulate retrieval of the appropriate information by memory association” [Barrows 84, p19].

It has since long been recognized in cognitive psychology that successful retrieval happens when the cues relevant to its retrieval are encoded with the information to be stored [Tulving 73]. Cognitive psychology has also found the “depth of processing” information prominent for learning and subsequently remembering a phenomenon [Craik 72; Craik 75]. In the medical domain, Neame emphasized that elaboration in learning time explicates which information can be a useful cue for future retrieval of an experienced patient situation.

In order to promote understanding, it is necessary for the learning to ‘elaborate’ on the knowledge, to turn it over in his head and establish its relationship to other information both newly acquired and already existing in his memory. Thus the inference appropriate to medical education seems to be that learning should be clinically oriented, preferably based on patients. It is also important that the information to be learned is structured when learning time if it is to be remembered at the same setting [Neame 84, p34].

According to Barrows and Tamblyn [Barrows 80], the role of group discussions is that the students explicate which information is relevant or important and in which way. This leads to encoding such information as retrieval cues for future use. The underlying reason is that through discussions, students explain which cues and why they are important, by linking them to other meaningful concepts. A support to this hypothesis is due to Chi [Chi 89], who investigated the role of self-generated explanations that the students generate while studying worked-out examples of mechanics problems. She suggest that the “good learners” are the ones who generate thorough self-explanation of the to-be-learned subjects while “poor learners” generate insufficient explanations. The role of explanations has been emphasized also in AI. In chapter 10 we elaborate how the role of explanations are emphasized in the framework of the Creek system [Aamodt 94].
2.8 Summary

The lessons we have learned from the literature on medical problem solving and instructional research literature help us to understand the requirements for a knowledge base system that supports human medical diagnosticians.

Domain specific knowledge is a characteristic of a quick and high quality diagnosis. Domain specific knowledge includes both biomedical and contextual knowledge, packaged together. Accessibility of knowledge is an important requirement. It is not enough that, as Chi notes, a person has knowledge. The knowledge must also be accessible in order to be used. Experts have this capability; although they have an extensive body of knowledge compared to novices, their knowledge is more accessible. "Whenever knowledge is relevant, experts appear to access it efficiently." [Ericsson 91]. This becomes possible, according to Schmidt et. al.'s theory, since the interrelated parts of knowledge are encapsulated in scripts and instances. In these 'packages' lower-level concepts such as symptoms and patient's personal information (i.e., context) are more often referred to in favour of causal, underlying, biomedical principles [Schmidt 93]. This makes it easy to access by using information at hand, which also is usually in terms of lower-level concepts.

The three-staged memory development theory can be considered, at the same time, a learning model. According to this model, learning happens through transformation from one structure to others, starting from network-like structure and ending with illness-scripts and patient instances. This is a continuous, life-long process.

The network-like structure of knowledge embodies explanatory knowledge which is resorted to when atypical cases are encountered by experts, and more often by less experts. So, both script-type knowledge and network-like knowledge are need to be captured in a knowledge based system. In this thesis a hybrid explanation-based case-based ([Kolodner 93]; [Aamodt 94b]) demonstrator is designed that exemplifies both biomedical and contextual knowledge.

We agree with Higgs Jones who "regards knowledge as a construction of the human mind seeking to make sense of the world, rather than something that is discovered" [Higgs 95, p 10]. This approach is also reflected in AI, by the case-based reasoning paradigm. Aamodt [Aamodt 91] calls this "sustained learning" which implies the continuity of constructing new memories. This will constitute an important foundation for our study of context knowledge and our managing of context knowledge in an AI perspective.
Chapter 3

Abductive inference - an evidential approach

3.1 Abduction as plausible reasoning

Abductive inference is typically relied upon in imperfect domains, in the face of incomplete or inconsistent information as well as in cases where the domain does not provide a strong theory. As such, it is frequently used both in everyday and in expert-level reasoning and is consistently stated to be practiced by scientists, detectives, and physicians. It is ubiquitous in daily life, although we may not be aware of employing it. For example, upon seeing your neighbour walk into his house wet, you may immediately think he has been swimming. However, swimming is not the only possible reason for one being wet. That is, your conclusion involving swimming is not guaranteed to be true. There could certainly be other reasons such as that he walked home in the rain without an umbrella, somebody poured water on him from the window of a building when he was passing by, or children sprayed him with a hose. One of these possible explanations will be the right one this time. Such reasonings that make up everyday life seems of a common-sense type. People reason quickly, even when there exist several alternative conclusions that one may have adopted. They do not perhaps even need enumerating and considering all possibilities in order to come up with one. You may, for example, immediately adopt the “he was swimming” explanation if there is a swimming pool in the backyard, and he was wearing a swimming suit.

3.2 History of Abduction

Similarities and discrepancies between different forms of inference have long been investigated in philosophy. The word abduction was first used by Peirce [Peirce 58] in order to define a distinct inference type which can be traced back to Aristoteles.

Even though the roots of abduction trace so far back, its definition and the scope is still controversial and usually vague. Peirce himself, who used the term “abduction” as a distinct type of inference for the first time, gave expressions belonging to two different views, which will later (section 3.3.2) be referred to as the ‘evidential’ and ‘methodological’ views. Thus, his writings present two different accounts of abductive inference. However, he did not make clear the distinction between his two motivations underlying each account. Neither did he make a connection between his two approaches. His two accounts, therefore have commonly been interpreted as conflicting alternatives, labelled as Peirce’s ‘early’ and ‘late’ accounts.
Studying the logic of argument and the logic of scientific inquiry are two important tasks that philosophers for a long time have been involved in [Burks 77]. The logic of argument "treats rules for deriving conclusions from premises and relating evidence to hypotheses and theories, judging these rules to be valid or fallacious, correct or incorrect, reliable or unreliable" [Burks 77, p 16]. The logic of scientific inquiry on the other hand, studies how these rules are useful within a larger scope (i.e., their function within an overall process). This approach evaluates these rules in terms of applicability, simplicity, and utility in the scientific inquiry process [Feibleman 72]. We have noticed that abductive inference has been referred to in discussions related both to the 'logic of argument' and to the 'logic of scientific inquiry'.

In AI, abductive inference was taken up in 1973 by Pople, in the context of diagnosis, and has gained increasing interest since then. The first examples of the computation of abduction were highly logic oriented [McDermot 88]. Later other approaches emerged. For example, the mental models of Johnson-Laird and Minsky's schema-based approaches influenced the way researchers modelled abduction. Our approach is analogy based [Keane 94], and in this respect is similar to Holyoak and Thagard's view [Holyoak 95].

3.3 The link between philosophical and computational accounts of abductive inference

The philosophy of mind is mainly interested in understanding how the mind works. In particular, the flexibility and quick reasoning of humans in everyday events have attracted many philosophers. Abductive inference was assumed to be the only type of inference which could explain some kinds of everyday problem solving. Philosophers who studied the method of scientific inquiry states that science also involves abductive inference.

Computer science and, in particular, artificial intelligence, has recently been concerned with abductive inference, especially in connection with expert-level reasoning. Our interest in abductive inference is mainly from a computational point of view. However, we realize that in order to achieve our goal related to abductive inference, we need to investigate the philosophical accounts of this type of inference, since philosophy elaborated both the nature of abductive inference and its use within a larger context, under the head of the 'logic of argument' and the 'logic of inquiry' respectively.

3.3.1 The computational point of view

As we expressed in chapter 1, our interest in abductive inference is mainly motivated by investigating the contextual knowledge (which relates to the 'logic of argument' dimension) and its use (relates to 'logic of inquiry' dimension) in diagnostic reasoning.

Low quality of conclusions is a well-known problem in computational models of abductive reasoning. In this research we attempt to lay the groundwork for improving the computational models of abduction by augmenting them with the notion of context.

For the purpose of developing a computational account, we study abductive inference at the knowledge level. Alan Newell distinguished between two levels of modelling in knowledge-based systems: knowledge level and symbol level [Newell 82]. A knowledge
level model is a high level description of the behaviour of a system, and the types of knowledge needed for such a behaviour.

At the knowledge level, a knowledge-based system is generally described in terms of three types of concepts, i.e., domain knowledge, tasks and methods. An analysis should be made concerning what to model and how to model regarding abductive inference at the knowledge level.

Intelligent behaviour can be explained in terms of two main components, the content and the process. Any knowledge-based system which can exhibit intelligent behaviour needs explicitly to represent these two components as well. The content includes the static aspects while the process reflects the dynamic aspects of a particular type of problem solving intended to be modelled at the knowledge level. The necessity of the link between content (the domain knowledge) and process (the method, the strategy) has been emphasized both in AI and in various fields of cognitive psychology. For example, educational psychologists studying medical curricula and the development of expertise emphasize that the

knowledge content should be learned in connection with the procedure that uses it. The link between the content and the process has been given prominence in our model at the knowledge level. We realise this link on the basis of a task-centered approach, where a task constitutes a media where the content and the process are coupled.

Modelling the diagnostic task includes modelling all its subtasks. Tasks are described at the knowledge level in terms of content-related concepts (such as input and output of the task), and the process-related concept such as method and strategy. In our view, abductive inference is a low level task underlying diagnostic reasoning. Figure 3.1 illustrates how a task connects the content to the process in which it is used for the accomplishment of the task.

We are concerned with the issues of modelling and implementing abductive inference in relation to diagnostic knowledge-based systems. The interest of the computational approach in the notion of inference lies in that it is a building block of an intelligent artifact that is ultimately to be developed. After a survey of philosophical accounts of abductive inference, we found interesting connections between a possible computational
account of abductive inference and its philosophical accounts provided in the literature. Even though the motivations behind philosophy and AI are different, we were able to utilize a philosophical account of abduction in developing a computational model of abductive inference. In chapter 5 we try to point to these possible links between various philosophical accounts of abduction and our computational account. We describe how philosophical approaches have provided us tools when developing a computational account involving contextual aspects. We propose that one of the two main approaches to abductive inference in philosophy, which recently has been referred to as the evidential approach, provides us with abductive patterns which we utilize when analysing content-related aspects of abductive inference. On the other hand, the other main approach, namely the methodological approach, provides a basis for investigating the process-related aspects of abductive inference. The evidential and the methodological approaches are elaborated, respectively, in this and the next section.

3.3.2 The philosophical point of view

We reviewed various philosophical accounts of abduction in pursuit of its definition/description, scope, and function. When doing this, we made an attempt to impose a structure on the tangled accounts by extracting the motivations around which these accounts can be clustered. We will partly explain the diversity of accounts in the literature by emphasizing the existence of two rather different main motivations which remained mostly implicit until recently. These motivations conveyed/expressed themselves respectively in what Fann [Fann 70] calls the evidential and the methodological approaches. These motivations also provide a criterion to evaluate the existing accounts, reflect the aspects of abduction that may mean something with regard to the question of what fails in these accounts, and guide the development of a better account.

Of these, the former view studies inference in isolation, while the latter studies it in the context of an overall application task (i.e., complex real world problems). In the evidential view, abductive inference is a justification pattern [Flack 96]. The methodological view, on the other hand, investigates abductive inference as information processing and deals with its use in complex tasks such as problem solving (e.g., diagnosis) or question answering. As such, the methodological view is concerned with the specific role and use of any particular inference type in a particular complex task, in combination with other inference types that are utilized in the same task. In a sense, the evidential approach adopts a static view, and the methodological approach a more dynamic one.

3.4 Inference as an evidential process

In philosophy, an inference is defined in terms of premises and a conclusion. In our interpretation this corresponds to input and output of a task description at the knowledge level. The philosopher’s intention to classify inference types has inspired an extensive abduction study. The underlying reason for classifying inferences was the desire for understanding the nature of various inference types. Inferences are classified first according to the different aspects of the conclusions they draw. That is, the motivation was to classify inferences with respect to their conclusions. However, researchers adopting this goal focused on dif-
ferent characters of the conclusion, and accordingly classified inference types along dif-
ferent criteria, such as ‘validity’ or ‘productiveness’ of the conclusions.

In this section we review inference classifications that have been constructed by an evi-
dential view. The selection of the different accounts to be included in the review is made
in order to structure- because such a structure is not obvious- the variety of inference
accounts as found in the literature. We organise and present different approaches accord-
ing to the criteria on which they -as far as we can see- founded their distinction. We lay
special stress on codifying the criteria that governed each classification. So, the following
subsections are arranged according to the classification criteria.

3.4.1 Classifying inferences according to validity of conclusions they
draw

A ‘traditional’ classification of inferences (i.e., before Peirce) is concerned with the ‘cor-
rectness’ of the conclusions drawn by different types of inference. This approach classifies
inferences based on their ability to draw ‘secure’ conclusions. Hence, inferences are partiti-
ioned into sound and fallible types. In this account, only deductive inferences were sound
and inductions were fallible. Abduction was not yet recognized and therefore did not
appear in this classification. So, the main - and perhaps the only- concern was a conclu-
sion’s degree of certainty.

3.4.2 Classifying inferences according to productiveness of conclusions

In his early work, Peirce also classified inferences. He emphasized, however, the conclu-
sion’s “value in productiveness” [Peirce 58, 8.384]. The subject matter was the question of
whether an inference creates new ideas. Peirce classified induction and abduction as
“ampliative” inferences, which create new ideas, and the deduction as “explicative”,
which does not. The difference between the two ampliative inference types was that
abduction hypothesizes unobservable entities from observed ones induction creates gen-
eralizations on the basis of some examples. The three distinct forms of reasoning are exem-
plified as follows (the example can be considered a classic and traces back to Aristoteles):

Deduction:
All beans from this bag are white. (rule)
These beans are from this bag. (case)
Therefore, these beans are white. (result)  (I)

Deduction is the inference of result from the rule and the case.

Induction:
These beans are white. (result)
These beans are from this bag. (case)
Therefore, all the beans from this bag are white. (rule)  (II)
Induction is the inference of rule from case and the result. In Peirce’s statements, “Induction is where we generalize from a number of cases of which something is true, and infer that the same thing is true of the whole class. Or, where we find a certain thing to be true of a certain proportion of cases and infer that it is true of the same portion of the whole class” [Peirce 58, 2.624]

Hypothesis (abduction):
All the beans from this bag are white. (rule)
These beans are white. (result)
Therefore, these beans are from this bag. (case)  (III)

Abduction is the inference of case from rule and the result. Peirce labelled this inference also ‘hypothesis’ or ‘retroduction’. Again in Peirce’s words, “Hypothesis is where we find some surprising fact which would be explained by supposing that it was a case of a certain general rule, and thereupon adopt that supposition. This sort of inference is called ‘making a hypothesis’” [Peirce, 2.623].

Peirce’s classification of inference types according to this view is illustrated in figure 3.2. In this view hypothesis and induction are both ampliative inferences, nevertheless each has its own logical form. We will see in the next chapter how Peirce modifies his abduction account when he adopts a methodological view. In AI terminology, we may say that his classification can be related to the notion of ‘knowledge refinement and extension’ or ‘development of a domain theory’ which is an important characteristic of problem solving with an incomplete domain theory. Denecker et.al, in their paper on the difference between abduction and induction, argue that the former “derives concrete hypotheses representing scenario knowledge while the second derives general theories representing general domain knowledge” [Denecker 96]. They grounded their argument on that human experts organise their knowledge in a hierarchy, where the most abstracts are at the top and the most concrete ones are at the bottom level in the knowledge hierarchy. They propose that abduction generates hypotheses taking place at the bottom level while induction generates hypotheses referring to knowledge at the upper levels in the knowledge hierarchy.

![Diagram of Peirce's classification of inferences]

**FIGURE 3.2 Peirce’s first classification of inferences**

### 3.4.3 Classifying inferences according to the explaining-ability of their conclusions

In the ‘abduction community’ there is an ongoing discussion about what an abduced conclusion explains. Some researchers consider inductive generalizations as abductive conclusions. The basis for this is the argument that such generalizations explain the evidence
on which they are based. So, we may consider the notions of ‘explanation’ and ‘explaining’ as important in this respect. Consequently, we may consider the conclusion’s ability to explain the premises as another dimension along which inferences are categorized. However, there are various approaches as to how these notions guide a classification of inference. According to Harman [Harman 65] and Josephson [Josephson 94; Josephson 96], a conclusion’s ability to explain the premises is vital when deciding which inferences have an abductive character. Harman identified ‘enumerative induction’, and Josephson the ‘inductive generalization’ as a specific type of abductive inference. However, they disagree on what ‘explaining’ means regarding the relationship between the premises and the conclusion. As this is an issue that must be elaborated for better understanding the nature of abductive inference, we will take it up in section 3.4.4.

3.4.3.1 Harman’s view

Harman states that the ‘inference to the best explanation’ corresponds approximately to what Peirce has called abduction. In his work on “inference to the best explanation” [Harman 65] he analyses, for example, the common characteristics of how a detective and a scientist reason. He states “when a detective puts evidences together and decides that it must have been the butler, he is reasoning that no other explanation which accounts for all the facts is plausible enough or simple enough to be accepted”. Or, “when a scientist infers the existence of atoms and subatomic particles, he is inferring the truth of an explanation for various data which he wishes to account for”. He proposes both detectives’ and scientists’ inferences as instances of the inference to the best explanation. Another type of inference, which he also considers a special case of inference to the best explanation, is enumerative induction. Enumerative induction “inferences from observed regularity to universal regularity or at least to regularity in the next instance” [Harman 68]. His reason for accounting enumerative induction as a special kind of inference to the best explanation is that it infers “from the fact that a certain hypothesis would explain the evidence, to the truth of that hypothesis” [Harman 68]. His classification is illustrated in figure 3.3.

![Figure 3.3 Harman's classification of inference types. The criterion is hypothesis “explaining-ability” of the evidences.](image)

In Harman’s approach we may also find the traces of a methodological approach, since it investigates abductive inference as an explanatory inference which is invoked when an explanation of surprising observations is required, for example, in cases of detective and scientific reasoning. As we will see in the next chapter, Peirce also adopted a methodological view when he intended to study the method of scientific inquiry. The notion of “infer-
ence to the best inquiry" is important from a particular aspect: it implicitly refers to the scope of abductive inference. It implies that abductive inference covers the whole explanation process. As will be seen in the next chapter, Peirce's methodological account differs from Harman's in this respect.

3.4.3.2 Josephson's view
A working definition of Josephson's abductive inference is "inference to the best explanation". In our interpretation this is a rather general and vague definition of abductive inference which includes the whole process of explaining the surprising observations. As such, its scope is very broad and includes different types/forms of reasoning as well as actions in the physical world. In fact, in Josephson's approach the two main forms of inference are abduction and prediction. Prediction comprises the process of inferring the expectations from a hypothesis. Deduction is viewed as a type of prediction task, and induction as an abductive type (see figure 3.4). On the other hand, prediction can be evoked as a subtask of abduction, and prediction can evoke abduction as a subtask. For example, hypothesis verification is a subtask of abduction and is performed through prediction. Prediction uses abduction for assessing the situation. Josephson's version of abduction has the following pattern:

D is a collection of data (facts, observations, givens)
H explains D (would, if true, explain D)
No other hypothesis can explain D as well as H does
Therefore, H is probably true

(IV)

FIGURE 3.4 Josephson's classification of inference types. Prediction is a distinct type of inference which works from hypothesis to data (from [Josephson 92].

As can be seen from his justification pattern, Josephson augments the pattern of abductive inference by adding the 'assumption' that 'no other hypothesis can explain D as well as H does'. When compared with Peirce's abduction pattern, from a knowledge level view, this pattern is superior to Peirce's as it is better specified. According to this pattern, justification of the conclusion is based on two conditions; the ability of conclusion (H) to explain the premises (D), and that the H explains D better than other possible inferable conclusions do. In a later chapter, we will, however, illustrate why this pattern does not adequately describe abductive inference.
3.4.4 What does an abductive conclusion explain?

What makes Harman think that enumerative induction, and Josephson that inductive generalization are specific types of abductive inference? Similarly, what is common to the tasks of story understanding and diagnosis, with respect to being involved in abductive inference? The notion of “explanation” seems to be at the heart of an answer that may be satisfactory. However, there are disagreements as to what explains what in abductive inference. Harman describes ‘inference to the best explanation’ as the inference ‘from the fact that a certain hypothesis would explain the evidence, to the truth of that hypothesis’ [Harman 65]. So, in his view, for an inference to be abductive, the hypothesis should explain the evidence. This view has been criticized by Ennis [Ennis 68], and recently by Josephson. The discussion emerged partly because different meanings were attributed to the notion of explanation, by Harman and the others. Also the explicit references are not made with regard to ‘what is it that explains’ and, more importantly, ‘what is being explained’.

3.4.4.3 What does “inference to the best explanation” explain?

Harman writes “I claim that, in cases where it appears that a warranted inference is an instance of enumerative induction, the inference should be described as a special case of another sort of inference, which I shall call ‘the inference to the best explanation’” [Harman 65]. In his account, enumerative induction “infers from observed regularity to universal regularity”, and has the form: From the fact that all observed A’s are B’s we may infer that all A’s are B’s. Harman describes ‘inference to the best explanation’ as the inference “from the fact that a certain hypothesis would explain the evidence, to the truth of that hypothesis”. He notes that this rule makes the reasoner able to select the best explanation among several possible ones, thus abductive inference provides a way to determine when a hypothesis is better than another, or when it is best.

Though Harman argued first that inductive generalizations should explain the evidence, he later accepted that it was a mistake to say that “induction always infers an explanation of one’s evidence” [Harman 68] and modified his view on an answer to Ennis’ critics, where he attributed his mistake to the ambiguity of ‘explanation’.

Ennis agrees that warrantability of enumerative induction depends on, as Harman notes, a number of factors other than observed regularity, but strongly argues that “the enumeratively induced proposition should be the best explanation” of the evidence. Ennis noted that ‘All A’s are B’s’ together with ‘S is an A’ does explain why the belief that S is a B is justified, but does not always account for the fact that S is a B. Sometimes it does and sometimes it does not, depending on its content [Ennis 68]. So, he argues that enumerative inductions do not necessarily explain the evidence. They explain why the reasoner believes that the hypothetical generalizations are true.

Josephson agrees with Ennis in this objection. In his view, the relationship between a sample and a generalization, and that between an effect and a cause is not the same. Nonetheless, he claims that there is an explanatory relationship between an inductive generalization and the samples on which it is based. The difference is that the relationship between a sample and a generalization involves a frequency consideration, while the relationship between an effect and a cause is a causative one.
Josephson notes that

"All A's are B's" cannot explain why "This A is a B" because it does not say anything at all about how its being an A is connected with its being a B. The information that "they all are" does not tell me anything about why this one is... A generalization helps to explain the events of observing its instances but it does not explain the instances themselves.

The word 'explanation' is actually rather ambiguous, as Harman says.

3.4.5 Ambiguity of the term 'explanation'

Clarification of the meaning of 'explanation' is important from a computational view of abductive inference. There may be various types of explanations of a surprising fact, such as structural or causal ones. However, in general only one type of explanation is relevant and useful for a certain task. At any time, only a particular meaning of a surprising fact is sought. This aspect is emphasized by Leake [Leake 95]. He illustrates that the relevance of an explanation is dependent on several factors, the needs of the explainer being one of the most important ones.

It seems that various interpretations of the meaning and the scope of the notion 'explanation' or 'explaining' has caused disagreements regarding the relationship between abduction and induction. The difficult question was what an inductive generalization explains, i.e., how this 'explaining' is different from explaining a set of effects by finding their cause.

So, despite its ubiquitous usage both in everyday life and in science, the word 'explanation' is by no means precisely defined. Achinstein's book [Achinstein 83] on 'the nature of explanation' contributes to a better understanding of this concept. We distinguish between its two usage, namely as the act of explaining, and the product of explaining. The act of explaining involves utilizing various relations connecting explanata and explanandum, via other concepts. Explanandum represents the phenomenon to be explained, and explanata is the phenomenon that embodies this explanation.

3.4.5.4 The 'relationship'

We may consider an explanation (the chain) as a relation between the conclusion and the premises, for example the hypothesis and the observations. In abductive inference, the reasoner formulates a hypothesis so that the observations are connected to the hypothesis by certain types of relations. The relations may be structural, causal, or functional if the observations are facts. If the observations are samples and the hypothesis is a generalization, then the relation is somewhat different. This difference we intend to illustrate by relying on the notion of 'meaning'.

Meaning, intension and extension. Explaining something is giving it a meaning. The 'meaning' of a concept is captured by its consequences, according to Peirce's view of 'pragmatism'.

The term 'intention' derives from Frege's term 'sense'. His usual example is that the two concepts 'morning star' and 'evening star' do not have the same meaning, even though they denote the same object [Maida 85]. The difference between intension and extension
is referred to as the one between sense and reference. Some words with the same extensional representation may be represented by more than one node in a conceptual graph, each of which represents different intensional concepts [Woods 75].

Lyons [Lyons 83] states that the extension of a term is the class of entities that the term defines, while properties defining the class constitute its intension. For example, the extension of ‘tree’ is the set of all trees, while its intension is what a human conceives when she classifies something as tree.

We may utilize the notions ‘intension’ and ‘extension’ in order to better understand what lies under the intuitive notion of ‘explaining’. Clarification of ‘explanation’ is important for the discussion of whether inductive generalisations are specific cases of abduction, i.e., explanatory reasoning. An explanation can be either intensional or extensional. An intensional explanation is made in terms of relations such as ‘causes’, ‘has-colour’, ‘pre-disposes’, etc. On the other hand, an extensional explanation is expressed in terms of relations such as ‘has-instance’.

### 3.4.6 The relationship between abduction and induction

Consider the question “what does ‘apple’ mean to you?” An answer may be ‘a red, round, hard fruit’. There may also exist other such fruits. So, you need to add more discriminating features in your description, like that it grows on trees, tastes such and such, and so on. Children may have a habit of using apples when they play war. For them, ‘apple is a hard fruit used when fighting’. Another answer may be to show some apples and say ‘these are apples’. So, depending on the interests of the questioner and the situation, ‘apple’ may mean different things, and its meaning can be captured in different ways. It is well-known that meaning is context dependent. That meaning is context dependent, and therefore there may be several meanings of something, is important for our argumentation and understanding of the relationship between abduction and induction. The following are important points for understanding this relationship:

1) The meaning of a thing may vary across situations.

2) From the components of meaning of something, one can ‘guess’ what the ‘something’ is, in a certain task context. This process calls for abduction.

3) The doctrine of pragmatism uses the notion of meaning for predicting the consequences of a hypothesis. The consequences of a hypothesis collectively comprise the meaning of the hypothesis. Usually only some of the consequences, i.e., parts of the meaning, are known at hypothesis generation time.

4) The meaning of something in a certain task context consists of a number of other things which have a special relation, or a set of special relations to that ‘something’

5) The implication of this is that if the reasoner is given parts of the meaning of a hypothesis, and the types of relations between these parts and the hypothesis, the hypothesis itself can be guessed.

The relation between the meaning of a hypothesis and the hypothesis itself is different in a cause-to-effect type of inference than in an inductive generalization [Öztürk 97b]. We think the question of whether these are of the same type of inference is not that important when the subject is put in this way. What is more important is what the implications of this difference are for the research in knowledge based systems.
3.5 Summary

In the evidential view, inferences are considered as justification patterns which are represented by premises and conclusions. In this view, justification patterns are used to classify inferences. Basing on different criteria various classifications have been made. The main criteria used are the validity of the conclusions, the productiveness of the conclusions, and the explaining-ability of the conclusions.

Using the evidential accounts of abductive inference, in this thesis we consider abductive inference as producing invalid/fallible conclusions that explain the given premises. We have find out, however that the existing justification patterns of abductive inference does not meet our needs. First, though we agree that the abductive conclusions are fallible, we believe that their likelihood of validity can be increased by modifying the existing justification patterns. This we attempt to do in chapter 5. Second, the evidential view is not concern with when and how to use each inference type in a problem solving process, such as diagnosis. So, evidential view is not sufficient for developing a computational model of abductive inference.

In the next chapter we elaborate on the second view - the "methodological" view - which may be considered as the study of abductive inference in a larger scope, with its relationships to other types of inference in complex problem solving situations. In philosophy this complex problem has, in general, been ‘scientific inquiry’. However, for our purpose, it is analogous to studying abductive inference for any other complex problem, for example diagnosis. First, regardless of whether the complex task is scientific inquiry or diagnosis, abductive inference should be described as a task at the knowledge level. Second, the uses of abductive inference in scientific inquiry, and in diagnostic as well as detective reasoning, have been commonly declared to be similar. In this way, the insight into abductive inference in scientific inquiry can largely increase our understanding of abductive inference in diagnostic problem solving.

In the next chapter, before starting to view abductive inference from a methodological view, we elaborate on how we can relate these philosophical views with the knowledge-based paradigm. In a sense, this will shed light on reasons for appealing to philosophical accounts of abductive inference when our starting point was knowledge-based systems involving abductive inference.
Chapter 4

Abductive Inference- a methodological approach

4.1 Peirce’s method of scientific inquiry

It has been commonly agreed that there are similarities between the way that a scientist invents new ideas, the way that a detective identifies a criminal, and the way that a physician establishes the aetiology of an illness. Therefore, understanding various approaches to the strategy of scientific inquiry will shed light on our understanding of diagnostic tasks.

In his later period, Peirce’s underlying motivation for investigating inference changed radically. His main concern became that of investigating which types of inferences the mind was utilizing when discovering scientific theories. That is, he focused on the logic of science, or more correctly the logic of scientific strategy. This led to his investigation of the relationship between various types of inference involved in scientific inquiry, and to the notion of “three stages of inquiry” corresponding to three types of inference.

The following quotation on the relation between abduction and induction hints that he was not concerned with the characteristics of the conclusions being inferred this time, but rather with the procedure that infers these conclusions: “... induction, hypothesis, and analogy, as far as their ampliative character goes, that is, as far as they conclude something not implied in the premises, depend upon one principle and involve the same procedure. All are essentially inferences from sampling”. (Peirce 6.40) (italic is ours). This inference type is formulated as follows:

\[
\text{The surprising fact } C \text{ is observed,} \\
\text{But if } A \text{ were true, } C \text{ would be a matter of course;} \\
\text{Hence, there is reason to suspect that } A \text{ is true.} \quad (V)
\]

An inconsistency between Peirce’s early and late view arose, at the surface, because he viewed abduction as the only “ampliative” type of inference in his later theory. He redefined the notions of abduction and induction (see figure 4.1). In fact, it is induction which underwent a radical change. In the present account, abduction was extended to cover part of what has been called induction before, and induction is restricted to ‘confirming a hypothesis’. In fact, in his later account, Peirce seems to propose that any hypothesis that is entertained as possibly true, whether it is an unobservable entity or a generalization, is arrived at by abduction. This has parallel with Harman and Josephson’s argument that inductive generalizations are reached by abductions.
From the AI point of view, we may say that Peirce was studying the method for realizing the very particular application task of scientific inquiry. In accordance with his later definition of various inference types, Peirce formulated their roles in scientific inquiry: Abductive inference generates new knowledge, that is, a hypothesis, deduction infers the consequences than can be inferred from that hypothesis, and induction is used in order to judge how far these consequences accord with experience. This account shows that the strategy underlying scientific inquiry consists of a hypothesis generation stage followed by a hypothesis testing stage. Thus, Peirce’s view of scientific strategy can be sketched as follows:

- generate hypotheses
- test the hypotheses

The scope of abduction is precisely limited to the first phase. It is worth noting that Peirce confines his attention to scientific inquiry and his methodological approach is specifically focused on this task. However, the way he studies the method of science as well as the results from this study show several parallels with studying a set of other tasks. For example, our computational approach to diagnosis benefits from Peirce’s methodological approach to science.

![Diagram showing inference, abductive, deductive, and inductive reasoning](image)

**FIGURE 4.1 Peirce’s classification according to the methodological view**

Even though Peirce signalled that he recognized an intuitional and psychological dimension in abductive inference, he did not elaborate on this. An important characteristic of our account of abduction, presented in the next section, is that we have also investigated psychological aspects of abduction and diagnosis. We put a very special emphasis on the act of ‘making guesses’ in relation to formulating hypotheses. We identify experience and the notion of similarity as important psychological concepts having an important role in the abductive process.

### 4.2 Josephson’s method of diagnosis

In Josephson’s abduction model, the whole scientific inquiry method is conceptualized as an abductive inference. He describes abductive inference as covering both generation and acceptance of hypotheses. Therefore, in his account, abduction is a higher level inference than in Peirce’s account.

According to Josephson, generation and evaluation of hypotheses cannot be separated. Josephson ensures the integration of generation and evaluation of hypotheses by putting both within the scope of abduction. In his view, an abductive strategy for medical diagnosis can be sketched as follows:
• generate elementary hypotheses
• form a composite hypothesis using the elementary ones.

This strategy differs radically from Peirce’s as this does not refer to an explicit testing process. Josephson states that a characteristic of this strategy is that it integrates generation and evaluation of hypotheses. Nevertheless, in our view, an explicit distinction between generation and testing of hypotheses does not necessarily entail that the generation of hypotheses lacks an evaluation stage. We propose that Peirce’s logic of scientific inquiry may also involve a ‘distributed’ evaluation. For example, in our account, evaluation is distributed between the generation and the testing phases. The generation mechanism can have an inherent evaluative power, for example, by the use of experience which also encodes contextual factors. This we elaborate deeply in later chapters. The second type of evaluation is through testing. The first kind of evaluation is done on the basis of incomplete information, available at the moment, while the second kind involves evaluation on the basis of information gathered in the future. The testing part, therefore evaluates the agreement of a constructed hypothesis with deliberately gathered information, while the generation part evaluates whether or not the construction of a certain hypothesis is justifiable on the basis of partially available data and information.

4.3 How contradictory are Peirce’s two classifications?

We have seen how Peirce’s account of abductive inference and of its relationship with other types of inference has shown a change in parallel with his changing motivations.

An inconsistency between Peirce’s early and late view arose because he first viewed both abduction and induction as creating new ideas, that is being ampliative while he viewed abduction as the only “ampliative” type of inference in his later theory. What made him change his early classification? His writings are fragmentary and unorganised, and therefore has led to confusions regarding, in particular, abductive inference. Nevertheless, Peirce’s ideas were possibly more systematic than his writings, and his two abductive inference accounts are not so contradictory as they seem, at the surface. In the early classification abductive conclusions consisted of unobservable entities, while inductive conclusions were generalizations. This made them different according to Peirce’s early view because the criterion was related to the characteristics of the conclusions and the premises (see figure 4.2). This figure differs from figure 4.1 in that it distinguishes between ampliative and explicative inference types. In the methodological view, abduction seems to be extended to cover parts of what has been called induction before, and induction is restricted to ‘testing a hypothesis’. Even though he does not explicitly state, Peirce’s writings hints to his opening of the door of abduction to inductive generalizations in the later period (see figure 4.3).

To quote from Peirce,

Any proposition added to observed facts, tending to make them applicable in any way to other circumstances than those under which they were observed, may be called a hypothesis.... By a hypothesis, I mean, not merely a supposition about an observed object, as when I suppose that a man is a Catholic

63
priest because that would explain his dress, but also any other supposed truth from which would result such facts as have been observed, as when van't Hoff, having remarked that the osmotic pressure of one per cent solution of a number of chemical substances was inversely proportional to their atomic weights, thought that perhaps the same relation would be found to exist between the same properties of any other chemical substance. The first starting of a hypothesis... is an inferential step which I propose to call *abduction* [Peirce 6.524-6.525].

He would, we think, call in his early view only the first example as abduction, and the second example of generalization he would name ‘induction’. Hence, in his later account Peirce seems to propose that any hypothesis that is entertained as possibly true, whether it is an unobservable entity or a generalization, is arrived at by abduction.

![Diagram of inference types](image)

**FIGURE 4.2** Peirce’s early classification

So, the criterion for his later classification is that of whether two inference types involve the same procedure. The requirement for judging two inferences to be of the same type is that they involve the same kind of process. If we paraphrase with the terms that usually are used by the knowledge level community, we may say that the criterion which made Peirce put inductive generalizations and effect-to-cause type of inferences into the same group is that he feels both use the same ‘problem solving method’. This implies that, in Peirce’s later view, two tasks that can be realized by one and the same method may be considered to be of the same type. So, he recognizes abductive and inductive generalizations as being the same sort of inference, as far as we understand, based on the criterion that they involve the same kind of process.

Peirce did not make explicit these links between his two accounts, and this evolution of his ideas are not clear from his writings. At this point, we want to draw attention to the nature of inductive generalizations. Inductive generalizations involve a ‘generation’ process; that is, generation of rules similar to the following:

\[
\text{All observed A's are B's} \\
\text{Therefore all A's are B's}
\]

The reason that this type of inductive generalization is counted as abductive must be that the phenomenon that after observing that all the A’s are B’s, the reasoner would not necessarily infer that all A’s are B’s. This may be only one of several possible relationships between A’s and B’s. Another example is that observing that the sun, a warm yellow thing
rises every day, one may make a generalization on the colour of the sun, that is 'sun is yellow', while another possible generalization may be 'sun rises every day'. Regarding verification, however, the generalization-hypothesis is given. For example, 'sun is yellow' or 'sun rises every day'. The similarity between the actual hypothesis, whether the first or the second one, and the experience determines how to modify the hypothesis. Regarding the underlying method of making a generalization for the first time and testing a previously generated hypothesis seems to be different. The first one, that is generation of the generalization in the first place probably involves a procedure similar to that of a cause-to-effect abduction, since both involve a formulation of a hypothesis which is not known in advance. In the second type, which we referred to as inductive verification, the hypothesis is known a priori. The result is either to reject the hypothesis if conflicting evidence is observed, or keep it. So, no new knowledge is produced.

\[
\begin{array}{c}
\text{inference} \\
\text{abductive} & \text{deductive} & \text{inductive (verification)} \\
\text{cause-to-effect} & \text{inductive} & \text{generalization}
\end{array}
\]

**FIGURE 4.3 Peirce's late classification**

A distinction between inductive generation and inductive verification is clarified by Feibleman [Feibleman 72]. For Peirce it was very important to separate what Feibleman called 'inductive verification' (which was what Peirce called 'induction' in his late account) as a distinct type of inference. Instead of declaring inductive generalizations as another type of inference, based on the idea that their processes may be similar, he combined hypothetical inference with inductive generalizations, and called them both 'abduction'.

Peirce makes this change in his view without remarking that he changed his classification criterion and, therefore, modified his early classification. Apparently, the reason for changing his classification was that the early classification did not support his account of the method of scientific inquiry.

An important concern of Peirce was to justify the method of science. The way he chose to do this compelled him to differentiate between two types of induction: inductive generation and inductive verification. As the conclusions of abductive reasoning are fallible, they need to be further tested. Only after testing, can the conclusions be accepted or refuted. He proposed inductive verification as being used for testing abducted results. In section 4.4.1.1, we take up Peirce's three kinds of inductive inference, namely crude, quantitative and qualitative types.

He tried to clarify the process underlying cause-to-effect type inference and inductive generalizations. The transformation of (II) and (III) in section 3.4.2 to the common form (V) in section 4.1 happened then.
4.4 The justification problem

Peirce modelled scientific inquiry as composed of abductive, deductive and inductive processes, employed in this order. That is, if he should model the method of science at the knowledge level, he would describe its method as consisting of an abductive, a deductive and an inductive subtask, performed in that order. Along these lines, it is worth remembering that induction in this context is not the same as induction in the "classical" sense. It has a much narrower sense, and includes only inductive verification. As to justification of the conclusions obtained from scientific inquiry, Peirce preferred the way of justifying a conclusion by justifying the method of drawing that conclusion. That is, knowledge is justified on the basis of the justification of the strategy. Regarding justification of the strategy of inquiry, Peirce argued that it is self-corrective. A self-corrective method, because of the nature of being self-corrective, is justified. When a method is justified, its conclusion also becomes justified. For Peirce, the matter, in this way, was transformed into the matter of proving the self-correctiveness of three-staged method of inquiry.

4.4.1 The notion of self-correctiveness

Self-correctiveness of a method indicates that the method is able to detect when its conclusions are not satisfactory, as well as to modify its conclusions. According to [Laudan 73] self-correctiveness of science claims that

1) The scientific method is such that, in the long run, its use will refute a theory T, if T is false;
2) Science possesses a method for finding an alternative T' which is closer to the truth than a refuted T.

So, in this view, scientists formulate a hypothesis, and by comparing its results with observation, improve on it. This implies that for science's being self-corrective, a mechanism is necessary for replacing or modifying the hypothesis when needed.

4.4.1.1 Peirce and self-correctiveness

According to [Gower 97; Laudan 73], in Peirce's view, what makes the method of science self-corrective is that the inductive part is self-corrective. In Peirce's words:

The true guarantee of the validity of induction is that it is a method of reaching conclusions which, if persisted in long enough, will assuredly correct any error concerning future experience into which it may temporarily lead us. (Peirce 2.769)

So, he reduced the problem of justification of his three-staged model of scientific method to the problem of justifying induction. He tried subsequently to show that induction is self-corrective. What was problematic was to satisfy the second condition of self-correctiveness, that is, the claim that method of science specifies a mechanism for replacing the refuted hypothesis with another one that is closer to the truth. Peirce managed to show that induction provides mechanisms for verifying or refuting hypotheses. However, he failed to completely show that induction, defined in his late work as testing of a hypothesis, is able to provide an alternative hypothesis, when the current one is refuted. It was possibly because he distinguished between three types of induction: crude, quantitative and qualita-
tive inductions (see figure 4.4). He refers to quantitative and qualitative inductions also 'gradual' induction, as these types of induction maintain hypotheses as long as new observations are made.

**Three kinds of induction.** Peirce's account of induction is by no means clear. We will now try to present an interpretation regarding various types of inductions that can be encountered in Peirce's writings. First, he identifies three kinds of induction; crude, quantitative and qualitative. A very important point to be mentioned, before continuing, is that he apparently considers the verification of inductive generalizations, not the generation of these 'inductions'. So, the term 'induction' refers to inductive verification of some hypotheses. More clearly, crude, quantitative and qualitative inductions are related to verification of different types of hypotheses. According to Peirce, crude induction is the weakest kind and deals with universal hypotheses having the form "All A's are B's". Such hypotheses are empirical generalizations about the trend of future experience, and are based on the presumption that "future experience... will not be utterly at variance with all possible experience" [Peirce, 2.756]. For example, we make detailed plans for our future without thinking that we may die tomorrow or the day after. Another example is that we believe the sun will rise tomorrow as it has done every day, but we may abandon this belief if it stops to do so. This type of inference is indispensable in everyday life but, according to Peirce, its use in science is very limited.

"Quantitative induction" is the strongest type of induction. It makes claims about a population."We simply propose that what observation shows to be true about a sample is also true about the population" [Gower 97]. If all the observed beans from this bag are black we propose that "all beans in that bag are black" while if 90 per cent of the observed beans are black then we propose "90 per cent of the beans in this bag are black". Quantitative inductions are statements concerning probability. Peirce also refers to this type of induction as "statistical induction".

**FIGURE 4.4** Peirce distinguishes between three type of induction: crude, quantitative and qualitative types

The third type is "qualitative induction" which, according to Peirce, plays a major role in scientific inquiry. It is founded upon the presumption that if a certain hypothesis is true, certain characteristics or qualities that follow from this hypothesis can be predicted to be true. So, testing of a hypothesis involves a similarity judgment between the actual world and the predictions made on the basis of the hypothesis. Depending on the degree of the
match between the predictions from the hypothesis and the observations in the actual world, the hypothesis is given a degree of truth. Peirce suggests that this type of induction “is simply an induction respecting qualities instead of respecting things” [Peirce, 2.706]. It is this kind of induction, in Peirce’s view, that is used in testing hypotheses in science.

4.4.1.2 Other views on self-correctiveness

Some researchers attempted to show self-correctiveness of science on the basis of the argument that the method of empirical science is identical to the methods used in mathematics, which are self-corrective. For example, in division, the quotient produced at each stage is more accurate than in the preceding stage. This claim is discredited by the argument that the domain theory in mathematics is rather different from that in empirical science, which makes it impossible to use the same method in both disciplines.

In summary, the self-corrective property of scientific methods is adopted by the majority. Nevertheless, a proper model of the scientific method and a clear account of how this model is self-corrective still remains to be explained.

Proponents of a self-corrective theory focused particularly on conceptual changes, that is replacing hypotheses by others, rather than modifying them. At the same time they failed to show how hypotheses are replaced; they succeed, ironically, only to show how hypotheses can be modified by quantitative inductions. It should be noted that the inquiry method can be self-corrective in two ways. Sometimes, in particular when hypotheses are of a quantitative type, a modification in their probabilistic expression would be self-corrective, while qualitative types of hypotheses may need a replacement for being corrected. This distinction leads us to a slight modification in Peirce’s model of scientific method, as we will see in 4.4.3.

4.4.2 Self-correctiveness of only induction?

Peirce was not able to provide a convincing account of how all induction types are self-corrective. In particular, qualitative induction is claimed to be a crucial problem in this regard for Peirce [Laudan 73; Gower 97]. Because “by using it we can test and perhaps justify a hypothesis, but if testing reveals that the hypothesis is faulty, there is no logical machinery to suggest a replacement which might have fewer faults” [Gower 1997]. Thus, qualitative induction, thus does not satisfy the requirements of self-correctiveness. Yet, qualitative induction is the type that plays a major role in science.

On the other hand, quantitative induction was not problematic with regard to self-correctiveness. Its conclusion can be modified easily when new instances arrive, since what is modified is a probability value.

Peirce has been criticised as merely showing self-correctiveness as “changing probabilities rather than changes in theories” [Laudan 73, p 298] who emphasizes that “it is the self correcting nature of science, not the self-corrective nature of a “puerile” rule, which should be our main concern”. This implies that, Laudan accuses Peirce of denying a need for the justification of abduction. As the inductive verification part does carry the whole justification responsibility, the abductive part does not need to be justified. The most important consequence of this is that abductive inference does not possess an evaluative power. This is possibly what Josephson also criticizes when he proposed a strategy for the
diagnostic task which was markedly different from the strategy for science by Peirce. Josephson stressed that his strategy integrated evaluation and generation of hypotheses.

We propose also a strategy for diagnosis consisting of generation and testing of hypotheses, similar to Peirce's three staged science method. However, in our account, abduction has an evaluative power by itself.

4.4.3 Revision of Peirce's self correctiveness claim: A circular, progressive method

Our view differs from Peirce's view in two respects:

- The method of science is not sequential but cyclic. That is, when needed, an abductive process follows the inductive one. The reason why Peirce failed to offer a new account of self-correctiveness of the method of scientific inquiry was that he did not appeal to abductive inference when induction renders a hypotheses unsuitable.

- The justification of the whole method can not merely be based on the justification of induction, and thus on the conclusion. Justification of the overall method should be based on justification of the composition of basic methods. An implication of this is that the overall method is justified when the following two conditions are maintained:

  1. various inference processes involved in the method must all be justified
  2. a justification should also be made as to the connection between these steps. That is, it must be guaranteed that the process continues (elaborated in Section 4.5).

Peirce's scientific method can be self-corrective only as a whole consisting of abductive, deductive, and inductive processes. This claim compels us to an appeal to a circular application of the abduction-deduction-induction sequence, where the need for a correction is identified by induction, and the correction itself occurs through abduction (see figure 4.5). The conclusions generated by abduction are of cause-to-effect type (i.e., qualitative hypotheses), since the mission of science is to explain surprising facts.

In science, abductive inference generates a fallible conclusion, 'deduction' infers its consequences, and induction monitors how these predictions accord with the reality. If there is sufficient agreement, then the hypothesis may be accepted. Otherwise, the inductive step gives feedback, and a new cycle starts with a new abductive process which replaces the hypothesis refuted by induction. A cyclic model of the abduction-deduction-induction triple has been employed by [Ramoni 92] in AI. They applied Peirce's scientific inquiry method to medical diagnosis, where they employed a new abductive stage when induction rendered the focal hypothesis incorrect.
We also argue that justification of a conclusion is not sufficient for justification of the whole method by which the application infers the conclusion. For example, is a diagnostic method that takes 40 years to perform justifiable? Even if the diagnosis is correct, the method is not justified. Nobody is willing to wait 40 years. Similarly, the fact that a method is self-corrective is not a sufficient criterion to justify it. It is not enough that the method ultimately finds the correct conclusion. Justifications of various kinds other than justification of the conclusion are involved in justification of a diagnostic method. Implication of this is that abduction as well as induction must be justified.

Regarding the justification of abductive inference, philosophical accounts of inference provide us with some criteria that can be used in order to justify the selection of certain hypothesis for further testing. Abductive inference has been regarded as involved in two activities. The first is to formulate some explanatory hypotheses. Once formulated, the hypotheses are ranked according to their plausibility, and the ones ranked higher in the order are selected for the test. The first activity, namely the nature of the phenomenon of formulation of hypotheses, is not elaborated in logic. The second activity, on the other hand, has been studied in detail. As Hookway explains, what the logic of abductive argument provides us with is the criteria for ranking and the selection activities in abduction.

Therefore, we have to explain how we are good at making guesses about the nature of phenomena; and we have to clarify the rules we follow in ranking explanations as more or less plausible. The logic of abduction is concerned with the second of these issues; in what circumstances are we justified in the opinion that a particular hypothesis is worthy of inductive testing?” [Hookway 1992, p. 224]

Formulation of hypotheses has been considered an ‘intelligent guess’ by Peirce, and is related to the humans having a ‘natural instinct for truth’ (Peirce 7.220). This implies the psychological aspects of abduction, which constitute a rather important dimension in our work. As will be seen in later chapters, we propose experience and context capture, at least partially, what is referred to as an ‘educated guess’.

Peirce provides some criteria for justification of selection of the hypotheses which are worthy of testing. First, the hypothesis to be tested should explain the observations that triggered the inquiry. Second, it must be empirically testable. This is the same thing as saying it must satisfy the pragmatic principle. Other rules are that we should favour theo-
ries that are simple and natural, and those that appeal to our sense of plausibility. The economical considerations are, of course the leading considerations, including 'money, time, thought, and energy' [Peirce 5.600].

Even though it seems that Pierce accepts that the logic of abduction should be concerned with criteria that determine the justifiability of an abductive leap, he did not provide an abductive pattern exhibiting these criteria. Josephson attempted to use some of these criteria in his pattern presented in section 3.4.3.2. We also propose a pattern which points to the role of context in abduction.

4.5 Pragmatism and continuity of the scientific method

According to Peirce, a hypothesis, in order to be admissible, should be capable of being subjected to experimental testing. This point is intimately connected to his doctrine of pragmatism. He defines pragmatism as follows: “Consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object” [Peirce 5.2].

Pragmatism has been asserted to be important also in contexts other than scientific inquiry. For example, in language comprehension, the ambiguity of meanings are resolved by using pragmatic knowledge. In pragmatics, the surrounding sentences are utilized in order to distinguish the intended meaning of a word or a phrase. That is, on encountering a word with several meanings, the reader may select one of these based on the previous words or sentences. Of course, the reader may check whether it was the right meaning by further reading. Peirce, in fact, formulated his pragmatism as a principle of meaning which states that “In order to ascertain the meaning of an intellectual conception one should consider what practical consequences might conceivably result by necessity from truth of that conception; and the sum of these consequences will constitute the entire meaning of the conception” [Peirce 5.9]. This statement is based on his idea that “the possible practical consequences of a concept constitute the sum total of the concept” [Peirce 5.27].

The reason underlying the need for testing is that, at the time of hypothesis generation, we have only some parts of what the target hypothesis means -i.e., practical consequences, for example some symptoms which are only a part of what a disease means. In order to be sure of the equivalence between what the hypothesis means and what the facts are, we need to compare them, by gathering necessary information (in the testing stage).

4.5.4 A connection between pragmatism and justification of the method

It is tempting to make a connection between Peirce’s doctrine of pragmatism and his claim that the scientific method is self-corrective.

Pragmatical aspects/concerns, assure continuity in the life time of the scientific method. That is, pragmatism imposes the link between hypothesis generation and hypothesis testing stages. It assures that the inquiry method gets completely applied, not blocked somewhere after an abductive step.
From the AI point of view, if we distinguish between the processes required to perform the inquiry task, and the global control structure that is imposed on those processes, we can see that pragmatism plays a determining role in this control structure. According to pragmatism, the hypothesis-generation subtask should be followed by the hypothesis-testing subtask.

4.5.5 The active role of the reasoner and judgment of the inquiry method

In order to clarify the role of pragmatism in the scientific method, we made a distinction between the process and the control structure that is imposed on these processes. A control structure can be imposed by someone or something. In scientific inquiry it is the scientist, the reasoner, who deliberately imposes that control structure. So, if the scientific method is a continuous and a circular process, it is the reasoner who ensures this continuity and circularity. The reasoner insures that the inquiry is completed, that is an adequate hypothesis is accepted. The justification of the scientific method resides in the circular application of the abduction-deduction-induction sequence. So, the problem is to show that this is a continuous and a cyclic process. Both continuity and circularity are conditions that can be satisfied by the reasoner.

There are two particularly important connection points between subprocesses of the overall method. One is right after the abduction. The reasoner should ensure that the task continues after the abductive process, by a testing stage. This is where pragmatism enters the scene, as we mentioned above. The second point is where an induction step appeals to an abductive process. This is when the induction renders the hypothesis inadequate, that is, when the hypothesis is refuted. What causes the reasoner to replace the failed hypothesis by a new one? That the abducted hypothesis was not as is expected, indicates that the reasoner failed to achieve her goal of explaining the surprising observations. Human beings, by their nature, insist in achieving their goals. It is that psychological habit which leads the reasoner to a new attempt at generating a better hypothesis.

These habits underlie the circular application of abduction, and thus constitute a step forward toward a proof of the self-correctiveness claims.

4.6 Summary

The ideas expressed in the preceding sections imply another claim: Logic alone is not sufficient in order to understand scientific inquiry and other similar tasks. Other considerations such as psychology and social context are also important in this regard.

Figure 4.6 summarizes the explanation offered here regarding the continuity and circularity of the inquiry method, where continuity between abductive hypothesis generation, and hypothesis testing is grounded in the pragmatic principle, while the circularity from qualitative induction to abduction is based on a goal-driven behaviour inherent in human nature. Our computational method incorporates the relationship between pragmatism and goal-directedness.
FIGURE 4.6 The circularity and continuity of the inquiry method are maintained by pragmatism and goal driven behaviour.

In the next chapter, we will utilize the insight gained from studying the philosophical approaches to inference in service of modelling inference at the knowledge level, as the primary step in the construction of contextualized knowledge-based systems.
Chapter 5

A knowledge level account of inference

5.1 Introduction

In order to discuss relevance of context to abductive inference, from an epistemological point of view, a framework for knowledge analysis and modeling is needed. Recent research in the area of knowledge modeling and engineering has produced several methodologies for analysing and modeling knowledge and information at a conceptual and implementation-independent level. This level of system description is often referred to as the knowledge level, after Newell's influential paper [Newell 82], while the level of implementation and representational constructs is referred to as the symbol level. Influential examples of methodologies that support knowledge analysis and modeling at the knowledge level are the Generic Tasks approach [Chandrasekaran 92], the COMMET methodology [Steels, 92], and CommonKADS [Breuker 94]. So far, context has not been systematically studied within a knowledge level framework. By building on results from each of the three example methodologies, we have found this level of analysis useful in order to understand the nature and the method of abductive inference and its connections with contextual knowledge. We will first outline the aspects important for knowledge level modeling. We will then identify the components of a knowledge level model, and then illustrate the role of inference in this model.

In this chapter we form a link between philosophical and computational approaches to inference. We show how philosophical accounts of logical inference as described in chapters 3 and 4, can be used to analyse knowledge, in particular contextual knowledge, at the knowledge level (see Figure 5.1).

<table>
<thead>
<tr>
<th>Philosophical approach to inference</th>
<th>How the philosophical approach guided accomplishment of our research goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidential approach</td>
<td>Determining the input, output and epistemological requirements of a task, i.e., task description</td>
</tr>
<tr>
<td>Methodological approach</td>
<td>Decomposing the main task into subtasks and locating the inferential ones in the task structure. Justification of that decomposition.</td>
</tr>
</tbody>
</table>

FIGURE 5.1 Philosophical approaches to inference have guided us in modelling inferences and various reasoning tasks at the knowledge level.
In modeling abductive inference at the knowledge level, logical accounts of the second stage of abduction is particularly important, that is, ranking the hypotheses, and selecting the ones to be tested first.

As Hookway explains [Hookway 1992, p. 224, italics ours]:

In other words, we might think of a number of theories of the phenomena that puzzle us and we order them, testing first those that rank higher in this ordering. The progress of science would be held up, first, if the true theory did not appear on the list at all; our imagination might fail us so that the true explanation of the phenomena does not occur to us. Secondly, if the correct theory received a very low ranking, we might never get around to testing it. Therefore, we have to explain how we are good at making guesses about the nature of phenomena; and we have to clarify the rules we follow in ranking explanations as more or less plausible. The logic of abduction is concerned with the second of these issues; in what circumstances are we justified in the opinion that a particular hypothesis is worthy of inductive testing?"

So, he considers two reasons for why a generated hypothesis may turn out to be wrong; the correct hypothesis is not generated, or it is generated but not considered worth to testing. Logic is concerned more with the criteria used in ranking and selection of hypotheses. However, the reasons for formulating certain hypotheses, in the first place, and not the others have been explained in other ways than by using only logic. For example, Peirce referred to minds as having a capacity for guessing right. He states, with respect to intelligent guessing, that "the human mind is akin to truth in the sense that in a finite number of guesses it will light upon the correct hypothesis", that is we have a "natural instinct for truth" (Peirce 7.220). On the other hand, Hookway presumes that "we cannot provide any reason for our best guesses, and they do not result from self-controlled logic" [Hookway 92, p 224]. These issues are related to the psychological aspects of abductive inference which we study in relation to the notion of experience, in later chapters.

In this chapter we deal with the logical aspects of abductive inference, and how the logical accounts of it have been useful for developing a computational model of abductive inference. It has been acknowledged, in artificial intelligence as well as in cognitive science, that context facilitates a selective and focused processing of information. In this chapter we present context-sensitive knowledge level” [Chandrasekaran 86] account of abductive inference.

5.2 Knowledge level in practice

In contrast to Newell’s rejection of structure, practical knowledge level descriptions give a special importance to structure. In Van de Velde’s view [van De Velde 93], what is important in a practical knowledge level model “is not only the knowledge that the agent seems to be using but, more importantly, the structure within which this knowledge is being used for achieving goals". So, the trend has been to capture both the content and the structure of knowledge at the knowledge level.
The importance of the structure of knowledge has been noticed in other disciplines as well. Schmidt and et al. [Schmidt 93] recognizes the ability of structuring the medical knowledge and being able to properly use those structures as a characteristic which distinguishes medical experts from novices. Their research aimed at identifying the distinctive characteristics of high performance medical experts, for the purpose of determining the curriculum that is most adequate for medical studies. In chapter 2, we have elaborated this subject in connection with the explication of the knowledge in medical domain.

A diversity of perspectives have emerged for describing a knowledge based system at the abstract level. Some examples are 'hierarchical classification', 'generic task', and 'problem solving methods'. These vary in their focus. Also, terms which do not occur in the original knowledge level view have been introduced, such as task, task structure, inference structure, and knowledge roles.

The notion of knowledge level has a large impact on two related fields: knowledge acquisition and re-usability of model elements.

**Knowledge acquisition and the knowledge level.** The knowledge engineer builds a model, together with the human expert, of the knowledge that is involved in generating expert behaviour. The knowledge acquisition community has been seeking a methodological framework that meets the demands of doing this job. This involves finding ways to structure the knowledge they get from human experts, as well as to perform computations of that knowledge in a systematic way. The theoretical basis of such a framework lies in the knowledge level descriptions. Independent from the particular implementation, this level provides a list of the types of knowledge needed to be extracted from human experts, as well as how various types can be related to each other. That is, it gives clues as to how to structure and organize extracted knowledge into a 'whole', analogous to a natural expert's memory.

**Reusability and the knowledge level.** Another important issue that has recently been addressed in AI is finding the commonalities between different domains and problem types. This aim has been reflected in the form of a search for tasks and methods that can be re-used across a diversity of domains and problems. Several researchers have been occupied with classifying tasks and methods (which are inevitable elements of practical knowledge level models) for the purpose of creating task or method libraries. The idea is that model elements used in one application may often be adequate elements in other applications. The intention is to support knowledge acquisition with predefined sets of model elements [Wielinga 93]. For example, classification, data-retrieval, plan selection, abductive assembly are all considered re-usable subtasks which can be used across various complex tasks.

Yet another aspect of knowledge level notion is that, in order to implement systems that solves problems in open and weak domain where knowledge has a vital importance, one needs an abstract model that provides guidelines for controlling the design and implementation of the system. Otherwise one falls into the trap of arbitrariness and serendipity. This thesis would be a step toward an abstract, reusable model of contextual knowledge and its use in diagnostic domains.
5.2.1 Some practical perspectives

We will roughly divide practical approaches to knowledge level modeling into two main groups: domain-model oriented and task oriented approaches. These groups differ in their focus on the elements of the knowledge model. They also differ in their terminology. However, some terms such as domain model, task model, and problem solving method are commonly used in almost all practical approaches.

**Domain-model oriented perspective:** A domain model is described in terms of domain objects and the relations between these objects. It offers a vocabulary and an ontology which enables the expression of knowledge pertinent to the domain. Understanding and predicting the behaviour of the system is dependent on the content and the structure of this model. For example, in a medical domain, the anatomical parts, the pathology underlying the organs, and the way these organs work may all be included in the domain model.

This approach emerged as an alternative to heuristic based approaches [Davis 83]. Heuristic based expert systems suffered from steep performance degradation as the first examples of expert systems relied only on heuristic rules. While heuristic systems only implicitly represented the structure and behaviour of the target phenomenon, domain models explicitly describe that knowledge. A big step toward the separation between the domain knowledge and the inference structure, as well as explicitly representing both, is taken from Davis’s [Davis 84] emphasis of the necessity of a deep model of domain knowledge. Research following this line focused on the ‘domain model’. Domain model oriented approaches can be considered one of the earliest applications of the knowledge-level notion, and focused on the theoretical account of the content and structure of the knowledge.

**Task-oriented approaches:** These approaches show variance in their use of different terms as well as applying different meanings to the same terms. Also they use different terms to refer to the same concept. So, there is sufficient grounds for confusion. Something typical in this kind of research is the utilization of tasks as reference points for the analysis of the domain model, as well as for the methods that accomplish the tasks. That is, they start by determining the tasks that the system requires to accomplish. Common to all approaches in this group is that the task decomposition has been utilized in order to organize knowledge.

“Task” is one of the terms that has been given very different meanings within the knowledge level community. As Chandrasekaran [Chandrasekaran 86] analyses, it may be used as a problem instance, a class of problem instances, the basic subgoal which includes a high level description of method, as a part of generic methods, or a sequence of operators that the system performs. This implies that there are two directions regarding the meaning of “task”. One trend is to couple the task to a method. The other trend separates task from method. Separation of tasks and methods means that there is not a 1:1 relationship between a task and a method. For example, in Chandrasekaran’s generic task approach, a generic task consisted of a couple of tasks and a method. Thus, there is a predefined method for each task. Chandrasekaran’s present “task-structure” approach [Chandrasekaran 93] distinguishes between a task and a method as to allow dynamical selection in run time, among possible methods that can realize a task. In this approach, thus, Chandraseka-
ran abandoned his task-method couples in favour of the idea of multiplicity of methods for a task.

Other terms often used are ‘method’, ‘operators’, and ‘inferences’. Methods are the ways to accomplish a task and decompose a task into its subtasks.

In the KADS methodology, an application task is decomposed into a set of generic tasks. Task knowledge is organised in a task structure which corresponds to a task tree, having inference structures at the leaves. The inference structures are independent of the domain model. KADS-I identifies four layers of description: domain layer, inference layer, task layer, and strategy level. A task in the task layer is described as consisting of a set of inferences. The inference layer describes the roles of objects in the problem solving process. That is, inferences link the domain layer to the task layer. KADS attempts to identify generic classes of such roles. Each inference plays a “role” that is necessary for achieving the task goal, for example, “classify” or “select”.

If the sequence of operators or inferences, that is the control regime, is coupled to the task, the system suffers from inflexibility. This is because the way to accomplish the task is (in a sense) predefined. In order to add flexibility to a system, a problem solving method is integrated into the system description. A task can be accomplished by several problem solving methods. Thus, there may be several lines of reasoning in order to accomplish a task. This issue is elaborated by Benjamins [Benjamins, 93].

In cases where a task is attached to only one method, the types of domain knowledge that are required for task accomplishment lie in the task description, while this knowledge is moved to the method description in cases where a task can be realized by several methods.

Chapters 10-12 of this thesis presents a set of methods that collectively implement the knowledge level account our ‘contextual knowledge and its use’ in medical diagnosis just to experiment with this account. The method-set is just one example, and other methods can be developed that also implement the context account.

Another task-oriented approach is “components of expertise” [Steels 90]. In this approach, the domain model is augmented by case models which are created as products of specific problem solving sessions. Problem solving methods can refine and expand the case model after each problem solution.

5.3 The connection between various abduction accounts

Within the knowledge level community, inferences are considered at the knowledge level in relation to terms such as task and method. In this thesis inferences are considered as low level tasks which can be realised by dedicated methods.

If we apply an evidential approach at the knowledge level, inference types would be distinguished to the extent that their conclusions differ with respect to a certain criteria. As we have seen in figure 4, Peirce classified inferences as to the characteristics of their premises (input) and conclusions (output). In this classification, abduction and induction differ in their having different types of input and output. This explains why Peirce identified abductive, deductive and inductive inferences as distinct types; if he would describe abductive and inductive inferences at the knowledge level, he would describe a distinct
task for each of these inference types. Such an ontological approach considers inferences as black boxes of which inputs (premises) and outputs (conclusions) are known, but what is inside the box is not interesting (see figure 5.2). Such a pure evidential view says nothing about the process- the method- at the knowledge level description of inferences.

![FIGURE 5.2 Task description, according to pure evidential view. This conceptualization considers only what is concluded from what.](image)

However, as Peirce changed his view, his criterion for distinguishing between two inference types became that of whether two inference types need the same type of processing. And then he considered abductions and inductive generalizations as being of the same kind since he conceived their processing as identical. So, if Peirce is required to describe inferences at the knowledge level, in his late period, he would identify abduction and inductive generalization as the same type of task, based on their being realized by the same method. This indicates that ‘method’ is also an important parameter in addition to input and output, for the description of tasks. We propose to integrate Peirce’s evidential and methodological approaches to inference and transform them into a task description at the knowledge level, as visualized in figure 5.3.

![FIGURE 5.3 Task description, integrating evidential and methodological view. This conceptualization considers also how the output is concluded from the input.](image)

This way of thinking reveals some possible benefits of utilizing logic as a tool for analyzing which types of knowledge should be used for describing inferences at the knowledge level.

### 5.3.2 The role of logic in analysing knowledge

Newell mentions in his original paper [Newell 82] that logic may have an important role in analysing knowledge. However, he does not elaborate on this, and thus gives no hint as to what he really had in mind. The idea that logic can be a useful tool for a knowledge level modeling conforms with our intention of utilizing the philosophical accounts of abductive inference as an inspiring guide for the purpose of modeling ‘tasks’ at the knowledge level.
Even though Peirce did not remark on the difference in his motivations as such, in our interpretation he first constructed an inference classification (i.e., a general inference ontology) and then investigated whether and how to use this classification when modeling the pattern of reasoning underlying scientific inquiry (i.e., the task structure of inquiry). Hence, Peirce was first engaged in studying the logic of argument while he later became more and more interested in studying the logic of the strategy of scientific inquiry as well. As his first classification did not afford his account of the logic of inquiry, he modified his inference classification. This led him to modify the definitions of abduction and induction.

A common aim at the knowledge level is to describe the tasks at various levels of abstraction in sufficient detail. Inference, in our view, is a kind of task, or rather subtask. The issue is to identify the type of knowledge that can properly describe a task.

We consider logic of an argument as a means for describing a task at the knowledge level. On the other hand, logic of inquiry is a specific example of a more general logic of an application task and reflects task-method correspondence and task decomposition at the knowledge level. It deals with the role of various inference types in a complex application task, and equally important, how these inferences are carried out.

Our concern is to model the logic of diagnosis which we consider to be similar to the logic of inquiry. When doing that, we put a special emphasis on a special type of knowledge, namely contextual knowledge which is one type of knowledge that should be used a natural part of task descriptions at the knowledge level.

### 5.4 The analysis of abductive inference patterns

In this section we want to make a comparison between two abduction patterns which are respectively due to Josephson and Peirce, according to their value as task descriptions at the knowledge level. Let us first rewrite the abductive inference patterns as proposed by Peirce and Josephson.

Peirce’s:

The surprising fact C is observed,
But if A were true, C would be a matter of course:

Hence, there is reason to suspect that A is true.

(1)

Josephson’s:

D is a collection of data (facts, observations, givens)
H explains D (would, if true, explain D)
No other hypothesis can explain D as well as H does

Therefore, H is probably true

For us, the most striking difference between these two patterns is that Josephson seems to try to offer in the premise an answer to ‘why’ question, which is important for some concerns- elaborated in the next section- at the knowledge level. He attempts to capture this by two means:
1) formalizing the relation between the input and output, such as "explains". One primary condition for justification of plausibility of the conclusion is that the desired relation exist between itself and the premises, in (II). An important point is that the explanatory relation is of causal type.

2) specifying another condition that should be satisfied in order for H to be counted as probably true (i.e., it is superior to the alternative hypotheses), in (III).

These two items constitute two prominent parameters that can be utilized in order to explain the rationality underlying the abductive reasoning pattern which maintained the formulation and preference of H. These are, therefore, arguments which can be used in order to justify the abductive inference.

5.5 The 'why' question

As we mentioned, the knowledge level model describes the behaviour of an agent (e.g., an expert, natural or artificial) in terms of domain entities.

Three main points of Newell's knowledge level notion, according to van De Velde [van De Velde 93] are:

1. Observation creates a 'what' model of agent behaviour, e.g., a series of episodes of the agent's activities.

2. Mechanization creates a 'how' model of agent behaviour, e.g., in terms of agent structure and local laws of interaction.

3. Rationalization creates a 'why' model of agent behaviour, e.g., in terms of world knowledge and principle of rationality.

According to the principle of rationality "The agent will select an action to perform next which according to its knowledge leads to the achievement of one of its goals".

van De Velde emphasizes also that the practical models of knowledge level so far, address the 'what' and 'how' questions, but the 'why' question largely remained to be investigated.

As is seen in Peirce's pattern, the reason why one could consider 'A' as probable is because 'If A then C is matter of course'. What does it mean to be a 'matter of course'? When does something count as being 'matter of course'? What are the criteria? All these questions are left unanswered in Peirce's account. Josephson, however, attempts to explicate this vague expression of 'being a matter of course', by replacing it with 'explains'. This partly makes the 'why there is a reason to suspect that A is true?' question easier to handle. Why A?, or 'why A is a matter of course when C is true?'. Because it explains C; there is an explanatory relation between them. It is exactly this relation which facilitates the answering of a 'why' question. For example, why is the conclusion (output) correct? How can it be justified? Peirce's pattern does not imply any particular type of relationship between A and C. Why it is natural to think A may be probable when we know C, is not mentioned in his pattern. In this sense, Josephson's justification pattern of abductive inference is superior to Peirce's as it does not concern merely the characteristics of the conclusion (e.g., whether it derives 'new' knowledge) but also the type of the relationship between the given (input) knowledge and inferred (output) knowledge.
In fact, the relationship should possibly be more specific than in that of Josephson's pattern, such as "causally explains" or "structurally explains". For example, if the explanation is a causal one, only a certain set of all relations represented in the domain ontology will be included in the explanation chain. Similarly if the explanation is a structural one, another set of relations will be utilized. The relations help to discover what is relevant and thus provide a sort of focus. In our account, the particular relation set comprises a part of what we call, in the following chapters, "task perspective".

Another distinction of Josephson's pattern originates from (III) in his pattern: 'no other hypotheses explain C better than A'. These kinds of 'assumptions' or 'constraints' govern the process of inferring the conclusion, given input knowledge. In this sense, we can consider Josephson's pattern as more adequate than Peirce's from a computational point of view. In other words, it seems to fit better as a starting point for a knowledge level model of abductive inference.

5.6 Abductive inference revisited from a knowledge level view

At the knowledge level (i) the components of an application task needs to be identified first and then (ii) it is described how these knowledge components are used for achieving the underlying goal of that application task.

Peirce's pattern of abductive inference (see (1) in section 5.4) may provide a source for extracting the content of a task description for abductive inference, that is, for (I). So, this pattern can be transformed to a task description, as illustrated in figure 5.4. The task is described in terms of input, output, and domain knowledge requirements. The input to the inference process, which is described by the pattern, is depicted by the letter C, and corresponds to the observed, to-be-explained facts, while the output is depicted by the letter A and corresponds to the hypothesis generated by the inference which the pattern represents.

![FIGURE 5.4 Task description extracted and transformed from Peirce's justification pattern](image)

Alternatively, we can use Josephson's justification pattern (see (2) in section 5.4) for the same purpose, that is for describing abductive inference. This pattern also can be transformed into a figure (figure 5.5) which describes the abductive inference in terms of again input, output and domain knowledge. However, the pattern explicates the need for more domain knowledge than Peirce's pattern does. In addition, there are some conditions related to the rationality of selecting H. If needed, the plausibility of output can be justified on the basis of its being able to better explain the input compared to its alternatives.
‘Why’ questions are important with respect to the ‘principle of rationality’. Even though the principle of rationality has been one of the key aspects in the knowledge level paradigm, it has been taken as a matter of course once the other aspects of the paradigm are realized. Therefore, not many attempts have been made to explicitly model the features that may constrain, or bias goal achievement. We consider this issue as a part of the principle of rationality and propose that the answers of ‘why’ questions need to refer to:

- the imposed relation between the output and input (e.g., causally explain),
- focal domain concepts,
- contextual knowledge

Along these lines, we attempt in the next three chapters to elaborate why contextual knowledge is important and should be involved in the task descriptions. We also categorize different types of contextual knowledge and determine the locus of their effects at the task level. None of the existing abduction accounts give the necessary, explicit importance to the notion of context. We therefore attempt to compensate for this absence, by extending the knowledge-level model with contextual knowledge, as illustrated in figure 5.6. The general role of contextual knowledge can be captured by the notion of relevance, which may be seen as a component of rationality. The figure emphasizes that a model or a pattern of abductive inference should include contextual knowledge.
5.7 The abstraction level of abductive inference

There is a main task in each domain. For example, the main task in diagnostic domains is to find the fault causing the surprising observations. This task has its subtasks, and these subtasks, in turn, have their own subtasks. The methods determine how a task is to be decomposed into its subtasks, and in what sequence these will be carried out. It is a well-recognized strategy to decompose a complex information processing task into smaller subprocesses that can more easily be accomplished (see figure 5.7). These subtasks are required to be identified before the actual processing starts. Thus, they must be recognized at the knowledge level.

![Diagram showing relationships between task, method, and action](image)

**FIGURE 5.7** Relationships between the knowledge level components task, method, and action. In the leaves are the immediately executable cognitive or action primitives.

The question addressed in this context is which subtasks or which subprocesses may be involved in a particular process. For example, as we have seen in the preceding chapter, Peirce decomposed the scientific inquiry process into two subprocesses: hypothesis generation and hypothesis testing. At this level, he matched the inference types with the subprocesses in which they have been used.

Researchers disagree about the abstraction level of abductive inference as a task. For example, Josephson puts abductive inference to the top level in the task tree, that is, abductive inference corresponds to the application task. According to him, abductive inference involves everything from evoking hypothesis to acceptance of a hypothesis as the best explanation. As such, testing of a hypothesis also falls into the scope of abductive inference. On the other hand, for Peirce abduction is a subtask of the main application task, that is, he places it at a lower level than does Josephson. In our view, abductive inference is a subtask of diagnosis or inquiry. It is not a top-level task, as is considered by Josephson. Thus, our approach, in this particular aspect is more similar to Peirce's. For us, abductive inference is the first step of an inquiry (both a scientific and a diagnostic inquiry) and thus excludes testing. Our view into abductive inference is shown in figure 5.8.
5.7.3 Abductive inference covers only the first stage

In order to emphasize that abductive inference covers only the generation - and partial evaluation- of hypotheses, and that it does not cover confirmation by tests we may modify its logical pattern of inquiry as follows.

D is a collection of data
H explains D

H is the best hypothesis for further consideration

It is plausible to test H first

This pattern has a narrower scope than Josephson's pattern. It is not complete as it does not sufficiently offer evaluative elements. However, it shows clearly that the abductive inference involves only the first stage of inquiry or diagnosis.

The last point we wish to deal with involves another modification regarding the pattern of abduction. In our account, the condition that the hypothesis explains data is not sufficient. There may be data/information that the hypothesis does not explain, but should be consist-
ent with. These we refer to as contextual elements which either explain the hypotheses or together with hypotheses explain D. So, the form of the abductive pattern becomes more complete when augmented with contextual considerations, Cxt:

- D is data available so far
- H explains D together with Cxt
- H coheres with Cxt
- **H is the best hypothesis for further scrutiny**
- It is plausible to test H first

Notice that Cxt includes various contextual information of which some parts predisposing or triggering the onset of the diagnostic entity (i.e., hypothesis), and some other parts taking part in the appearances of the manifestations, D.

On the other hand, the overarching task, e.g., inquiry or diagnosis, can be formulated as follows:

- D is a collection of data
- H explains D together with Cxt
- H coheres with Cxt
- **H explains D better than its alternatives**
- H is probably true.

Notice that the scope of this formula is larger than that of the above (abductive) formula since the last one includes testing.

### 5.8 Summary

Peirce’s late abductive pattern seems to be an attempt to integrate some evaluative elements into the logical form of abductive inference. However, the pattern does not provide sufficient basis for a plausibility evaluation. Josephson’s abductive pattern is more adequate in this respect. It is an improvement on Peirce’s, nevertheless it is not sufficient for a thorough evaluation. For example, it does not give any hint as to why all possible hypotheses are not generated. There must be an evaluative filter that hinders formulation of an infinite number of possible hypotheses. His pattern obviously disregards non-explanatory elements. As can easily be seen from the premises, only the data to be explained is included in the premises. However there may be very many hypotheses capable of explaining the observations. We suggest the use of contextual knowledge as a means for filtering out some implausible hypotheses. Contextual knowledge can play an important role as a focusing mechanism providing for relevance criteria. This is explained in more detail in the following chapters.

In the broad framework of scientific inquiry, Peirce relied on induction in order to confirm plausible conclusions of abductive inference. Since he failed to account for the justification of abductive inference, he dumps the entire justification on induction. This is why he argued for the self-correctiveness of the inquiry method based only on the justification of inductive verification. However, in scientific inquiry, as well as in diagnosis, justification of the conclusion does not prove the method which has drawn these conclusions. The
method as a whole should also be justified. Therefore abductive inference also needs to be justified.

Justification of abduction involves justification of the plausibility of its conclusion, which we see also as justification of the reasoner’s rationality. That is, justification of believing why the conclusion is plausible. Evaluation of the plausibility of the conclusion is distributed through the process of abductive reasoning. In an abductive process, first a number of hypotheses are evoked, in other words, activated or formulated. Then these are elaborated by using deep knowledge which makes possible refutation of some implausible ones. The hypotheses are then ranked according to some criteria. So, there are several subprocesses which may be evaluated as to their rationally. For example, we suggest the use of contextual information, usually easily available in the situation, for formulation of plausible hypotheses. This needs to be reflected in the pattern of abductive inference. For example, this would lead to a modification in the second statement of Josephson’s pattern as in the following:

\[
H \text{ explains } D \text{ in context } C_{\text{current}} \text{ (would, if true, explain } D)\]

In our view, evaluation of an explanatory hypothesis takes two forms:

- **evaluation of the rationality of the reasoner**
  1. the explanatory hypothesis should meet the goal, i.e., the type of hypothesis must be imposed by the goal. For example, should it be a 'causal' or 'structural' explanation?
  2. the candidate hypothesis must explain the so far available surprising facts,
  3. the candidate hypothesis should agree with contextual information.
  4. the candidate hypothesis should not conflict with the past experiences

- **evaluation of the conclusion**: the conclusion is evaluated by testing it on the basis of new evidence. This refers to confirmation or rejection of the hypothesis.

In the next chapter, we start investigating the notion of context. We propose that context is important for ensuring the relevance of the generated hypotheses. Context is also important for developing efficient computational methods for realizing abductive inference. In particular, in relation to psychological aspects of abductive inference, context constitutes a bridge between abductive inference and the notion of intuition which has been stated to originate from episodic experiences.
PART II

Relevance of context
Chapter 6

A methodology for modeling context

6.1 Introduction

We have developed a methodological framework that originates from Newell’s knowledge-level paradigm in order to systematically study the notion of context. In this chapter we describe this methodology.

There is very little work done for a thorough and systematical conception of context in AI. In particular, we have not found any knowledge level account of context. Recently, however, some initial work has been done (e.g., International and interdisciplinarly conference on modeling and using context, Brasil, 1997). Therefore, we did not have much to build on. In this chapter we show how a task-view provides important points of leverage in recognizing and subsequently modelling the categories of contextual knowledge. We also show various types of contextual influences in problem solving.

6.2 Distinction between the elements and the roles of context

In attempting to make a theoretical account of context, we have noticed that a large number of researchers have studied context with reference to a specific area or a specific problem, isolated from other context studies. An important question is whether there may exist some aspects that are shared by several researchers or research communities that study context effects.

We have found that distinction between the elements of the context and the roles of the context provides helpful in answering this question. Studies show that the identification of context elements heavily depends on the type of task and domain in question. On the other hand, the role that context plays can be generalized over specific tasks and domains. We will therefore start out by claiming that the general role of context does not show much variance across domains or tasks, whereas the elements of a situation that play these roles do.

In the next section we present our view on the roles of context not with respect to a particular task or process, but at more abstract level. These issues are related to the elements of context which are analysed in chapters 7 and 8. The exact roles of context at the process level, and the general strategy to investigate the roles of context at that level are presented in chapter 9.

6.2.1 The role of context - relevance and focus

The notions of relevance and focus capture the essential aspects of context roles. Relevance refers to the appropriateness and usefulness of a response in a particular environ-
ment. Reasoning from natural language is dependent on the social context. For example when you ask “How is your family” to a 14-year old school girl you mean by “family” her mother and father. If you ask the same question to your colleague whose wife you know is sick, you mean his wife. The meaning of a statement may be different in different contexts. It is our ability to quickly shift from one context to another that makes our everyday life bearable.

At a more detailed level, context is important for the generation and evaluation of explanations. People asking why an airplane crashed to the ground will not be satisfied with the answer, ‘because of gravity’, even though this is not wrong. A possible acceptable answer would convey an anomaly that occurred, for example, in the engine of the plane.

Generally, principled knowledge (i.e., textbook knowledge) does not change across its users or the situations in which it is used. However, the use of principled knowledge is relative to the context in which it is applied. Relevance is, therefore, directly proportional to the quality of the solution produced for a problem. In problem solving, there exist several lines along which to reason, and often several alternative solutions to a problem. Context plays an important role in choosing the most relevant candidate. For example, recommending an angioplasty in a hospital which lacks the necessary instruments will not be useful and can be said to be a ‘low quality’ conclusion.

In addition to relevance, the other main role of context in problem solving and learning is its focusing ability. Focus is important, for ensuring efficiency of the problem solving process while maintaining relevance. At any point, the attention of a problem solver is focused on particular issues and also particular aspects of that issue. This is necessary because the amount of information that reaches us is huge. Neither our cognitive capacities nor our time is enough to process all that information in detail. We focus our attention in two ways: concentrating on certain goals, and focusing on a portion of knowledge and information that may be related to that goal. Context plays a significant role in concentrating on the most adequate goal among several ones, as well as in focusing on the most ‘relevant’ portion of information. When there exist several possible solutions, the one which is most relevant to and consistent with the current context is preferred.

6.2.2 Elements of context
The elements of context are identified and explicated, first as high-level context ontology. The ontology can be shared across various diagnostic domains as it promotes the intrinsic structure of contextual knowledge relevant in the diagnostic domains. The low-level contextual elements pertinent to a particular diagnostic domain (e.g., medical, car mechanics) will be elicited by instantiating the context ontology.

6.3 Research goals
We investigate the notion of context along two dimensions: as a knowledge type, and as a means for focusing attention.

The view of context as a knowledge type sets up the goal of identifying elements of the context for a particular domain. Context has usually been referred to as atomic, like a
closed box. We attempt to look into the box, and break it into meaningful and useful pieces, in the form of knowledge types.

The other dimension is concerned with context as a means. So, its effects are of interest in this view. The questions pertaining to this view are what is contextual knowledge for? What kind of roles and effects may it have? Where and when may these effects be the matter of fact? And how may these effects be realized?

We may now list four research subgoals which we will investigate in the following sections:

- categorizing contextual knowledge types, i.e., defining a context ontology
- categorizing contextual effects
- mapping context effects to the context ontology

6.4 A Task-centered methodology for studying context

Modelling the process of diagnosis involves identifying the subtasks which, when accomplished, lead to the achievement of the diagnostic goal, i.e., finding the fault causing the abnormal observations. It is exactly these subtasks that have a central role in modelling a shift of attention. This is because the subtasks are the loci where a shift in the focus of attention may be invoked. Knowledge level analysis can be performed from different knowledge perspectives, i.e., a task perspective, a method perspective or a domain knowledge perspective. The tasks, in turn, identify the need for knowledge, including contextual knowledge. We start by locating the subtasks of the whole diagnostic process where contextual knowledge is useful. After locating the points of contextual knowledge use, we try to understand the way these influences come into existence. The last step is in attempting to analyse the types of knowledge being utilized at these points.

The process of modelling contextual knowledge can be considered as consisting of two activities: (i) typing/categorizing contextual knowledge, and (ii) identifying what kind of role each type plays, and where/when this happens during problem solving. This leads to two perspectives that are not mutually exclusive, but closely interrelated. They consider context as, respectively,

- a means for triggering a shift in the focus of attention.
- capturing a special type of knowledge and information, and

The first perspective leads to the study of how contextual knowledge can be categorized, starting with the context ontology. The second perspective takes, as a starting point, an application task (in our case diagnosis) and identifies the subtasks of the overall diagnostic process where some contextual influences are anticipated, as well as the role of contextual knowledge in each subtask.

The "focus of attention" perspective leads to the study of the relationship between a task and the type of context that it relies on. We may also say that the first perspective adopts a static view, while the second one is more concerned with dynamic aspects of context, i.e., its use.
A methodology for modeling context

A discipline that provided us with much support and insight in developing a methodology for investigating context has been medical research. The link between process and content, emphasized by Barrows [Barrows 94], in teaching medical expertise, has guided us toward the way we study the contextual knowledge, a continuous attempt to keep the link between the reasoning process and the disease process.

The idea of the dependence between the knowledge and its use, proposed in AI by Chandrasekaran [Chandrasekaran 92], agrees with Barrows’ emphasis on the link between process and content in connection with problem-based curricula in medicine.

According to van Heijts et. al., [van Heijst 97] the following activities are identified in the construction of a knowledge-based system: (i) construct a task model for the application; (ii) select and configure appropriate ontologies, and if necessary refine these; (iii) map the application ontology onto the knowledge roles in the task model; (iv) instantiate the application ontology with domain knowledge.

Inspired by the way cognitive psychologists have studied context empirically, and from van Heijts’ approach to knowledge-level analysis and modeling, we have developed a methodology which is general enough to be used, at least, across various diagnostic application. The methodology can be expressed as following four stages:

- construct a model of diagnostic ‘process’ knowledge. This serves to explicate all the relevant application subtasks (i.e., diagnostic subprocesses),
- identify the loci of contextual effects (i.e., the subtasks where context may have an influence)
- identify the role that context plays at each locus (i.e., the way it effects each subtask)
- identify the source of context that can play these roles.

The ‘loci’ here corresponds to ‘task’, and the ‘source’ to the domain knowledge -more specifically its contextual portion- in figure 3.1 Related to ‘source’ is the generic ontology we develop and presented in chapter 8. Chapter 7 grounds the basis for this ontology.

Figure 6.1 illustrates the entities, such as loci and source, that our methodology explicates when applied to the medical domain. In the figure, the task-tree is a part of process knowledge, and the contextual domain knowledge is part of the general domain knowledge (i.e., content knowledge). ‘Generate-hypothesis’ is the name of a task, while ‘age’ is a concept belonging to contextual domain knowledge. What we illustrate in the figure is that the ‘generate-hypothesis’ task is a loci where context affects the process. The knowledge type (i.e., source of this context effect) is ‘age’, which is an element of the general domain knowledge, thus ‘content’. The description belonging to each task in the task tree refers to the core and the contextual knowledge that is needed for its accomplishment. Thus, this figure also provides insight into how we realize the link between the ‘process’ and the ‘content’ knowledge.
A METHODOLOGY FOR MODELING CONTEXT

FIGURE 6.1  The task descriptions (not shown in the figure) of the tasks in the task tree refer to both core knowledge and source of contextual effects, in terms of the entities in the knowledge base. This figure partially shows the way how contextual elements are integrated in our system.

In chapter 7 we survey studies from cognitive psychology on context, and illustrate how a global interpretation of these experiments reveals two dimensions which possibly shaped the way these experiments are set up. We may see from the experiments that the two parameters which have been used for detecting context effects are the various kinds of contextual elements, and various kinds of learning tasks. So, tasks may be a starting point for investigating the elements and effects of context. We, therefore, attempt first to determine the subtasks of the diagnostic task. We do this by applying the method of scientific inquiry, as proposed by Peirce, to the diagnostic task and determine the subtasks underlying medical diagnosis. These subtasks comprise a set of loci from which we can point to various context effects.

6.5 Categorizing contextual knowledge types

We include contextual elements both in the generic ontology and in the domain knowledge. In particular, we are interested in identifying different types and setups of contextual knowledge. This will make it possible to refer to, and to utilize contextual information of various types and at different abstraction levels.

Context research in cognitive psychology has studied both theoretically and empirically the context influences. As we will see in chapter 7, the empirical studies of context are centered around experiments which study context effects on a particular task. The tasks are most often cognitive tasks such as word learning and remembering in various contexts, face recognition, or eyewitnessing. Any context that is used in such experiments embraces a set of contextual information. The experimenters observed whether the use of a chosen set of contextual elements did or did not affect the reasoner's cognitive performance. The aim of the early experiments was only to show context effects in general. A global look into the empirical results indicates that different tasks use different contextual informa-
6.6 Categorizing context effects

The main intention underlying the attempt to categorize the context effects is to develop a framework within which context effects may be mapped out. This research task will be accomplished when we are able to determine which type of contextual information, and in which stage, context influences an overall problem solving process.

Though we are mainly interested in explaining context effects in diagnosis, as most context studies in psychology go hand in hand with understanding perception, we start by understanding several context effects, which we will then transfer to diagnosis.

Cognition involves an immense number of skills and processes. The acquisition and use of skill and process knowledge, according to the "information processing" view, consist of a number of separate stages. Researchers have proposed various models of perception in order to describe the subprocesses where context influences are determined. Figure 6.2 (From [Reed 91], p. 5) illustrates the stages that are commonly included in the information-processing model. This model we use in order to understand the context effects observed in different areas such as perception, memory, language, concept formation, or concept categorization.

It has been recognized that perception has a selective nature. This must be so, because otherwise the human cognitive system would be overloaded with information. An implication of this is that attention must be selective. So far there is agreement among researchers. However, theories vary as to the stage at which this selection occurs. For example, does selection occur before or after pattern recognition? We will not go into the detailed discussions that one can encounter in the literature, but rather give some ideas about how some cognitive psychologists relate context to human cognition. Then we turn to diagnosis and locate the stages of diagnosis at which context effects may be postulated.

![Diagram of information-processing model](image)

**FIGURE 6.2 Stages of an information-processing model (from [Reed 91, p.5]).**

External stimuli bombard human sensory receptors, where they are converted to signals. These signals are stored in the sensory store for a brief period. At this stage what differentiates between signals is their physical characteristics, such as intensity and location. When they come to the pattern recognition stage, other characteristics are used to recognize them. When we recognize a familiar pattern we utilize the knowledge that we previ-
ously stored. Some theorists claim that humans have a rather low capacity of pattern recognition. So, there must be filtering to determine which signals are to be recognized. The mission of the "filter" is thus to eliminate some signals when many patterns arrive. Other theorists argue that many patterns are recognized simultaneously but not all of them are stored. "Selection" determines which are to be stored. Thus, "the filter limits the amount of information that can be recognized at one time, and the selection limits the amount of material that can be entered into memory" [Reed 91, p 6], i.e., learned.

This implies that attention is possibly focused in several loci, in order to ensure a balance between the amount of things to be learned and the human cognitive capacities. Context is claimed to have an influence in the filtering mechanism as an attention focuser. For example, the person may be instructed to pay attention to a man’s voice rather than a woman’s [Thomson 86]. This instruction is a form of context, in this example. Another context effect may be seen when the pattern recognizer has many alternative identities. For example, when the pattern recognizer has to choose between cat or lion as the identity of the figure on a picture. It may base its selection on the biases implied by the context, such as information about adjacent objects. "Cat" will possibly be the right identification if the picture is taken inside a normal home.

We want to stress that our general approach is not to duplicate cognitive process models of human behaviour when constructing computational decision support aids. We try, rather, to borrow some ideas from human context utilization and, on this basis, construct a plausible model for an AI system. In the following chapters, we develop a similar model, but for diagnosis and use that model for explaining the context effects in diagnostic procedures. The task of locating and categorizing context effects can not be accomplished without making a model of the process that one intends to deal with.

### 6.6.3 Context effects vary

In 'empirical' tasks such as scientific inquiry and diagnosis, we may summarize two important roles of context as situating a problem

- in the reasoner's memory, and
- in the physical world.

The former may intuitively be related to what Peirce calls 'abductive talent'. The latter is related to the pragmatic aspects and utility considerations of problem solving. At the very top level, we classify tasks involved in real world problem solving as cognitive and physical tasks. Both types of tasks utilize contextual information, but they use different types.

Cognitive tasks are triggered by observations made in the real world, and physical tasks are guided by memories. It is just these connections between memory and the real world that shape the course of problem solving. These connections are context-sensitive. For example, when observations trigger a cognitive task, which parts of the memory that will be evoked is not independent of the context in which the observations occur. Similarly, when it comes to acting in the real world, our knowledge about the situation determines what types of actions will be of benefit. This implies that context serves as a bridge connecting mind to reality. Figure 6.3, illustrates the parts of both memory and the real world have contextual components. In humans, the mind possesses a model of the real world. The real world is contextualized, that is, there is no single ideal world, but several versions
of it. The inclusion of contextual aspects into the mind’s model of the real world, makes the mind closer to reality. At least the gap between the real world and its model gets smaller by enhancing the model with contextual components. So, an artificial system which models the mind’s model of real world will also greatly benefit from context. A contextualized artifact will be more comparable with, in an indirect way, the artifacts in the real world than one which is not contextualized.

![Diagram of inquiry process](image)

**FIGURE 6.3 Our knowledge about the world is contextualized**

A system performs one of two types of tasks at any moment, i.e., either a cognitive or a physical task. Cognitive tasks are performed in the mind, based on the information available in the real world, often by using memory. Physical tasks, on the other hand, imply interaction with the physical world by way of taking some actions. The real world is contextualized and is there independent of the mind. That is, it is there regardless of whether the mind is capable of using it or not. A model which is not capable of recognizing contextual information that the real world provides cannot properly cope with such a rich world.

As stated in chapters 4 and 5 in connection with the role of abductive inference in inquiry, inquiry means formulating hypotheses at the first place that explains a surprising phenomenon in the real world. This is a cognitive task and utilizes memory in order to find stored memories that can shed light on the phenomenon. In case the memory is tuned by contextual components, then there is a large possibility to find a memory which can, to a substantial degree, help to explain the newly observed phenomenon. Testing hypotheses entails physical interaction with the world. When taking actions in the real world, the mind should recognize the realities of the world it is operating in, its resources and possibilities. The actions which are wisest to take are those that best agrees with reality (see figure 2.1).

### 6.7 The map between context elements and effects

Different application subtasks render different context effects which are maintained by different contextual elements. For example, some tasks involve only memory, such as recognizing somebody. You may easily recognize the bus driver whom you sometimes see when you take the bus from your job. On the other hand, you may have difficulty remembering him upon seeing him at the cinema. So, the task of recognizing people you know in
connection with their job, is conditioned on the location you required to recognize them. The location-context has a facilitating effect on your recognizing people who you know in connection with their job. You associate location with such people. Another, rather different example is the task of drinking water from a glass. Here, the type of glass and whether it is completely full, or half full becomes important for how you drink water. So, for the drinking-water task, context may have a more physical importance, e.g., you may decide to make different motions in different contexts. Therefore, for a knowledge level model, it is important to make explicit links between a task, and the elements that facilitate or inhibit the performance of that task.

The preceding research tasks are assumed to establish a model of the application tasks, and to specify where and which type of contextual information is expected. What remains is to explain how these effects happen, that is, in what way contextual elements are facilitative or detrimental at each stage that they are believed to have influence. Such an elaboration is necessary before developing the model of an AI system that illustrates context-guided reasoning. Based on the importance of past experiences in problem solving and learning, we adopt the case-based reasoning paradigm to explore our context-sensitive approach to problem solving. This is elaborated in chapter 10.

6.8 Summary

This chapter highlights two concepts along which the notion of context can be investigated, its elements and its role. The role of context is twofold. It provides the loci where a shift in the focus of attention is triggered. It also provides a means to focus attention. We have also presented a methodology that supports a systematic modelling of context in a domain. In later chapters we will discuss how we applied this methodology in our work.

The term ‘context’ has, in recent years, been used in order to account for a wide range of behavioural phenomena. Much empirical research is also devoted to context. However, the focus of the research was not on understanding the phenomenon itself, but context has rather been used as a tool for investigating other topics [Wickens '87]. An example is the line of research investigating the structure of semantic memory, where the question is whether words are stored as a set of interlinked features. Another line of research investigates the nature of forgetting and remembering. So, empirical research on memory has been the natural medium for context research. In the next chapter, we summarize the empirical and theoretical results provided by cognitive scientists on the study of context.

In order to clarify the elements of context in any domain, a need arises for a context ontology. We will develop a context ontology in chapter 8.
Chapter 7

Context Studies in Cognitive Psychology

7.1 Cognitive science and AI

Cognitive psychology has been a pioneer in extensively studying the notion of context both from empirical and theoretical perspectives. We gain much inspiration form these studies. In fact, cognitive psychology and artificial intelligence have had a rather close relationship in general, and our work illustrates a synergy between these two disciplines, in diagnostic problem solving. As Fisher and Yoo assert, “The common wisdom is that psychology provides a specification of intelligence, and AI provides tools for implementing this specification” [Fisher 93, p 220].

A primary difference between AI and cognitive psychology is that psychology studies the intelligence of human beings while AI studies intelligence in systems in general, “intelligence by whatever means......The underlying theme is that psychology addresses intelligence as it actually is (already, in people), where AI pursues intelligence in all its possibilities” [Haugeland 84].

Cognitive psychology have arranged experiments for studying the cognitive system with regard to context. Experiments are means for observing which variables contribute to the cognitive task under consideration. Thus, as the scope of the task gets bigger, the number of variables that may influence the experiment gets larger, and the experimenter may not think of everything that may be important. So, experiments are not usually arranged to study cognitive processes with a large number of contributing variables. We will see in this chapter that this has also been the case for the study of context. Cognitive psychology has primarily focused on studying contextual effects on relatively isolated (and simpler) tasks, such as verbal learning and face recognition, rather than the whole cognition, such as problem solving.

Regarding the question of how experiments can contribute to doing better AI, Polson [Polson 84] writes that psychology can provide guidance in the identification of useful components for AI systems. That is exactly the way we utilize psychological studies of context in this work. We draw upon the experiences from cognitive psychology, even though the tasks and domains for which cognitive psychology studied contextual influences do not quite fit with ours.

In this chapter we selectively review specific studies that demonstrate the role of context in memory tasks such as learning and remembering. We draw upon empirical studies and theoretical analyses which, in fact, are conducted in areas where much less knowledge is required and manipulated than what is needed for real-world problem solving. This is because context has only been investigated in a narrow sense. However, these studies have been an inspiration source in several ways, such as pointing to how and where to seek context effects, and which variables that can be important for manipulating context.
7.2 Empirical research on memory

Cognition involves perception, comprehension, memory, problem solving, decision making, and planning and controlling actions. Cognitive psychologists reported significant contextual effects in memory. It was, in fact, memory studies that alerted some influential researchers to the importance of context [Tulving 73; Eich 80; Baddeley 82a-b; Smith 86; Thomson 88]. Therefore, most empirical and theoretical context investigations are related to memory tasks.

In this section we will present a repertoire from studies on learning and remembering, and then natural evolution in cognitive psychology. A main issue addressed in memory research has been various types of remembering, such as free recall, cued recall, and recognition. These differ from each other by the number of cues provided at retrieval time, and by the significance and specificity of the cues. In a typical recall test, subjects are given lists of words and subsequently asked to write down as many words as they can remember from these lists. The test becomes cued recall when some contextual information is associated with the words to be learned, and these are kept constant or changed at retrieval time. On the other hand, in a typical recognition test subjects are given words and asked to decide whether these are on the list that they were shown previously.

Across a significant number of published studies, we found large agreement on the context effects on memory. Nevertheless, conflicts exist regarding the context effects on recall and recognition. In fact, the relationship between recall and recognition has been a very heated subject and continues to take a prioritized place in research agendas. Towards the end of the chapter, the formulation that we adopt for the relationship between recall and recognition will be presented. An interpretation that we found close to ours regarding the various types of context on recall and recognition will also be elaborated.

We will first give examples from empirical studies pursuing context effects in remembering. This will involve the basic experimental lines that contributed to the evolution of current theoretical accounts of context. The experiments usually aimed at studying either recall or recognition. Both used verbal learning, including both nonsense syllables and words, as well as face recognition, as primary domains. For each experiment, the interpretation of the experimenter has been included, without commenting this interpretation. An experiment and its interpretation have later been referred to whenever other researchers interpret the results of the experiment differently.

A paradigm called 'reinstatement' is utilized in these experiments. In this paradigm, memory testing is arranged in the same context as the one that existed at learning time, or in a deliberately manipulated context. When remembering is found to be dependent on the reinstatement of the context, then memory is said to be context dependent.

Researchers declared that environmental context has a particularly strong effect on recall. The internal state (e.g. the mood, influence of alcohol) of the subject also is reported to effect recall.

A definition of context in controlled experiments is "background surrounding what is presented to the participant to be remembered as the experimental material". Investigations of context effects on memory involve many aspects suspected to constitute a context. These include state and mood of the subject, aspects related to the location of the experiment,
instructions given at the start of the experiment, incidental characteristics of the target, etc. The interpretations placed on these experiments conflicted with each other, both regarding context, and otherwise. In order to resolve these conflicts, a more global theoretical accounts is developed. We will try, below, to reflect this evolution towards the theories of memory.

**State-dependent memory experiments:**

A group of experiments aimed at investigating the effects of a change in the internal state of the subject. These are centered around “state-dependent learning” and “mood and memory” experiments.

State dependent memory experiments (e.g., [Eich 80]) manipulated subject's internal state by using pharmacological means, e.g., alcohol and marijuana. The state of the subject is the context in which the memory is studied. Subjects experienced an event when they were drunk and the test of memory occurred either when they were drunk or sober. Results suggested that state-dependent learning is observed to some degree on tasks which called for recall, but such effects are not observed for tasks involving recognition. As a somewhat eccentric example, drunk people did not remember events when they sobered up, but did recall them when they became drunk again.

Broadly similar results are observed with mood-dependent memory tests, where hypnosis is used to induce happy or unhappy mood states. Mood-dependent performance on recall is determined. That is, recalling with the same mood as in the learning stage was almost double of that of in an opposite mood. For example, subjects recalled unpleasant experiences faster when they were depressed [Bower 81]. These experiments seem to indicate that state and mood could be regarded as effective contexts, and that reinstatement of such contexts assist recall.

**Environmental context effects:**

A series of experiments showed how learning is influenced by environmental contexts even when there is no obvious or intended relationship between learning material and the environmental context. For example, findings showed that if subjects learned a list of words in one room and tried to recall them in another, they remembered less than if learning and recall occurred in the same room [Smith 78]. Even stronger context dependency is obtained from experiments with divers in two different environmental contexts (below and above water). Subjects could remember several words when the environment was the same (either on land or under water) in the learning and recall stage, than when it happened in opposite conditions [Godden75]. Such findings were not observed for recognition.

**Effects of the physical properties of the target:**

Experiments involving aspects incidental to the to-be-learned word, such as the colour or the font in which the words are printed in the learning stage showed strong context effects upon recognition [Smith86].
The distinction of intrinsic and extrinsic context types:

Attempting to interpret the findings of these experiments, researchers distinguish between extrinsic and intrinsic context types. The criteria on which they base their distinction is the target itself. Extrinsic context embraces everything in the situation that is not the to-be-learned stimulus. Thus, the colour of the experiment room as well as the state or mood of the subject are counted as extrinsic contexts. Intrinsic context comprises the characteristics of the to-be-learned target itself, such as the type of the font in which a word is printed.

Geiselman and Bjork [Gieselman 80] suggested, in light of the experiments mentioned above, that extrinsic memory does not appear to affect recognition, whereas intrinsic memory does affect both recall and recognition. However, later experiments conflicted with this account, since they observed extrinsic context effects also on recognition. An example of this is face recognition.

Face recognition:

Experiments on face recognition demonstrated that environmental context elements also may have an influence on recognition. In a number of studies, the face learning process is modified by instructions, such as “assess the honesty of the face”, or “assess the relationship between two faces”. Also experiments on eyewitness memory showed consistent effects of extrinsic context on recognition. The results showed that changes in context such as face associates and background cues have resulted in decreased recognition accuracy. So, the environmental type of context, in contrast with earlier accounts, could have a role in recognition. Thus, the distinction between context types that affect recognition and the types that do not has no relation to the extrinsic context, as was previously anticipated.

The situation showed that a given feature, such as extrinsic context has an effect on recognition in some cases (e.g., face recognition), and not in other cases (e.g., the experiment with divers, and state-dependency experiments).

In order to make theoretical sense out of these contrasting empirical findings new theoretical studies have been developed. Two important approaches that will be elaborated below are Tulving’s Encoding Specificity [Tulving 73], and Craik and Lockhart’s Level of Processing principles [Craik 72, Craik 75].

7.3 Theoretical Research on Memory

Attempts have been made to explain “What the context influences”. The view adopted by Hewitt and others in order to account for contextual effects on recall and recognition was too stimulus-oriented. According to this account, being either extrinsic or intrinsic to the stimulus determined whether the contextual information could affect recognition or not. When this view failed to explain the variety of the findings, attention was turned to others aspects. The underlying process was suggested as having a vital responsibility for the retrieval performance. A particular point is that the compatibility of processes invoked during learning and retrieval time is the mystery which explains a context element’s varying effects upon recognition. Tulving’s Encoding specificity principle emphasizes the encoding process during learning.
Encoding Specificity Theory:

Tulving [Tulving 73] suggested his encoding specificity theory based on a distinction between episodic and semantic memories. The theory maintains that, “An encounter with a word results in the creation of a unique trace....” [Watkins 75], and implies that a subject does not learn, for example, words but experiences. This principle, consequently, asserts that a better retrieval is conditioned on the closer resemblance between learning and retrieval situations, and thus makes clear the role of context elements both in learning and retrieval. Hence, a cue will only be effective if it is specifically encoded at the time of learning.

The emphasis is on two aspects: The existing retrieval cues and the characteristics of the past encoded episodes. What improves the retrieval is the degree of match between these two. The statement that, “Specific operations performed on what is perceived determine what is stored, and what is stored determines what retrieval cues are effective in providing access to what is stored” is accommodated also by experiments reported in [Barclay 74]. They showed how the word PIANO could be encoded differently when its different aspects are emphasized at the time of encoding. The subjects were given sentences such as “the man tuned the piano”, or “the man lifted the piano”. The cues “something melodious” and “something heavy” are used at retrieval time. The recall of piano was higher when the cue was “something melodious” and the learned sentence was “The man tuned the piano”, and similarly when the cue was “something heavy” and the sentence was “The man lifted the piano”.

This implies that only specific aspects of piano was encoded each time, and the combination of the retrieval cue and the original encoding was necessary for the recall.

One implication of the encoding specificity research is that what is encoded is a record of the item interpreted within its context. Thus, the context in which encoding and retrieval take place plays a central role in determining what is remembered [Smith 94.]

While the Encoding specificity theory suggests the strong relation between the encoding and retrieval, Craik and Lockhart’s Level of processing elaborates the underlying factors of encoding.

Level of Processing hypothesis:

Craik and Lockhart’s levels of processing hypothesis [Craik 72] explains how the nature of the process affects memory encoding. It elaborates on the encoding specificity theory in that it notices that the depth of the process underlying learning determines what to encode. Craik and Tulving showed that processing a word for its meaning (e.g., having to answer questions about the meaning of the word) led to a vastly better recall than considering whether it is printed in upper or lower case. [Craik 75]. Experiments by Winograd supported Craik and Lockhart’s argument. He observed better recall when subjects judged the honesty of a person than when they answered a question about the nose size of the person [Winograd 78]. Craik and Lockhart suggested that semantic processing induces a deeper level of processing than phonemic processing (e.g., rhyme with another word), while phonemic processing itself induces deeper processing than physical features (e.g., font of the printed word) do.
Baddeley’s interpretation of depth of processing provides two assumptions: “first, encoding many distinctive features will help recognition, and second, deeper semantic and elaborative encoding will tend to lead to the encoding of more features” [Baddeley 90, p 169]. That is, deep encoding enhances memory because it creates memory traces that are more discriminable than items that have been encoded shallowly.

**Interactive vs. independent Context**

Realizing that Hewitt’s context taxonomy fails to account for some of the empirical findings, and being inspired by the above-mentioned two principles (i.e., Encoding specificity and levels of processing), Baddeley suggested a distinction between *interactive* and *independent* context types. Baddeley’s formulation is based on the role of context in learning, as well as in retrieval. In his account, context may be encoded in two different ways, interactivity and independently. “He argued that interactive context “influences the memory trace directly by affecting the way in which the stimulus is encoded” [Baddeley 82], whereas independent context and stimulus are processed without interaction between them. Independent context bears a purely arbitrary relationship with the material learned. As such, it does not determine the interpretation of the material. “The concept of interactive context places the emphasis on the *processing* carried out by the subject rather than on the *characteristics* of stimulus material” [Baddeley 82].

When context is independently encoded, it is stored “together with” the trace of the stimulus, without actually making any change on the trace [Baddeley 90]. An example of independent context is the environment where the learning and remembering of words happen. On the other hand, semantic cues are interactive context for word meanings. The experiment by Tulving and Osler is a good example of this distinction: the subjects were presented the word ‘city’ with separate cues ‘dirty’ and ‘village’. The concept ‘city’ is semantically very rich, and aspects of it that are encoded with the cue ‘dirty’ versus ‘village’ are quite different. When encoded with ‘dirty’, its semantic features like traffic fumes, garbage and dust are activated, whereas features like quiet, clean and friendly are encoded when the cue is ‘village’. So, the two cues lead to different traces being stored.

The empirical findings mentioned so far can be explained in terms of the distinction of interactive vs. independent context; independent context effects only recall while interactive context assists both recall and recognition.

### 7.4 Recall Models

There exist mainly two contrasting recall models in recent psychology literature. One is proposed by Kintsch and referred to as generate-recognize model. It suggests a two-staged recall. The first of these stages involves an *access* process where candidates are generated, and the second involves a *distinction* process where the target is selected between the candidates. The last stage is also called recognition. Thus, in this account, recognition is the second stage of recall. Knowledge is represented in an abstract semantic network, and learning occurs through tagging nodes in this network. The original generate-recognize model does not consider episodic memory structures.

Another approach denies a two-staged recall and such a relationship between recall and recognition. On the contrary, it assumes that processes underlying recall and recognition
are almost the same. Watkins and Tulving state that “There is no fundamental difference between the processes of recognition and recall; the term recognition is used to describe a retrieval environment in which the nominal copy of the encountered word is physically present, and the term recall when it is not” [Watkins 75]. An important difference between these two approaches is that the latter assumes that humans learn experiences, not, for example, words. Consequently, it considers an episodic memory where the unique traces of experiences are stored.

Tulving used, as evidence, some empirical findings, against the two-staged model. For example, Thomson and Tulving showed that when a word is learned in the context of a weak association, a subsequent recall was better with this association than with a stronger association. In this experiment strongly associated pairs of words are selected, such as WHITE-BLACK and WOMAN-MAN. One of the words in such a pair is then paired with another, weakly related word. For example BLACK is paired with train. The new pairs (e.g., BLACK-train) are presented to the subjects at the time of learning. At retrieval time, the recall of the target words (e.g., BLACK) was superior when the weak cues (e.g., train) were used, compared to non-cued recall. On the other hand the performance of a recall cued with a stronger associate (e.g., white) was not significantly different from non-cued recall. This finding is in contrast to the generate-recognize approach to recall, because if the knowledge was represented only as an abstract semantic network, the strong cues would always be better cues than weaker ones.

Another finding that weakens the generate-recognize model is the words that can be recalled but cannot be recognized. In these experiments subjects learned target words in the context of associated words. In subsequent cued recall tests, subjects could recall target words when the cues were presented; they recalled all the words without error on two successive tests. But, they failed to identify 10% of the target words in the recognition test.

Proponents of episodic memory interpret these two important findings as dismissing a two-staged recall model. They argued that if recognition is the second stage of recall, an item that can be recalled must have passed also recognition stage, and should be recognizable.

A theoretical account that manages to explain all these tangled findings was put forward by Baddeley [Baddeley 90]. He suggested a two-staged recall, where the first stage (i.e., generation of candidates) occurs in episodic memory. This implies a modification in the knowledge representation that is adopted by two-stage proponents. Jacobs [Jacobs 90] also advocates a two-staged model with episodic memory. In this new approach, the access to the candidates, that is, the generation stage, fetches traces from the episodic memory. These candidates are subsequently entered into a recognition process in semantic memory. This is the model we adopt in order to ‘generate hypotheses’, and it has a key importance in our work.

This model can explain the phenomenon of recalled but non-recognizable words. First, the contexts that assist the two stages of recall are not the same. The independent type of context assists only the generation step, but not the subsequent recognition stage since these are not encoded in the traces. And, these candidate traces are indeed identified in the recognition stage. Hence, recognition performance depends on whether good candidates are generated. In a recognition test, however, the subjects are presented with candidates
directly, that is, the subjects do not generate these candidates. When the candidates are not compatible with what the subjects have in their memory, they can not be recognizable. For example the traces encoding *train* and BLACK interactively, and *white* and BLACK interactively will not be the same. When the subjects learned BLACK in the context of *train*, and subsequently were shown white-BLACK, they may not be able to recognize BLACK. This is because they don't generate their own candidates in a recognition test. If they had generated the train-BLACK candidate, they would then have been able to recognize it.

### 7.5 Only interactive context affects recognition

A memory can be invoked by an associative process. According to the literature, both independent and interactive cues are important for this association. If as Tulving, and later many others claim, namely that this association happens on episodic memory, then it is reasonable that independent context elements are also effective in this process. This is because only subjective (episodic) parts of the memory may include any combination of cues. On the other hand, semantic memory is more objective and shows less variance across individuals. Since recognition happens in semantic memory, only interactive context elements which are connected to core domain context in an interactive, well-formulated manner will obviously be of importance there.

An immediate conclusion is that the processes that utilize episodic memory may be influenced by the independent type of context as well as the interactive type, while the processes that utilize semantic memory are affected only by the interactive type of context elements. This conclusion may have a theoretical significance, but its practical results are more important for our purpose. In tasks where the independent context is often known to be different at encoding-time and retrieval-time, an artificial system taking into account independent context will be negatively effected by this variance in the independent context. For example, in medical diagnosis, it is not very often that a physician encounters a beautiful female patient. If the aim of the physician is not to remember exactly that patient, using the cue 'beautiful' may decrease the probability that he remembers the episode related to his diagnosing this woman upon encountering an ugly male suffering from the same disease.

Fortunately, we have enough control over an AI system and what it does and does not take into consideration. As opposed to a human expert, we can make an artifact forget irrelevant independent cues.

#### 7.5.1 Context influence and ambiguous words

Although the context effects have been demonstrated by experiments in a wide variety of areas in cognitive science, a general cognition theory specifying the various loci of context intervention has not yet been developed.

For example, the intervention of context in language is undeniable, but the following question has long since been raised by psychologists and linguists, yet its answer is still controversial: "Whether context has an effect at the time of accessing lexical information in the mental lexicon, or at a later stage, after all meanings of the ambiguous word have been accessed, regardless of contextual information" [Tabossi 87]. Even though there are
agreements on the existence of context effects in interpreting the correct meaning of a word, there is disagreement as to when context assistance happens. There are mainly two views concerning the point at which context assist the interpretation of an ambiguous word. The context-dependent view claims that “initial lexical activation is affected by prior context so that only the contextually constrained meaning is activated” [Paul 92]. The context-independent view, on the other hand, claims that all meanings are activated and only then does context resolves ambiguous word meanings.

**Context effects in activation:** Most context-dependent views base their account on a semantic network model of memory where words are connected with each other by various relations. In this view, activation of a word’s meaning happens in a ‘long-term memory’. Each word’s meaning is represented in this memory by a collection of semantic features [Tabossi 87]. Supporting this claim, [Kellas 91] further claims that a meaning can be described as a particular constellation of semantic features activated at a given moment within a specific context. Different contexts activate different features of the target word. Hence, at each time only a portion of the features of a word are active. The activated features serve to define the correct meaning of the word. The context that is established prior to the encounter of the word activates a portion of these features. The degree of overlap between these already activated features and presently activated features of the target word determine the significance of context. This approach is supported by some experimental results [Paul 92]. Similarly, as the hypothesis space is often very large in diagnostic domains, we investigate whether the context-driven approach in linguistics also can be applied to diagnostic problem solving. From an AI point of view, we are concerned with how contextual information may limit the search in the problem solving space. In other words, how can context possibly lead to a more focused, and thereby, more efficient, abductive reasoning process.

**Context effects in the selection of a meaning:** Either with or without context effects in activation, a selection process is usually necessary. The need for a selection process is obvious in the case of context-independent activation. In context-driven activation as well, it is usually not possible to limit the activation to a single meaning. Thus, a selection process that follows activation may be needed in both cases.

Which hypotheses do we prefer among a multitude of plausible ones? What is it which makes us think that a particular hypothesis is more worthy of further consideration than its alternatives? These issues have been discussed in chapters 5 and 6. The matter now is to find a relationship between these reasons and the notion of context. Finding this relationship involves an analysis of the types of knowledge that have possibly been accounted as reasons for preferring a particular hypothesis. In this prospect, context is a particular type of knowledge which can be regarded differently from what we refer to as ‘core domain knowledge’, based on its role in solving a problem.

The ability of selecting relevant and ignoring irrelevant information characterizes expertise. An expert, in order to do her job well, should, in various situations be capable of focusing her cognition. The ability to control cognition is proportional to the ability to eliminate irrelevant knowledge. An expert should use all necessary knowledge and only this knowledge. The amount of information and knowledge activated unnecessarily in memory decreases sharply the efficiency of the process, and may even hinder its completion. The activation of knowledge can be limited in various phases. A large amount can be
activated at the outset, and may gradually be limited. Or it can be, to a large degree, limited at the very start.

This problem is similar to asking how contextual information can assist the medical doctor when generating tentative diagnoses. Does a physician activate all disease hypotheses which the (usually scarce) complaints and signs of the patient may, in some way, be related to, or does context constraints the number of diseases to be generated at the first place?

7.6 Summary

The following are the main lessons we derived from theoretical and experimental studies in cognitive psychology. These provide a basis for our context view, which will be presented in the rest of the thesis.

An episodic memory is a collection of specific memories. Context is an important component of such memories which when changed, impeds the recall of a memory.

Various types of context elements may accompany the target phenomenon within an episode. Experiments illustrate that both 'state of the mind', and the environment and target-related characteristics can play the 'context' role.

Another series of experiments and subsequent theories emphasize a distinction between the two ways of encoding contextual knowledge. Baddeley refers to these as independent and interactive encodings.
Chapter 8

Towards a content theory of context

8.1 Introduction

As should be evident from the previous chapters, context has been attributed to different roles and effects in different areas of research. So, it will be difficult to come up with a satisfying general definition. The notion has mostly been used in language and linguistic, and in psychological studies. Dictionary definitions consider context first and foremost with respect to language. It has been defined as the part of the text which precedes and follows a target word and gives this term its meaning. In a broader definition, context is the whole set of conditions accompanying the occurrence of an event.

In a wide range of areas, context has been declared to influence the cognitive processes. In any process, two kinds of information sources available to a reasoner are the contextual and so called core information sources. In the cognitive psychology experiments surveyed in chapter 7, a word to be learned was the core information, while the colour it is printed in was a contextual information.

A thorough context account needs a description of context as well as a specification of the difference between contextual and core sources. It has been a difficult task to place a boundary between contextual and core knowledge. This is probably also the main reason why context does not have a definition that is acceptable in general.

As can be seen from the experimental research in cognitive psychology, context has been used in three different ways. First, especially in early context investigations, context is equated with the environment in which the target is observed. For example the room in which the words are read. Second, context is often concerned with the accompanying features of the target, such as the colour of the print. And finally, context is attributed to the mood and state of the reasoner. For example, the reasoner may be drunk or sad. In all these approaches, context is thought to influence the changes in the reasoner’s learning performance.

Our model of context is based on a combination of these three approaches with some extensions and some refinements. In our definition, context is the collection of a special type of information, namely implicitly assumed facts about a problem or a statement, and the goals and intentions of the reasoner. Context is not the conclusion, nor the problem. Yet, it influences the reasoning from the problem specification to the solution.

8.2 Two views on context

At the start of our context investigation, we decomposed this context investigation task, into two more easily achievable main subtasks. Accordingly, we tried to analyse what the notion of context ‘means’ from two distinct but closely related perspectives. According to these perspectives, context is viewed
• as a knowledge type capturing portions of the focus of attention and
• as a means for triggering a shift in the focus of attention

The former perspective leads to analysing and categorizing contextual knowledge, while the latter imposes an analysis and categorization of contextual effects in different areas and tasks.

We utilize the first view in pursuit of studying the context as a knowledge type and investigate how this knowledge can be categorized. This lead to the construction of a context ontology. Then we turn to the second view and elaborate the correspondence between different types of contextual knowledge and the various subtasks of our diagnostic method. The second view takes, as a starting point, a task (in our case diagnosis) and investigates which type of contextual knowledge or information is used in each stage, and in what way various types of contextual information assist these stages. As such, this research subtask deals with the nature of the relationship between a task and the type of the context it relies on.

We may also say that the first perspective adopts a static view, while the second one is more concerned with dynamic aspects of context, i.e., its use.

8.3 Context as a knowledge type

From an epistemological standpoint, context is a type of knowledge. However, a general rule does not exist to determine which knowledge is contextual, and which is not. This is partly because a knowledge “piece” can be of contextual type only with respect to a reference task. Our top-level task is diagnosis which involves an explanation process. Depending on what type of explanation we are seeking and what needs to be explained, we may identify the contextual knowledge. So, we begin viewing context as a type of knowledge by presenting a criterion with which we can hopefully distinguish contextual knowledge core knowledge. The main principle is that contextual knowledge contributes to the explanans explaining the explanandum in an adequate manner. Contextual knowledge is not a part of what is to be explained (the ‘input’ of the task, in knowledge level terms). It is neither the conclusion (the ‘output’ of the task) to be drawn. Notice, for example, that drug abuse is a contextual knowledge w.r.t. the ‘diagnosis’ task in medical domain, as it may predispose some diseases, but will be a core domain concept when it comes to a ‘prevention’ task. This is because in diagnosis, the overall task of the reasoner is to explain the anomalous observations in terms of a disease concept. In a prevention task, a habit may be the conclusion (output) as it may be what the reasoner is seeking for; the patient should stop drug abuse since it is what predisposes her disease.

If we consider the entire semantic network in a medical domain, an explanation between a disease and an observation is only a part of a rather extensive explanation which may also include an explanation/connection between, for example, the habit “drug abuse” and the disease “endocarditis”. However, in the scope of a diagnostic task, the reasoner is not always required to come up with the whole explanation of how she found the explanation. Her goal, in the course of diagnosis, is to come up with a disease label and its connection to the findings. This goal does not necessarily include the connection between habits and
diseases, though this would be rather common in a ‘prevention’ task. Hence, the basic aim is an explanation chain between a disease and the observed anomaly, not the explanation of how the reasoner generated the target explanation. However, if the reasoner is required to explain how she came to this conclusion, then she should explain in what way contextual knowledge ‘drug abusement’ governed her reasoning. In other words, “Drug abuse” does not appear in the explanation where the disease endocarditis explains the observed fewer, sweating, etc. Both drug abuse and fever are features of the target object of the reasoning process. However, fever is “core” information while drug abusement is target related “contextual” information. Figure 8.1 shows how “enabling conditions” of a disease are considered, in our approach, as a type of contextual knowledge. These consist of characteristics of the target (i.e., the patient) but are different in their role, from the symptoms and signs observed with the target.

The reasoner uses all possible and easily available types of knowledge in order to achieve her goal. Contextual knowledge helps the reasoner to justify her beliefs and reasoning (i.e., explains to herself). So, in our account, contextual knowledge plays a prominent role in justification of plausible reasoning, such as abduction.

![Figure 8.1 The boundary between target-related-context and core cues in medical domain.](image)

8.3.1 Context ontology

This section briefly reviews the criteria used in the literature for distinguishing between various categories of contextual knowledge. We seek a starting point for our categorization of contextual knowledge. There must be some criteria according to which contextual categories can be identified at the highest level.

We summarize the context types discussed in the cognitive psychology experiments under three main groups: attributes of the environment, attributes of the target, and attributes of the subject (i.e., reasoner). Our starting point for a context ontology is that problem solving is a deliberate process of which two basic elements are the agent and the external situation. This implies that we do not study problem solving separate from the problem solver, as has traditionally been done in AI. Recent trends in diverse communities support also
our stance [Nardi 96; Baddeley 90; Leger 93]. For example, Nardi criticizes proponents of
the “situated action” approach for their insistence on the “situation” as the primary deter-
minant of the action [Nardi 96]. She questions “How do we account for variable responses
to the same environment or ‘situation’ without recourse to notions of object and con-
sciousness” (page 89). She considers the difference between three individuals going on a
nature walk; a bird watcher who looks for birds, an entomologist who studies insects as
he walks, and a meteorologist who gazes at clouds. Even though the “situation” is the
same, they behave differently. The difference, according to her, originates from “the sub-
ject’s object”.

FIGURE 8.2 The context ontology, emphasizing the active role of the reasoner

The behaviour of an agent (e.g., a person) is shaped by two important factors: its own per-
sonal characteristics or state of mind, and the characteristics of the problem with which it
is occupied. Our interpretation of the findings from experiments described in the prece-
ding chapter highlights two important factors which help to maintain context typing. The
first is the nature and demands of the process that takes place during learning and remem-
bering. The second factor is the ground facts that happen to exist in a situation. We pro-
pose a distinction between internal and external context types in order to reflect these
factors, where internal context relates to the former, and the external one reflects the latter
factor. The key criteria for this distinction is the deliberate activity of the reasoner. The
aspects regarding the reasoning agent are referred to as internal context. The type of cog-
nitive process behind the reasoning is partly decided by the agent itself. The selection of
the type of process can also be imposed by giving instructions. This was the case in the
experiments related to face recognition. The other alternative is that the subject itself
decides how to process the material. This is particularly important in rather complex tasks
such as problem solving, where the agent essentially replaces possible instructions with its
own choices shaping its own cognitive behaviour. In our account, the agent’s decisions
are shaped by its goals, hypotheses, and expectations, i.e. the internal context.

External context elements originate outside the reasoner, and consist of two distinct
groups of elements: those related to the target and those related to the environment. Exter-
nal context elements basically stay static during problem solving. That is, external context
comprises the static facts in the problem solving situation. For example, in a clinical set-
ting, the agent is the clinician (the reasoner), the target is the patient case, and the environment is the place where the diagnosis and the treatment occur. At the next level of specialization, the internal and external contexts are divided into interactive and independent types - in agreement with Baddeley’s distinction (elaborated in chapter 8). We also differentiate internal-context into interactive and independent, while Baddeley concentrates on external context. That is, he identifies what external context elements are independently or interactively encoded. The top-level ontology is shown in Figure 8.2. Here, the two types of external context, target-related and environment-related, are modelled explicitly.

Notice the difference between our criteria and Hewitt’s regarding the classification of context elements. Our classification takes the agent (i.e. the reasoner) as the main criterion, while Hewitt’s distinction is based on the target as criterion; we are reasoner-centered while Hewitt is target-centered. Hence, in our work, the distinction between internal and external context is relative to the agent, while Hewitt’s distinction is relative to the target. Our target-related context resembles Hewitt’s intrinsic context while our environment-related one resembles his extrinsic context.

Some examples may help to clarify our approach to context typing. First, the independent type: In a medical diagnostic setting, what the patient is wearing is of type target-related context, whereas the colour of the examination room is of type environment-related context. Whether the clinician has short hair is of type independent internal context. Regarding interactive context, for diagnosis and therapy tasks, goals and predictions (hypotheses) of the clinician are of type internal-context, while the pregnancy condition or previous diseases of the patient are target-related. The characteristics (i.e., the conditions and constraints) of the place where the patient-clinician encounter occurs are of type environmental-related context. For example, it is important whether the place is a well-established hospital or an emergency tent in a forest.

8.3.2 Roles of independent and interactive types of external context

Both independent and interactive types of external context constitute cues invoking associations to other entities in the memory (i.e., knowledge base). However their way of associating themselves to other entities are different. In the case of independent context, it is just ‘associated to’ another entity, without clarifying the actual meaning underlying this association. Since the real link between the two entities is often unknown. For example, in medicine, a correspondence between the patient’s gender and a certain disease is known (e., g., statistically), but the underlying reason for this correspondence is not yet known in medical science. This kind of statistical knowledge can be captured within episodic memory by encoding it in the way we called independently. On the other hand, interactive knowledge can be encoded both in episodic and semantic memory, because necessary relations are known. Consequently, independent context can not be utilized in all places where interactive types of context can. Interactive types of context can be utilized both in shallow and deep reasoning, while independent context type can only be used in shallow reasoning. This difference is important for us, because in this work we demonstrate a hybrid diagnostic system (named ConSID-Creek) that commutes between shallow reasoning (case based reasoning) and deep reasoning (explanation based reasoning). During
shallow reasoning, both interactive and independent context can be utilized by the system, while only interactive context is utilized during deep reasoning.

The independent type of context (in fact, independent type of any knowledge) is important, in particular, for weak and open domains.

8.3.3 Qualitative difference between internal and external context

It is important to note that the nature of internal and external context elements are qualitatively different. The existing accounts of context consider one of external and internal context, not both. However, a proper account of context should integrate both internal and external elements of context. External context elements are object domain entities. These are rather static and do not change in the course of the whole process. In other words, these are related to ground facts, and either to the target or the environment. The internal context effects originate from the reasoner. Factors such as the goals and interests of the reasoner are vital for the whole flow of the process chain. Also the conclusions that the reasoner draws for a task may constitute a context for the following task. For example, activated hypotheses are a context for the successor task. The notion of goal has drawn much attention in relation to its role in governing and modifying behaviour [Leake 93; Ram 91; Turner 94] while it has seldom been studied as a context element. However, the human communication community has emphasized the importance of the intention of the speaker as a context type. Also, animal behaviourists recognized what we call “internal context” as a separate type of context pertinent to signal reception, when studying animal communication behaviour [Leger 93].

Internal context may vary across reasoners attempting to solve one and the same problem, however, the external context will be constant for all the reasoners.

8.3.4 Context role in structuring memory

It has been reported that experts utilize contextual information to a much larger extent than novices [Devine 95; Custers 96]. The expertise assessments are usually based on speed and accuracy metrics. Expertise can not be explained only by powerful search algorithms, but the amount and type of information must also be of concern for an understanding of expertise.

Association is considered a law of mind [Peirce, 7.388]. Associative reasoning is given prime importance in the context of weak and open domains. Two types of association are commonly recognized in psychology. One is association by “resemblance” and the other association by “contiguity”. Association by contiguity happens when we remember “husband” upon hearing “wife”, that is when a concept suggests another one. Association by resemblance is the type of association when you remember a similar event upon experiencing a new one. Both types of association are closely related to the organisation of memory. It has been argued that memory organisation is subjective [Barsalou 92]. People find their own ways to link concepts together. Experts do this better in their own specialities than novices. The fact that experts utilize more contextual information and that experts better organize information may suggest that contextual information plays a substantial role in organizing the memory. Tulvings ‘encoding specificity’ hypothesis is sug-
gestive both for the subjectivity of memory organisation and for the existence of a relation between context and memory organisation.

The role of context as a knowledge type is that it contributes to clustering, partitioning, and organizing the knowledge of the world which a person may possess. It integrates in a sense, relativity and subjectivity into knowledge. The features of a piano related to "going to concert" (e.g., sound) may be clustered together, while its features related to being a movable object (e.g., size, weight, shape) may also be clustered together and therefore can be accessed quickly when other members of the group are given.

Custers [Custers 96] claims that the medical experience of an expert is organised for each disease as a package with three components, one of which is the enabling context for the disease. He further claims that experts are superior to novices because of their ability to make better associations to these packages.

8.4 Context as a means of focusing attention

Researchers from various communities studying fields such as the generation of explanations, user modeling in communication, natural language processing, and human-computer interaction have all acknowledged the importance of perspective ([Suthers 91]; [Mittal 93]; [Pichert 77]; [Lester 91]). For example, regarding text comprehension, Pichert and Anderson argue that "if, for whatever reason, people take divergent perspectives on a text, the relative significance of text elements will change". A presumption underlying the notion of perspective is that concepts are represented as a set of features, and not all the features of an item become activated each time the item is activated. Each episode that contributes to the encoding of a concept refers only to a subset of these features. The task context in which a concept is presented determines that particular subset of encoded features. The importance of taking into account the type of task to be accomplished has recently been recognized ([McCoy 89]; [Edmondson 93]; [Cahour 93]).

Our account of the relation between the notions of context and perspective establishes a chain that starts at the goal. It is commonly agreed that goals are of crucial importance for contextualized problem solving and learning (e.g. [Ram 91]; [Leake93]; [Bogdan 94]). The goal invokes a task, which in turn determines the perspective to be taken (see figure 8.3). The perspective pinpoints where the attention needs to be focused. A coherent set of aspects constitutes a focus of attention. For example, the feature weight of the piano is not activated with respect to a going-to-piano-concert event, but it is activated in other task contexts, such as when the piano is to be moved. Hence, the goal determines which aspects of a concept will currently be processed, including both non-contextual and contextual features and cues.

Problems may arise when the perspective is not taken into consideration explicitly. An example may be taxonomies combining multiple hierarchies. Biology is a research area presenting such hierarchies. Suppose that some biologists, being interested in an animal's habitat, may partition the 'animal' class into subclasses 'sea animal' and 'land animal'. Some others, more interested in the dietary habits of the animals, may partition 'animal' class into 'carnivorous animal' and 'herbivorous animal' subclasses. When both hierarchies are integrated in the same taxonomy, 'whale' for example should be represented as a
carnivorous animal as well as a sea animal. On a question related to the stomach tooth, whale should be considered a carnivorous animal rather than a sea animal. Thus, the knowledge that is currently relevant is directly related to the perspective that is currently adopted. The knowledge that stomach tooth may be interesting in the perspective of biologists studying the dietary habits would serve a bias toward carnivorous-herbivorous classification.

![Diagram](image)

**FIGURE 8.3** The explanation of how internal context serves as a focusing mechanism.

In our system, the task perspective (see figure 8.3) helps to determine the focal regions both of specific (episodic) and general domain knowledge. The focus of attention is switched when the goal of a reasoner changes. A change in the goal signals that a new task is to be performed. Attention parallels the epistemological needs of the reasoner when engaged in a dedicated task. Therefore, a task is the natural medium for discussing shifts in the focus of attention.

As a result, internal context has a role in determining which features, including both external context and core cues, that are relevant in a particular situation, and how relevant they are. In other words, internal interactive context determines the interactivity of external-context related features as well as non-context ones. Regarding external context elements, this influence is particularly important for tasks in which available information is incomplete, that is, core retrieval cues are not sufficient to access target concepts unambiguously. In such cases, context cues may strongly influence the degree of match between the present material and a past encoded material. Hence, internal interactive context is the main context type that imposes a perspective, which in turn determines the characteristics of both encoding and retrieval.

There has been an ongoing discussion [Brezillon 93] in the ‘context community’ about whether context captures the state of the mind or the situation. In our view, it captures both. A very important issue is to identify and explicate the link between the two components that are pertinent to the context phenomenon. Internal context reflects the state of the mind (of the reasoner) while external context reflects the situation (including both the target and the environment related elements). Internal context imposes the focus of attention, which is captured in terms of core domain concepts, cases, and the other type of contextual knowledge such as external context (see figure 8.4).
This relationship between the "state of the mind" context and "situational" context matches our concerns related to formalizing the link between the "content" and the "process" mentioned in preceding chapters.

8.5 Summary

We presented a high level context ontology. The ontology is based on

\( i. \) a distinction between internal and external types of context, and

\( ii. \) a distinction between the type of encoding

We investigate context from two points of view. According to the first view, we are concerned with context as a knowledge type. In the other view, context is a means for shifting the focus of attention. These two views are not independent.

Based on the study reviewed in chapter 7, the generic context ontology depicted in figure 8.2 is developed.
PART III

An architecture for context-sensitive diagnostic problem solving
Chapter 9

A context-sensitive iterative diagnostic model

9.1 Introduction

This chapter is a transition from a cognitive psychology study of context, philosophical research on inference (abductive inference in particular), research on diagnostic expertise in medical reasoning, and instructional research towards a view of medical diagnosis in artificial intelligence.

In the previous chapters the thesis have described a context-sensitive abductive inference pattern having an evaluative nature that supports the assessment of plausibility of inferred conclusions. Then a methodology and a high-level ontology is proposed to guide the acquisition of contextual knowledge for diagnosis. Also, a mechanism for establishing the focus of attention is presented. Combining these results from the previous chapters, this chapter presents the principles of a context-sensitive iterative approach that provides the grounds for developing diagnostic knowledge-based systems. Chapter 10 deals with the low-level details of the focus mechanism while Chapter 11 concentrates on methods that implement the context-sensitive abductive pattern as well as other inferences, though less detailed.

Diagnosis has not been considered solely as a classification task as has traditionally been done. It involves a planning task as well. So, diagnosis involves two rather generic tasks, both well-known in AI, namely generating an explanatory classification (i.e., a fault/disorder label), and planning the activity of gathering missing information that is required for solving the diagnostic problem.

In our conceptual model, the main task of diagnosis is accomplished by employing a “generate-and-test” method that decomposes the diagnosis task into two more easily accomplished subtasks: generation of explanatory hypotheses, and testing of these. This sketchy model, illustrating the relationship between the diagnosis task and the subtasks of generation-of-hypotheses and testing-of-hypotheses, is analogous to the model proposed by Peirce for conceptualization of scientific inquiry (see chapters 4 and 5).

In this chapter we propose a model of diagnosis which illustrates the roles of classification and planning tasks, and the relationship of these tasks to the generate-and-test method. We also show how the generate-and-test method can be conceptualized as an iterative application of the triple of abductive, predictive and inductive inferences.

In our view, the use of contextual knowledge and information comprises a key principle with respect to the issue of improving both the efficiency of the diagnostic process and the quality of the results. The general role of context is to provide a means for focusing attentions on the relevant aspects and features related to a problem. This is particularly impor-
tant for real world tasks where the domain knowledge is rather rich, the problems are typically ill-structured, and the domain theories are weak. So, we suggest that a knowledge-based system should be contextualized in order to be effective (related to focusing), and to provide solutions with high quality (related to relevance). In our context-sensitive iterative diagnostic approach, which we call ConSID, context plays a central role in triggering the shifts in the focus of attention as well as in capturing the focus of attention.

How can a knowledge based system be contextualized? Our approach conceptualizes cognitive processes as divided into stages which are captured as subtasks in the knowledge level model of an application problem. Each subtask is accomplished in order to achieve a goal established by the reasoner. Each cognitive or physical task needs to focus only on a portion - the portion relevant for the current goal - of cognitive (i.e., in the mental world) and physical (i.e., in the real-world) resources.

9.2 A context-sensitive iterative approach to diagnostic problem solving

Context-sensitive iterative diagnosis (ConSID) approach conceptualizes diagnosis as a context sensitive process integrating reasoning and action. The main principles on which the ConSID approach is based can be classified as follows:

1. diagnosis is realized by an iterative application of the abduction-prediction-induction cycle (see figure 9.1).
2. diagnosis is a combination of explanation and planning
3. the problem solving process is goal-driven, and a shift in the focus of attention is triggered by a change in the active goal.
4. the focus of attention is captured in terms of core and contextual entities
5. pattern recognition is an important component of diagnosis at the expert level

![Generate-and-test method](image)

**FIGURE 9.1** The Generate&test method can be conceptualized as an iterative application of abduction-prediction-induction cycle.

This approach to diagnosis is context-sensitive in two ways. First, the reasoning line is goal-driven (thus, internal-context driven). Second, it takes into consideration the contextual aspects of the grounded facts (i.e., external context) which allow the focus of attention to be captured in terms of contextual elements in addition to core domain concepts.
9.3 The view of a diagnostic label as an explanation

The hypothesis generation task involves abductive inference where a set of hypothetical explanations of the anomalous observations are tentatively provided.

As we mentioned before, an explanation, according to common understanding, has two components: explanandum and explanans. In our view, a third factor which deserves to be counted as a component is the explanatory relationship linking explanans to explanandum. According to the ConSID approach, a diagnostic explanation should

- explain the anomalous findings,
- explain them in terms of the desired explanatory relations, i.e., the relations which are determined in advance, for linking explanans and explanandum
- agree with the contextual features pertinent to the situation. This condition ensures that the explanans (the fault) is consistent with the contextual factors of the faulty situation. For example, in the medical domain the risk factors comprises a part of contextual features. This type of contextual information does not directly explain the observations, but it does, in a way, imply the fault.

In medical diagnostic tasks, the explanans, that is the explanatory hypothesis, is a diagnosis label, a disease. In this domain, an explanandum is a set of findings such as complaints, measurements and signs. An explanatory hypothesis (explanans) in this domain is linked to the findings it explains (explanandum) by various causal, associational and structural relations.

9.4 Iterative application of abduction-prediction-induction cycle

A diagnostic label is established by employing a strategy consisting of the generation of tentative explanatory hypotheses and the testing of these hypotheses on the basis of evidence which is incrementally gathered by performing actions in the physical world.

So, diagnosis is realized by a generate-and-test method which decomposes this task into two main subtasks, the first of which is the abductive generation of hypothetical diagnoses, and the other is the testing of these hypotheses by collectively employing predictive and inductive reasoning lines. This implies an iterative process because the hypotheses are generated with scarce information and through an abductive process which produces fallible results. As such, the hypotheses may fail to explain the newly gathered evidence. Consequently, a new hypotheses generation process may need to be invoked. This means a new cycle of hypotheses generation and of testing.

This model is similar to the model of scientific inquiry which consists of a sequential application of abductive, deductive and inductive inferences proposed by Peirce. However, an important difference is that we propose an iterative and cyclic application of this triple. We argue also that inductive inference is not corrective, but needs another abductive step which is corrective. In our view, the corrective step in the whole diagnostic process is the abductive step. However, an abductive step is not able to verify a generated
hypothesis by itself. The abductive step includes, however, a justification of why the reasoner generated a particular hypothesis.

9.5 Internal context utilization

Internal context elements include the goal of the reasoner, the reasoner’s current hypotheses, the hypotheses which were previously considered and rejected (after testing), and the strategy the reasoner chooses for testing hypotheses (e.g., confirmation or elimination of a hypothesis). All these are interactive-internal type of context (see figure 8.2).

The goal of the reasoner is related to its background, interests, motivations, as also proposed by [Leake 93] who points to importance of the goals of the reasoner in everyday-life explanations.

9.5.1 The shift of focus is goal-driven

Reasoning can be viewed as intentionally constrained information processing. The underlying mechanism for this constraint is goal dependent. In our system, the active goals of the problem solver determine the immediate task to be accomplished. A task, in turn is realized by a method that invokes either cognitive or physical actions. So, the reasoning line can be conceptualized as a chain of goal, task, and method sequences. For each goal there may exist several subgoals that the problem solver should, in turn, adopt. These subgoals are related to each other. For example, complex goals are decomposed into smaller, more easily and quickly achievable ones.

For a specific task, not all goal sequences can be valid. For example, the testing of a hypothesis may not precede the generation of that hypothesis. The valid goal sequences for a diagnostic task can be sketched as in figure 9.2, where g-determine-anomalies and g-explain-anomalies are the immediate subgoals of g-establish-diagnose, and the achievement of g-determine-anomalies precedes the achievement of g-explain-anomalies.

![Diagram](image)

FIGURE 9.2 Valid goal sequences in a diagnostic problem solving
The question to be addressed now is to put forth how the relevance and focus can be realized. At each point in reasoning, only a small portion of the huge knowledge base is relevant, and only that relevant part should be brought to the working memory. If we are able to pinpoint this relevant portion then we can focus on that portion. So, the notions of focus and relevance are rather interlinked. As our interests and needs change, what was relevant before the change may be irrelevant, while new aspects may become relevant. So, if we want to model the mechanism underlying a trigger of shift of focus, we have to be able to locate the points where and when a shift is required.

In our model, the points where the focus shifts correspond to the points where the goal of the system changes. That is, we adopt a goal-driven shift of focus.

This view parallels an emphasis placed by Barrows on the link between the ‘process’ and the ‘content’ with regard to the issue of finding the best way of teaching medical expertise to medicine students (see section 2.5). The link between a goal-task-method plane to a domain knowledge plane happens through tasks in our system. A task specifies the epistemological needs by referring to the domain knowledge.

So, in order to model focus shift in a diagnostic task, we need to identify the possible goal changes in this domain, and correspondingly the points where new tasks (or subtasks) are invoked. Thus, the tasks comprise possible loci where a shift in focus of attention may be natural at the knowledge level.

In the views of the knowledge acquisition and reusability communities (e.g., [Chandrasekaran 92; Breuker 94]), a goal is attached to a task whose accomplishment entails achievement of that goal. In our approach, each task imposes a perspective, as is shown in figure 8.3, which determines the type of hypotheses that will be generated and the aspects that can be relevant for the problem.

The reasoner sets up new goals dynamically, either via subgoaling, which means to continue in the same reasoning line, or by interrupting the current task and generating a new goal. The former alternative happens if everything goes as the reasoner predicted and expected, and the latter when expectations are failed, that is, when the reasoner is on the wrong track (e.g., wrong hypothesis is pursued).

9.5.2 Generated hypotheses provide further focus

Once some hypotheses are formulated, these guide how the problem solving behaviour of the diagnostician proceeds. The currently considered hypotheses determine the types of checks and measurements to be employed. The set of hypotheses determines the strategy for information acquisition. For example, if one of the hypotheses is far more promising than the others, it becomes a candidate for confirmation. After the strategy is decided, the hypothesis again serves as a focusing mechanism, since information is gathered that presumably contributes to the confirmation of that particular hypothesis.

9.6 External context utilization

An expert, after solving a problem, learns an episode where the involved fault process is explained. A fault process is the understanding of how a particular fault has occurred in
the situation. The fault processes that occur in the discourse of the considered diagnostic domain need to be analysed and transformed into fault models at the knowledge level. It is these fault models which the reasoning process uses towards the goal of finding the underlying fault, from a set of anomalous observations. A knowledge level model, therefore, should be able to appropriately represent the fault models. So, the task of determining the types of knowledge to be included in the knowledge base can not be isolated from the task of modeling the fault processes.

The relationship between the conceptualization of fault models and the diagnostic reasoning process is a reflection of the link between the 'content' and the 'process'.

Context is a special type of knowledge which is a necessary component in order to understand the faults in a real-world setting. Therefore, before determining the knowledge types at the knowledge level, a contextualized model of fault processes in the domain should be constructed.

**FIGURE 9.3** Enabling context factors 'enable' the faults, and 'outcome modifying context' influences the outcome. A problem is expressed in terms of enabling conditions, the fault, manifestations, and outcome modifying context. It is these 'problem expressions' that the reasoner stores in its memory. An outcome consists of a set of manifestations. Only the manifestations of outcome-3 and outcome-4 are illustrated in the figure. Manifestations within the dark region are the ones available in the new situation at the moment. The ones written with italics are common to both outcome-3 and outcome-4.

Figure 9.3 illustrates our general model of a disorder process. This model differs from most traditional ones as it integrates contextual aspects thoroughly into the model. It elicits the interactions between two main types of contextual knowledge, and the incidence of a fault and the way the fault manifests itself. The *enabling context* and the *outcome modifying context* play two important roles during the fault process - and during the explanation process as well. An explanation consists of a chain starting from enabling conditions, including the fault, and ending with the manifestations (e.g., "high fever") of the fault. The collection of all manifestations is the outcome. If we divide the life-span of the fault
process into 'before' and 'after' the occurrence of the fault, we may roughly highlight two roles of contextual factors played respectively by enabling and outcome-modifying contexts. The former is effective in the first stage, and the latter in the second stage of the fault's life span (see figure 9.3).

The influences of the former on the fault happens temporally 'before' the fault arises. The latter influences the way the fault manifests itself in terms of findings, i.e., the outcome (see figure 9.8).

A correspondence exists between this model and the way we consider the use of context in the reasoning process involved in finding the fault from observed anomalies of the object (e.g., a car, a patient, or an electronic board) under consideration. Figure 9.4 illustrates how this model of the fault process is reflected in our knowledge level model of context.

FIGURE 9.4 Context effects on the diagnostic process. The figure illustrates that various types of contextual elements play different roles. The goal determines the types of hypotheses that are relevant for the current task. As such, it constraints the search in the hypothesis space. Parts of target-related context (ctx-1) further narrows the set of hypotheses that have been found relevant for the current goal. The other part (ctx-2) of the target-related context influences the way a fault manifests itself. That is, it constraints the outcome. So, in general, the various contextual elements collectively narrow down the various parts of the domain knowledge that are used for the diagnostic reasoning process.

According to our methodology for modeling context, as formulated in chapter 6, we study external context in the following sequence: First, locate the subprocesses where context effects can be investigated at the levels from which we can refer to our context ontology (as defined in chapter 8). Second, analyse the nature of the interaction between the context element and the subprocesses they are utilized in. Finally, identify the type of context element in the ontology that is utilized in each subprocess.
9.7 Loci of context effects in diagnosis

Accounts of context have usually referred to context effects either on input, output, or on the memory (i.e., knowledge base). Figure 9.5 illustrates how a process can be visualized in terms of input, output, and memory. Note, for example, that classical contextual accounts of perception consider either a restriction on selection of input signals, or a limitation in the storage of the output of the pattern recognizer. On the other hand, in language, context effects have been studied in connection with the selection of word-meaning from memory.

![Figure 9.5 Context may interact with input, output, or memory.](image)

In diagnostic problem solving, there are several loci where context has apparent influences. We identify three main loci for context assistance corresponding to three different stages in a diagnostic process. On the basis of our approach to the logic of diagnostic inquiry we propose context influences in hypothesis generation, information-gathering strategy selection, and making a plan for gathering information (see figure 9.6).

![Figure 9.6 Stages of diagnosis which render effects of the external-context](image)

Based on the nature of the subprocesses, we can recognize the types of contextual elements that underlie the context assistance in each subprocess, in a specific domain. In each subprocess, different types of context elements are found to be functional. In the next chapter we elucidate the types of external context elements which influence each of these loci, in our example domain, namely medicine.

9.7.3 Context assistance in hypotheses generation

In the ConSID approach, the experiences, and the ability to find similarities between a current problem and the previous experiences, are the basis for formulating good hypotheses. As Hatano and Inagaki [Hatano 92] assert, “the construction of models by abduction is nearly impossible when one does not possess a usable old model. People tend to be at a loss when they cannot think of any model that can explain the observed unexpected events” (page 123). According to them, it is hardly possible that we can formulate hypothe-
eses which explain the surprising facts “unless we possess a familiar prototype” that we
can take as the starting point. The notion of ‘familiarity’ can also be found in Peirce’s
writings. He relates this notion to a human’s ‘abductive talent’ or instinct.
We will now elaborate on the relationship between context utilization and hypothesis for-
mulation by recalling previous episodes. An episode consists of a combination of a set of
contextual and core cues. Some of the contextual cues are interactively encoded with core
concepts and explanations, and some others are independent. Contextual elements from
both groups are used in associating past experiences to the new one.
We now analyse the nature of the relationship between the hypothesis formation process
and the contextual information that assists this process. This endeavour will shed light on
the matter of how context assists hypothesis formulation. This issue is rather important
with regard to the efficiency of abductive inference, since its efficiency is proportional
with the number of evoked hypotheses, and since we predict that contextual information
should have a constraining effect on the formulation of hypotheses. The smaller the
number of activated hypotheses, the more effective the abductive inference will be, under
the condition that the hypotheses set includes the best hypotheses. So, context should not
constrain the activation for any price: the hypothesis set should cover the most relevant
hypotheses.
We summarize our claim of the nature of the interaction between the context and activa-
tion process as a chart, illustrated in figure 9.7.
Evoking is based on how distinct the current case is and how similar it is to the past expe-
riences. Increased distinctiveness and a constrained similarity judgement collectively
facilitate ‘association by resemblance’ [Peirce 58]. The underlying context effects in
hypothesis formation are the following:
- context comprises a reference point for judgment of similarity between ‘core’ cues pre-
sented in the new experience and the past experiences
- context increases the distinctiveness of the available information
- a good association can be accounted for by the notions of similarity and distinctiveness
- a good association entails enhanced recall
- enhanced recall implies good recognition as well, in the sense that, the likelihood that
the recalled episodes’ is similar to the current situation/case, with regard to the hypoth-
esis, is high.
- a hypothesis dictated by a similar past episode may be more likely to explain the sur-
prising facts. The idea is that the hypotheses formulated by retrieving a similar episode
is plausible in the current situation since it has once been verified to be true in similar
situations.
According to Dreyfus and Dreyfus [Dreyfus 86] "... when a similar pattern is seen, the memory is triggered and the diagnosis comes to mind" (p 31). They explain the mental process of an expert by "experience-based holistic recognition of similarity produces deep situational understanding" (p 32). In our account, what makes an experience whole is its encoding of contextual elements, which fill the gaps between core cues. Consideration of an experience as a "whole" opens the possibility for an 'overall similarity' judgment. It is exactly the adjective 'overall' which makes the similarity notion practically useful.

Context as a reference point for similarity. Internal context, by imposing a perspective, determines which of the available cues are salient for the current purpose. That is, a goal modifies the relevance of pieces of available information. This is a top-down imposition of perspective. McGuire and McGuire [McGuire 81] investigated how classification is dependent on context. They showed that adjectives individuals used in order to describe themselves depend on the context in which they imagined themselves. For example, a tall, male, contact-lens wearing basketball player described himself as tall relative to his classmates, as a contact-lens wearer relative to his team-mates, as a basketball player relative to other athletes, and as male relative to his family. Even if he, all the time knew about his tallness, contact-lens-wearerness, and maleness, he found only one of these relevant in
each different context. Being a family member or a basket player provided a reference point for comparing himself with others. This implies that he judged his similarity to and distinctiveness from others in the context of a certain goal.

On the other hand, external context also creates a somewhat similar effect. However, this will be more bottom-up and implicit in the sense that sometimes a single external context element and other times a collection of such elements may imply a certain reference point. To give an example from medicine, the contextual information of 'drug abusing' and 'unemployment' (thus, possibly bad nutrition) may collectively be associated to immunosupression, which may constitute a reference point for the search for hypotheses, even if it is not explicitly expressed. However the combination of drug abuse and unemployment, at least for medical experts who experienced the relationship of these two factors with diseases related to immunosupression, will constitute a bias in their search of the hypothesis space.

**Distinctiveness:** To be distinct means to be distinguishable, which in turn, means to be more easily accessible and recognizable. If something is not common, that is, if it has properties that make it different from similar concepts, it can easily be picked out among many others in the same set. For example, rare symptoms immediately trigger particular disease hypotheses, while common ones such as fever, or sweating can be associated with many diseases.

Distinctiveness can be increased by contextual information in two ways, both affecting the strength of the total association:

1. By *quantitatively* increasing the total information
2. By *qualitatively* modifying the salience of core cues. For example "the eyes are filled with tears" is a rather salient symptom in *spring time*, since allergic conditions arise then. In winter time, however it is not that obvious whether it may indicate something special.

The first item highlights the *additive* role of the contextual features where contextual features as independent indices would influence the total relevance of information provided by the new case as a whole. The second item implies the *synergistic* effects of the contextual features on the relevance of the non-contextual features, that is, findings.

A highly distinct phenomenon would have similarity with a small number of other phenomena, while less distinct ones may share a lot with very many others.

In our context terminology, both independent and interactive type of contexts can increase distinctiveness of an episode. However, in the same way as the similarity between contextual information in learning and remembering time illustrates a positive effect, the reverse, that is, variance in contextual information in learning and remembering time may also have a detrimental effect. Therefore, we select carefully which independent context element to include in the system.

Saying that something is similar to another thing does not mean much unless one states in what respect they are similar. As claimed by cognitive psychologists, similarity requires a "frame of reference" [Medin 93]. The most accepted method for measuring similarity is based on matching aspects or features. According to Medin et al., "The only way to make similarity nonarbitrary is to constraint the predicates that apply or enter into computation
of similarity" (p 255). By providing a reference point for a similarity assessment, context supports a constrained search in memory. The search is constrained since the similarity is assessed in a certain context, not arbitrarily and extensively. This ensures a 'relevant' and focused similarity assessment.

The hypotheses evoked in such a way are elaborated in the following recognition stage, which consists of both elaboration and selection phases.

9.7.4 Context in testing a hypothesis
The process that follows hypothesis generation shows differences in language understanding and in scientific inquiry or diagnosis. Testing in the latter group happens in the real world. In scientific inquiry, controlled experiments are done, and in diagnosis some measurements and examinations are performed. Conversely, in language understanding, a selected meaning is tested against the rest of the text that follows. Nevertheless, there is something common to all tasks: they include a process of making predictions about the consequences of to-be-tested hypotheses, and a process of assessing whether the expectations are fulfilled in the case.

9.7.4.1 Selection of the strategy for gathering information
The question taken up in this section is 'how to use the set of formulated hypotheses further'. The matter of what to do with this set of hypotheses is dependent on what intention the reasoner has.

Each evoked hypothesis has a certain confidence value. However, it is not sure that these values are correct, especially because such plausible reasoning is made on the basis of partial information. More certainty needs more information; this is gained with the process following hypothesis selection. Even the most confident hypotheses can be rendered implausible in the light of new information. This is why abductive inference is fallible.

Diagnostic reasoning employs various strategies for gathering information. The strategy to be employed is selected dynamically, according to the nature of the activated hypothesis set. There are mainly three strategies in diagnostic reasoning:

- confirming a hypothesis,
- eliminating a hypothesis, and
- differentiating between hypotheses.

Based on the relationships between hypotheses in the set, with respect to the criteria we just mentioned, one of these strategies is applied. Selection of the strategy reflects the priority criteria. These priority criteria may be based on utility and policy considerations. For example, what happens if a life-threatening, but possibly less likely hypothesis is not considered. Or, what do we gain or lose from considering the less likely hypothesis first. The policy adopted is dependent on the location where the diagnosis happens. For example, concerns of the medical expert will be different in an ambulance, in a war zone, and in a well established hospital.

Of the above listed three strategies, the first one is generally used when one of the hypotheses in the set is much more plausible than all the others. The second can be used if a less
plausible but highly life-threatening hypothesis exists in the set. The last one is employed when none of the hypotheses is distinguished with respect to being either very plausible or highly life-threatening. So, the selection of a strategy for gathering information is dependent on,

1. the characteristic of each hypothesis with respect to its plausibility
2. the relative plausibility and urgency

From our context perspective, the relations between the hypotheses in the hypotheses set serve as a context for strategy selection. Policy and utility concerns also constitute a context for strategy selection.

9.7.4.2 Contextual role in prediction
Peirce’s pragmatical principle requires that a hypothesis (or its meaning in his view) should be able to be tested. That is, its ‘practical conditions’ should be testable. Implications of the pragmatical principle are rather demanding; first and foremost these practical conditionals must be finite and determinable. However, determining all the ramifications of a concept is neither possible, nor necessary. Contrary to Peirce’s view, we do not view the prediction of consequences as simply deduction. There must be some mechanisms in order to determine the practical conditions, and only the conditions that are essentially needed for refuting or confirming a hypothesis. These conditions constitute a small portion of all possible consequences. What makes the role of context significant is that, in each context, only some of these conditions are usually relevant. Furthermore, sometimes a practical condition follows from the hypothetical concept only together with another concept. For example, the outcome of a disease presents itself in various ways across different patient contexts. The way a disease develops after its onset depends on other factors related to the situation of the patient; outcome is context dependent. For example, if the physician made the tentative diagnosis of endocarditis, then she expects low blood pressure as a consequence. But if the patient is suffering from chronic high-blood-pressure then the physician will not expect a low value, and a high blood pressure value will not reduce her suspicion of endocarditis. Such context which we have called “outcome modifying context” in the preceding sections, modifies the outcome of the disease. So, in weak domains (which diagnostic domains are) we cannot deduce from a hypothesis its true consequences. The deduction of consequences should be a contextualized process, and this we call prediction instead of deduction in order to differentiate it from traditional deduction. So, Peirce’s ‘deduction’ of consequences of a hypothesis is not as straightforward as he implies.

9.7.4.3 Context and information gathering - planning
Suppose that the tentative hypothesis for the cause of a car’s problem is ‘engine does not start’. In order to confirm this certain checks should be done. However, if ‘engine does not start’ is not the only plausible hypothesis, a strategy to discriminate between ‘engine does not fire’ and ‘jammed carburettor valve’ is adopted. The information required in order to distinguish between these hypotheses will possibly be different from what is needed in order to confirm ‘engine does not fire’. So, the information gathering strategy constitutes a
context for the planning of information gathering activities. This context determines the relevant information required to be gathered. A second type of context captures the preconditions related to the applicability of various actions that gather the required piece of information. These preconditions may be of various kinds. Some are related to the target, others to the environmental conditions. In the preceding chapters we have referred to the former type of context, as target-related context, and the latter type as environment-related context. The environment related context includes the physical resources required by the actions, such as instruments, and medical.

We may say that the contextual factors, more specifically external-context related context factors capture requirements related to the resources needed for either cognitive or physical activities of the reasoner. In the explanation task the resources are of epistemological type, while the planning task is facilitated by considering the physical resources.

### 9.7.5 Evaluative power of contextual knowledge

As we have mentioned in relation to abductive inference, evaluation of the reasoner’s reasoning is important for efficiency and relevance aspects of abductive reasoning. We emphasized also that evaluation of the hypothesis alone is not sufficient. Evaluation and generation should be integrated, as has also been proposed by [Josephson 94] and [Leake 92].

The question we try to answer in this section is a somewhat more abstract form of the question ‘how is it that only a limited number of diseases comes to the mind of a medical expert, based on rather scarce information which could, in fact, be related to many diseases’. There must be an implicit evaluation mechanism which eliminates some hypotheses from being activated in the first place.

We integrate evaluation into generation, or in other words, we apply a distributed evaluation by way of utilizing contextual knowledge. In this approach, a kind of implicit evaluation in diagnostic reasoning is made when formulating/evoking hypotheses. By using contextual criteria we prevent all possible hypotheses from being evoked. This happens because hypotheses are evoked in a particular context. That is, the hypothesis activation is conditioned on coherence with the context. One implication of this is that hypotheses are attempted to explain the surprising facts, but only under the condition that they also agree with the contextual facts and considerations. As these contextual concerns are taken into consideration as early as in the phase of evoking hypotheses, the hypotheses evoked can be assumed to have gone through an evaluative filter. The reasoner does not even access the hypotheses which are at odds with the contextual factors.

The activated hypotheses are subsequently evaluated and assigned a plausibility value. The evaluation in this step, takes common sense knowledge into consideration. The plausibility value of each hypothesis reflects also its plausibility as to contextual knowledge. Some implausible hypotheses are eliminated in this phase. For example, the activated hypothesis ‘pregnancy’ may turn out implausible in deeper investigation if the patient is ‘male’.

In the strategy selection stage we consider the hypotheses in the hypothesis pool all together. The activated hypothesis pool is a component of internal context in our model. This gives the possibility to compare alternative hypotheses. On the basis of some strat-
egy-selection criteria, we select a strategy to continue with. The strategy-selection criteria is shaped by utility/policy considerations. An example criteria is that if there is a hypothesis calling for emergent interference then the “eliminate hypothesis” strategy is adopted. If there is a hypothesis that seems to have a much higher plausibility value than that of all other alternatives, then prefer to “confirm” it. Otherwise select the “hypothesis discrimination” strategy. In this work we consider only the first one involving a selection of a hypothesis for testing. Testing is the last step of evaluation. These hypotheses may be eliminated or confirmed.

9.8 Why does agreement with context imply plausibility

The basic fact is that some enabling conditions may gradually make some changes in the object under diagnostic consideration, which, in turn, enables the occurrence of a fault. The fault itself will subsequently make some changes on the object which will present itself in the form of symptoms and signs, i.e., the outcome (see figure 9.8).

It is tempting to build a connection between this process and the reasoning process where a diagnostician formulates fault hypotheses. The strength of the relation between the enabling context and the fault may vary across faults. Enabling context may be, for some fault types, a necessary condition for the incidence of the fault, and in some other cases it may only slightly increase the likelihood of the fault, compared to absence of these contextual factors. Therefore, when we know that the enabling conditions of a phenomenon do exist we may expect the phenomena itself also be true. Similarly, if we knew that some consequences that follow from a phenomena are true we may, through association, think that the phenomenon itself is also true. So, both consequences and the enabling conditions of a phenomenon are the familiarity sources that lead us to the guess of a phenomenon.

Peirce asserted that the best hypothesis must render surprising facts likely and must be likely in itself.

We argue in this section that the likelihood of a fault in general is different from its likelihood in particular contexts.

Based on Peirce’s statements, it would not be irrational to say that a “good” hypothesis must be likely in itself and must be likely to explain the surprising facts. We think that a hypothesis is plausible when we think it seems likely for explaining the facts. There are two questions that immediately arises:

- what does it mean that a fault “renders the surprising observations likely”?
- what does it mean that a fault “is likely in itself”?

In our interpretation, the first part of the statement “H explains D in the context C_{current}” (which has been elaborated in sections 5.7 and 5.8), namely “H explains D” has a close relation with the first question while the second part, “in context C” to the second question above.
What does it mean that a fault "renders the surprising observations likely"? What should a diagnostician take into consideration in order for this condition to be satisfied? The answer may be that the hypothesis formulation process should formulate hypotheses that have causal associations with the surprising facts.

However, this is not that straightforward, because there exists no knowledge that simply conveys that something like "in case of fault A this and that consequence follows", which when applied gives a perfect result in determining whether fault A is the correct diagnosis. In other words, the relationship between a fault and its outcome is not one that always brings you to the same outcome from a certain fault. There are ‘outcome modifying context’ factors that shape the outcome.

The triplet of enabling-condition, fault and outcome may be seen as a local knowledge network where invoking some portions leads to invoking other parts. When a piece of critical information on the local network is provided, the partial picture becomes complete and the whole network becomes apparent. We may conceptualize the process of developing a fault as a chain connecting the nodes of ‘enabling-conditions’, ‘fault’ and ‘outcome’ (see figure 9.9).
An interpretation of hypothesis formulation is that the diagnostician is given parts of the information describing the outcome and is required to find the ‘fault’ node. This is the classical approach which has commonly been adopted in computational models of diagnosis. An alternative is that the expert may also be given parts of the enabling conditions as well. Then, finding the node in the middle, namely the fault will be easier as the number of possible fault nodes connecting both the given enabling conditions and the outcome may dramatically be reduced. This can be considered as the part of the mechanism underlying, in Peirce’s word, a “skilled guess”.

As Arthur C. Doyle writes,

“From a drop of water,” said the writer, “a logician could infer the possibility of an Atlantic or a Niagara without having seen or heard of one or the other. So all life is a great chain, the nature of which is known whenever we are shown a single link of it. (italics are ours)

9.8.6 Distributed evaluation

Related to evaluation aspect of abductive inferences, our approach proposes a distributed evaluation of hypotheses. The evaluation is first and foremost distributed between hypothesis generation and hypothesis testing stages.

A) Evaluation in hypothesis generation: This represents evaluation during abduction. This concerns evaluation of a hypothesis on the basis of incomplete information, that is information available so far. It happens in two stages of abduction.

Evaluation in hypothesis formulation: This is the combined evaluation of:

(i) what Peirce calls “how likely is the hypothesis itself”
(ii) how likely does the hypothesis render the observations.

The first one in our account corresponds to the agreement between the contextual features and the hypotheses. The second is the measure of how well a hypothesis explains the observed findings.

The inclusion of contextual aspects in the hypothesis formulation stage is a nice property, because the activation of hypotheses will then involve an early evaluation. This includes an evaluation with respect to the disease hypothesis being consistent with contextual features. This is the evaluation of whether the disease hypothesis ‘is likely by itself’, which is proposed by the Peirce as one of the two conditions as the basis for the plausibility of a hypotheses. As such, formulation of hypotheses includes an intertwined generation and evaluation processes. This early evaluation obviously takes place only on the basis of the information available so far.

Evaluation in hypothesis selection: This is evaluation of which hypotheses are worthy of testing, or which hypotheses are more worthy of testing. The statement implies a comparative evaluation of activated hypotheses. At this stage, pragmatical consideration in addi-
tion to likelihood concerns provide the necessary criteria, with respect to "testing". If the utility of testing a hypothesis is higher than the others, then the hypothesis is tested early.

B) Evaluation in hypotheses testing: This is evaluation during induction, which happens in light of all available information, after the information needed for evaluating the generated hypotheses is gathered. So this involves evaluation of whether the observed consequences are able to confirm or evaluate the hypothesis. This is evaluation by testing, which evaluates the truth of the hypothesis (the solution) while evaluation in the generation phases justified the reasoner's generation of certain hypotheses.

So, even though we do not adopt Josephson's view that abduction includes the whole process of diagnosis, and limit abduction to hypotheses generation, we agree with him that generation and evaluation of hypotheses are isolated by integrated processes.

The context-sensitive similarity assessment is a step towards grasping the notion of 'overall similarity' suggested by Dreyfus and Dreyfus [Dreyfus 86, p.210].

9.9 Summary

This chapter presents a context-sensitive iterative approach to diagnosis. This approach considers the reasoning of an expert as a pattern recognition process, where the formulation of diagnostic hypotheses relies on experience. The use of experience as a shortcut requires the ability to identify the similarity of the new problem with previously experienced problems. This indicates that context plays a significant role in this similarity assessment.
Chapter 10

ConSID-Creek - domain and task model

10.1 Introduction

ConSID-Creek is a system based on the ConSID problem solving approach, which is combined with the Creek framework ([Aamodt 94] [Aamodt 94b]). It is a knowledge-based system architecture designed to support medical diagnosis. Important issues emphasized by the system are the dependence of classification on planning, the relationship between specific and general domain knowledge (which implies a reasoning method that commutes between these two types of knowledge structures), and the augmentation of both types of knowledge by contextual elements. Its distinctive characteristics are:

- It integrates classification and planning
- It is a hybrid case-based explanation-based reasoning system.
- It behaves in a context-sensitive manner.

We discuss, throughout the chapter, what type of modifications a more traditional model of domain knowledge needs in order for a knowledge-based system using that knowledge to be able to display context-sensitive problem solving behaviour. First of all, domain knowledge base needs to explicitly model contextual knowledge. We employ the methodology developed in this thesis and presented in chapter 6 for analysing and modeling contextual knowledge and its use in ConSID-Creek.

Two different kinds of knowledge that most knowledge-based systems contain are 'content' and 'control' knowledge, the first embodying the knowledge about the world, and the other capturing the knowledge about how the knowledge about the world can be exploited in order to achieve a goal. A system is described, at the knowledge level, in terms of domain knowledge, tasks and methods. The domain knowledge corresponds to the content while the tasks and methods capture control knowledge. In the first part of this chapter, we deal with the domain knowledge. Thus we analyse and model the 'content' knowledge pertinent to medical diagnosis. The term content, in this connection, refers both to the structure of knowledge and its richness with respect to both amount and type. In the second part of the chapter, we model the tasks underlying clinical reasoning.

The 'problem solving method' of ConSID-Creek combines two pioneering reasoning paradigms, namely model-based and case-based paradigms. In our view, the basic strategy employed in clinical inquiry is a generate and test strategy. The subtasks which comprise this strategy may employ either a model-based or a case-based reasoning paradigm. For example, the subtask of generating hypotheses is a typical abductive task and in ConSID-Creek has been realized by a case-based method. The impact of viewing the medical diagnosis process as a combination of model-based and case-based reasoning is that the
knowledge base needs be structured in a way that allows the application of such a hybrid problem solving framework. The methods and mechanisms underlying the system will be elaborated in the next chapter.

Section 10.3 introduces the components of ConSID-Creek at the knowledge level: domain, task and method model. Sections 10.4 and 10.5 introduce the domain model. The starting point for analysing and modelling the content knowledge in the medical domain is to conceptualize the disease process. This is because an important portion of the content knowledge captures various disease processes. Our ‘target’ is the disease and more correctly, the ‘diseased’ person. A disease process is represented as a network where the nodes correspond to concepts such as a disease category, findings, and the pathological states that represent intermediate states that the body of the patient may be in. Pathological states often correspond to syndromes that can be described in terms of a collection of findings. Before everything, we have conceptualized a contextualized model of the disease process. This supports us in constructing the model of the disease mechanism in a context-sensitive manner at the knowledge level. In the next section we present the context-sensitive disease process which constitutes the model we use in ConSID-Creek. We incorporate the high level context ontology which we developed in section 8.3.1 into the medical diagnostic ontology, in section 10.5.2.

In our approach, the most important characteristic distinguishing between the reasoning of an expert physician and that of a novice is the content of the knowledge. Experts own a large, contextualized episodic memory while novices possess a limited memory. Consequently, novices and experts employ different reasoning processes; experts employ a pattern recognition strategy for typical problems. At the knowledge level, these characteristics of content are reflected by incorporation of episodic knowledge structures into a contextualized general domain knowledge model. Chapter 12 introduces how an expert’s pattern recognition strategy can be implemented as a hybrid case-based and explanation-based reasoner.

Section 10.5.4. describes the way case-specific and general domain knowledge are incorporated in the system. Sections 10.6 analyses ConSID-Creek’s control knowledge. The distinction between the content and the control knowledge by no means implies their being independently modelled. On the contrary, there is a close link between them: control knowledge uses content knowledge. So, content knowledge should comprise the domain knowledge used by the control knowledge. Section 10.7 emphasizes the close connection between the content knowledge (i.e. domain model) and the control knowledge. Section 10.8 introduces the task model of the system.

10.2 Methodology for Studying Context

This section summarizes the methodology that has been developed in this research (see section 6.4), in the pursuit of modelling a context-sensitive diagnostic problem solving behaviour. The methodology emerges from a task-oriented approach and involves the following subprocesses:

- construct a model of the diagnostic ‘process’ knowledge. This serves to formalize all the involved application subtasks (i.e., diagnostic subprocesses),
• identify the **loci** of contextual effects (i.e., the subtasks where context may have an influence)

• identify the **role** context plays in each locus (i.e., the way it affects each subtask)

• identify the **source** of context that plays this role

This methodology strongly emphasized the interdependence of content and process knowledge, i.e., of domain, task and method models described in the next section.

10.3 Components of knowledge level; Domain, task and method

In line with Luc Steel’s components of expertise view, ConSID-Creek’s expertise is described at the knowledge level in terms of three concepts: domain, task and method.

**Tasks:** A task is associated with a goal to be achieved. It describes what to accomplish. A task is either a reasoning (e.g. T-generate-hypotheses) or an action (e.g. T-measure-blood-pressure) task. Reasoning tasks are represented as full-fledged objects in the system and each task is characterized by:

• the goal which it achieves when accomplished

• the method it utilizes. A method is a particular way of accomplishing a task, and may generate a set of specific subtasks to be pursued. Complex tasks are decomposed into their subtasks by means of methods.

• the kind of information it uses as input

• the output to be produced

• the perspective it imposes.

Except for the primitive ones, the reasoning tasks are decomposed into subtasks. A reasoning task can be realized by one or several methods. However, in this thesis, a task is accompanied with a single method. So, there is a one to one correspondence between a task and a method. An action task in ConSID-Creek, on the other hand, is realized by several methods. For example, T-measure-blood-pressure can be accomplished by either M-measure-bp-manually, or M-measure-bp-by-supervision-method-X.

A task model illustrates the decomposition of tasks iteratively into its subtasks, until reaching primitive, directly achievable subtasks. Each subtask in the model corresponds to a stage in the diagnostic problem solving process. The medical problem solver is engaged in one subtask at each time.

**Methods:** These describe how to accomplish a task and to apply domain knowledge. That is, they model the control flow. There are two types of methods: *reasoning* methods and *action* methods, corresponding to the two types of tasks that we mentioned above. Reasoning methods are of two types:

• **Task-decomposing** methods decompose task into subtasks

• **Primitive** methods are directly applicable
Domain models and experience models: Different kinds of domain knowledge can be modelled. We emphasized, for example, structural and fault (i.e., pathological) models. This part of domain knowledge capture general knowledge, the domain theory. Experience models capture the concrete problems learned after solving problems. An example experience from ConSID-Creek system looks like the following:

(case#44 (has-goal generate-hypothesis)
  (has-location well-established-hospital)
  (has-name tore)
  (has-age 76-year-old)
  (has-sex male)
  (has-occupation cattle-raiser)
  (has-findings(high-fever sweating chills weight-loss nausea
    productive-cough mild-cheast-pain heart-murmur hematuria
    hypokalama hyponatremia dyspnea-on-exertion low-bp high-hrt)
  (has-solution endocarditis)
  ....
)

In this knowledge structure (‘case’ in ConSID-Creek terms) the slot ‘has-findings’ collects the information type ‘core features’ while the slots ‘has-sex’, ‘has-occupation’, ‘has-age’ ‘has-name’ specifies the target-related context, and ‘has-location’ is of ‘environment-related’ type of context.

10.4 The ‘disease process’ and the ‘content’

The decision of what medical knowledge to include in a system cannot be taken in isolation from the decision of how to model the general ‘disease process’. Thus, the key to explicating the knowledge elements lies in a thorough model of the disease process. In order to construct a context-sensitive domain knowledge base, we need a context-sensitive model of the disease process. Also, in order for the reasoning process to display a context-sensitive behaviour, the content of the knowledge base should include contextual aspects.

![Diagram](attachment:diagram.png)

FIGURE 10.1 (a) Traditional model of the disease process. (b) Our model of the disease process, which allows multiple models for the same disease. Multiplicity of the models of a disease process is illustrated by the difference in the colour of the squares representing the contextual factors and the findings. That is, as the contextual factors varies the findings may vary as well, for the same disease.
This is because there should be a correspondence between the domain concepts that are used in the model of the disease and the reasoning processes. Figure 10.1 illustrates a traditional model of the disease process (a) and our context-sensitive model (b). Two types of contextual knowledge are particularly important for modelling a disease process. These are ‘enabling-context’ and ‘outcome-modifying-context’, illustrated in figure 10.1-b. The former precedes the onset of the disease, while the latter, plays a role after the onset of the disease, and on the manifestation of the outcomes. This model leads us towards a model of domain knowledge that embodies contextual aspects as well as core domain features.

10.5 The content knowledge in ConSID-Creek

Content knowledge consists of two kinds of knowledge, the general domain knowledge and the specific knowledge, also referred to in this work as case-specific knowledge.

10.5.1 General domain knowledge

General domain knowledge is represented as a deep, detailed model of knowledge, and therefore is often referred to in the literature as model-based knowledge, or first-principle knowledge.

The model-based representation of domain knowledge in ConSID-Creek consists of five main components:

- findings (e.g., the complaints of the patient and observations of the medical expert),
- contextual features (e.g., habits, occupation, gender of the patient),
- contextual states (e.g., young-newlywed-woman)
- intermediate pathophysiological states (e.g., syndromes),
- disease classifications.

This model is inspired from the CASNET [Weiss 78]) system and the Galeno-2000 project [Aamodt 91b], but differs from both CASNET and its followers, and Galeno-2000 in that it includes contextual aspects and opens the way for common sense reasoning. Model based approaches typically structure medical knowledge in the form of causal-associational networks. One of the early and most elaborate examples is CASNET, which has been a leading example for its followers. The disease process in this dominant model-based approach is described in terms of three types of domain elements. These are observations of the patient, pathophysiological states, and diagnostic and therapeutic categories. Observations are the direct evidence, pathophysiological states are intermediate constructs, and categories are the solution classes. In this picture, observations are the least abstract concepts, while diseases are the most abstract ones, with pathological states being in the middle. In causal-associative approaches, the mechanisms of the disease processes are described in terms of a ‘cause and effect’ relationship between pathological states. As is implied, traditional approaches do not deal with contextual elements, even though the effects of contextual factors in a disease process are by no means rejected.

In this work, we complete the picture of disease process modeling by augmenting the traditional model with contextual components. So, in addition to disease categories, patho-
logical states and findings, we have an additional knowledge type: contextual knowledge. Figure 10.2 illustrates general domain knowledge in an abstract level, after its augmentation with contextual elements. In the figure contextual states and contextual findings include both enabling context and outcome modifying context elements. The findings and contextual features are associated with states, which are causally associated to each other, forming a network that explains the mechanism of the diseases. So, ConSID-Creek models disease processes as a causal-associational network.

Another entity in the domain knowledge is an ‘action’. Actions may be of type therapeutic or information-gathering. Since the treatment task is not implemented in our system, the domain knowledge does not include such actions. The information-gathering actions, on the other hand, are used in the planning task. These realize information gathering tasks. In the Creek framework, all concepts are represented as frames. The representation language is CreekL and, the domain knowledge is represented as a network of frames. A frame is composed of a number of slots, each of which consists of a facet and a facet value. Thus, each action is represented as a frame, with a content

- the info-gathering task it realizes (realizes-task slot)
- the preconditions that needs be fulfilled on order for this action to be performed (has-preconditions slot)
- the Lisp function to be executed (has-funct slot)

![Figure 10.2 The three levels of the concepts related to the disease process.](image)

The preconditions consist of a list of conditions that the actions need in order to be fulfilled. An example:

(m-venous-press-by-catheterization (realizes-task (t-gather-venous-press )
(instance-of (complex-act))
(has-precondition (cardiolog wire catheter sproeyte roentgen-machine )
(has-funct (m-venous-press-by-catheterization)
...... )

For this action, preconditions are the existence of a set of equipment (wire, catheter, scope, etc), and existence of a cardiolog. The value of the has-funct slot is the name of a lisp function which when executed acquires the data for venous pressure.
10.5.2 The domain ontology

The disease process (see figure 10.1-b) guides the construction of an ontology (i.e., types of knowledge required), and a vocabulary, as well as facts (propositions) capturing the domain knowledge. Together these provide a basis for modeling the reasoning process. In other words, the understanding of disease onset and how they change the body of the diseased person become the basis of how to solve diagnostic problems. We identified a deficiency of the traditional model of disease process (figure 10.1-a). Our attempt to improve this model ended up with an augmentation of the ontology and vocabulary of traditional diagnostic systems, with contextual knowledge. Figure 10.3 illustrates a portion of the ontology of ConSID-Creek which extends the Creek ontology [Aamodt 91], which refers to both the content and the control knowledge.

![Ontology Diagram]

**FIGURE 10.3** Integration of our high level context ontology into Creek general ontology

At the highest level, the Creek ontology defines two types of terms: entity and relation. We keep this distinction in our system. The entities in our system are augmented by contextual components. Creek considers core domain entities only.

The set of relations in Creek are extended to cover relations that are needed to reason about and with contextual entities. The contextual relations are ‘enabling relations’, predisposes
and triggers; the ‘outcome modifying relations’, e.g., modifies, interacts-with; ‘internal context related relations’, e.g., is-accomplished-by, is-realized-by; ‘action-precondition relations’, e.g., cannot-tolerate, has-available-equipment, has-available-medicines; and ‘feature relations’ such as has-age, has-recent-therapy, has-recent-travel, and has-habit.

Figure 10.4 illustrates a portion of the instantiation and use of the domain ontology in ConSID-Creek. As is seen, the enabling context consists of independent-patient-ctx and interactive-patient-ctx in the figure. Drug-abuse is an interactive type of context when generating hypotheses since the details of its interaction with endocarditis is known: Abuse of drug may imply the use of contaminated needles which, in turn, may cause the entrance of bacteria into the body, which may lead to bacteremia, which may trigger factor for endocarditis. The interaction of age, on the other hand, with myocardial infarction is not clearly explained. It is statistically known that there is a relationship between these, but this knowledge is used without knowing the details of the path between old-age and myocardial infarction (MI). Hence, age is an independent type of context for MI. This is typical for weak domains such as medicine. Such knowledge cannot be easily embodied in a deep knowledge base, but can be stored as independent context within cases.

![Diagram](image)

**FIGURE 10.4** A portion from the knowledge base illustrates various relationships between concepts, both from the ‘content’ and the ‘control’ knowledge. The shaded area reflects our view to the disease process which we illustrated in figure 10.1-b. The unlabelled arcs represent has-subclass relations.

In chapter 6 we distinguished between the roles and elements of context. The main role of context is to invoke a shift in the focus of attention. The context type that plays this role is the internal-interactive type. The other role is capturing the focus of attention - together with core domain factors- which are played by the external context elements, either independent or interactive.
10.5.3 Specific knowledge-case base

ConSID-Creek consists of a case-based explainer and a case-based planner in addition to an explanation based reasoner which supports the case based modules.

As referred to in chapter 7, the principle of encoding specificity [Tulving 73] emphasizes the role of "experience" in learning and remembering. According to this principle, an experience is something more than, for example, a word to be learned and includes also the context in which the word is learned. The encoding specificity theory itself does not concentrate on the elaboration of what this context consists of. The impact of the encoding specificity principle on our research is that we emphasize the necessity of representing experiences and the importance of context for capturing experience. We conceptualize the representation of what is learned as an experience, that is, as a combination of context and core domain elements in the form of 'cases'. Thus, we have augmented Creek's case model with contextual components.

In our clinical diagnostic problem solving model, two types of cases are considered; explanation cases and plan cases. The first type comprises an encoding of the explanation that connects the relevant features to the disease. The second is the encoding of a plan that includes a sequence of diagnostic actions, such as defining an examination protocol, selecting the way to perform a particular test etc. Both explanation cases and plan cases include a has-case-type slot with value correspondingly explanation-case and plan-case. The contents of the two types of cases are illustrated in figure 10.5.

Explanation cases:

The content of an explanation case is similar to the illness scripts of Schmidt and Boshuizen [Schmidt 93]. We have augmented their cognitive model with outcome-modifying context and used this model for the purpose of representing individual cases, instead of 'illness scripts'. The enabling context includes the factors that either predispose or trigger the onset of the disease, while the outcome modifying context influences the way a disease manifests itself.

An explanation case is indexed by core-findings, enabling context, and outcome-modifying context. It provides a solution that explains the anomalies in the core findings within the given context.

Plan cases:

A plan case includes a slot named 'info-gather-strategy' which implies both what the to-be-tested hypotheses are and what can be done with these hypotheses. A hypothesis may be confirmed, eliminated, or may need to be distinguished from other plausible alternatives. For example, the system may try to eliminate the hypothesis 'bacterial endocarditis'. The value of info-gather-strategy is a symbol that starts either with the term 'confirm', 'eliminate' or 'differentiate', which is followed by one or several diagnostic hypothesis labels. Some examples of concrete information goals are confirm-endocarditis, eliminate-myxoma, or differentiate-endocarditis-myxoma. This is an important clue as to what information may need to be gathered. The information requirement is also obviously dependent on what is already known. However, the info-gather-strategy is important when determining which past plans may be relevant in the current situation. Therefore, it is an important index in plan cases. Another important index is has-info-goals which indicates the information to be gathered. The info-gather-strategy is used as the first filter which can
quickly eliminate all but plans with the same strategy. The third kind of index is related to environmental factors, called has-location. Environment is the place where information gathering happens. The selection of actions that gather information happens in a certain environment. The environment determines whether an action can be included in a plan or not. An action has some preconditions. If these do not fit into the situation (i.e., the patient and environmental situation), then the action can not be performed. Patient related context specifies the situation of the patient that may be important when selecting and performing information gathering actions. For example the age of the patient, the allergical conditions of the patient, or other general health conditions (e.g., suffering from renal insufficiency which may hinder use of some particular medical) of the patient may be important.

<table>
<thead>
<tr>
<th>explanation-case</th>
<th>plan-case</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>has-goal</strong></td>
<td><strong>has-goal:</strong></td>
</tr>
<tr>
<td>generate-hypotheses</td>
<td>generate-plan</td>
</tr>
<tr>
<td><strong>enabling context:</strong></td>
<td><strong>information gathering strategy:</strong></td>
</tr>
<tr>
<td>drug abuser, ....</td>
<td>confirm-aortic-stenosis</td>
</tr>
<tr>
<td><strong>findings:</strong></td>
<td><strong>information goals:</strong></td>
</tr>
<tr>
<td>sweating, ....</td>
<td>single-s2 ekg-lv-strain......</td>
</tr>
<tr>
<td><strong>outcome modifying ctx:</strong></td>
<td><strong>environmental context:</strong></td>
</tr>
<tr>
<td>tuberculous, ....</td>
<td>well-established-hospital.</td>
</tr>
<tr>
<td><strong>solution:</strong></td>
<td><strong>patient context:</strong></td>
</tr>
<tr>
<td>bacterial endocarditis</td>
<td>allergy-against-antibiotical</td>
</tr>
<tr>
<td></td>
<td><strong>control schema:</strong></td>
</tr>
<tr>
<td></td>
<td>action-1 action-2,......</td>
</tr>
</tbody>
</table>

**FIGURE 10.5 Components of explanation and plan cases**

In order to face the least possible planning failures because the planned actions cannot be performed in the current situation the preconditions are required to be used as indices, and thus determine which plans are plausible. So, plans include three types of indices

- information gathering strategy,
- information goals,
- plan context

Another type of knowledge included in plans are plan-schemas. A plan schema consists of a set of actions. Past plans also encode the information goals achieved, as well as by which action each is achieved, and the observed/measured result of that action.

All the three types of indices are important, since the degree of need for adaptation on a retrieved plan is proportional to the mismatch between these indices and the corresponding information for the current problem case.

The adaptation of a case can take place in different forms. The retrieved plan may include an information goal which is not relevant in the current case. That is, the plan may gather information which is not needed in the current situation. This may happen when the information goals in the retrieved cases don’t perfectly match with the corresponding information in the current case.

Another important source of adaptation is the actions included in the retrieved plan. This happens when the actions of the retrieved case require resources which are not available in the new case. Similarly, adaptation is needed when an action suggested by the retrieved
case can not be applied because of the differences between the patients with regard to specific conditions, such as allergy, etc.

In ConSID-Creek we aim at retrieving cases that are applicable in the current situation; an approach that agrees with the pragmatical view. Other approaches investigating retrieval of cases that need least adaptation depends either on various kinds of similarity assessment ([Kolodner 89]; [Fox 94]) or on the use of adaptation rules during retrieval [Smyth 95]. Our approach investigates how the use of contextual knowledge influences the similarity assessment for retrieving cases that needs less adaptation.

10.5.4 Integration of general and specific knowledge

Figure 10.6 illustrates a portion of ConSID-Creek’s domain knowledge. The system incorporates a case base with deep, explanation-based knowledge as is proposed in the Creek framework. It augments both models with context.

FIGURE 10.6 A portion from ConSID-Creek knowledge base, incorporating cases and model-based knowledge, both having contextual components.
This leads to a contextualized case representation which is also a step towards a representation of common sense knowledge, which medical experts ubiquitously employ in their work. The concepts that are used in the representation of cases are concepts represented in the content ontology. As we will see later, this is important for a hybrid system commuting between case-based and model-based reasoning paradigms.

10.6 The control knowledge in ConSID-Creek

What we refer to as a control knowledge model resembles a knowledge-use model (van de Velde 93) which aims at modelling, at the knowledge level, how domain knowledge can be used. In our approach, the explication of the dependence between the reasoning processes and the domain knowledge has served an important role in developing the methodology for analysing and modeling contextual knowledge and its use in diagnostic process.

![Diagram of control knowledge](image)

**FIGURE 10.7 The control knowledge is expressed in terms of goals, tasks and methods**

The conceptual model of reasoning process is reflected as control knowledge at the knowledge level. It is captured in terms of goals, tasks and methods. A goal may have several subgoals. Similarly, a task can be decomposed into several subtasks. A goal can be achieved by performing a task which can be realized by a method. In fact, there may be more than one method capable of realizing a task. However, in this work we suppose that there is a one to one correspondence between a task and a method. That is, a task can be realized by one dedicated method. The control knowledge also is a network representing the relationships between various goals, tasks and methods. Figure 10.7 illustrates a portion from the control knowledge, at a high level of abstraction. The main task in ConSID-Creek is T-all-ConSID-Creek of which one of the two subtasks is in turn 'make-diagnose'. This subtask is realized by a M-generate-and-test method which decomposes the diagnosis making into its two subtasks: T-generate-hypotheses and T-test-hypotheses. Generation of hypotheses is an abductive task and is realized by a case-based retrieval method.

10.7 The map between the control and the content knowledge

Medical research has recently drawn attention to the link between control and content, and that these two should be learned simultaneously, in relation with each other. In our system
the reasoning processes are represented as tasks. The content knowledge is captured in the domain model. The task descriptions refer to content knowledge. This suggests that tasks are keystones of a mechanism for focusing on the relevant concepts of the domain knowledge. The idea, in brief, is that the model of the diagnostic reasoning process is not developed independently from modeling the disease process.

### 10.8 The loci, role, and source of contextual knowledge in the ConSID-Creek system

As noted earlier, while internal context elements usually display inter-process influences, the influences by external context elements can be investigated in the frame of separate processes, that is, local to processes. The modelling of context in ConSID-Creek is performed by applying the methodology for studying context, presented in chapter 6.

<table>
<thead>
<tr>
<th>subtask (loci)</th>
<th>role of context (role)</th>
<th>type of context playing that role (source)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-formulate-hypotheses</td>
<td>provides cues needed for activating most similar episodes/cases</td>
<td>external context: primarily target-related (i.e., patient-related) ones. These features serve mainly as enabling factors and outcome modifying factors in the disease process. Internal context: determines which type of hypotheses/cases are to be activated as there may be various kinds (e.g., explanation-cases, or plan-cases)</td>
</tr>
<tr>
<td>T-elaborate-formulated-hypotheses</td>
<td>pinpoint which features and aspects should be elaborated for deep similarity assessment</td>
<td>Internal context: The task perspective determines the lower-bound (which includes elements from the external context e.g., habits), upper bound (which consists of a hypothesis for 'hypothesis elaboration'), and the relevant relations for the spreading activation. The task perspective partitions also the features (expressed as lower bound) into meaningful combinations illustrating which features may be matched/compared in order to make a similarity assessment constrained into meaningful comparisons.</td>
</tr>
<tr>
<td>T-select-most-plausible hypotheses</td>
<td>gives possibility to compare alternative hypotheses. Determines which of the activated hypotheses should be tested first.</td>
<td>Internal context: the set of activated hypotheses, the plausibility of these hypotheses.</td>
</tr>
<tr>
<td>T-determine hypothesis-testing-strategy</td>
<td>explicates the strategy for gathering information. Depends on the relation between the hypotheses which have been selected for testing.</td>
<td>Internal context: Depending on the differential pool, an information gathering strategy is adopted. A strategy that is smart to pursue when the hypotheses to be tested is determined. For example, confirm one of the activated hypotheses, eliminate a certain hypothesis, or distinguish between a number of hypotheses from the pool.</td>
</tr>
<tr>
<td>T-generate-expectations</td>
<td>modifies consequences (outcomes) of a hypothesis.</td>
<td>External context: mainly target-related. Modifies the outcome of the fault (e.g., disease) hypothesis (i.e., internal context). This type shows a synergistic effects on the consequence of a fault.</td>
</tr>
<tr>
<td>T-gather-info</td>
<td>puts constraints on the actions. Determines which actions are possible, or wise.</td>
<td>External context: The resources in the environment (environmental context e.g., available equipment), the condition of the target (e.g., patient-related context). For example, special conditions of the patient that should be considered when selecting a test, or therapy.</td>
</tr>
</tbody>
</table>

**FIGURE 10.8** Loci, role and source of context effects in the diagnostic task.
Following this methodology we discuss context-related aspects of the system in this order: first, locate the subprocesses where context effects can be investigated at the levels from which we can refer to our context ontology; second, analyse the nature of the interaction between the context elements and the subprocesses they are utilized in, and finally, identify which type of context elements from that ontology are utilized in each subprocess. Figure 10.8 summarizes the context influences in each locus that we recognize. We have identified six distinct main loci for context assistance corresponding to six different stages of a diagnosis process. These effects are summarized in the figure. The loci correspond to the subtask of the ConSID-Creek task model introduced in the following sections. The third column in the figure, i.e. 'source' of context refers to the entities in the domain model, again introduced in chapter 10.

After determining the main loci of context influences we elaborate, starting from section 11.4, the context influences in each loci. Before continuing with the details of the context effects at specific loci, we describe the general technique by which we realize triggering and capturing the focus of attention in the next section.

### 10.9 The Task model in ConSID-Creek

We refer to the overarching task of diagnosing and treating a disease as T-all-ConSID-Creek. This task is accomplished when its three subtasks are accomplished (see figure 10.9). The first acquires problem features. The second establishes a diagnostic label that explains the anomalous features, and the third treats the disease that the diagnostic label points to. In this thesis we do not deal with the task that proposes how to treat the disease.

![Figure 10.9 The top level task tree in ConSID-Creek](image)

T-make-diagnosis is the name of the main problem that ConSID-Creek attempts to model and solve. T-treat has not been detailed at all in this work. T-acquire-problem-features is considered at a rather simple level, at which the task model looks like figure 10.10. The diagnostic task is elaborated in the rest of the chapter. The methods that realize its subtasks are detailed in the next chapter.

![Figure 10.10 Problem acquisition consists of acquiring information related to core findings and contextual findings](image)
10.9.5 The task model of diagnosis

In this section we present a model of a task tree for medical diagnosis (illustrated in figure 10.11). The overall task T-make-diagnosis is triggered by the goal of explaining chief complaints of the patient, g-establish-diagnosis. Its two subtasks are T-generate-hypotheses and T-test-hypotheses. The first manages to generate a set of hypotheses which will be tested by the second subtask, T-test-hypotheses. It involves gathering more information in order to establish an ultimate- hopefully- diagnosis. This task is also decomposed into its subtask in a way similar to that of T-make-diagnosis. The reason is that both T-make-diagnosis and T-make-info-gathering-plan are realised by the same method as we will examine in the following chapter. So, a T-generate-XX and a T-test-XX task are common subtasks for accomplishing T-make-diagnosis and T-make-info-gathering-plan.

![Diagram](image)

**FIGURE 10.11 The task model for medical diagnosis**

*T-generate-hypotheses* takes as input the problem description and outputs a hypothesis or a set of hypotheses that the system proposes as plausible explanations of patient problems.

*T-test-hypotheses*: takes as input these plausible hypotheses and tests whether these really explain all the observed anomalies. If not, the task of T-generate-hypotheses is reinvoked.

**The subtask of T-generate-hypotheses:**

The subtask model decomposing T-generate-hypotheses is shown in figure 10.12. Its main subtasks are T-formulate-hypotheses and T-select-to-be-tested-hypotheses.

*T-formulate-hypotheses*: on the basis of the problem description including core-findings and contextual features, a set of hypotheses are formulated by associations. Thus, input: problem description, output: a set of disease hypotheses. This corresponds to the first stage of recall in cognitive psychology terminology.

*T-select-to-be-tested-hypotheses*: From the formulated hypotheses the most plausible hypothesis or a set of hypotheses is selected. The hypotheses that are activated on the basis of associations are elaborated. The ones that are not plausible are eliminated. This task is decomposed into two subtasks: T-elaborate-formulated-hypotheses, and T-select-most-plausible-hypotheses.
**T-elaborate-formulated-hypotheses:** This corresponds to the second stage of recall. The hypotheses are elaborated in order to assess how well each explains the anomalous findings and agrees with the contextual features.

**T-select-most-plausible-hypotheses:** This task selects a smaller set of hypotheses for further scrutiny on the basis of how strong each hypothesis is a candidate in the current situation. It also determines which hypotheses are smart to test early with regard to utility concerns.

**FIGURE 10.12 The task model of T-generate-hypotheses**

**Subtask of T-test-hypotheses:**
The task model of T-test-hypotheses is shown in figure 10.12.

**T-determine-hypothesis-testing-strategy:** Takes as input a set of hypotheses that are found plausible enough to pursue. This task establishes a strategy as to what kind of information is needed to reach the ultimate diagnosis.

**T-determine-info-needs:** Given the information gathering strategy and the currently available information, this subtask determines the information that needs to be gathered. It outputs a plan description. This is accomplished by the subtasks T-generate-expectations and T-determine-tobe-gathered.

**FIGURE 10.13 Task model for T-test-hypotheses**
**T-generate-expectations:** Takes as input the hypotheses set (to-be-tested ones) and reasons towards a set of consequences that would be true in case each of the hypotheses were the correct diagnosis. The output is a set of consequences.

**T-determine-to-be-gathered:** Takes as input the expectations computed by T-generate-expectations, and the currently available information and computes what else needs be true in order to reach the ultimate diagnosis. A plan description is constructed including knowledge on what information is intended to be gathered, and the external context elements relevant for the task.

**T-gather-info:** is the high level task for gathering information. Its subtasks are T-establish info-gathering-context and T-make-info-gathering-plan.

![Diagram](image)

**FIGURE 10.14 The task model of T-make-info-gathering-plan**

**T-establish-info-gathering-context:** Specifies the environmental context after taking into consideration possible interactions between environment-related context and patient-related context.

**T-assess-gathered-info:** evaluates the current hypothesis in the light of new evidence. This is not implemented in the prototype.

**T-make-info-gathering-plan:** makes a plan that gathers the information determined by T-determine-to-be-gathered. Its two subtasks are T-generate-plan and T-test-plan (shown in figure 10.13).

**T-generate-plan:** is accomplished by T-formulate-plans and T-select-to-be-performed-plan. The last subtask, in turn, is decomposed further into other subtasks; T-elaborate-formulated-plans and T-select-most-plausible-plan.

**T-test-plan:** Tests the selected plan. It is decomposed into three simpler subtasks: T-check-preconditions, T-modify-plan, and T-execute-actions.

These subtask are realized through methods elaborated in chapters 11, and 12 presents some example excerpts from ConSID-Creek problem solving sessions that illustrate the task sequence executions and the outputs of some of these tasks.
10.10 Summary

This chapter describes the domain knowledge as composed of content and control knowledge. Each of these components are analysed and structured. The content knowledge is a rich network consisting of general and case-specific knowledge which are tightly interconnected. The integration of content and control knowledge is emphasized based on the idea that content knowledge is there to be used by the control knowledge.

The ConSID-approach imposes a context-sensitive reasoning process. The way we implement the context-sensitive processing of information is elaborated in the next chapter. In this chapter, the contextual aspects of content knowledge are described. Context knowledge analysis is based on the contextualized model of the disease process as described above. It provides grounds for modelling a context-sensitive content knowledge.

ConSID-Creek can be considered as a combined refinement and extension of the Creek framework, since it takes the Creek ontology as the basis and clarifies the distinction between the content and control knowledge as well as augments it with contextual components.

In this chapter we introduced a high level model of the control knowledge. The goal instantiation is illustrated in figure 9.2, though not in detail. The task structure is modelled in section 10.9. The methods are dealt with in the next chapter.

In the next chapter, we discuss how ConSID-Creek realizes the use of context knowledge in an implemented demonstration system. We determine the process - loci in the medical diagnosis where context effects are anticipated, the role that context plays in each locus, and the type of contextual knowledge that plays this role. The mechanism underlying the triggering of a shift in the focus of attention is described. The role of contextual knowledge in capturing the focus of attention is also explained there.
Chapter 11

ConSID-Creek
Methods and Mechanisms

11.1 Introduction

The preceding chapter dealt with the domain and the task models in ConSID-Creek. This chapter deals with ConSID-Creek from a third aspect, that is the method aspect. We describe the methods underlying the system’s diagnostic behaviour. In this way, we also show how the ConSID approach can be used in knowledge-based systems field.

A complete implementation of the conceptual framework for contextualized abductive inference presented in parts I and II needs a huge work. In the scope of this thesis we attempt to develop a simplified, minimal demonstrative application in order to experiment around the theoretical results presented in the preceding parts of the thesis.

The system’s main diagnostic method is M-generate&test which decomposes the T-make-diagnosis task into T-generate-hypotheses and T-test-hypotheses. These subtasks are further decomposed into simpler subtasks. We consider M-generate&Test as performing a sequence of abductive, predictive and inductive tasks (as described in chapter 9). In this chapter we will show what each of these tasks involves and how these relate to the subtasks of the CBR cycle.

M-generate&test is used by the system in two places, related to classification and planning tasks and are correspondingly used when

- generating and testing explanatory hypotheses: M-generate&test-hypotheses
- generating and testing plans for gathering needed information: M-generate&test-plan

We describe in this chapter the mechanism and methods underlying the context-sensitive behaviour of ConSID-Creek. Since ConSID-Creek is based on the ConSID approach (presented in chapter 10), it implements context sensitivity in two ways:

- the focus of attention is captured in terms of contextual elements in addition to core domain elements.
- the line of reasoning is internal context-driven and a change in the goal triggers a shift in the focus of attention.

Before introducing the methods and mechanisms that operationalize ConSID-Creek we will briefly summarize the Creek methods. ConSID-Creek methods are built upon the CBR methods that Creek provides. A distinguishing characteristic of Creek is that its CBR approach features knowledge intensive case-based problem solving and learning. Its typi-
cal application areas are complex diagnosis and interpretation tasks (e.g. [Grimnes 96]). General domain knowledge is extensively used in its problem understanding, similarity assessment and case adaptation. A Creek reasoner falls back on general domain knowledge whenever it fails to reason solely from case-specific knowledge. This typically happens when the problem situation is completely novel, or when the reasoner needs to justify its own reasoning (e.g., explanation of similarity between two cases). As such, a Creek reasoner may be considered to commute between explanation-based and case-based reasoning modes, thus between deep and shallow knowledge types.

11.2 Creek’s subtasks and methods

The central subtasks involved in the Creek framework are retrieve, reuse, revise and retain. Figure 11.1 shows CBR subtasks. In this thesis we put a special emphasis on retrieval and illustrate how retrieval realizes the abductive subtask of generation of explanatory hypotheses (as well as information gathering plans).

![Diagram of CBR process and explanation engine](image)

**FIGURE 11.1 The CBR process and explanation engine (from [Aamodt 95])**

As a problem solving paradigm, the basic steps in CBR are:

**Retrieve:** is the process of recalling a past case that resembles the current problem. Past cases reside in the case memory. The retrieval process is based on the similarity of the new case to the past cases. Determination of the features of the new problem that will be considered when looking for similar cases is an important issue influencing the quality of the retrieved cases.

**Reuse:** The solutions of the retrieved cases are utilized. Although the retrieved cases are similar to the new case, they may show differences and thus may need be adapted before using.

**Revise:** this step evaluates the result of the reuse step and gives feedback to the system.

**Retain:** Updates case memory. This may happen through building a new case or generalizing the old case.

Creek provides methods that operationalize knowledge-intensive case-based reasoning. These methods are implemented as Lisp functions. There are mainly three basic methods that underly all CBR subtasks: activate, explain, and focus (see figure 11.1).

**Activate:** activates the relevant portions of the knowledge base.
**Explain:** explaining derived consequences and new information within the activated knowledge structure

**Focus:** focusing towards a conclusion that conforms with the task goal.

This “activate-explain-focus” cycle is a general mechanism that has been specialized for each of the four major reasoning tasks of the CBR cycle, as illustrated in Figure 11.1.

### 11.3 Mechanism for focus of attention

The mechanism for focusing attention has two aspects that need be modelled: focus capturing and focus shifting. We adopt a task-oriented, context-driven approach to these issues. Before going further we want to clarify a possible confusing regarding the use of the term ‘focus’. The ‘focus’ used in this section is not the same as the one in the preceding section.

The mechanism for focusing of attention features the filtering out of irrelevant cases and the concentration on only a small set of cases. A good focus on the case base requires specification of the type of cases relevant for the current goal, as well as the identification of relevant features (i.e. information) to be used as indices for the retrieval.

The goal, being an element of the internal context, invokes a task which specifies the relevant portions of both the case base and the general knowledge base. The relevant portions are specified at the task frame through a set of slots. The collection of these slots imposes a perspective on the knowledge base, indicating the entities and relations that may be utilized for the task accomplishment. Some of these specifications deal with the case base, and some with the general knowledge base. Depending on its characteristics, each task may need one or several of these specifications.

The perspective for tasks that rely on case-based reasoning is captured by

- the case goal
- relevant features

The perspective is specified in the related task frame through the following slots:

- **has-goal**: indicates what type of cases can be utilized, as the case base may include cases serving different goals. In the current implementation two types of case goals are represented; generate-hypothesis and generate-plans.

- **relevant-feature-slot**: specifies which slots of the new case frame refer to relevant findings and contextual features. In a sense, these slots identify the types of findings and contextual features that the current task may need to utilize.

Figure 11.9 illustrates how perspectives are specified in the related task’s frame on two examples, one involving the case base and the other involving explanation based knowledge.

The focus on general domain knowledge happens via filtering out of the irrelevant concepts and relations. This is achieved by activating relevant concepts and relations through
a spreading activation mechanism. In ConSID-Creek spreading activation is realized as a task-oriented context-directed method.

The perspective for tasks that rely on deep, semantic network of knowledge and therefore employ a spreading activation is captured by

- relevant spreading activation relations
- upper bound
- lower bound (in the form of partitions)

These are also specified in the task frame, through slots has-rlv-upward-spreading-rels, has-rlv-downward-spreading-rels, has-rlv-upper-bound-for-focus, and has-slot-partition (see figure 11.2).

![Figure 11.2 Relevant upper and lower bound nodes and relations constraining the spreading.](image)

**has-rlv-upward-spreading-rels** and **has-rlv-downward-spreading-rels**: are the only relations that are to be utilized during spreading activation on the general domain base. Examples to has-rlv-upward-spreading-rels are caused-by, subclass-of, and triggered-by, while examples to has-rlv-downward-spreading-rels are causes, predisposes, and has-instance.

**has-rlv-upper-bound-for-focus**: constitutes the concepts which are the upper bound of the spreading activation. This may be, for example, a disease hypothesis under consideration (for generation of candidate disease hypotheses), or an action hypothesis (for generation of plan candidates).

**has-slot-partition**: specifies the lower bound features. This is structured into partitions. A single lower bound feature may take place in more than one partitions. A partition constitutes a meaningful combination of relevant findings and contextual features. Each partition includes a set of features of which similarity may make sense for the task of explaining similarity between two cases.

In ConSID-Creek the focus of attention is captured in terms of core findings (i.e., through the value of has-findings slot), internal context elements, and external context features (e.g., through the values of has-recent-surgery, has-habit, and has-psych-sit slots).

The mental trace starting from the goal, passing through the task and perspective, and leading to a focus of attention on a portion of memory (i.e, a set of core findings and external contextual features) has been realized on the basis of a task-oriented perspective. Figure 11.3 visualizes this mental trace.
FIGURE 11.3 The explanation of how internal context serves as a trigger for attention-focusing mechanism.

The upper bound is generally described in terms of internal-interactive-context, for example, the current goal or activated hypotheses, while lower bounds are specified in terms of core concepts, such as findings, and external contextual factors relevant for the dedicated task. Figure 8.4 shows the relationship between internal and external context elements. Figure 11.4 illustrates two examples on how the task perspectives are specified in terms of domain concepts. The first example, i.e., T-formulate-hypotheses corresponds to the first subtask of the abductive hypotheses-generation task.

<table>
<thead>
<tr>
<th>task name</th>
<th>task-perspective specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-formulate-hypotheses</td>
<td>has-goal generate-hypotheses relevan-feature-slot has-findings has-age has-sex has-occupation has-habit has-recent-surgery has-prev-srgery has-previous-disease has-other-disease has-recent-therapy has-appearance</td>
</tr>
<tr>
<td>T-select-to-be-tested-hypotheses</td>
<td>has-rlv-upward-spreading-rels caused-by subclass-of implied-by isnot-tolerated-by cannot-tolerate instance-of associated-with predisposed-by triggered-by consistent-with occurs-during co-occurs mutually-exclusive-with . . .</td>
</tr>
<tr>
<td></td>
<td>has-rlv-downward-spreading-rels causes has-subclass implies isnot-tolerated-by cannot-tolerate has-instance consistent-with occurs-during co-occurs mutually-exclusive-with associated-with predisposes triggers drinks . . .</td>
</tr>
<tr>
<td></td>
<td>has-rlv-upper-bound-for-focus: &lt;to-be-tested hypotheses&gt; has-slot-partition (has-finding) (has-occupation has-econ-situation has-psyc-sit has-habit) (has-recent-disease has-habit has-recent-surgery has-recent-therapy has-prev-disease) . . .</td>
</tr>
</tbody>
</table>

FIGURE 11.4 Examples on how a task frame describes the way to focus on relevant entities in the knowledge base.
This task is accomplished by using only the case specific knowledge, i.e., the case base. The second example is for evaluating the activated hypotheses, and selection of the most plausible ones. The evaluation of a hypothesis is done against the deep domain knowledge, which is represented as a semantic network of concepts, connected by relations. The task perspective of each task specifies the portions of the knowledge base that will be activated. The details of each task will be elaborated later in the chapter.

### 11.4 ConSID-Creek methods

In this and the following sections, we describe the methods that realize the diagnostic task in ConSID-Creek. The main diagnostic method is ‘M-generate&test’, which invokes two subtasks of which T-generate-hypotheses is realized by M-generate-hypotheses-by-CBR, and T-test-hypotheses by M-test-hypotheses-by-CBR-reuse (see figure 11.5). Section 11.11.1 details the M-generate-hypotheses-by-CBR while section 11.11.2 details M-test-hypotheses-by-CBR.

![Diagram of the methods for T-make-diagnosis](image)

**FIGURE 11.5 The methods for T-make-diagnosis**

A characteristic of the ConSID approach is that it features expert’s diagnostic reasoning as a pattern recognition process. As described in chapter 3, we view pattern recognition as realized by a generate and test method, which is referred to as M-generate&Test in this chapter. We view diagnosis as realized by an iterative application of the M-generate&test method in ConSID-Creek (see figure 11.6).
Moreover, the application of M-generate&test happens in a nested manner, since the system considers diagnosis as an integration of a classification and a planning task (which is a characteristic of the ConSID approach, introduced in chapter 10) where both classification and planning are realized by the M-generate&test method. In the system, the planning task is a subtask of the classification task. This is because classification involves a hypotheses testing process following hypothesis generation, and the testing process involves a planning task. As such, the 'M-generate&test' in a classification (i.e. M-generate&test-hypotheses) task which generates diagnostic hypotheses invokes another 'M-generate&test' (i.e. M-generate&test-plan) for purpose of generating plans (i.e., for testing-related actions) candidates.

### 11.5 Method for T-generate-hypotheses

*M-generate-hypotheses-by-CBR*: In ConSID-Creek, both generation of explanatory hypothesis and generation of plans for gathering information involve an abductive task. Such abductive tasks are proposed in the literature (e.g., [Benjamins 93]; [van Heijst 97]) to be realized by various methods. These methods are usually investigated under two groups, the empirical methods and the model-based methods. The distinction is made on the basis of the strategy used and the type of knowledge (both content and structure) involved in the abductive task. In ConSID-Creek, the abductive subtask is realized by a CBR retrieval method. This method can basically be classified as an empirical method, as it gives priority to the use of associative, short-cut knowledge as much as possible.
As shown in figure 11.7, M-generate-hypotheses-by-CBR decomposes T-generate-hypotheses task into T-formulate-hypotheses and T-select-to-be-tested-hypotheses subtasks. This method realizes a three-staged recall process where the first stage is accomplished by T-formulate-hypotheses, and the second and the third stages by T-select-to-be-tested-hypotheses. T-formulate-hypotheses relies on shallow knowledge while T-select-to-be-tested-hypotheses relies on deep explanatory knowledge.

This view of hypothesis generation resembles the recent generate-and-recognize theories of cued recall in cognitive psychology (described in chapter 7). The generation process in these theories corresponds to 'activate' in Creek architecture while recognition corresponds to the combination of 'explain' and 'focus'. An important difference is that in cued recall the task is to retrieve the exact copy of the current 'situation', while in our retrieval, the retrieved episode (i.e., case) is usually only similar to the current situation, not exactly the same.

The method M-generate-hypothesis-by-CBR is proposed to realize a context-sensitive abduction of hypotheses, and thus implements the abductive pattern described in chapter 5. The pattern that we proposed in section 5.7.3 as better conveying the nature of abductive inference is repeated below:

- D is data available so far
- H explains D together with Cxt
- H coheres with context Cxt
- H is the best hypothesis for further scrutiny

(Cxt is the outcome-modifying ctx)
(Cxt here is enabling ctx)

It is plausible to test H first

This "conceptual model" of the abductive pattern is converted to the following "design model". This model describes how we realize abductive inference within the case-based reasoning paradigm.
case Case-n is the new case presenting the collection of data D
(1)
case Case-p is a past case which explains data similar to D
(2)
case Case-p happens in a context similar to Cxt
(3)
case Case-p is the case with most similar D and Cxt to those of Case-n’s
(4)
H is the solution proposed by the past case Case-p
(5)
It is plausible to test H first.
(6)

Given Case-n (i.e., case-new), Case-p (a previous case) and other previously stored cases
which seem similar to Case-n are activated first.

![Diagram of methods for T-generate-hypotheses]

**FIGURE 11.8 Methods for T-generate-hypotheses**

### 11.5.1 Method that realizes T-formulate-hypotheses

*M-activate*: The method first focuses its attention to a portion of the knowledge base.
Then it realizes the generation part of the three-staged recall model (i.e., a type of generate-and-recognize model of recall) we proposed in the previous parts of the thesis.

The method operates only on the case base. Consistent with the current goal of the reasoner, it identifies the cases that are relevant for the current task are the ones having ‘has-goal = generate-hypothesis’.

The frame of the task invoking this method specifies the part of the case base to focus on:

(t-formulate-hypotheses
(has-method m-activate-hypotheses)
(has-goal generate-hypothesis)

For the time being the system uses M-activate in two places, when generating hypotheses
(i.e., T-formulate-hypotheses) and when generating information gathering plans. Other
possible goal types related to the use of case retrieval could be related to prevention and
therapy proposal and administration. So, the method uses the ‘internal-interactive’ type of
context element (i.e., the goal) for the ‘early’ focus purpose. A further focus is made for reminding computation.

The method that formulates hypotheses is based on combined remindings from a set of features from one or more cases [Aamodt 91]. The features that are useful for this task are a subset of the features represented in the knowledge base:

(t-formulate-hypotheses
 . . . .
 (relevant-feature-slot (has-findings has-age has-origin
 has-occupation has-habit has-sex has-prev-disease
 has-recent-disease has-recent-surgery has-family-history
 has-condition has-life-pattern has-appearance
 has-medical-in-use has-recent-therapy)

This time elements of patient-related external context is used in focusing. The features can be either independently or interactively encoded in the semantic knowledge base. Now, the epistemological constraints are put through a task-driven mechanism, and reminding starts.

Reminding: The combined reminding is calculated by using the relevance factors of each feature in this set. The relevance factor of a feature for a case is reflected through its importance and predictive strength for that case.

In case representation the features that have been relevant for the solution of the problem are explicitly specified. Accompanying each relevant feature is its relevance degree, which is expressed through two qualitative measures: ‘importance’ and ‘predictive strength’. The importance of a feature indicates how important the existence of the feature is for making a diagnosis for a case while the predictive strength expresses how strong a role the feature plays in predicting a particular solution [Aamodt 91]. Possible values for importance are necessary, characteristic, informative, and irrelevant. Values used for predictive strength are sufficient, strongly-indicative, indicative, and spurious. The following example from ConSID-Creek illustrates how the description of a relevant finding within a case may look like

(case#44 (has-goal generate-hypotesis)
 (has-location well-established-hospital)
 (has-sex male)
 (has-occupation cattle-raiser)
 (has-findings high-fever . . . .)
 (has-solution endocarditis ))
 (has-relevant-feature
  (cattle-raiser :importance informative :predictive-strength indicative)
  (high-fever :importance characteristic :predictive-strength strongly-indicative)
 . . . .)

The past cases of which combined remindings reach a predefined threshold are activated. The set of relevant features determines the combined reminding. Figure 11.4 illustrates the inputs and outputs of the M-activate method. As is seen in the figure
In reminding calculations, combination of the importance and predictive strength measures is converted to a quantitative value on the basis of the following table:

<table>
<thead>
<tr>
<th>predictive strength</th>
<th>importance</th>
<th>Relevance factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>sufficient</td>
<td>&lt;any&gt;</td>
<td>1.0</td>
</tr>
<tr>
<td>strongly-indicative</td>
<td>necessary</td>
<td>0.95</td>
</tr>
<tr>
<td>strongly-indicative</td>
<td>characteristic</td>
<td>0.90</td>
</tr>
<tr>
<td>strongly-indicative</td>
<td>informative</td>
<td>0.85</td>
</tr>
<tr>
<td>indicative</td>
<td>necessary</td>
<td>0.80</td>
</tr>
<tr>
<td>indicative</td>
<td>characteristic</td>
<td>0.65</td>
</tr>
<tr>
<td>spurious</td>
<td>characteristic</td>
<td>0.30</td>
</tr>
<tr>
<td>spurious</td>
<td>informative</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Importance and predictive strength is only assigned to relevant features, not all the features of a case. In ConSID-Creek these values are determined together with medical experts, but if the learning component of the system had been implemented, the system would come up with a proposal for these values that, in turn, would be confirmed or modified by the human expert. The details of this support is described in [Aamodt 91].

Activation starts with recalling a set of cases. Following is the steps illustrating how this happens:

1. Determine the reminding strength of cases by the relevant core features
2. Determine the reminding strength of cases by the relevant contextual features
3. Compute the combined reminding strength of the cases by core and contextual features
4. Select the cases Cr of which combined reminding strength is over a threshold T
For reminding calculations the system checks which features in the current case have been relevant in past cases. This is found by looking at the frame of the particular feature. In the
frame’s ‘relevant-feature-in’ slot the cases are listed up for which the feature has been relevant. An example, illustrating the cases where ‘high-fever’ has been relevant:

(high-fever
  (has-relevant-feature-in
    (case\##44 (importance characteristic :predictive-strength strongly-indicative))
    (case\##17 (importance characteristic :predictive-strength indicative))
    (case\##39 (importance characteristic :predictive-strength indicative))
    ...
  ...
)

Reminding calculation is made on the basis of checking how many of the core and contextual features in the current problem is relevant in past cases. The past cases of which reminding is over a threshold are activated in M-activate. This stage operates on a syntactic match and serves as a filter that narrows down the number of hypothesis to be further considered (i.e., in M-explain).

The formulation of reminding calculation is as in Creek. A minimal requirement for a case to be reminded of is that it has the same goal (i.e., an element of internal context) as the current case. Before the reminding calculation starts, the cases having goals different than that of the current one are eliminated. For example, in hypothesis generation solely the cases that have value ‘generate-hypothesis’ of their ‘has-goal’ slots are relevant:

(case\##44 (has-goal generate-hypothesis)
  ...
)

Cases of which reminding strength is over a threshold are activated. The reminding strength has two components; core-feature based and contextual-feature based reminding. So,

\[
\text{reminding strength} = \text{reminding-strength}_{\text{core-features}} + \text{reminding-strength}_{\text{context-features}}
\]

This is a rather simple formula, yet as we discuss in chapter 13 reveals context effects in activation. In the same chapter we discuss alternative formulas emphasizing the distinction between core and context feature effects on reminding. Experimenting with alternative formula may give more insight into the differences in the effects of the core and context information in reminding.

Reminding strengths originating from either of core and contextual features are calculated as: the ratio of relevant features multiplied by a combined relevance factor calculated from the relevant features pointing to the case where

\[
\text{ratio of relevant core features} = \frac{\text{nr. of relevant core features}}{\text{nr. of all core features in the case}}, \quad \text{and} \quad \text{ratio of relevant contextual features} = \frac{\text{nr. of relevant contextual features}}{\text{nr. of all ctx features in the case}}
\]

We want to emphasize that what we mean by relevant contextual features we mean the contextual features of the case that have been relevant in the past, and has been specified as relevant for the current task (e.g., T-hypothesis-formulation).

As to combined relevance factors, both core-originated and context-originated relevance factors are computed in the same way: as average (arithmetic middle of the relevance fac-
tors) for all the (either core or contextual) features pointing to the case. For example if a case has three contextual features with relevance factors 0.2, 0.65, 0.43 respectively, the combined relevance factors for context will be equal to the result of \( 0.2 + 0.65 + 0.43 \)/3.

After reminding computation is finished, the cases of which combined strength is over threshold T will be returned as activated cases. The system will next explore the hidden similarities, by using semantic knowledge.

11.5.2 Method for T-select-to-be-tested-hypotheses

\textit{M-evaluate-and-select-hypotheses-by-deep-knowledge}: Decomposes the T-select-to-be-tested hypotheses task into T-elaborate-formulated-hypotheses and T-select-most-plausible-hypotheses.

\begin{verbatim}
(t-select-to-be-tested-hypotheses
 (has-subtask (t-elaborate-formulated-hypotheses)
 (t-select-most-plausible-hypotheses))
 (has-method (m-evaluate-and-select-hyp-by-deep-knowledge))
 .........)
\end{verbatim}

11.5.2.1 Method for T-elaborate-formulated-hypotheses

\textit{M-explain}: The activated cases are constructively elaborated. \textit{M-explain} has two components that respectively elaborate the similarity of the new case with past cases on the basis of core- and of contextual features respectively.

It returns a summary of the similarity between the new case and the activated cases:

\begin{verbatim}
(#s(focus-struct case-name case' #44 ..... sum 2.58496)
 #s(focus-struct case-name case' #31 ..... sum 0.65) .....)
\end{verbatim}

The way \textit{M-explain} works is illustrated in the following pseudo code:

\begin{verbatim}
M-explain (new-case activated-cases)
   establish-comparison-context
   for each activated-case Ca
      c-list1 = generate-compare-list (Ca new-case core-features)
      c-list2 = generate-compare-list (Ca new-case contextual-features)
      check-for-direct-match (c-list1 Ca)
      check-for-direct-match (c-list2 Ca)
      check-for-paths (c-list1 Ca)
      check-for-paths (c-list2 Ca)
   return explanation-structure
end
\end{verbatim}

\textit{establish-comparison-context}: looks at the value of \textit{has-slot-partition} of the current task. The partitions constitute the constraints indicating the kinds of features that can be meaningfully compared with each other. For example, the task T-select-to-be-tested-hypotheses identifies eight partitions:

\begin{verbatim}
(T-select-to-be-tested-hypotheses has-slot-partition

171
The list of core features to be compared:

((has-findings) (high-fever headache sweating joint-pains splenomegaly arthritis) (chills sweating joint-pains high-fever))

The list of contextual features to be compared:

(((has-occupation has-econ-sit has-psych-sit has-health-cond has-habit has-condition has-life-pattern has-life-pattern)

check-for-direct-match:
Now the system will make a total similarity assessment including both syntactical and deep comparison. Check-for-direct-match makes a comparison similar to that of done by M-activate. It removes from the comparison-list c-list the features that directly matched. The rest will be subject to deep comparison (by check-for-path). It returns expl-struct which is a summary of similarity between the two cases so far. An excerpt exemplifying a direct match is shown below:

((#(expl-struct slot (has-findings) from high-fever to high-fever method direct-match ref 0.65))

check-for-paths: elaborates the similarity of the available cues with each of the activated episodes, by explaining why things can be considered similar. In a sense, it justifies in detail why the reasoner decides that a certain past case is similar to the new case. In this way the reasoner is capable of finding a similarity between two episodes which may not seem very similar at the first glance. As input check-for-paths receives c-list which consists of a set of lists consisting of a couple of features, i.e., (feature1 feature2) - where feature1 is the feature of new case, and feature2 - of the activated case which it will check
whether there is any connection (i.e., path). The way check-for-paths works is shown in the following pseudo code:

```
check-for-paths
  for each (feature1 feature2) element of c-list
    upper-bound= value of 'has-relevant-upper-bound' slot of the current task
    lower-bound= feature1 and feature2
    active-concepts=perspective-goal-spreading-activate (upper-bound lower-bound )
    expl= find-path ( feature1 feature2 active-concepts)
    strength = get-total-path-value (expl)
    expl-struct= update with expl and strength
  end
  return expl-struct
```

As already mentioned, the deep evaluation of similarity occurs on the deep knowledge (i.e., semantic network) and relies again on the ‘focus of attention’ mechanism proposed in the previous chapters of the thesis. That is, it employs a context-directed spreading activation (referred to as, in the pseudo code, perspective-goal-spreading-activate) of concepts and relations which are relevant for the ‘explain’ task.

**perspective-goal-spreading-activate**: realizes the context-driven focus on the semantic knowledge. Activates a set of concepts by employing a spreading activation algorithm. The algorithm uses a beam search, originally implemented by Lockert [Lockert 93], constrained by the perspective specified in the frame of the current task:

```
t-select-to-be-tested-hypotheses
  . . . . . .
  (has-rlv-upper-bound-for-focus (value (hypothesis))
  (has-slot-partition ((has-findings)
    (has-recent-surgery has-recent-therapy has-prev-disease
      has-habit has-recent-disease)
    (has-occupation has-econ-sit has-psych-sit has-health-cond
      has-habit has-condition has-life-pattern)
    (has-age) (has-sex) (has-family-history) (has-medical-in-use))
  (has-rlv-upward-spreading-rels (value ((caused-by subclass-of implied-by
    isnot-tolerated-by cannot-tolerate instance-of
    consistent-with occurs-during co-occurs mutually-exclusive-with
    associated-with predisposed-by triggered-by))
  (has-rlv-downward-spreading-rels (value ((causes has-subclass implies isnot-tolerated-by
    cannot-tolerate has-instance consistent-with occurs-during
    co-occurs mutually-exclusive-with associated-with predisposes triggers)) )
```

The algorithm is based on the spreading activation method originally developed in the Creek framework. The algorithm in CONSID-Creek is an augmentation and modification reflecting the focus of attention imposed by the current task perspective. The terms “upper
bound" and "lower-bound" that are used in the following algorithm (see figure 11.11) corresponds to the ones we will introduce when describing how the task perspectives capture the focus of attention (in section 11.3). The algorithm performs a spreading activation constrained by the current task via task perspective which is represented by "upper bound", "lower-bound" and "explanation relations". The first two constitute the nodes from which the spreading starts, and the latter one specifies the set of relations along which spreading takes place. The spreading involves activation in two directions.

From the nodes represented by "upper bounds" downward activation (i.e., towards the nodes represented by "lower-bounds") is invoked, while from the "lower bounds" an upward activation occurs. The lower-bound nodes correspond usually to the features observed/measured in the world such as symptoms, while upper bound concepts are more abstract ones inferred during the problem solving process, such as a diagnostic hypothesis generated. Controlling spreading activation by these bounds and relevant relations leads to both activation of certain concepts and inhibition of other concepts.

As an example, Figure 11.19 illustrates spreading activation when the system attempts to explain the similarity between acupuncture and drug-abuse. The context-guided spreading activation activates only the features within one partition at a time and thus limits spreading of activation.

![Diagram](image)

**FIGURE 11.10** Spreading activation for comparing drug-abuse and acupuncture. The shadowed area is activated. The bold arcs indicate the explanation chain illustrating the existence of similarity between acupuncture and drug-abuse.

Let us assume that the patient under consideration is a drug abuser, and one of the activated episodes is on a patient having recently acupuncture therapy. These is no similarity between these things at the first glance. But if the hypothesis is endocarditis, these become similar, as both trigger bacteremia, and consequently endocarditis. However, this can only be detected by relying on the semantic memory.
Similarity assessment reflects the confidence value of the hypotheses proposed by the activated episode. The most similar episode is assumed to possess the most confident hypothesis. Obviously this is just a guess, nevertheless a ‘skilled guess’, in Peirce’s term. So, the matter is to make this similarity judgment as elaborate and extensive as possible. In our approach, the use of contextual elements improves the quality of similarity judgements.

1. Activate the node or nodes representing the "upper bound". Call the resulting node set N1.

2. Spread activation exhaustively from the nodes in N1 along the relations that are members of the relation set has-rlv-downward-spreading-rels in the system and are task specific”. Call the resulting set of activated nodes N2.

3. Activate the node, or nodes representing a partition in the list referred to as "lower-bounds". Call the resulting node set N3.

4. Spread out from the nodes in N3 along the relations that are members of the relation set has-rlv-upward-spreading-rels. Mark the nodes in N2 that are hit by such a path, and call the resulting node set N4. Activate all nodes lying on paths between N3 and N4. Call the resulting set of nodes N5.

5. The focus of attention is established by taking the union of the node sets N2 and N5. Call this node set F.

\[
\text{FIGURE 11.11 The context guided spreading activation}
\]

\textbf{find-path:} produces an explanation structure which looks like:

\begin{verbatim}
(EXPL-STRUCT SLOT (<partition1>) (<features to be compared>)
(<path connecting the features>) <strength of the path>)
\end{verbatim}

For example, the following explanation structure represents a path between ‘shepherd’ and ‘cattle-raiser’ which are the values of the ‘has-occupation’ slot of the new case and an activated case, respectively:

\begin{verbatim}
Total value is: (((#st(expl-struct slot (has-occupation has-econ-sit has-psych-sit
has-health-cond has-habit has-condition has-life-pattern)
from shepherd to cattle-raiser method
(shepherd (raw-milk drinks 0.9) (cattle-raiser dranked-by 0.9)) rf 0.81) . . . .))
\end{verbatim}

An explanation structure is produced simply by propagating activation (i.e., spreading activation) from each of the concepts along the set of explanation relations defined (collectively by has-rlv-upward-rels and has-rlv-downward-rels) for the current task. The explanation structure produced represents a path between the two concepts.

Each explanation-relation has a default explanation strength as the value of its has-explanation-strength slot. For example, explanation strength of causes is equal to 1.0:

\begin{verbatim}
(causes
(has-inverse (caused-by)
(instance-of (causal-relation))
. . . .
\end{verbatim}
The explanation strength of all relations connecting two concepts are combined to form a resultant strength for the complete explanation. The explanation strength of a starting concept is assumed to be 1.0. The total explanation strength of a chain is equal to the multiplication of the strengths of all the relations included in the explanation chain connecting the given two concepts.

Now, assume that the system will explain the similarity between concepts ‘shepherd’ and ‘cattle-raiser’ in the context of hypothesis ‘brucellosis’. Brucellosis is an infectious disease triggered by ‘brucella melitits’ which usually is received through ingestion of raw milk. The following explanation chain is produced which means that shepherd and cattle raisers are found similar in the context of brucellosis since both usually drink raw milk:

\[(\text{shepherd (raw-milk drinks 0.9)} \ (\text{cattle-raiser dranken-by 0.9})) \ \text{rf 0.81}\]

The system does not generate this path in the context of another disease hypothesis, for example ‘aortic stenosis’ which has nothing to do with ingestion of raw milk. This is because when the upper bound is ‘aortic stenosis’ the concepts like ‘raw-milk ingestion’ and ‘raw-milk’ does not take place in the focus of attention. That is, perspective-goal-spreading-activate does not spread the activation to these concepts since it works in an interactive-internal context (i.e., generated hypothesis) sensitive way.

11.5.3 Method for T-select-hyp

\textit{M-focus}: Selects one or a small set of retrieved cases as to-be-tested hypothesis. The method uses two criteria for ranking the hypotheses and selecting the to-be-tested one: its strength (computer by M-explain as described in the preceding section), and utility and policy considerations. Regarding the utility assessment, for the time being the system considers only whether a hypothesis is a ‘life threatening’ disease and places such hypotheses higher up in the ranking. Policy aspects are not included in the current demonstrator. In brief, interactive-internal context determines the ranking of the cases- the plausibility assessments of hypotheses done by the reasoner, and the utility considerations of the reasoner.

11.6 Methods for T-test-hypotheses

11.6.4 Method for T-determine-hyp-testing-strategy

M-select-strategy: The selection of a strategy involves evaluation of the hypotheses in the differential pool (i.e., the ones proposed by retrieval, in particular ‘M-focus’) with respect to each other. For example, if there are two hypotheses which are difficult to distinguish between by using the currently available information, more information that helps the reasoner to resolve this situation needs to be gathered. So, the strategy depends on the interrelationships between the hypotheses in the pool. Another factor is related to the existence of hypotheses of which early consideration may provide high utility. For example, life threatening diseases should be pursued first. This method is not properly implemented in the system. If it were, it would output a strategy consisting of two components. The first could be either confirm, eliminate, or distinguish. The second consists of one or more disease names. For example, confirm-endocarditis, or distinguish-endocarditis-atrialmyxom. Currently, the demonstrator applies only a ‘confirm’ type of strategy. It generates the strategy in the following way: appends the term ‘confirm’ to the hypothesis name proposed by the case number one in the ranked list (i.e., the output of M-focus).

11.6.5 Method for T-determine-info-needs

M-determine-info-needs-by-CBR-reuse: Decomposes T-determine-info-needs task into two subtask: T-generate-expectations and T-determine-to-be-gathered as seen in figure 11.13. The method first determines the manifestations (i.e., core findings) that would outcome if the diagnostic hypothesis was really true.
Then it determines the manifestations about which the reasoner does not have information at the moment. These constitutes to-be-gathered information.

11.6.6 Method for T-generate-expectations

*M-generate-expectations-from-retrieved-cases*: In line with Josephson [Josephson 94], we call the inference type involved in determination of the consequences of a disease, prediction not deduction (as discussed in chapter 3). In strong domains, determination of what can be expected from a hypothetical fault involves a deduction of consequences of the disease. However, in weak domain theories which cannot be captured as a one-world theory, the expectations can not simply be deduced as there are several various versions of how a fault manifests itself. This is because the expectations so inferred could vary, for the same hypothesis, across various patients. The reason for this variation is the outcome modifying context. The disease process model would look similar to figure 10.1-a, had it been a strong domain. In a weak domain, however, it looks rather like in figure 10.1-b. It is the combination of a disease and the outcome-modifying context that synergistically determine the outcome of the manifestations. For example high-fever is a symptom appear with pulmonary disease. However, in case of a patient with pulmonary disease having the context of ‘has-recent-disease= Cushing’s syndrome’ high-fever will often not be encountered. This is because of the coexistence of these two disorders.

In our view prediction nicely fits within the ‘reuse’ subtask of the case based reasoning paradigm. We use retrieved cases for finding out which manifestations may be expected if the solution of the case is adopted as a hypothesis. The mapping between the context of the current problem and the context involved in a retrieved past case is important for the selection of the version of manifestations that can be expected and thus, provides guidance for the rest of the problem-solving process. Since the hypothesis formulation methods presented in the previous sections use both enabling- and outcome modifying contextual (both are being of type patient-related context) elements, the to-be-tested case is sensitive to the outcome modifying context. So, the task of determining expectations from a hypothetical diagnosis can be realized by looking into the consequences provided by a retrieved case. The method simply outputs the ‘has-finding’ slot of the to-be-tested case.

11.6.7 Method for T-determine-tobe-gathered

*M-infer-missing-info*: Based on what is already known and what is expected from a hypothesis (i.e., the output of M-generate-expectations-from-retrieved-cases), the system finds out what information needs to be gathered. The information is the discrepancy between the expectations, (i.e., predicted findings) and the knowns(i.e., the observed findings). Some of the discrepancies, especially the ones that are expected but not observed may be explained by the known facts that are not on the expected findings list. In that way some of the expected findings can be explained by a closer look at the similarities between expected findings and known findings. This may be possible when a phenomena presents itself in seemingly different ways. In this way some of the expected findings that are activated and that would ultimately play the role of information goals will be explained away. This facility is not implemented in the current demonstrator.
Suppose that the findings (i.e., Manifestations) of the new case be represented as M(new-case), and the findings manifested by the solution of a retrieved case are M(retrieved-case). The findings present in the retrieved case but not in the new case constitute the missing information. The difference

M(retrieved-case) - M(new-case) constitutes the ‘Info Goals’, a type of internal-context that serves as input to the method of gathering information, namely M-gather-info-by-applying-plan.

The missing information comprises the information needs or information goals of the reasoner which will constitute a part of the indices that are used in plan cases.

11.7 Method for gathering information

M-gather-info-by-applying-plan: This realizes a planning process, planning of the actions for gathering the information expressed in terms of information goals. In ConSID-Creek, planning is viewed as a combination of plan generation and plan testing/executing subtask just as in classification (i.e., generation of diagnostic hypotheses). So, planning is also realised by an ‘M-generate&Test’ method. The generation of plan candidates realized by a CBR retrieval (M-generate-plans-by-CBR). Execution of plans is viewed as CBR reuse which involves a rather primitive adaptation of plans on the fly. Invokes two subtasks: T-establish-info-gathering-context and T-make-info-gathering-plan.

![Diagram showing the method for gathering information]

11.7.8 Method for T-establish-info-gathering-context

M-infer-info-gathering-context: Before starting to making a plan, the reasoner explicates the context in which the plan will be realized. From the current problem description (i.e., new-case), the interactions between the various context features have been determined. The interactions can be among patient-related contexts or between patient-related contexts and environment-related contexts. M-infer-info-gathering-context first determines environmental context (by calling M-establish-plan-environmental-ctx), the patient context relevant for planning of information gathering (by calling M-establish-plan-patient-ctx), and then computes a final context that describes what is both available in the current environment an (the place where the plan is being made and applied) and can be employed on
the current patient when his/her special conditions (i.e., patient context) is taken into consideration (see figure 11.15). The pseudo code for M-infer-info-gathering-context:

```
M-infer-info-gathering-context
  list1 = M-establish-plan-environmental-ctx
  list2 = M-establish-plan-patient-ctx
  interact list1 list2
```

M-establish-plan-environmental-ctx returns the features of the place where health care is provided. The demonstrator has a limited variety of such places: well-established-hospital, forest, ambulance, and village hospital each of having different resources. The resources of a care place is captured in the semantic knowledge through its ‘has-available-medic’, ‘has-available-equip’, and ‘has-available-personal’ slots. Also the task frame describes these as the set of relations to be used by perspective-goal-spreading-activate method:

```
(t-establish-info-gathering-context
  (has-method (m-infer-info-gathering-context ))
  (has-rlv-spreading-environment-relation (has-available-medic
    has-available-equip has-available-personal )
  (has-rlv-environment-info (has-location))
  (has-rlv-spreading-patient-relation (cannot-tolerate) )
  (has-rlv-patient-info (has-prev-disease has-recent-disease has-allergy
    has-recent-surgery has-age has-sex))
```

An example to what may return is:

```
(has-available-medic medic-gr-5 anesthesia antibiotics)
(has-available-equip xr-machine echomachine manometer thermometer
  scope catheter sproeyte wire de-machine sx-qt-machine ewd-machine
  stethoscope)
(has-available-personal skilled-xr-person cardiolg)
```
On the other hand M-establish-plan-patient-ctx returns the features describing the patient context which may intervene with the use of particular equipment and chemicals used when gathering information. The following is the frame that describes the perspective for guiding the context-sensitive spreading activation algorithm:

(t-establish-info-gathering-context
  (has-method (m-infer-info-gathering-context))
  (has-rlv-spreading-environment-relation (has-available-medic
    has-available-equip has-available-personal))
  (has-rlv-environment-info (has-location))
  (has-rlv-spreading-patient-relation (cannot-tolerate))
  (has-rlv-patient-info (has-prev-disease has-recent-disease
    has-allergy has-recent-surgery has-age has-sex))
)

M-establish-plan-patient-ctx returns the following, meaning that the patient does not tolerate antibiotics in group 1:

(antibiotic-1)

The last thing M-infer-info-gathering-context does is to find the results of a possible interaction between the environmental context and patient context. It returns the features describing the environmental context after the patient’s general situation (e.g., antibiotics allergy) is taken into consideration:

((has-available-medic medic-gr-5 anesthesi antibiotic-2 antibiotic-3)
  (has-available-equip xr-machine echomachine manometer thermometer
    scope catheter sproyet wire de-machine sx-qt-machine
    ewd-machine stethoscope)
  (has-available-personal skilled-xr-person cardiolog))

The system explicates the application environment after excluding antibiotics in group 1. This is done by replacing ‘antibiotics’ with the specific types that can be used, as there are some types that cannot be used.

As can be noticed environmental and patient-related types (both interactive type) of context is used by this method.

11.8 Method for T-make-info-gathering-plan

M-generate-and-test-plan: Decomposes the task into T-generate-plan and T-test-plan sub-tasks.

11.8.9 Methods for T-generate-plan

M-generate-plan-by-CBR: This method operates in the same manner as the method M-generate-hypotheses-by-CBR outputs a set of cases proposing diagnostic hypotheses. It decomposes the T-generate-plan task into T-formulate-plans and T-select-tobe-performed-plan and outputs a plan generated on the basis of past plans that have been used for the similar situations. T-formulate-plans is realized by M-formulate-plans-by-retrieval, and T-select-tobe-performed-plan by M-evaluate-and-select-plan-by-deep-knowledge.
11.8.9.2 Method for T-formulate-plan

M-formulate-plans-by-retrieval: Works in the same way as M-activate which is used for formulation of diagnostic hypotheses. Activates similar past plans. The similarity of situations are assessed through the indices:

- information gathering strategy
- information goals
- environmental context (the place where actions take place)

These indices are specified by the task frame T-formulate-plan as the value of relevant-feature-slot (i.e., info-gather-strategy, has-info-goal and has-location) as shown below:

(t-formulate-plans
  (has-goal (generate-plan))
  (relevant-feature-slot (value (has-info-goal info-gather-strategy has-location)))
  . . . .
  (has-method (value (m-formulate-plans-by-retrieval)))

The first two of these indices specified by ‘relevant-feature-slot’ are of type interactive internal context. As to external context only environmental-context is used by this method.

11.8.9.3 M-evaluate-and-select-plan-by-deep-knowledge

Decomposes the task into the subtasks T-elaborate-formulated-plans and T-select-most-plausible-plan. T-elaborate-formulated-plans is realized by M-explain, and T-select-most-plausible-plan by M-focus.

M-explain: Evaluates in depth the activated plans. The special emphasis is on which of the activated plans are used in similar environmental and patient contexts. Relies on explanatory reasoning. Focus of attention is imposed by perspective that directs how the spreading activation should take place. The perspective is specified by T-generate-plans. The upper bound is external-context.

(t-select-tobe-performed-plan
  (has-subtask t-elaborate-formulated-plans
    t-select-most-plausible-plan ))
  (has-method m-evaluate-and-select-plan-by-deep-knowledge )
  (has-rlv-upward-spreading-rels (causes has-subclass cannot-tolerate
    has-instance has-function has-precondition associated-with
    mutually-exclusive-with associated-with available-medic-of
    available-personal-of ) )
  (has-rlv-downward-spreading-rels (caused-by subclass-of instance-of
    isnot-tolerated-by function-of is-precondition-of associated-with
    mutually-exclusive-with associated-with has-available-medic
    has-available-personal ) )
  (has-slot-partition ((has-control-schema has-available-medic
    has-available-equip has-available-personal)
  (has-info-goal)
  (info-gather-strategy))
  (has-rlv-upper-bound-for-focus external-context-entity)

182
The lower bound consists of a subset of patient-related context features and the environmental features. The subset of the patient-related context is restricted to ones that may mean something for information gathering activities. These are specified by the slots has-age, has-recent-disease, has-recent-surgery, has-allergy-against has-medical-in-use. The partitions are organised as follows: (has-location) (has-age) (has-recent-disease has-recent-surgery uses-medicals has-allergy-against has-medicals-in-use) (has-available-medic) (has-available-equip) (has-available-personal).

**M-focus**: Ranks the activated cases in terms of accordance with the current planning context determined by *M-infer-info-gathering-context*.

### 11.8.10 Methods for T-test-plan

**M-test-plan-by-CBR**: Is realized by three subtasks: T-check-preconditions, T-modify-plan, and T-execute-action (see figure 11.16). It goes through each action consisting of a “plan-schema” of the retrieved plan and attempts to perform the actions. If the first task returns a fit between the preconditions of the action under consideration and the current plan context, then T-execute-action is employed, otherwise T-modify-plan is employed. T-check-preconditions is realized by M-compare, T-modify-plan by M-substitute-action and T-execute-task by M-ask-user.

![Figure 11.16: Methods realising T-test-plan](image)

**M-compare**: This relies on simple, syntactic checking. It returns true if all the preconditions of the retrieved plan are found and has the same value in the current plan context.

**M-ask-user**: If the preconditions are okay, the action is executed. This happens by asking the user to enter the value of a core finding.

**M-substitute-action**: Realizes a very simple substitution. Is employed when the precondition of the to-be-performed action does not agree with the current situation. Examines the deep knowledge in order to check that the considered information goal can be realized by other action methods than the one whose preconditions do not match. If preconditions of some alternative action methods match the plan context, the one with less cost is selected.

**Plan execution-reuse**: Adaptation of a plan is not extensively modelled in ConSID-Creek. This is because our concern is primarily to illustrate context influences in specific tasks, and not to implement a complete diagnostic system. We intend to improve efficiency in a particular task on the basis of context use. Adaptation happens only in the form
of substitution of actions that are not possible to employ in the current situation with others that achieve the same information goals and at the same time fit the current situation.

For example, if a past plan measures intraocular pressure by a schiotz tonometer, in order to be able to perform the action of m-measure-intraocular-pressure-by-schiotz-tonometer the schiotz tonometer should be available in the location where the diagnosis takes place. However, if the schiotz tonometer is absent in the current situation, but an applanation tonometer is available, then an action using that instrument should replace the action that uses the schiotz tonometer, namely m-measure-intraocular-pressure-by-applanation-tonometer. A good plan is one that needs only the available instruments. Though it may not be difficult in some cases to replace an action which requires unavailable instruments, with another action which is applicable at the current location, it may sometimes be difficult or even impossible to replace the actions proposed by a retrieved plan. Replacement of an action may effect the other parts of the plan as well. To avoid such problems, ConsID-Creek tries to retrieve plans that propose actions that are possible to perform under the conditions imposed by patient-related and environment-related context.

Thus, features that are used in plans as indices are of both internal context type (information goals) and external context type (preconditions such as the medicine, equipment used for executing the plan, and the conditions of the patient which may have an affect on the selection of a particular action over its alternatives, such as age or a previously established chronic disease).

11.9 Method for T-assess-gathered-info

M-assess-by-induction: The last inference step in the diagnostic inference cycle is induction, which involves an assessment of available information, with respect to the question of whether a hypothesis agrees with the known findings. As each new piece of information arrives, it is assessed as to whether it supports or refutes the focal hypothesis. This is, at the same time a review of whether the case is useful for the current problem or was a bad 'guess' (i.e., hypothesizing) and should be replaced by another one.

1. Activate: activates the nodes in the knowledge base corresponding to the recently gathered information.

2. Explain: explains how each new fact effects the hypothesis being tested, by relating each such fact to the hypothesis. This constructs an explanation chain between each information piece (a finding) and the hypothesis, which shows how the information affects the verification or refutation of the hypotheses. This explanation process should take into consideration the 'outcome modifying context' in order to correctly explain the relationship between a finding and a hypothesis. See figure 10.1-b to review our account of the effects of such contextual elements on the outcome (i.e., findings) of a disease.

3. Focus: Considers all the explanations constructed in (2) collectively, and determines what they collectively mean regarding verification or refutation of the hypothesis.
11.10 Summary

In this chapter the methods that collectively implement the context framework and the CONSID approach has been presented. The principle method is CBR, which promotes retrieval of hypotheses and plans, and reuse that uses retrieved cases for predictive inference.

The heart of our context-sensitive similarity judgment is that both independent and interactive type of contexts can increase distinctiveness of an episode. ‘Activate’ picks up the past cases that are distinguishable as being more similar to the new case. Distinctiveness corresponds with being easily differentiable from alternatives. This implies that independent context is also effective in the activation stage, as this stage relies on shallow associations which do not require interactive encoding. On the other hand, the elaborative similarity- ‘explain’ judgment needs deep knowledge, and thus interactive encoding. Consequently, only interactive context is effective in the ‘explain’ task.

‘Explain’ is a refinement and justification of similarity. This is inherent in abductive reasoning, since it is characterized by fallibility and its reasoning needs be justified, even though this is different from verification of the ultimate solution/result.

So, ‘activate’ relies on both the interactively and independently encoded knowledge, while ‘explain’ relies only on the interactively encoded type.

A context-sensitive spreading activation lays in the heart of the explanation-based reasoning process.
PART IV

Evaluation, conclusion and future work
Chapter 12

Example

12.1 Introduction

We will in this chapter show excerpts from a run of the implemented ConSID-Creek system, with comments. The comments and the traces will exemplify how the various parts of CONSID-Creek are intended to behave. We show three aspects of the way the notion of context is handled in the system. These three aspects of context are

- the typing of context elements
- the encoding of context elements and
- the role of context elements

We distinguish, at the top level, between the internal and external types of context elements. Goals, generated hypotheses, and information gathering strategies are of internal type. External context is divided into two main groups: patient-related context and environment-related context elements. Each type is further decomposed into their subtypes as discussed in chapter 8.

Some elements of the context are interactively encoded while others are independently encoded. In order to be able to represent both, appropriate knowledge structures should be chosen. This we do by combining case-related, specific knowledge with explanatory knowledge. Independently and interactively encoded knowledge is modelled and used in different ways. In this chapter, we illustrate how each is handled in the system. In brief, independently encoded context elements are only used by the processes relying on CBR while interactively encoded ones can be used by both processes that rely on CBR and those that rely on explanation-based reasoning.

Different context elements play different roles in the system. Goals, a specific type of internal context, are responsible for the line of reasoning. A goal imposes - via perspective- a shift in the focus of attention. Other types of internal context, as well as external types are used by the mechanism for explicating and capturing the focus of attention.

In general, context plays a role in narrowing down the search space that ConSID-Creek deals with. In particular, ConSID-Creek provides

- a “selective” case activation and
- a “directed” spreading activation mechanism
- context-sensitive explanation (i.e., similarity elaboration)
12.2 The patient reminds the system....

A 35 year old, unemployed man developed a high fever, anemia and chills over several day's time. The medical history of the patient includes recent acupuncture therapy.

The patient, 'a 35 year old man having recent acupuncture therapy complaining from fever, chills and anemia' reminds the ConSID-Creek system of other patient cases it solved before. One of the cases is a '27-year old male having a history of intravenous drug abuse, complaining of fever, sweating, chills, anemia and dyspnea'. This patient had the diagnose of infective endocarditis. Another patient the system 'recalls' is a '30-year old woman complaining from fever and chills. She had a recent dental surgery'. Yet another past episode is about 'a 42-year old woman with fever, chills, anemia, heart murmur, papular rash, diffuse lymhadenopathy and mild polyarthritis'. These rememberings lead to the generation of some hypotheses on the cause of patient complaints.

A similarity between 'a 35 year old man having recent acupuncture therapy' and 'a 27-year old male having a history of intravenous drug abuse' is important in the situations where the patient complaints from fever and chills. The similarity lies in the fact that both intravenous drug abuse and acupuncture therapy are risk factors which may lead to bacteriaemia. Bacteriaemia and fever can easily be associated with the disease endocarditis by a medical practitioner. The system's ability to establishing a diagnosis relies on its experiences which are represented as cases in the knowledge base.

12.3 Generation of diagnostic hypotheses

At this stage, the goal of the system is g-make-hypotheses, that is generating some hypotheses that explain the available information, the features.

Generation of hypotheses is realized by case-based retrieval methods in ConSID-Creek. The involved tasks are T-formulate-hypotheses and T-select-to-be-tested-hypotheses. The methods that realize these tasks are correspondingly M-activate, and the combination of M-explain and M-focus.

12.3.1 Formulation of explanatory hypotheses

The task T-formulate-hypotheses (described in chapter 10) accomplishes activation of hypothesis by applying CBR method M-formulate-hypotheses-by-CBR (described in chapter 11). We will now look into ConSID-Creek’s hypothesis formulation behaviour.

As we have mentioned in previous chapters difference between expert and novices have been observed in the quality of the formulated diagnostic hypotheses (i.e., differential diagnosis) [Elstein 78]. If the definite hypothesis is generated at the first encounter with the patient, it is almost sure that it will be recognized later in the diagnostic process. It is common that the inclusion of the definite hypothesis in the differential pool at the start is prevalent. We will now try to show that in ConSID-Creek, both independent and interactive context elements play a significant role in the initial formulation of hypotheses. Remember that the medical domain is an open domain and therefore the expert works from scarce information at the start. On the other hand, contextual information is rela-
tively easy to gather. It is usually available or inferrable in the early stage of the diagnostic process.

ConSID-Creek imitates a pattern-recognition process in diagnosis which relies on similarity in overall appearance of the new case to past cases. This makes it possible to use context knowledge. Again, in the literature it has been shown that experts are able to use available context information while novices are not [Regehr 94]. ConSID-Creek imitates the expert behaviour and is capable of using a certain piece of context information if it has cases using that piece in its case base.

12.3.2 M-Activate-hypotheses: Case-based recall

Suppose that the new case ‘an unemployed man complaining from high fever, chills, and anemia who has a history of recent acupunptur therapy’ is entered into ConSID-Creek as ‘case-new’. The location where the diagnostic process takes place is a well established hospital (i.e., environment-related-context).

In CreekL language, macros are defined for some functions to make terminal input more convenient. We use some of these in the example presented in this chapter. We explain these before going further:

‘#>’ is used in order to type in the value of a slot in a frame. It takes arguments in the form: <frame name>.<slot name>.value

‘#>L’ is used in order to print out a frame. It takes an argument in the form: <frame name>

A single ‘>’ indicates a prompt from the computer.

In the rest of the chapter we focus on a concrete patient problem and show excerpts from ConSID-Creek runs which are accompanied by text that elaborates and explains the excerpts. A difference in font between what the system prints out (in garamond font, size 11) and our explanatory text (in times font) is made in order that the reader can easily distinguish ConSID-Creek’s responses from our comments. The system’s output is slightly edited for increasing the readability.

In Allegro Common Lisp running CreekL we start CONSID-Creek with the following lisp expression:

```
>(start-ConSID-Creek)

Please enter the features of ‘case-new’, the new case.

#>case-new.has-findings.high-fever
#>case-new.has-findings.chills
#>case-new.has-findings.anemia
#>case-new.has-recent-therapy.acupuncture
#>case-new.has-occupation.unemployed
#>case-new.has-location.well-established-hospital
#>case-new.has-sex male
```
Following is the frame representing the new case:

\[
\text{(case-new (has-findings (high-fever chills anemia)))}
\]
\[
\text{(has-location (well-established-hospital))}
\]
\[
\text{(has-recent-therapy (acupuncture))}
\]
\[
\text{(has-occupation (unemployed))}
\]
\[
\text{(has-sex (male))}
\]
\[
\text{(has-age (30-35-year-old))}
\]

The system invokes the task \text{t-make-hypotheses}, which in turn invokes its subtasks.

\[
\text{Current task to be performed : t-make-diagnosis}
\]
\[
\text{Current task to be performed : t-generate-hypotheses}
\]
\[
\text{Current task to be performed : t-formulate-hypotheses}
\]

At this stage the system 'recalls' a set of cases whose solutions can be considered as candidate hypotheses for diagnosis. The cases similar to case-new in the case base are activated. The solutions of the activated cases are proposed by the system as the working hypotheses.

The system identifies at each time (i.e., when a new goal is established and thus a new task is invoked) a subset of currently available information as currently relevant. The task representation for 't-formulate-hypotheses' specifies which type of information and knowledge is relevant. This includes the specification of the type of case (i.e., the ones having "has-goal = generate-hypothesis") and the features (i.e., the ones pointed in \text{relevant-feature-slot}) relevant for the accomplishment of the case. The task frame for T-formulate-hypotheses specifies only the explanation cases, and only the patient-related context elements, in addition to core findings (pointed to from the slot 'has-findings') as relevant for this task. This is an example of how the focus of attention is identified in terms of core and contextual features. Now we want to look at the frame representation of the t-formulate-hypotheses:

\[
>\#lt-formulate-hypotheses
\]
\[
\text{(t-formulate-hypotheses}
\]
\[
\text{(has-method (m-activate-hypotheses))}
\]
\[
\text{(has-goal (generate-hypothesis))}
\]
\[
\text{(relevant-feature-slot}
\]
\[
\text{(has-findings has-age has-origin}
\]
\[
\text{has-occupation has-habit has-sex has-prev-disease}
\]
\[
\text{has-recent-disease has-recent-surgery has-family-history}
\]
\[
\text{has-condition has-econ-sit has-life-pattern}
\]
\[
\text{has-appearance has-medical-in-use has-recent-therapy)}
\]

---

1. This type of cases are also referred to as 'explanation cases' in the thesis.
So, the similarity assessment used for activation of cases takes into account both contextual and core domain features. The activation of cases is based on a direct, superficial match, and involves only the case base, not the deep, general domain knowledge.

The system first displays the past cases for which the core features has been found relevant in the past:

(anemia
  ((case\#36 0.55) (case\#9 0.65) (case\#6 0.65) (case\#62 0.65) (case\#5 0.65)
   (case\#26 0.55) (case\#25 0.65)))
(chills
  ((case\#heidi112 0.65) (case\#heidi111 0.65) (case\#9 0.65) (case\#6 0.65)
   (case\#62 0.65) (case\#46 0.65) (case\#5 0.65) (case\#57 0.65) (case\#27 0.65)
   (case\#26 0.65) (case\#25 0.65) (case\#19 0.55) (case\#44 0.65)))
(high-fever
  ((case\#43 0.9) (case\#42 0.9) (case\#31 0.65) (case\#32 0.65) (case\#36 0.65)
   (case\#9 0.65) (case\#6 0.65) (case\#62 0.65) (case\#46 0.65) (case\#5 0.65)
   (case\#57 0.65) (case\#21 0.65) (case\#27 0.65) (case\#26 0.65) (case\#19 0.65)
   (case\#29 0.9) (case\#39 0.65) (case\#17 0.65) (case\#44 0.9)))

This means, for example, that the finding “anemia” was found to be relevant in cases #36, #9, #6, #5, #62, case#25 and #26 in the past. The numbers following the case identification reflect the degree of relevance of a feature for that case. For example, “high-fever” was most relevant in cases #43, #42, #44, and #29 among the cases in which it has been found relevant.

The system then displays the contextual features found to be relevant and the past cases in which they have found relevant:

((acupuncture
  ((case\#26 0.9) (case\#44 0.85)))
(unemployed
  (case\#26 0.2)))
(male
  ((case\#46 0.2) (case\#27 0.2) (case\#26 0.2) (case\#25 0.2)
   (case\#19 0.2) (case\#44 0.2)))

Based on combined reminding formula given in section 11.5.1, the activation strength of cases are computed and displayed, ranked with respect to the activation strengths:

Activated cases are:

((case\#26 1.05 6) (case\#44 0.8 4) (case\#5 0.65 3) (case\#9 0.65 3) (case\#6 0.65 3)
 (case\#62 0.65 3) (case\#25 0.43 2) (case\#46 0.43 2) (case\#57 0.43 2) (case\#43 0.3
  1) (case\#27 0.43 2) (case\#29 0.3 1) (case\#42 0.3 1) (case\#31 0.216 1) (case\#32
  0.216 1) (case\#heidi112 0.216 1) (case\#heidi111 0.216 1) (case\#36 0.4 2) (case\#19
  0.4 2) (case\#17 0.216 1) (case\#39 0.216 1) (case\#21 0.216 1))
Each case name is followed by two numbers, the first of which indicating a measure of activation strength and the second indication the number of features the case shares with the new case. Now the system displays the activated hypotheses.

Activated hypotheses are:

(endo
carditis sarcoidosis pernicious-anemia pulmonary-tuberculosis atrial-myxom 
stills-disease brucellosis lupus-erythematosus pulmonary-tuberculosis juvenile-rheu-
matoid-arthritis)

These hypotheses proposed by the activated cases where case26, case44, case27, 
case29 and case25 propose endocarditis, case5 sarcoidosis, case9 pernicious-anemia, 
case6, case46 and case62 pulmonary-tuberculosis, case57 atrial-myxom, case43 
stills-disease, and case42 brucellosis. Hypothesis endocarditis comes out as the strongest 
reminded hypothesis based on the superficial feature match.

Let us now look at the particular effects of core- and contextual features on these remind-
ings. If we we had turned the context-sensitivity off and run the system for activation of 
past cases the result would be different. In the following, the activation is repeated taking 
only the core features into consideration:

Activated cases are:

((case5 0.65 3) (case62 0.65 3) (case6 0.65 3) (case9 0.65 3) (case26 0.616 
3) (case44 0.516 2) (case57 0.43 2) (case25 0.43 2) (case27 0.43 2) (case46 
0.43 2) (case19 0.4 2) (case36 0.4 2) (case42 0.3 1) (case29 0.3 1) (case43 
0.3 1) (case32 0.216 1) (case12 0.216 1) (case31 0.216 1) (case1012 0.216 1) (case1012 0.216 1) (case17 0.216 1) (case39 0.216 1) (case21 0.216 1))

All the available core information (i.e., high fever, chills and anemia) have been found 
relevant in many past cases: case5, case6, case62, case9, and case26. Case44, 
case57, case25, case27, case46, case19, case36 all have two of these features as 
relevant features. If we look at the reminding strengths of the activated cases, we see that, 
hypotheses sarcoidosis (proposed by case5), pulmonary tuberculosis (case6 and 
case62), and pernicious anemia (case9) are indistinctive as their reminding strength 
have equal numerical value(see above). That is, the currently available core findings are 
not distinctive enough. As can be seen, all of the three findings have been highly relevant 
in many cases of which solutions propose different hypotheses, so the system is not capa-
bale of distinguishing between the hypotheses.

A difference worth noticing is that the differential pool established with and without con-
textual features look rather different. The first three hypotheses of the differential pool 
generated when only using core features do not include the hypothesis generated and 
ranked as number one (i.e., case26) when contextual features are also taken into consid-
eration.
On the basis of the available features in the new case, the important differences between case#26 and for example case#5 originates from the contextual features. The system displays, on request, the contents of case#5 and case#26:

```
> #Lcase#5
(case\#5 (has-goal (generate-hypothesis))
   (has-findings (high-fever) (chills) (anemia) (heart-murmur)
    (papular-rash) (mild-polyarthridis) (diffuse-lymhadenopathy)
   ... ...)
   (has-age (28-year-old))
   (has-sex (male))
   (has-origin (african-american))
   (has-solution (sarcoidosis))
   (has-location (well-established-hospital))
   ... ...)

> #Lcase#26
(case\#26 (has-goal (generate-hypothesis))
   (has-location (well-established-hospital))
   (has-age (30-40-year-old))
   (has-sex (male))
   (has-recent-therapy (acupuncture))
   (has-findings (high-fever) (sweating) (chills) (anemia)
    (dyspnea-on-exertion) (loss-of-vision) (high-hr) (cardiomegaly)
    (heart-murmur) ... ...)
   (has-allergy (allergy-against-antibiotic-1))
   (has-condition (malnutrition))
   (has-solution (endocarditis))
   ... ...)
```

The formulated hypotheses are ranked according to their plausibility. Plausibility is a measure of the overall situation. As Peirce proposes, it is a judgment of both the hypothesis’ likelihood in itself, and of how likely it renders the consequences to be true.

In ConSID-Creek plausibility is judged through similarity with respect to two considerations. The first corresponds to the first part of Peirce’s statement, the plausibility of the hypothesis itself. The system assesses this as ‘the plausibility of a disease given the patient-related context’. The second part corresponds to the second part of Peirce’s statement where the system assesses the likelihood of the disease’s manifesting the core findings.

It is worth mentioning that even though the cases case26, case44, case#27, case#29 and case#25 all propose ‘endocarditis’ as the hypotheses, they have not been activated equally strongly. The reason is that a disease may usually present itself in different ways. It is apparent that some of the manifestations (i.e., core findings) that are observed in case#26 are not observed in these cases even though all have the same solution. The reason is the interaction of ‘outcome-modifying’ context. For example in the patient case represented
as case#25 'high-fever' is not observed. This is because the patient suffers from Cushing
syndrom which suppresses the development of fever.

Since the details of the relationships between case features and the concepts in the general
domain knowledge are not important at this stage, both independently and interactively
related features supports activation. This reflects the findings presented in medical
research which emphasize the importance of a holistic pattern recognition process by
expert. "The case is diagnosed based on its similarity to a specific instance of the disease
that has been previously encountered, and similarity is, in turn, based on the totality
of features of the case, whether analytically relevant or not" [Regehr 94, p. s33] (italics are
ours). In the same way as the research in cognitive psychology declares the effects of
'indirect' context in recall (see chapter 8), Regehr stresses the role of features possibly
irrelevant for _analytical_ questioning in medical diagnostic pattern recognition. Such
contextual features we investigate under the heading of 'indirect context'.

The similarity of cases are considered with respect to their sharing exactly the same fea-
ture-value set. This is a limited, superficial similarity match. It considers only possible
direct matches between the features of the two cases, and does not take into account the
deep knowledge.

12.3.3 M-Explain: Assess deeper similarity

The current task now is to select the to-be-tested hypotheses (i.e. t-select-tobe-tested-
hypotheses). This task invokes the subtask t-elaborate-formulated-hypotheses which
relies on the explanation-based reasoning that operates on the semantic network of gen-
eral domain knowledge.

Current task to be performed: t-select-tobe-tested-hypotheses
Current task to be performed: t-elaborate-formulated-hypoteses

Entering explain...

Input-case is: case-new
Old case is: case#26
Hypothesis: endocarditis

The frame of the task t-select-tobe-tested-hyp specifies a perspective in terms of domain
relations and concepts that, when instantiated by the information pertinent to a particular
situation, promotes a directed spreading activation in the general domain knowledge base.
The perspective imposes the features to start spreading of activation with, and the rela-
tions along which the activation should spread. Thus, partitioning of relevant slots in this
way makes the similarity search in the deep knowledge base a focused, limited process.
Meaningless comparisons are avoided. For example, the core finding 'sweating' is not
compared with the contextual feature has-habit. As is seen, has-habit is put into the same
partition together with has-occupation, has-econ-sit, has-psych-sit, and has-health-cond.
The reason is that a relation may exist between habits and the occupations or a recent-dis-
ease.

The partitions in ConSID-Creek are intended to be arranged according to the way com-
mon sense tends to work. However, the partitions we set up here do not necessarily con-
stitute the most useful ones. Nevertheless, they illustrate the idea that common sense employs constrained, or more correctly, focused reasoning.

ConSID-Creek attempts now to find out whether there are deep similarities between the new case and the activated cases. It starts from identifying the features to be compared:

Input-case is: new-case

Old case is: case#26

The list of core features to be compared:

- ((has-findings) (high-fever sweating chills anemia dyspnea-on-exertion loss-of-vision high-hr cardiomegaly heart-murmur) (anemia chills high-fever))

The list of contextual features to be compared:

- ((has-recent-surgery has-recent-therapy has-prev-disease has-habit has-recent-disease) (acupuncture) (acupuncture))
- ((has-occupation has-econ-sit has-psych-sit has-health-cond has-habit has-condition has-life-pattern) (malnutrition) (unemployed))
- ((has-age) (30-40-year-old) (30-35-year-old))
- ((has-sex) (male) (male))

The list of features to be compared represented as a list of which each element consists of two components. For example, in the second element ((has-occupation has-econ-sit has-psych-sit has-health-cond has-habit has-condition has-life-pattern) (malnutrition) (unemployed)), the first component refers to a slot partition which points to comparable relevant-feature-slot, the second is a list of feature values that fill these slots in the activated case (i.e. malnutrition), and in the new case (i.e. unemployed).

All the core features available in the new case exist also in case#26, thus they directly match. Some of the contextual features also directly match, and some (e.g., malnutrition and unemployed) are not. The system explores whether there are any relationship between malnutrition and unemployed, and the ages in the two cases.

Attempting to find a path from malnutrition to unemployed

We will see now how the system attempts to find a similarity between features malnutrition and unemployed. When seeking a similarity between the features 'malnutrition' and 'unemployed' the spreading activation is restricted by the task perspective to the following relations and concepts. The upper bound is 'endocarditis', and the lower bound consists of 'malnutrition' and 'unemployed'. The focus of attention is expressed in terms of a set of relations and concepts:

Activated relations are:

- (associated-with instance-of has-instance subclass-of has-subclass triggers triggered-by predisposes predisposed-by caused-by causes)
Activated concepts to be used for similarity match are:
(unemployed adult degenerative-heart-disease predisposing-cardiac-lesion bad-econ-sit bad-psyc-sit drug-abuse bacteriemiae-triggering-habit bacteriemiae-triggering-factors dyspnea-on-exertion malnutrition immunosuppression bacteriemiae fever-reaction high-hr aortic-insufficiency vegetation-on-aortic-valve)

The system finds a similarity between the features malnutrition and unemployed. When we look at the system's reasoning about the similarity between the two features, we see that it involves a type of common-sense reasoning. The system displays the similarity discovered by employing deep knowledge as an explanation:

:from malnutrition :to unemployed
:method (malnutrition (bad-econ-sit caused-by 1) (unemployed associated-with 0.5) :rf 0325)

The system found a similarity between malnutrition and unemployed in the context of endocarditis. If we look at the focus of attention (i.e., activated concepts above), immunosuppression, bacteriemiae, bacteriemiae-triggering-factors, bad-econ-sit are all within the focus. The explanation of the similarity between malnutrition and unemployed is due to: 'unemployed is associated with bad economical situation' which in turn 'causes malnutrition', and 'malnutrition causes immunosupression' which is relevant for endocarditis. The result of the similarity evaluation between new case and case#26 is presented by the system:

Finished similarity explanation between new-case and case#26
Total value is: (((#s(expl-struct slot (has-age) from 30-40-year-old to 30-35-year-old method (30-40-year-old (adult subclass-of 0.8) (30-35-year-old has-subclass 0.8)) :rf 0.64) #s(expl-struct slot (has-occupation has-econ-sit has-psych-sit has-health-cond has-habit has-condition has-life-pattern) from malnutrition to unemployed method (malnutrition (bad-econ-sit caused-by 1) (unemployed associated-with 0.5) :rf 0.325) #s(expl-struct slot (has-sex) from male to male method direct-match :rf 0.2) #s(expl-struct slot (has-recent-surgery has-recent-therapy has-prev-disease has-habit has-recent-disease) from acupunctur to acupunctur method direct-match :rf 0.9) #s(expl-struct slot (has-findings) from anemia to anemia method direct-match :rf 0.55) #s(expl-struct slot (has-findings) from chills to chills method direct-match :rf 0.65) #s(expl-struct slot (has-findings) from high-fever to high-fever method direct-match :rf 0.65) (((has-findings) (heart-murmur cardiomegaly high-hr loss-of-vision dyspnea-on-exertion sweating) nil))

The system continues with elaboration of the similarity between the new case and the other activated cases. During this process, the similarity assessment between case#39 - which proposes endocarditis - and the new case produces the following explanation:

:from dental-surgery :to acupunctur
:method (dental-surgery (bacteriemiae-triggering-non-cardiac-surgery instance-of 0.8) (bacteriemiae-triggering-factors subclass-of 0.8) (bacteriemiae-triggering-recent-therapy has-subclass 0.8) (acupunctur has-instance 0.8)) :rf 0.36864)
The system finds a similarity between dental surgery and acupuncture in the context of again, endocarditis.

ConSID-Creek detects these similarities when it employs a deep similarity assessment where general domain knowledge including common sense knowledge in the periphery of medical domain is utilized. These similarities were not possible to detect in the initial surface similarity assessment which took place during ‘activation’ task. Detection of a similarity between malnutrition and bad-econ-sit is important when the hypothesis ‘endocarditis’ is considered as the cause of patient’s problems. The detected similarity between these features contributes to the total similarity between case#26 and the new case. In this way, similarity on a contextual basis improves the quality of the similarity assessment as it increases the plausibility of hypothesis that would not be detected otherwise with the same degree.

Additionally, there are some features which exist only in either the retrieved case #26 or the new case. The features that are observed in the retrieved case and not in the new case guide the system in deciding which information needs be gathered, as will be seen in the hypothesis-testing task.

The system hereof tries to explain the deep similarity of case-new to other activated cases. After the system has gone through assessing the deep similarity of all of the retrieved cases, the important task now is to select the hypothesis or the set of hypotheses which is plausible to pursue further, by a subsequent test.

Retrieved cases and their similarity with the new case:

(#s(focus-struct case-name case#26 sum 3.27) #s(focus-struct case-name case#25 sum 2.81) #s(focus-struct case-name case#44 sum 2.58) #s(focus-struct case-name case#5 sum 1.95) #s(focus-struct case-name case#6 sum 1.95) #s(focus-struct case-name case#9 sum 1.95) #s(focus-struct case-name case#27 sum 1.89) #s(focus-struct case-name case#46 sum 1.5) #s(focus-struct case-name case#62 sum 1.3) #s(focus-struct case-name case#57 sum 1.3) #s(focus-struct case-name case#heidi111 sum 1.3) #s(focus-struct case-name case#heidi112 sum 1.3) #s(focus-struct case-name case#39 sum 1.018) #s(focus-struct case-name case#17 sum 1.018) #s(focus-struct case-name case#19 sum 0.9524) #s(focus-struct case-name case#43 sum 0.9) #s(focus-struct case-name case#42 sum 0.9) #s(focus-struct case-name case#29 sum 0.9) #s(focus-struct case-name case#31 sum 0.65) #s(focus-struct case-name case#32 sum 0.65) #s(focus-struct case-name case#36 sum 0.65) #s(focus-struct case-name case#21 sum 0.65))

Finished retrieval.

The best case is: CASE#26
12.3.4 Selection of to-be-tested cases: M-Focus

At this stage the system evaluates the activated and explained hypotheses with respect to each other. Some will be selected in order to be further considered, according to their superiority to the others in the pool of activated hypotheses. First, the total similarity of all the cases after elaboration are displayed. The cases are ranked with respect to their similarity to the new case. We converted the system’s results to a table (table 1).

<table>
<thead>
<tr>
<th>case-name</th>
<th>total similarity</th>
<th>hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>case#26</td>
<td>3.27</td>
<td>endocarditis</td>
</tr>
<tr>
<td>case#25</td>
<td>2.81</td>
<td>endocarditis</td>
</tr>
<tr>
<td>case#44</td>
<td>2.58</td>
<td>endocarditis</td>
</tr>
<tr>
<td>case#5</td>
<td>1.95</td>
<td>sarcoidosis</td>
</tr>
<tr>
<td>case#6</td>
<td>1.95</td>
<td>pulmonary-tuberculosis</td>
</tr>
<tr>
<td>case#62</td>
<td>1.33</td>
<td>pulmonary-tuberculosis</td>
</tr>
<tr>
<td>case#57</td>
<td>1.33</td>
<td>atrial-myxom</td>
</tr>
<tr>
<td>. . . . . .</td>
<td>. . . . . .</td>
<td>. . . . . .</td>
</tr>
</tbody>
</table>

TABLE 1. Total similarity of recalled cases

The system selects for further consideration the cases which are most similar to the new case and the cases that need emergency intervention. The selected cases are displayed below. Of these cases, endocarditis is the most similar one, and also requires quick intervention.

   The most plausible case to be tested is: case#26
   To-be-tested hypothesis is: endocarditis (plausibility degree: 3.27)

Now, the second main task begins, namely, testing of the selected hypotheses.

12.4 Test of the selected hypotheses

The subtasks of this task determine the information needs and gather this information.

   Current task to be performed: t-test-hypothesis
   Current task to be performed: t-determine-hypothesis-testing-strategy

12.4.5 Strategy selection

The system now tries to select a strategy for testing the generated hypotheses. Case case#26 proposes ‘endocarditis’ as the candidate hypothesis. Since ‘endocarditis’ is a disease that must be treated immediately, and also since the most similar case (case#26) proposes endocarditis as a candidate hypothesis, a ‘confirm-endocarditis’ strategy is selected for further scrutiny:
The system at this stage tries to find a strategy for gathering information in an inexpensive and timely matter.

current-task to be performed: t-determine-info-needs

12.4.6 Determination of information needs

The new case description is

\[
(\text{has-findings} \ (\text{high-fever} \ (\text{chills}) \ (\text{anemia})) \\
(\text{has-location} \ (\text{well-established-hospital})) \\
(\text{has-recent-therapy} \ (\text{acupunctur})) \\
(\text{has-occupation} \ (\text{unemployed})) \\
(\text{has-sex} \ (\text{male})) \\
(\text{has-age} \ (30-35-year-old)) \\
\ldots
\]

For this case, the system has retrieved case#26 of which content looks like this:

\[
(\text{case}\#26 \ (\text{has-goal} \ (\text{generate-hypothesis})) \\
(\text{has-location} \ (\text{well-established-hospital})) \\
(\text{has-age} \ (30-40-year-old)) \\
(\text{has-sex} \ (\text{male})) \\
(\text{has-recent-therapy} \ (\text{acupunctur})) \\
(\text{has-findings} \ (\text{high-fever}) \ (\text{sweating}) \ (\text{chills}) \ (\text{anemia}) \ (\text{dyspnea-on-exertion}) \\
(\text{loss-of-vision}) \ (\text{high-hr}) \ (\text{cardiomegaly}) \ (\text{heart-murmur})) \\
(\text{has-allergy} \ (\text{allergy-against-antibiotic-1})) \\
(\text{has-condition} \ (\text{malnutrition})) \\
(\text{has-solution} \ (\text{endocarditis}))
\]

The selected information strategy is ‘confirm endocarditis’.

Task ‘t-determine-info-needs’ is decomposed into two subtasks: t-generate-expectations and t-determine-tobe-gathered. The first subtask proposes the core findings reasonable to expect in the current situation.

Current task to be performed: t-generate-expectations

The retrieved past case supports the determination of expectations.

T-generate-expectations:

The onset of each disease manifests itself in terms of symptoms, signs, and laboratory measurements. In our approach to the disease process (see figure 10.1-b) a disease may manifest itself differently across various patients. The patient-related context determines the way a disease manifests itself for a particular patient. The retrieved case#26 provides
the particular combination of context and manifestations. The system proposes that the
following manifestations (i.e., core features) would be expected in this patient case:

\[
(\text{high-fever}) (\text{sweating}) (\text{chills}) (\text{anemia})(\text{dyspnea-on-exertion})
(\text{loss-of-vision}) (\text{high-hr}) (\text{cardiomegaly}) (\text{heart-murmur})
\]

Of these high-fever, chills, and anemia are already available in the current situation. In
ConSID-Creek the differences between the candidate cases and the new case constitutes
expectations. Now, the system will establish information goals from expectations.

Current task to be performed: \text{t-determine-tobe-gathered}

\text{t-determine-tobe-gathered:}

To be gathered information is:

\[
((\text{sweating}) (\text{dyspnea-on-exertion}) (\text{loss-of-vision}) (\text{high-hr}) (\text{cardiomegaly})
(\text{heart-murmur}))
\]

The system sets up these as ‘information-goals’ which essentially comprise part of the
internal context. Now the external context for information gathering will be established.
This context, together with information goals and information gathering strategy will con-
stitute a case which serves as ‘new case’ in the planning stage.

Now the system investigates the possible interactions among various patient-related and
environment-related context features.

\subsection{12.4.7 Establishing the information-gathering context}

This task involves a preparation stage before the generation of plans. It constructs a plan
case description which specifies the particular information goals, information gathering
strategy and the ‘real’ environmental contextual features.

Current task to be performed: \text{t-establish-info-gathering-context}

This task first establishes the ‘real context’ that will be used for planning. It does this by
creating and adding to the description of the new case three slots labelled as \text{has-available-
medic}, \text{has-available-equip}, and \text{has-available-personnel}. These slots indicate the result of
the interaction between the patient-related and environment-related context of the new
case for which the system aims to generate a plan. In a sense, these three slots summarize
the consequences of ‘togetherness’ of the special combination of the location and the
patient.

In order to do this, the task frame specifies a perspective which guides the system to focus
on the relevant parts of the information and knowledge related to the external context. The
perspective is specified through four slots: \text{has-rlv-spreading-environment-relation}, \text{has-
rlv-spreading-patient-relation}, \text{has-rlv-patient-info}, and \text{has-rlv-environment-info}. The
system displays the frame:

\[
>\text{#t-establish-info-gathering-context}
\]

\text{t-establish-info-gathering-context}

\text{(has-method (m-infer-info-gathering-context))}

\text{(has-rlv-spreading-environment-relation)}
When establishing the real-context, not all of the available patient-related context is taken into account. For example, as far as we understand, the ethnic origin of the patient is not relevant in planning. Just to emphasize, this may not necessarily be correct. We do not intend to introduce a perfect system in the sense that it behaves in the medically most correct way. What we try to do is to introduce a mechanism which is able to reveal focused, selective processing of information, and a mechanism for the selection of actions in an effective and plausible way.

The slot ‘has-location’ summarizes a bundle of information, for example what equipment is available and can be used for information gathering purposes, and what medical and skilled personnel exist at the location. In fact, all the environment-related context is summarized under the ‘has-location’ slot when the patient case is described at the start. This is because we assume that resources (personnel, equipment, etc.) are dependent on location. This may not be entirely correct. In this case, the environmental context slots should be explicitly included in the case contents.

The system distinguishes between various types of locations where a patient problem may happen to be solved. For each type of location, it specifies the available personals, equipment, and medical. For example, for well-established-hospital, the environmental context is specified as follows in the system.

well-established-hospital
    (has-available-medic medic-group-5 medic-anesthesia antibiotics)
    (has-available-equip xr-machine echometer manometer scope catheter
    sproeye wire de-machine sx-qt-machine ewd-machine s-oscopy
    (has-available-personal skilled-xr-personal)

The system sets up the environmental context for the care place. The environmental context:

((has-available-medic medic-gr-5 anesthese antibiotics) (has-available-equip xr-
machine echomachine manometer thermometer scope catheter sproeye wire de-
machine sx-qt-machine ewd-machine) (has-available-personal skilled-xr-person-cardiolog))

The special conditions of the patient may interact with the choice of information gathering actions. Therefore the system explicates such conditions, if there are any relevant available:

The relevant patient context:
    (allergy-against-antibiotic-1 30-35-year-old male)
The real context describes what equipment and medicines are both available and applicable in the location when the patient’s context is considered. For example, although all kinds of antibiotics are available in a well-established hospital, the current patient cannot tolerate a group of antibiotics (i.e., allergy-against-antibiotic-1, cannot-tolerate, antibiotic-group-1). The antibiotics in the system are grouped into three categories: groups 1, 2, and 3. As antibiotics-group-1 is not usable because of allergy (see the description of new-case), only the ones in groups 2 and 3 are usable. Therefore, in reality, antibiotic-group1 is considered like unavailable in the environment. The conclusive environmental context when taking into consideration possible interactions is presented as follows.

The environmental context for this patient:

((has-available-medic medic-gr-5 anesthesi antibiotic-2 antibiotic-3)
(has-available-equip xr-machine echomachine manometer thermometer scope catheter sproeye wire de-machine sx-qt-machine ewd-machine) (has-available-personal skilled-xr-person cardiologi))

After the information gathering strategy is determined, the information goals are established, and the real context is set up. The system constructs a case (i.e., an auxiliary one) embedding these and refers to this new case as case-new during planning.

The case-new looks like the following:

> #aux-case
(aux-case
(has-goal (generate-plan))
(has-info-goal
(sweating) (dyspnea-on-exertion) (loss-of-vision) (high-hr) (cardiomegaly)
(heart-murmur))
(has-age (30-35-year-old))
(has-sex (male))
(has-allergy (allergy-against-antibiotic-1))
(has-recent-therapy (acupuncture))
(has-location (well-established-hospital))
(info-gather-strategy (confirm-endocarditis))
(has-available-medic (medic-gr-5) (anesthesia) (antibiotic-2) (antibiotic-3))
(has-available-equip
(xr-machine) (echomachine) (manometer) (thermometer) (scope) (wire)
(catheter) (sproeye) (de-machine) (sx-qt-machine) (ewd-machine)
(has-available-personal (skilled-xr-person) (cardiologi)))

Now the system makes a plan for gathering the needed information. T-make-plan is invoked, which further invokes its two subtasks: T-generate-plan and T-test-plan.

12.4.8 T-generate-plan

T-generate-plan is accomplished via its two subtasks: T-formulate-plans and T-select-to-be-performed-plan. First, a number of plan-cases which seem initially able to achieve the information goals in the current situation are activated.
current task to be performed: t-generate-plan
current task to be performed: t-formulate-plans

12.4.8.1 T-formulate-plans - Activate

Frame t-formulate-plans provides a perspective that imposes a selective activation of cases. On request it is displayed:

> #L-formulate-plans

t-formulate-plans
  relevant-feature-slot value (has-info-goal info-gather-strategy has-location)
  has-goal value generate-plan
  has-method value m-formulate-plans-by-retrieval

The indices used for activation are referred to by the slots has-goal, info-gather-strategy, info-goals and has-location. The first one constrains the type of plans to retrieve and reflects the current task. The following two are the internal type of context, and the last one is of type environment-related (i.e. external) context. The patient related context has not been used as indices in order to speed-up the process. The intention is to eliminate first the plan cases that do not match the current information gathering strategy, the information goals and the location. The information gathering strategy constitutes a high-level index that gives a general idea of what the system needs to know in order to achieve the strategical goal. Thus the match between information gathering strategy of a past case and the current case ensures that irrelevant "information gatherings" are avoided. Information goals further specify what information is needed but is not available at the moment. So, together these make the system capable of determining which plans agree with the needs and intentions of the reasoner/system.

The location is also a high level index which gives a general idea of the type of environment. It provides insight into which resources may be available in the environment.

Based on these three index groups, the system determines the relevant features for the task:

Relevant features:
  (well-established-hospital confirm-endocarditis sweating dyspnea-on-exertion loss-of-vision high-hr cardiomegaly heart-murmur)

The system subsequently displays the past cases in which these features have been relevant:

The features are relevant in the following cases:
(heart-murmur ((plan\[#5 0.75) (plan\[#alp1 0.75) (plan\[#3 0.75) (plan\[#2 0.75))))
(cardiomegaly ((plan\[#5 0.75) (plan\[#alp1 0.75) (plan\[#3 0.75) (plan\[#2 0.75))))
(high-hr ((plan\[#5 0.75) (plan\[#alp1 0.75) (plan\[#3 0.75) (plan\[#2 0.75))))
(loss-of-vision ((plan\[#5 0.75) (plan\[#1 0.75))
(dyspnea-on-exertion ((plan\[#3 0.75)))
Five cases are activated.

Activated cases are

\[
\begin{align*}
&\text{(plan}\#5\,0.69375\,7)\,(\text{plan}\#3\,0.6125\,6)\,(\text{plan}\#2\,0.51875\,5) \\
&\text{(plan}\#1\,0.51875\,5)\,(\text{plan}\#1\,0.33125\,3))
\end{align*}
\]

At this stage, plan\#5 seems to be the one most similar to the current case. We will see whether it will continue to be so when the activated cases are elaborated by also considering patient-related context, after the 'explain' stage.

Now the activated plans will be elaborated in order to see whether they are in 'deep' agreement with the current case. The one which seems most plausible to apply after elaboration will then be put into execution.

**12.4.8.2 Selection of the plan to be performed:**

- current task to be performed: t-select-tobe-performed-plan
- current task to be performed: t-elaborate-formulated-plans

**Elaboration of similarity - Explain.**

This process works in the same way as in elaboration of explanation cases (see section 12.3.3). Additional similarities between each activated case and the current case will be investigated.

The elaboration involves a knowledge-intensive process. Thus the attention needs to be focused. This is realized with the help of the perspective specified at the task frame t-select-tobe-performed-plan. The frame of t-select-best-plan-cand is as follows.

``````
EXAMPLE

The slot has-slot-partition determines how to make the deep similarity assessment by specifying the slots to be compared in the activated and the current case. It can be seen from t-select-to-be-performed-plan frame that three partitions are considered.

A past plan consists of an information-gathering strategy, information goals, the control schema referring to the actions which can achieve the information goals and the information gathering method used for executing the actions in the control schema.

Here we take a look into the system's assessment of the similarity between plan#5 and the current case:

Explaining the similarity between the new case and activated past case, plan#5

The list of features to be compared:

- ((has-control-schema has-available-medic has-available-equip has-available-personal)
  - (m-ask-dyspnea m-ask-sweating m-measure-hr-by-x m-ask-loss-of-vision
    actmurmur-a m-measure-cmegaly-xr) (antibiotic-3 antibiotic-2 anesthese
    medic-gr-5 stethoscope ewd-machine sx-qt-machine de-machine wire
    sproeye catheter scope thermometer manometer echomachine xr-machine
    cardiolog skilled-xr-person))
- ((has-info-goal) (high-hr loss-of-vision sweating heart-murmur cardiomegaly
  dyspnea-on-exertion) (heart-murmur cardiomegaly high-hr loss-of-vision
  dyspnea-on-exertion sweating))
- ((info-gather-strategy) (confirm-endocarditis) (confirm-endocarditis)))

At the end of a perspective-directed spreading activation, similarities that are not detected in activation stage are detected. For example:

- from m-measure-cmegaly-xr to xr-machine
  method (m-measure-cmegaly-xr (xr-machine has-precondition 1)) rf 1)

In brief, the system found a similarity between m-measure-cmegaly-xr (in the past case-control schema) and a xr-machine (in current case- the 'real context'). This could be detected because, the deep knowledge provided 'm-measure-cmegaly-xr.has-precondition.xr-machine'. The past case is proposing the action m-measure-fever-by-termo, and in the current real situation a thermometer is available. So, there is a consistency between what the past case proposes and what can be applied in the current situation.

After the elaboration of all the activated cases is finished, the system will now select one of the cases for further consideration.

- current task to be performed: t-select-most-plausible-plan
12.4.8.3 Selection of the most plausible plan candidate - focus

First, the activated cases are ranked according to their overall similarity, and displayed.

(#s(focus-struct case-name plan \#5 hits-with-rf 6 average-with-rf 0.76 sum 4.6)
 #s(focus-struct case-name plan \#3 hits-with-rf 5 average-with-rf 0.79 sum 3.95)
 #s(focus-struct case-name plan \#2 hits-with-rf 4 average-with-rf 0.8 sum 3.2)
 #s(focus-struct case-name plan \#1 hits-with-rf 2 average-with-rf 0.85 0.56 sum 1.7)
 #s(focus-struct case-name plan \#alp1 hits-with-rf 1 average-with-rf 0.95 sum 0.95) )

The best case is plan#5

This result indicates the importance of patient-related context which has not been taken into consideration in the activation stage.

Now the system is ready to execute the retrieved plan.

12.4.8.4 Reuse of past plans: T-test-plan

current task to be performed: t-test-plan

Let us first ask the system to display plan#5.

>!Lplan#3
(plan \#5 (has-goal (generate-plan))
 (info-gather-strategy (confirm-endocarditis))
 (has-location (well-established-hospital))
 (has-info-goal (high-hr) (loss-of-vision) (sweating) (heart-murmur) (cardiomegaly)
   (dyspnea-on-exertion))
 (has-observed-result (dyspnea-on-exertion)
 (loss-of-vision (acquired-by-action (m-ask-loss-of-vision))
 (has-observed-result (loss-of-vision))
 (sweating (acquired-by-action (m-ask-sweating))
 (has-observed-result (sweating))
 (cardiomegaly (acquired-by-action (m-measure-cmegaly-xr))
 (has-observed-result (cardiomegaly))
 (high-hr (acquired-by-action (m-measure-hr-by-x)))
 (has-observed-result (high-hr))
 (heart-murmur (acquired-by-action (m-murmur-by-stet)))
 (has-observed-result (heart-murmur))
 (has-control-schema (m-ask-dyspnea) (m-ask-sweating (m-measure-hr-by-x)
   (m-ask-loss-of-vision) (m-murmur-by-stet) (m-measure-cmegaly-xr) ...)

Plan#5 gathers the following data

(high-hr loss-of-vision sweating heart-murmur cardiomegaly dyspnea-on-exertion)

These are gathered by the following control schema:

(has-control-schema (m-ask-dyspnea) (m-ask-sweating (m-measure-hr-by-x)
 (m-ask-loss-of-vision) (m-murmur-by-stet) (m-measure-cmegaly-xr)
The information goals shared by plan#5 and the current case are high-hr, loss-of-vision, sweating, heart-murmur, cardiomegaly, and dyspnea-on-exertion.

Current task to be performed: t-establish-retrieved-plan-ctx

Subsequently, the system modifies the control schema of the retrieved case so that it gathers only the information needed in the current situation. An auxiliary case is constructed having the modified control schema. This case is referred to as aux-plan-case and represents the plan to be performed. Now the system displays this plan case:

(aux-plan-case
  (has-info-goal
    (high-hr) (loss-of-vision) (sweating) (heart-murmur) (cardiomegaly)
    (dyspnea-on-exertion))
  (has-control-schema
    (m-ask-dyspnea) (m-ask-sweating) (m-measure-hr-by-x) (m-ask-loss-of-vision)
    (m-murmur-by-stet) (m-measure-cmegaly-xr))))

Now the system starts performing this plan.

**Execute actions.**

Current task to be performed: t-execute-plan

The execution starts:

Do you have loss of vision? (yes or no): yes

Okey. action m-ask-loss-of-vision is succesfully executed

Current task to be performed: t-check-preconditions

Current task to be performed: t-execute-plan

Preconditions of m-murmur-by-stet are not satisfied, therefore rejected adaptation is needed.

Current task to be performed: t-modify-plan

The action m-murmur-by-stet has the equipment related precondition, it uses s-oscopy which is not available in the environment. Obviously, stethoscope would exist in a well-established-hospital’ but we pretend as if it does not, in order just to create a problem with the execution of the plan. Now, the control schema needs to be modified. This happens through substituting this action by another which is applicable in the current situation.

**Modify control schema.**

As can be seen from plan#5 (displayed above) m-murmur-by-stet is for measuring heart murmur. That is, it realizes the task t-gather-heart-murmur. The system checks now whether there exists other methods that realize this task.

This action realizes the task t-gather-heart-murmur

alternatives are (actmurmur-a actmurmur-b actmurmur-c)

preconditions of actmurmur-a: (de-machine)
preconditions of actmurmur-b: (sx-qt-machine)
preconditions of actmurmur-c: (ewd-machine)

The system finds actmurmur-c and actmurmur-b are applicable in the current situation. It randomly chooses actmurmur-a.

actmurmur-a can be executed
this action measures heart murmur. write the observed result: murmur
current task to be performed: t-check-preconditions
current task to be performed: t-execute-plan
measure-cardiomegaly-by-xr can be executed
This action is performed to see whether there is enlargement in the cardia.
Write the observed result: cardiomegaly
Okey. action m-measure-cmegaly-xr is succesfully executed

Now all the actions that have been planned are executed. The system should assess this new evidences in order to see that the generated hypothesis ‘endocarditis’ is the definite diagnosis.

current task to be performed: t-assess-gathered-info

ConSID-Creek in its current version stops here. The features that are expected but not exist in the retrieved plan are not gathered. These features can be gathered by retrieving another plan. However, in the current version this has not been implemented. The inductive task is not implemented either. Future versions should include the whole abduction-prediction-induction sequence, and more importantly should be made able to determine whether a new abduction-prediction-induction cycle needs be invoked.

12.5 Summary

In this chapter we illustrate how ConSID-Creek employs a task-oriented perspective mechanism that features the system’s focusing mechanism.

Selection of the elements that mediate he capture of focus are specified at the task level. Contextual elements comprise a part of attention. Through various examples, we showed how different kinds of contextual elements facilitate different tasks. The system displays different behaviour (i.e., delivers different results) in the existence and lack of contextual elements. We have seen that in the case of the existence of contextual elements, the similarity of contextual elements with past cases improves case-based reasoning, while difference in context impede the same process. For example, the difference in context renders a need for significant adaptation of retrieved cases. This we have seen in the planning task, where some actions in the retrieved plan were not possible to employ, thus the plan had to be adapted. Our adaptation is a rather simple one, yet it shows how differences in context indicate a need for adaptation. An implication of this is that the use of proper contextual elements in retrieval may decrease the need for adaptation in the tasks following retrieval.
Chapter 13

Evaluation

13.1 Introduction

In this chapter we evaluate our contributions with respect to the research goals and the research strategy we described in chapter 1. This is not an easy task both because the evaluation criteria are not standardized, and because the thesis spans a rather wide scope of disciplines, as well as issues in these disciplines. Moreover, one of our aims is to connect the theories and methods within various disciplines in order to develop a unified, global approach to context.

Any scientific research requires evaluation which means making observations of all aspects of the research. Evaluation is an interpretation of how well the research meets its goals and contributes to its field. AI research, when compared to other sciences, have less standards for evaluation [Cohen 89]. This comes partly from wide variety of research undertaken in this field. Therefore a distinction between kinds of AI research will yield some hints on the selection of evaluation foci. Two main research subtypes are research in conceptual level targeting ill-defined problems and transforming them to better-defined design/architecture oriented research subtype, and research aiming at proposing algorithms and methods for solving well-defined problems (i.e., system oriented research subtype). Evaluation metrics for the two kinds of AI will be different. The important criteria for evaluating system-focused AI is studying how well a program performs, while it is both insufficient and inappropriate for evaluating analytical AI research. The study of context in this thesis falls into the first group.

Our development on metrics for evaluation for progress is inspired by the work by [Cohen 89]:

...some of the most informative observations are not performance measures but, rather describe why we are doing the research, why our tasks are particularly illustrative, why our views and methods are a step forward, how completely they are implemented by our programs, how these programs work, whether their performance is likely to increase or has reached a limit (and why), and what problems we encounter at each stage of our research ([Cohen 89], p35)

The thesis is of more investigative and analytical nature than of an implementation-driven one. Obviously, producing new ideas and approaches is one thing and testing these is another one being a complementary counterpart. The ideal is that new ideas and approaches are created as well as fully implemented and exhaustively tested. However, as the efforts allowed in the scope of a thesis is limited, we could not both create new approaches to the context notion bringing to bear the research results from a set of disciplines, and test these extensively on a fully implemented system. We had to give priority to either of these. This thesis made its choice in the direction of the first alternative and put
its most effort on developing a holistic account of the context since we view it as a prerequisite for starting a well-founded implementation. Therefore, our evaluation criteria and method will be tailored to this choice.

How can we evaluate weaknesses and contributions? What is the metric with which we measure the degree of the achieved success? We can make evaluation along two dimension. The first is evaluating the outputs of the implementation of the ideas. The other evaluating the cognitive plausibility of the work including why we did the research in the way we did. Cognitive plausibility is important at least for two reasons. First, if the research is cognitively plausible, it lends to support that it is “on the right track”, since our best examples of intelligent behaviour are humans ([Cohen 89], [Turner 94]). Second, if the knowledge structure and knowledge processing structure of a KBS system are cognitively plausible it should be easier to elicit knowledge from human experts.

13.2 Research goal and aimed contribution

In parts I and II the results form the theoretical part of our work is presented and discussed. In part III we designed a system that employs the results from part I and II, that is, context ontology, the context modeling methodology, the abductive inference description, and a context-sensitive iterative diagnostic. In chapter 11 we propose a method (CBR-based) for realizing the designed system, and in chapter 12, we looked at how this particular implementation of parts of CONSID-Creek works. CONSID-Creek is implemented for demonstrative purpose in order to tentatively explore which ideas of context theory and CONSID seem to be realizable in practice, and which ones seem to be working and feasible. In this we might be able to raise new questions, identify deficiencies, and thus, problems for future research.

We start with a reminder of our goal and the research strategy in section 13.1. From these we extract a set of criteria along which we discuss the contributions of the thesis.

- Our intention is to make context a natural, necessary and explicit component of a knowledge based system. We claim it deserves a deliberate attention and effort, and needs to have a status other than optional and arbitrary.

- A contextual knowledge ontology for diagnostic domains is provided

- Guidelines for how to analyse and model contextual knowledge in a particular domain are provided through a methodology.

- A context sensitive model of abductive and predictive inference is described at the knowledge level. The pattern of such an abductive inference is elaborated.

- In the ConSID architecture, evaluation of abductive process and abductive conclusions, and its connection with other inferences in a diagnostic process is modelled. A mechanism for shifting and capturing focus of attention is provided.

- In a demonstrator implementation a method for realizing the use of the proposed methodology, context ontology and the inference models is illustrated.

In order to evaluate the thesis’ contributions we raise to ourselves a set of questions:

1. What implies to consider context as vital in diagnostic knowledge base.
2. How expressive, useful, and general is the developed ontology.

3. How good is the methodology? Is it really feasible and guides the explication of contextual knowledge and its use.

4. How well, general and detailed is the abductive pattern? Does it cohere with and use the proposed context ontology?

5. Does the mechanism of attention focusing really delivers selective information processing?

6. How well ConSID-Creek combines the context ontology, diagnostic strategy, and the abductive pattern?

We devote a section to each of these questions (sections 13.2-13.8). In section 13.9, we briefly describe other AI approaches to context. In section 13.10 we mention some necessary future work.

13.3 What implies to consider context as vital in diagnostic knowledge bases

As Cohen states, for evaluation it is important to describe why we are doing this research, and why our views and methods are a step forward.

Context has been recognized and acknowledged to have a prominent importance in computational systems. However, it is an overloaded concept. There does not exist a proper/sufficient account of what it is, what its scope is and what is not. It is, in a sense, like a black box. About the inside of this box, it has been said only fragmentary things, here and there. Moreover, it has been presumed to known what context is, and the majority of the efforts put, is put on representing it and reasoning with it. The thesis adopts a deliberate and determined attitude toward understanding the inside of context.

Our work looks into context as including various aspects and puts these aspects together in a meaningful holistic model. It lays down a context sensitive model of the fault process underlying the onset and progress of the fault process itself, at the first place. The model is an improvement on the existing fault models in AI. The results of Schmit and Boshuizen’s study on the differences between expert and novices’ memory constituted a basis for our fault model. According to this theory, the expert’s memory develops episodes where enabling context is one of the main components. We enhanced their model of episodic content with another component, the outcome modifying context. This type of context has been referred to in research on medical reasoning in connection with, for example, the interactions between other diseases the patient may have had.

The thesis, then, provides an ontology and a methodology that coheres with and reflects this fault model. The ontology and the methodology constitute collectively a sufficient basis for organising the knowledge acquisition process and support knowledge acquisition. As such the main contribution of the thesis is in the KBS area, in particular in knowledge acquisition an structuring field.
13.4 Ontology

The term 'ontology is particularly fashionable in the research communities of knowledge sharing and reuse. It has a number of definitions, as there is no agreement on what an ontology is and how it can be used. We use the term as it is used in European ESPRIT KACTUS project: "explicit, partial specification of a conceptualization that is expressible as a meta-level viewpoint on a set of possible domain theories for the purpose of modular design, redesign and reuse of knowledge-intensive system components" (Schreiber 95). The existing ontologies are either on inferences or core domain knowledge. This thesis intended to develop a contextual ontology for diagnostic domains. If it is argued for context-sensitivity of KBS systems, then one should start with construction of a contextual ontology complementary to core domain ontology.

We can now evaluate the way we performed this research task. As there was not much in AI to start with, we started from exploring context studies in other disciplines. There is much research done in natural language understanding (NLU) field, however the diagnostic problem show significant differences from the NLU problems. Moreover, we didn’t encounter any ontology developed in this field. We turned our eyes to cognitive psychology which provided us really the essential ideas for the development of the ontology presented in the thesis. The context ontology we presented in section 8.3.1 is an amalgam of our interpretations of the experiments conducted in cognitive psychology on context effects in recall and theories of memory, applied to diagnostic problem. As such, the ontology is consistent with the theoretical and experimental results of context study in cognitive psychology. This increases its cognitive plausibility.

The ontology emphasizes classification of contextual knowledge along two dimensions. The first dimension is related to the general, often-asked question 'is context the internal state of the mind or does it represent the state of the world?' Since our answer is 'both', we include both internal and external contextual elements into our ontology. This is a significant contribution in AI because the existing systems either emphasize parts of internal context (i.e., goal-driven systems), or include some arbitrary and most obvious fragments of external context (e.g., age, sex) in their domain knowledge. Our ontology interpret both goal, age and sex as parts of contextual knowledge, makes explicit the differences between these, and comes with a formulation of the connections between these two types.

Another important aspect of contextual knowledge that needs deliberate handling is the way it is encoded. Contextual knowledge can be either independently or interactively encoded. During knowledge acquisition, this factor should guide the way domain knowledge is constructed and modelled.

Both the distinction between internal and external type of context, and the two ways of encoding them are rather general theoretical results. These can be useful in all domains where goal-driven and need-driven reasoning and learning is emphasized. Similarly, the encoding type is also a common important aspect for all types of problems where the reasoning algorithms makes a distinction between general, explanatory knowledge and concrete, specific knowledge.
Usefulness: Another type of evaluation is made on the ontology's usability in practice, i.e., in developing KBS. We exploited this ontology to organize the knowledge acquisition process and to support the computational design of ConSID-Creek. It has shown to be able to guide these tasks. The human experts have felt comfortable with the way we communicate with them and had no significant difficulty to match their expertise with our questions guided by the ontology.

Expressiveness and generality: The developed ontology is assessed by evaluating its coverage with regard to the context sensitive disease process (figure 10.1) which has been evaluated by human experts. Based on the expert judgments, the distinction between enabling context and outcome modifying context seems reasonable. However these expert judgments are limited to the medical domain, and therefore can not make statements about the generality of the model. On the other side, the fact that the ontology is inspired from cognitive psychology studies and these studies are scattered through a wide scope can be taken a positive implication for the evaluation of its generality.

13.5 How good is the methodology?

Before everything else, we need a methodology for eliciting and modeling contextual knowledge. In this section we evaluate the methodology developed and presented in the thesis. In order to do this we need to answer whether the methodology was really usable, useful, and accomplishes the task it is intended for.

These question can be answered in two ways. The first is by evaluating the validity of the methodology with respect to the established methodologies provided by knowledge-level community. The second is by using it in the real world applications, during implementation of a KBS. In this thesis, the second alternative was not possible, as our work didn't involve an implementation of a fully detailed and tested system. None the less, an evaluation can be made limited to the demonstration system - ConSID-Creek.

In order to achieve the research subgoal of developing a methodology for analysing and modeling contextual knowledge we studied and evaluated the current knowledge analysis and modeling approaches. Inspired from these, which studied the core domain and task knowledge, we developed a methodology for contextual knowledge and its connection with task knowledge in diagnostic domains. We argue that he cognitive plausibility of the methodology stems from its agreement with the existing methodologies that are proved to be useful in practice.

We explored the usefulness and feasibility of this methodology in the medical domain both when consulting the medical experts, and when structuring the gathered knowledge into a domain model. It rendered to be practical to use with regard to knowledge acquisition from experts as well as in design phase.
13.6 How well, general and detailed is the abductive pattern? Does it cohere with and use the proposed context ontology?

An abductive pattern should explicate the criteria for evaluating the quality of an inferred conclusion and for justification of the efficiency of the process producing this conclusion.

Existing abduction accounts, both in philosophy and in AI remains insufficient in this respect. Recently a particular research community, named 'abduction-induction' community, started to explore and reinterpret the philosophical and AI accounts of abduction. A number of workshops are arranged for this purpose. The accounts that views abductive inference from a knowledge level view, in terms of task, method, domain knowledge triplet have gained support [Öztürk 97b]. This thesis follows this line and provides a knowledge level account of abductive inference bridging to the philosophical account in the literature. The quality of the abductive conclusions are related to the concept of plausibility which philosophical views state as having two key evaluation criteria:

- the abductive hypothesis should be likely in itself, and
- should render the manifestations likely

Our account matches the enabling context and outcome modifying context into these components, respectively. However a pitfall of our pattern is that it does not explicate these two types of context, but refers only to context, without elaboration:

D is data available so far
H explains D together with Cxt
H coheres with context Cxt
H is the best hypothesis for further scrutiny
It is plausible to test H first

For example, the pattern does not explicitly refer to different roles of context, i.e., enabling-context and outcome-modifying context. This is because we attempt to use the same pattern both in diagnostic hypothesis generation and plan generation. In order that it could be used in both tasks, it should be general. However this seems to conflict with our overarching goal of explicating context as much as possible. The solution we choose is to keep the pattern general at the knowledge level but make a distinction in the symbol level, when implementing. However, this is not an ultimate solution but needs at the first place a thorough theoretical study of whether the underlying inference mechanism is really the same for hypothesis and plan generation tasks.

A point worth special attention is in relation to, again, generality. The term 'explains' in the pattern is used in a looser meaning than a 'causal' one. That is, explanation can be made by using various relations, under the condition that for each task the system needs to know what these relations are (just to remind, these are defined in terms of spreading relations in ConSID-Creek). This clarification may make possible the use of this pattern across different tasks, and also sheds light to the discussions on the relationship between abduction and induction [Flach 97].

216
Exactly this relaxed use of 'explains' increases the generality of the pattern in different tasks. In the scope of this task, the pattern is attempted to use in two different tasks; in generation of diagnostic hypotheses, and in generation of 'plan hypotheses'. However, an experimental judgment of the generality claim of the pattern is not made thoroughly. This is partly because the 'planning of information gathering' domain is not modelled as thoroughly as the 'generation of diagnostic hypothesis' domain. The overarching task is the second one and the focus is the continuity of abductive-predictive-inductive inferences. Planning is taken as the subtask of diagnostic hypothesis generation.

Thus, the pattern is an improvement but needs further clarification and fine-tuning.

13.7 Generality of the attention mechanism, its effects on relevance and efficiency

In this section we first evaluate the generality of the attention mechanism adopted in the thesis, and the scope in which it can be employed. The attention focusing mechanism assumes the existence of a knowledge level description of each task involved in the problem solving process. As long as the assumption is correct, there is no obvious reason for disfavouring the applicability of the mechanism in other domains than the diagnostic domain.

The task-oriented approach is not new, and bases on the well-formulated and tested views -such as Generic task view of Chandrasakeran. Our contribution is extending the approach in the epistemological level, and using it as a medium for specifying the focus of attention. More specifically, we modify a task description so as to express the contextual knowledge related to its accomplishment. As cognitive psychology studies declare the contextual information/knowledge as a natural and necessary part of the memory, this is in fact natural that a focusing mechanism is context-sensitive. The cognitive psychology studies does not focus on a particular type of memory task, and therefore the mechanism sounds to be a general, and applicable domain-independently. Moreover the mechanism works with both shallow and deep knowledge which means that it is appropriate for using in connection with both associative and explanatory reasoning paradigms. In addition, the way attention is captured in this thesis can easily be changed by modifying its specification in the task level. For example, if the reasoning method implies that the attention can not be established by using relevant-slots (which was tailored to CBR) but by something else, it is completely possible to make the necessary changes in the description.

Regarding the shifts in the attention, the focus mechanism is triggered by a goal change. Goal-drivenness is not new neither in social science nor in AI. To mention some examples, in AI SOAR [Laird 86], SWALE [Leake 92], and AQUA[Ram 89] are all reflecting goal-driven behaviour.

A shortcoming in our focus mechanism is that the trigger of a shift in focus happens through predefined goals, more precisely a predefined goal-chain. The goal-chain of the reasoner is almost defined in advance and reveals a rather static behaviour lessening the significance of the goal in the reasoning and problem solving of the reasoner. Nevertheless, the mechanism allows a kind of 'reactive' shift in the focus if it is applied in a system behaving in reactive manner. In fact, for the mechanism it does not make any difference
whether the goals are reactively set up or predefined since it is dependent on the existence of a task description including the goal that invokes it. So, the restrictions of the applicability and scope of the mechanism does not come from the mechanism itself, but the way we conceptualized its use.

A more implementation oriented evaluation of the mechanism, and evaluation of whether it works and promotes selective information processing will be made in the next section when evaluating the demonstration system.

13.8 How well the ConSID approach integrates the methodology, ontology, the inference patterns and the focusing mechanism

ConSID’s methodology for modeling contextual knowledge presupposes a task structure and an ontology. ConSID relies on an explicit representation of the diagnostic task structure at the knowledge level. Construction of this structure is based on the literature in philosophy which have extensively studied the reasoning process underlying scientific inquiry. Scientific inquiry is declared to be have a similar ‘thinking’ process as diagnosis. Also, studies in medial reasoning filed and the interviews with medical experts gave insight into the diagnostic reasoning.

The methodology guides the knowledge engineer towards mapping the task structure and context ontology in order to explicate what kind of contextual knowledge is needed in what task. Abductive inference is perhaps the heart of the diagnostic reasoning. We gave a special importance to the task involving this type of inference. An abductive pattern is developed in order to identify the role of context in this inference. That is, epistemological resources that the inference relies on is provided by the pattern.

In addition the task specifies also, its focus of attention through a set of abstract concepts and a set of relation types. So, the ConSID approach fully utilize the methodology and the ontology. Its reasoning line reveals selective and focused use of epistemological resources by using the focus of attention mechanism presented in previous chapters.

13.9 ConSID-Creek

This section explores and evaluates the correspondence of the implemented method and the proposed theory. In order to evaluate the realizability and utility of the context theory and the ConSID approach we implemented parts of it. This implementation we refer to as ConSID-Creek. The principle reasoning method is knowledge-intensive case-based reasoning (more concretely, Creek). We want to stress that this is just one method to realize the theory, and we believe that it may possibly be realized by other methods as well.

We will first elaborate why we think CBR is a good choice as a key method. Our reasons to believe that this method is appropriate for implementing the context theory are:
1. The ConSID approach proposes that expert’s memory is arranged as a collection of episodic memories. CBR provides case structure that is perfect for representing episodes.
2. ConSID favours pattern recognition as the main reasoning method for expert reasoning. Case-based reasoning paradigm is known to imitate this method well.
3. Knowledge-intensive CBR is appropriate to encode contextual knowledge both independently and interactively.
4. Creek can easily be modified into a task-oriented approach.
5. Creek supports implementation of context-sensitive focus mechanism, both in episodic and in semantic knowledge.
6. Creek framework is suitable to explicitly representing the internal and external context.
7. Creek’s activate, explain and focus primitive methods are suited to realize the ‘generate-and-recognize’ model of recall (for abductive inference) adopted in ConSID.
8. CBR supports context sensitive prediction through retrieved cases.

The application domain of ConSID-Creek is basically infectious diseases and the heart diseases. The knowledge base includes 10 different diseases. The case base comprises 39 cases that are acquired from two medical experts and from the literature. The knowledge base includes 32 disease hypothesis cases and 7 plan cases. The semantic knowledge is also acquired in the same way. The emphasis is given to the generation of disease hypotheses, while planning is of secondary importance.

From an implementation point of view ConSID-Creek comes in two parts: hypotheses generation and hypothesis-testing. Hypothesis testing involves planning for information gathering. However, planning by itself is not a research subject in this thesis. We have just explored whether the ConSID approach could apply the most straightforward type of planning, that is precondition checking, by employing the use of internal and external context notions.

An implementation can be evaluated either by comparing it with the behaviour of other implementations, or by using human experts for evaluation.

In our case, both of the general domain knowledge and the case base was too small for exposing it to extensive human judgment and expecting a trustable and informative assessment. Nevertheless we have illustrated to human experts the details of how ConSID-Creek behaves with respect to the context use for which it has been assessed to be akin to human expert’s use of contextual knowledge and in general as behaving reasonable.

We have also evaluated the system by comparison, not with other systems but by turning on and off the contextual components of the system. As we have said, ConSID-Creek is based on the methods of the Creek system applying a knowledge-intensive CBR method. We observed the behaviour of the differences between the two systems with respect to hypothesis generation as Creek implements only this part.

The important aspect of the ConSID approach is its use of a three-staged recall model, and its use of contextual knowledge in order to improve the performance of the recall results.
According to the statements of medical experts and of educators studying the expert and novice differences, a well-known phenomena is that experts do generate correct diagnostic hypotheses early in the diagnosis process. Once the correct hypothesis is ‘recalled’ it is most possibly ‘recognized’ later during the rest of the process. Therefore including the correct hypothesis, at the very start, into the differential pool is a kind of guarantee for ending up with the correct diagnosis.

In the system we explored whether contextual information may increase the possibility of recalling the correct hypotheses at the start.

We conducted two types of experiments related to M-activate and M-explain. The results are described in the following.

The activation with and without context:
We have experimented with the context effects on the activation. We have first run the system when the knowledge base did not include contextual knowledge (i.e., patient-related context). And then run the system with contextual knowledge included in the KB. The observations revealed that the use of context effected the reminding strength of the cases providing hypothesis.

Two observations are particularly important:
First, a difference is repeatedly observed in activating hypothesis when external context is used and not used. Both the activated hypotheses themselves showed difference, as well as their strengths of reminding.

Second, the difference diminished as more and more core features were available in the new case. Especially when the core features having high relevance degree in past cases the difference between employing contextual knowledge and not employing has decreased. So, the effect of using external context is directly proportional to the ambiguity of the available core information. However, we have not transformed this observation into quantitative assessments determining when exactly the use of external context becomes unnecessary. This needs be done in order to better understand when context is most useful.

It can be argued that patient-related-context in activation reveals only an ‘additive’ effect and is not different in its effect from the core findings. This is partly correct. We have mentioned in section 9.7.3 two kinds of context effects; additive and synergistic effects. This example falls into the first group, that is the effects mentioned here has an additive character. However this issue has a more pragmatical aspect. When considered from an epistemological perspective, contextual information is usually more available, or easily reachable compared to core domain information. Therefore when other type of information is scarce, because of its availability it plays and important role.

Explanation with and without context:
Also in explanation difference is observed when the context is turned on and off. When the context turned off, the ‘M-explain’ method have processed too far many information pieces (i.e., it processed features irrelevant for the current task), and activated larger portion of the semantic network compared to when the context is turned on. In other saying
the context-sensitive spreading activation seems to have capacity to focus on relevant parts of the knowledge base. This should have bearing on efficiency. However we did not measure the time dimension of ConSID-Creek.

The explanations that facilitate the similarity assessment when the similarities were not obvious at the first sight revealed quality differences. When the context is turned off paths connecting two concepts are proposed which were irrelevant for the current internal context, though they may be of high relevance in other context. When irrelevant explanations are used in reminding calculations, the quality of the similarity assessment are negatively impacted.

These evaluations have not yet converted into numerical measures, and are not sufficient to claim a definite success on behalf of context use. However, the observations made so far, and the responses of the human experts are encouraging and gives reason to continue the research in these lines in the future.

There are some shortcomings especially in application of the context theory and ConSID approach. Below we will mention these.

The relevance of contextual features are elicited from human experts in terms of importance and predictive strength. This has been a difficult task for the human experts because contextual knowledge is less explicated compared to core knowledge by the experts, but it is in their heads compiled. We have not checked one expert’s relevant judgments with other experts, because of time constraints in the thesis scope.

Connected to this issue again, we don’t have much idea about whether and how the importance and predictive-strength factors of core findings are dependent on the known contextual information. We had thoughts about that there could be a dependence between relevance of core findings and contextual features. This was, however, not easy to find out because the human experts couldn’t just make definite judgments of how the presence and absence of contextual information affects their judgment of importance and predictive strength of core features. So, though we mentioned only additive effects of external context in activation stage there may be synergistic effects resulting from a possible relation between the relevance factors of some core findings and the contextual features present in the situation. This remained to be explored.

Creek’s method is used for similarity computation after adapted to include contextual factors. When calculating the activity strength the degree of effects of core and contextual information is not exclusively experimented with. For example, in activation different degrees of effect could be given to core and contextual information bulks. We tested the system by only giving equal degree of effects. In the current implementation activation strength is equal to the sum of activation strength collectively of contextual information and of core information.

In addition, other experiments could also be done in order to find out the best ways of combining the effects of core and contextual information. Some alternatives may be:

(i) a set of hypotheses based on only core features and another set based only on core features are identified. Their union is computed modifying the reminding value of hypotheses that take place in both hypothesis sets.
(ii) another way to combine core and contextual features could be to take the intersection of the two hypothesis set.

(iii) yet another way might be to formulate hypotheses only on the basis of core features and then operating, only this set of hypotheses making another formulation, this time based only on contextual features. In this we could find out which method for recall would be most feasible.

Before doing this elaboration it will be difficult to argue how the context theory and CONSID can best be realized as an implementation.

### 13.10 Related work

The existing knowledge based systems that deal with context in one or another way gives different meanings to context and its role as well. The prevalent trends can be grouped as ones where:

- the notion of context is associated only to the reasoner's intensions, needs, and wishes. These are usually referred to as the common designation of goal-driven approach. Prominent examples in this group are Soar [Laird 86], SWALE [Leake 92], and AQUA [Ram 89].

- the notion of context is associated only with what we labelled external context. Examples are MYCIN [Shortliffe 76], and [Console 93].

- the notion of context which gives context a rather wide meaning including goals, environmental conditions, and in addition what we call core features. Context corresponds, in fact, to what we name as 'the focus of attention'(e.g.,[Halasz 92]).

Common to these approaches is that they do not model contextual knowledge at the knowledge level, as the main focus of these work are not the knowledge level but more system level.

They are most concerned with the behaviour of the reasoner. An important difference between this group of research and our work is that we are concerned also with the knowledge acquisition and structure and the construction of reusable components (both at the domain knowledge as well as inference and task level).

Another dimension is the explicitness of context. Again here, it is not obvious what should be explicit. In our work, the model of context is explicit, while in some approaches, its explicit representation is the main focus. The best-known example to explicit representation of context is CYC, and some logic-oriented approaches. CYC deals with the validity of propositions, and when (i.e., in which context) is a proposition valid. It tries to make propositions valid.

Our work does not stick to such a strict validity understanding. Diagnostic domains are weak and open domains, so we do not target to make all propositions valid as this is not possible achieve.

In this section we briefly describe a few well-known AI systems that in some way take context into consideration.
Soar:
Problem solving in Soar occurs as a search in problem spaces (Newell 1980). A problem space is comprised of states corresponding to situations in the domain and operators that transfer one state to another. Problem solving consists of employing a sequence of operators which starts from an initial state and ends with a desired state. Reaching a desired state is considered to mean achieving a goal.

A goal in Soar has three slots: problem space, state, and operator. A goal with its three slots comprise a context, in Soar terminology.

Achievement of a goal means instantiating these slots. That is, goal satisfaction involves a set of decisions related to the selection of problem space, state, and operator. Satisfying a goal begins with selecting the space. Subsequently, the initial state and the operator to apply to the initial state are selected.

In this framework, the term 'goal' is sometimes used interchangeably with 'context'. In that sense, Soar is only interested in internal context, more specifically with the internal interactive context in ConSIC-D-Creek's context ontology. Thus, a holistic approach to context is lacking in the Soar framework. This is apparently because Soar is particularly interested in 'shifts in the focus attention', but not with the context as a knowledge type. We may say that while Soar concentrates on the cognitive agent, it pays little attention to the environment in which the agent acts. So, regarding the debate of whether context captures state of the mind or the situation, Soar seems to choose the former. Nevertheless, Soar publications do not express a deliberate decision on why they ignore this situational component. In fact, Soar is not interested in the question of various dimensions of context. We believe this is because the Soar framework is primarily concerned with general problem solving strategies, and probably therefore pays less attention to contextual knowledge as a special type of knowledge, but addresses the issue of how context influences the selection of portions of domain knowledge. The Soar publications do not show any effort to analyse how that domain knowledge could be augmented by contextual elements.

CYC:
Microtheories of CYC have been used as a means for storing information which if stored in a uniform knowledge base, would lead to inconsistency [Guha 90]. That is, knowledge is partitioned into microtheories which allow representations of multiple worlds. So, CYC, in contrast to Soar, is more interested in the representation of contextualized knowledge. A microtheory, in a sense, resembles Creek's cases.

In CYC, to say that the assertion worksFor (Blair CycGroup) is an axiom of the JobMt, one can write

\[ \text{ist} (\text{JobMt worksFor} (\text{Blair CycGroup})) \]

where 'ist' stands for 'is true in'. This notation is used to state that an axiom is true in a certain context.

JobMt corresponds to a 'perspective' in our work, which impose a focus on a portion of domain knowledge. In our work a perspective imposing such a focus (e.g., JobMt of Cyc) is described at the task level. Hence, the current goal, via task, indicates where to focus. Correspondingly, a shift of focus happens through setting up goals. Goals are modelled in
our system while Cyc does not emphasize the internal context as a type of knowledge to be modelled at the knowledge level.

In fact, neither CYC nor Soar provides a thorough context model at the knowledge level.

**Medic and Orca:**

Medic and Orca adopts a schema based approach where contextual knowledge is represented in the form of schemas [Turner 94]. The focus of Turner’s work is the representation of context. Our approach differs from Turner’s work primarily in that we use cases instead of schemas to capture episodic contextual knowledge. Maintaining unique episodes as cases, instead of generalizing them to schemas, avoids the abstraction away of details that may be useful for the problem-solving process. In our view, it is left to the problem solving process to decide what to reuse of a past experience, and make the necessary generalizations on the basis of the actual problem situational instead of a predicted one. This enhances system flexibility.

### 13.11 Summary

In this chapter we evaluated both our context ‘theory’, and a partial demonstration of this theory in medical domain. We argued that the theory is useful and that parts of the theory works and promises improvement on the diagnostic reasoning process and conclusions. The main shortcomings of our evaluation are:

- We cannot empirically argue that our approach is better than others since we make no comparison. This is because what we want to measure is not a system and its performance, but the theoretical contributions in the line of context understanding and conceptualization and how well this theory is realizable.

- We cannot empirically argue that our approach is able to achieve clinical quality diagnosis since we do not use medical experts for an extensive quantitative testing. This is because, it would be difficult to compare our results with human experts in that we would not be able to find out whether a difference between the decisions of a human expert and ConSID-Creek’s is originating from the shortcomings of the context part or the original Creek methods that our implementation bases on. Besides, the system is simply not suitable for such a thorough test as the case base is small, and the core domain knowledge is not fully constructed as our focus was the contextual knowledge.

- We made a set of tests to see whether parts of the context theory works. We have observed promising results especially in two respects and believe it is reasonable to continue with further implementation and testing of the theory. In particular,
  - contextual knowledge improves the quality in activation of cases
  - context-sensitivity of the focus mechanism increases the degree of relevance of the generated explanations during similarity assessment.

These two collectively have a positive impact on the conclusions of the abductive task. More precisely, the set of generated hypotheses shows difference when the context-sensitivity is turned on and off, regarding both which hypotheses are generated and the way they are ranked.
Chapter 14

Conclusion

In this chapter, we discuss some of the contributions our research makes to artificial intelligence and cognitive science and look at some future directions to ‘context’ research.

The possible contributions to AI can be formulated in connection with exploring how to increase the efficiency of reasoning processes and the quality of the produced solutions. However, before starting to implement a system that reveals definite improvements in these respects, a thorough research in the ‘knowledge level’ is vital. Our work can be classified as contributing most to this kind of research. The developed context framework is characterized with that:

1. contextual knowledge model agrees with the theories and experimental findings in cognitive psychology

2. the methods for context-sensitive, selective processing of knowledge agrees with the theories on ‘inference’ patterns in philosophy, theories on attention, memory and recall processes, and recent ‘situated’ trends in cognitive psychology

14.1 Contributions

The most important contribution of this research to AI is in the field of knowledge analysis and modeling because the thesis addresses basically the question of epistemological adequacy. In order to reason correctly, one needs to know which knowledge to use and how to use this knowledge. This research targets to define the important aspects of contextual knowledge, identify its elements, types, its relations to other types of knowledge/information, its role with regard to a diagnostic process, and elaborate its use, first by human, and then demonstrate how all these can be used in KBS context.

The thesis explicates a holistic model of context knowledge, including its components pertinent to all dimensions of a reasoning process: the reasoner, the target, the environment. The choices made are not arbitrary, but are made with the motivation of capturing contextual knowledge as exhaustively as possible. The holistic model should serve as a source of inspiration for context awareness and encourages to look at context deliberately. The context methodology together with the ontology also shows how to discover the context elements and roles from a context-sensitive model of the problem under consideration.

In order that a context understanding can be reflected in a computational system, first of all the understanding should be modelled. This we do at the knowledge level, because we claim that a model that can be reused should be modelled at the knowledge level. Both the ontology and the methodology developed for contextual knowledge should be reusable. In the same line with the core domain knowledge (which is the only focus at the knowledge level accounts so far) the contextual knowledge also needs to be modelled in this level. Because, otherwise only its fragments can be used in a KBS. In addition, reuse becomes
impossible. Though we targeted diagnostic domain, the developed ontology abstracts from the context studies on human in a large variety of tasks and, therefore, be reused in other domains than diagnosis as well.

14.1.1 Epistemological adequacy
In this thesis we make a distinction between process and content related knowledge. The building blocks of process knowledge are the tasks, which are connected to methods. The process knowledge captures the dynamic part of a computational system while the content knowledge captures the more static part. Our contributions has been in the line of extending both parts with contextual components. In fact, we construct a thorough content knowledge and explain how this content is used by processes.

A distinctive characteristic of the Creek framework is the knowledge intensive character of its case-based reasoning approach. In our knowledge model, both deep knowledge and case knowledge is augmented by contextual elements. In the demonstration system we instantiated the contextual knowledge model in the medical domain. An excellent advantage of case representation is that it can embed any cluster of information. However, the modelling decisions about the content of a case are rather important. Most of the existing applications include only core domain knowledge, ignoring the context. So, even though the nature of case representation is very suitable to represent knowledge within its context, this has not been fully utilized so far.

In deep knowledge, we modelled the contextual knowledge from the periphery of the core domain as well as within the core domain itself. So, we prepared the necessary knowledge basis which may be needed for context-sensitive processing in AI systems. Epistemological adequacy is the first requirement for such systems.

Acquisition of contextual knowledge: The knowledge engineer responsible with the knowledge acquisition task, should understand how the experts in the domain perform their jobs in order to explicate the contextual knowledge that is naturally, and often implicitly used. In order to explore knowledge one needs to understand the basic concepts in the domain. The high-level context ontology constitutes the basis for acquisition of contextual knowledge. The contextual knowledge pertinent to the chosen application domain is acquired by instantiating the generic diagnostic contextual concepts in the ontology.

14.1.2 Adequacy of information processing
As described throughout the thesis, we adopted a generate-and-test reasoning strategy for realizing a diagnostic task. This strategy decomposed the overall tasks into abductive, predictive and inductive subtasks. We argued that context effects all these subtasks. However, we dealt only with the first two of these and showed how context may improve these processes. A particular emphasis is put on the abductive subtask.

We demonstrated context effects in abduction by demonstrating its effects in

- accessing (i.e. formulating) relevant (i.e., high quality) hypotheses
- elaborating the explanatory coverage of the activated hypotheses
- evaluating the plausibility of hypotheses

These have been declared as important factors for abductive reasoning [Josephson 94]. Our abduction view features a distributed evaluation. This means that hypothesis formulation integrates an evaluative process. The plausibility of a hypothesis is assessed within a certain context during formulation and elaboration.

Our abduction view differs from Josephson’s [Josephson 94] in several respects, which we covered in preceding chapters. In brief, we distinguish between evaluation of a hypothesis against the findings, and the evaluation of the plausibility of generating a hypothesis. The last one is evaluating a hypothesis against the memories, that is, the background knowledge (both semantic and episodic). In order to realize that we developed an analogy-based abduction account. This implies a new approach to abductive process.

In ConSID-Creek, prediction is also a context-sensitive process. Since our knowledge model is able to represent several possible worlds in the form of cases, the system can explore consequences of a hypothesis within a particular context. This supports an efficient information gathering process, since determination of information needs can be limited to relevant consequences.

### 14.1.3 Basis of the context framework

This thesis attempts to answer the question of ‘what the notion of context means to diagnostic problem solving’, from two distinct but closely related perspectives. According to these perspectives, context is viewed

- as a knowledge type capturing portions of the focus of attention and
- as a means for triggering a shift in the focus of attention

The former perspective leads to analysing and categorizing contextual knowledge, while the latter imposes an analysis and categorization of contextual effects in different tasks.

Consequently, two main contributions of this thesis are

- analysing and modeling contextual knowledge
- exploring the relationship between contextual knowledge and the diagnostic process

We studied the contextual knowledge both as a part of general domain knowledge, and of episodic, case-specific knowledge. We studied also the connections between these two types of contextual knowledge.

Our starting point has been the differences between experts and novices regarding the performance issue. This led us to investigate the reasons underlying these differences.

### 14.1.4 Expert/novice differences

We all know that experts are superior to novices with respect to establishment of the correct diagnosis. This observation has drawn special interest from educational psychologists. They want to understand why this is so. Why the experts are better at finding the
explanation of observed anomalies? A group of scientists anticipated that experts use different reasoning methods than novices. Another group argued that the difference between experts and novices is due to the difference in types of knowledge they use. We argued in this thesis that experts use both different reasoning methods and different types of knowledge. They apply a 'pattern recognition' method that relies on a special structure of knowledge: episodic knowledge. Episodic knowledge consists of two main types of knowledge: core domain knowledge and contextual knowledge, encoded together.

14.2 CBR and External Context

Inspired by recently emerging trends in the research on clinical reasoning and educational psychology we adopted in ConSID-Creek a context-sensitive model of a disease process and a context sensitive view of the diagnostic process. 

Context-sensitive disease process: (see section 10.4). In this model, there is a relation between the contextual features (referred to as the enabling context) and the onset of the disease. Additionally, the relation between the diagnostic label and the consequences is not absolute but depends on contextual features (which we referred to as outcome-modifying context). Both types of contextual features mentioned here are of type external, patient-related context.

Context-sensitive diagnosis process: Context-sensitive similarity of two cases implies that these cases can be represented by the same specific model of the disease process (see figure 10.1). The aim underlying the retrieval of similar cases is to retrieve a model of disease-process which facilitates the interpretation of the information provided in the new case. Note that a disease-process is a contiguous process (starting with enabling context and ending with anomalous findings), which may be captured as an episode, a script. Context is an important element of such a script both as a feature by itself, and as a factor influencing the salience and relevance of other features (i.e., core features). So, a context-sensitive approach to similarity allows a better assessment of whether a past case is similar enough to the new case to facilitate the explanation of the new situation.

Explanation cases in the system represent a concrete disease process consisting of various components including enabling context, disease, outcome modifying context and findings (i.e., consequences). As such, this case model is able to represent the entire process, including all the related parts with regard to the issue of identifying a disease given a set of anomalous observations. The effects of context in similarity assessment in the system is twofold: additive and synergistic effects.

A case is most strongly reminded when it is pointed to by as much relevant information (i. e. features) available in the current situation as possible. However, as the domain is characterized by ill-described problems, not all relevant findings are known at the outset. So, merely the similarity between currently available findings in the new and a past case will not be reliable enough. Therefore, from two directions, that is starting from contextual information as well as from findings, we may infer the disease with more certainty. This explains the importance of contextual factors in the activation of hypotheses; or, more specifically, the justification for why using contextual elements.
14.3 Multi-directionality of inference to the diagnosis

The traditional computational model of the abductive approach to diagnosis is an inference from core findings to a disease. Abduction is realized by CBR retrieval in ConSID-Creek, where it involves simultaneous inference from enabling context to diagnosis and the inference from collectively core-findings and outcome-modifying-context to disease. This is because a case wraps all the information and knowledge components necessary to take into consideration. ConSID-Creek supports holistic similarity assessment, which improves the retrieval of very similar cases.

14.4 Characteristics of diagnostic problems

We elaborated the reasoning methods that can suitably be used in diagnostic domains. The real-world diagnostic problems are characterised by ill-structured problems and weak domain theories. In these cases strong methods are usually not adequate. Weak methods such as those involving abductive inference are typically more suitable for such problems.

We further explored the relationships between the types of methods used in diagnostic problem solving and the type of knowledge they require. The relationship between abductive inference and contextual knowledge then became a touchstone in our work.

Starting from the resemblance between scientific inquiry and diagnostic inquiry, we identified two main subtasks underlying diagnosis: hypothesis generation and hypothesis testing. In our diagnostic model, abductive inference corresponds to the hypothesis generation subtask. This is the task wherein a hypothesis that explains the anomalies is adopted. So far we have said nothing about the method that realizes abductive tasks. This thesis argues that abductive tasks are realized by pattern recognition methods.

An expert’s generation of diagnostic hypothesis involves a pattern recognition method where the past experiences similar to the current diagnostic problem comes to the mind of the expert. It is this process that underlies the mystery of expert’s ability to include the correct hypothesis into the diagnostic pool generated at the very start. This implies that an important reasoning task involved in diagnostic process relies on ‘remembering’. Recognizing that, we turned to the research on learning and remembering in cognitive and educational psychology. This revealed that we have been even more aware of the relationship between learning/remembering and the contextual knowledge, and subsequently between contextual knowledge and abductive inference.

Diagnostic situations are ambiguous since the reasoner is often faced with incomplete information, especially at the start. The ambiguity may entail that too many hypotheses cross the mind of the reasoner. This is typically the case in abductive reasoning. As Peirce states, abductive inference involves a guess, albeit an educated guess. In this thesis we explore the question of what leads to this ‘educatedness’. Educatedness avoids generation of arbitrary and too many hypotheses. Our hypothesis is that educatedness has a close relation to experience and the ability to use additional information available at diagnosis time. By ‘additional knowledge’ we refer to contextual knowledge; the use of contextual knowledge reduces ambiguity.
14.4.5 Impacts of context in a diagnostic system

Taking contextual aspects into consideration implies a need for changing the models of diagnostic thinking as well as a need for another kind of domain model than the traditional one. Our demonstration application domain is the medical domain. The implication of contextual knowledge on the model of a disease process is prominent. Accordingly, the diagnostic reasoning model based on the disease model needs a significant revision. This thesis is a step towards a context-sensitive diagnostic domain knowledge model, and a context-sensitive diagnostic reasoning model.

The disease process can not be modelled as a strong domain model where the relationship between disease and its consequences is conceptualized as a one-world model. Our model conceptualizes the disease process as consisting of these knowledge types: enabling conditions, a disease entity, outcome modifying context, and the consequences. Accordingly, the diagnostic reasoning process takes into consideration all of these.

14.5 A high-level context ontology

An episodic memory is a collection of specific memories. Various types of context elements may accompany the target phenomenon within an episode.

The thesis presents a high level context ontology which is based on

i. a distinction between internal and external types of context, and

ii. a distinction between the type of encoding

iii. various ‘roles’ of contextual knowledge

Experiments (in cognitive psychology) illustrated that both ‘state of the mind’, and the environment and target-related characteristics can play the ‘context’ role.

Another series of experiments and subsequent theories emphasize a distinction between the two ways of encoding contextual knowledge. Baddeley refers to these as independent and interactive encoding. We argued that independent context can not be utilized in all places where interactive types of context can. Interactive types of context can be utilized both in shallow and deep reasoning, while independent context type can only be used in shallow reasoning. This difference is important for us, because in this work we present a hybrid diagnostic system that commutes between shallow reasoning and deep reasoning. During shallow reasoning, both interactive and independent context can be utilized by the system, while only interactive context is utilized during deep, knowledge-intensive reasoning.

14.6 The methodology for analysing and modeling contextual knowledge

Modelling the process of diagnosis involves identifying the reasoning subtasks which, when accomplished, lead to the achievement of the diagnostic goal, i.e., finding the fault causing the abnormal observations. It is exactly these subtasks that have a central role in
modelling a shift of attention. This is because the subtasks are the *loci* where a shift in the focus of attention may be triggered.

We developed a methodology in order to analyse and model the contextual knowledge at the knowledge level. This methodology can be presented in the form of a four-step process:

- construct a model of the diagnostic ‘process’ knowledge. This serves to explicate all the relevant application subtasks (i.e., diagnostic subprocesses),
- identify the *loci* of contextual effects (i.e., the subtasks where context may have an influence)
- identify the *role* that context plays at each locus (i.e., the way it affects each subtask)
- identify the *source* of context that can play these roles (i.e., referring to the ontology).

14.7 (Purpose-oriented) logical aspect of abductive reasoning

A frequently asked question is whether abductive inference has a logic. Philosophers answer yes, and provide a logical formula for abductive inference. We believe that in order to answer this question we need to clarify what kind of logic we are talking about. For example, is it a logic concerned with the preservation of truth, as in deductive logic? According to Herbert Simon, we ‘commonly call a process ‘logical’ when it satisfies norms we have established for it; and the norms derive from our concern that the process be efficacious or efficient for accomplishing the purpose for which it was established’ [Simon 73]. So, the very first thing to do, in pursuit of the answer of ‘whether generation of hypotheses has a logic’, is to identify the kind of logic we are dealing with. The kind of logic we refer to in this thesis is, in Simon’s words, ‘purpose oriented’ logic. The purpose of hypothesis generation is to generate a plausible hypothesis given a set of anomalous observations, and to do this efficiently. In fact, abductive inference has been referred to as ‘inference to the best explanation’, though a clear statement is not made about the characteristics of the explanation.

The question is now converted/reduced to the issue of establishing the norms that may help to judge whether the reasoning process delivers a satisfactory result. A logical formula may indicate the criteria or conditions on either the conclusion of an inference or the inference process that draws the conclusion. In this thesis we argue that contextual information, more specifically, the agreement between the contextual information and the inferred conclusion constitutes an important criteria for judgement of both the conclusion and the inference process. The inclusion of contextual aspects/norms in the logical formula of abductive inference is a step toward a computational model of abductive inference that generates high quality conclusions in a time-efficient manner. This is particularly important in diagnostic support systems since they often rely on abductive reasoning.

This thesis is a step towards establishing a basis for ‘plausibility’ of an abductively generated hypothesis. We argue that the use of contextual knowledge strengthens the plausibility of a generated hypothesis. Thus judgment of the quality of a hypothesis should take into consideration the contextual knowledge. And, rather importantly, this judgement can partially be done during the hypotheses generation stage. The important point to notice here is that our diagnostic reasoning account promotes a distributed evaluation of hypoth-
es; evaluation of hypotheses is not limited to the hypothesis-testing stage but an evaluation takes also place in the hypothesis-generation stage.

In our account, the condition that the hypothesis explains data is not sufficient. There may be data/information that the hypothesis does not explain, but should agree with. These we refer to as contextual elements which either explain the hypothesis, or together with hypothesis explain information/data. So, the form of the abductive pattern becomes more complete when augmented with contextual considerations:

\[
\begin{align*}
D & \text{ is data available so far} \\
H & \text{ explains } D \text{ together with } Cxt \\
H & \text{ coheres with context } Cxt \\
H & \text{ is the best hypothesis for further scrutiny} \\
\text{It is plausible to test } H & \text{ first}
\end{align*}
\]

On the other hand, the overarching task, e.g., inquiry or diagnosis, can be formulated as follows:

\[
\begin{align*}
D & \text{ is a collection of data} \\
H & \text{ explains } D \text{ together with } Cxt \\
H & \text{ coheres with context } Cxt \\
H & \text{ explains } D \text{ better than its alternatives} \\
H & \text{ is probably true.}
\end{align*}
\]

Notice that the scope of formula (ii) is larger than that of (abductive) formula (i) since the second one also includes testing.

### 14.8 Psychological aspects of abductive reasoning

This thesis contributes to the discussions on whether abductive inference has a logic and/or has any psychological basis. We do this by studying abductive inference at the knowledge level, and on the basis of a task-oriented approach where we relate the logical aspects of the abductive reasoning to the task concept, and the psychological aspects to the method concept. So, in our view, abductive inference has both a logical and a psychological aspect.

The logical form we provided in this thesis is context-sensitive and provides the criteria for evaluating the abductive process and its result. The method connected to the abductive task, on the other hand, elaborates how this task can be achieved. The method involves remembering, and relies on similarity assessment between two episodes (i.e. cases).

We summarized our claim of the nature of the interaction between the context and activation process as a chart, illustrated in figure 10.8.

Evoking hypothesis is based on how distinct the current case is and how similar it is to the past experiences. Increased distinctiveness and a constrained similarity judgement collectively facilitate ‘association by resemblance’ [Peirce 58]. The underlying context effects in hypothesis formulation are the following:
conclusion

- context comprises a reference point for judgment of similarity between ‘core’ cues presented in the new experience and the past experiences
- context increases the distinctiveness of the available information
- a good association can be accounted for by the notions of similarity and distinctiveness
- a good association entails enhanced recall
- enhanced recall implies good recognition as well, in the sense that the likelihood that the recalled episodes are similar to the current situation/case, with regard to the hypothesis, is high.
- a hypothesis dictated by a similar past episode may be more likely to explain the surprising facts. The idea is that the hypotheses formulated by retrieving a similar episode is plausible in the current situation since it has once been verified to be true in similar situations.

14.9 Synergistic and additive effects of contextual features

In CBR research, various indices have been used across different domains. However, the research on index and similarity has not given the required importance to the contextual features as a special kind of index. In our view, the neglect of contextual features in similarity assessment is a highly detrimental to the quality of similarity assessments, as well as the efficiency. The reasons are twofold:

1. The contextual features modify the evaluation of other type of indices, the core findings
2. The contextual features as independent indices would influence the total relevance of the collection of indices as a whole.

The first item implies the synergistic effects of the contextual features on the relevance of the non-contextual features, that is, findings. For example, a finding may be more or less important for a certain diagnostic label in different contexts. Ignoring this effect of contextual features on the importance and predictive strength of the findings on the solution leads to inflexibility.

The second item highlights the additive role of the contextual features on the relevance of the total information provided by the new case. This is similar to disregarding some available, relevant information which may decrease the distinctiveness of the problem, as we considered in the preceding chapter. Ignorance of this aspect of contextual effects will have negative effects on both quality and efficiency.

According to Dreyfus and Dreyfus, an appropriate account of pattern recognition which characterizes expertise, should not model the situation in terms of “objective features” but “as an unanalyzable whole” [Dreyfus 86]. In our understanding, by ‘unanalyzable whole’ he means the effects of ‘togetherness’ of a set of features. The contextual features play a special role in the togetherness respect and serve as a glue converting separate findings into a meaningful whole. This happens through their synergistic and additive effects on findings and together with findings, respectively.
14.10 Holistic case representation

A key notion that case-based reasoning is based on is ‘similarity’. The quality of the conclusions reached by performing CBR is extensively dependent on the quality of similarities detected between the new case and previous cases. So far the similarity assessments employed by various CBR variations consider only a portion of the entire phenomenon of the disease process. This was the portion that, for example considered the explanation path connecting disease to findings. Our similarity account takes, as a basis, the entire disease process, and thus applies a more holistic approach, by including contextual aspects.

The discussion on the use of contextual aspects in language is better known and examined than the one in diagnostic systems and in connection with the case-based reasoning paradigm. The discussions in language understanding are centered around the question of when context is useful; its role in deciding which meaning of an ambiguous word to activate, or in selecting the appropriate meaning among the activated ones.

14.11 A context-sensitive, iterative approach to diagnosis

This thesis discusses a diagnostic problem solving model, called context-sensitive iterative diagnosis (ConSID), which is particularly appropriate for open and weak domains. ConSID conceptualizes diagnosis as a context sensitive process, integrating reasoning and action. The main principles on which the ConSID approach is based can be classified as follows:

1. diagnosis is realized by an iterative application of the abduction-prediction-induction cycle (see figure 9.1).
2. diagnosis is a combination of explanation and planning
3. the problem solving process is internal context-driven, and a shift in the focus of attention is triggered by a change in the active goal the reasoner.
4. the focus of attention is captured in terms of core and contextual entities
5. pattern recognition is an important component of diagnosis at the expert level

According to the ConSID approach, a diagnostic explanation should

- explain the anomalous findings,
- explain them in terms of the desired explanatory relationship, i.e, the relationship, which is specified in advance, for linking explanans and explanandum
- agree with the contextual features pertinent to the situation. This condition ensures that the explanans (the fault) is consistent with the contextual factors of the faulty situation. For example, in the medical domain the risk factors comprise a part of contextual features. This type of contextual information does not directly explain the observations, but it does, in a way, imply the fault.
14.12 A context-sensitive, CBR-based method for diagnosis

We described the mechanism and methods underlying the context-sensitive behaviour of ConSID-Creek system. ConSID-Creek is based on the ConSID approach and employs context sensitivity in two ways:

- the focus of attention is captured in terms of contextual elements in addition to core domain elements.
- the line of reasoning is goal-driven and a change in the goal triggers a shift in the focus of attention.

The system's main diagnostic method is M-generate&test which decomposes the T-make-diagnosis task into T-generate-hypotheses and T-test-hypotheses. M-generate&test is used by the system in two places in our model of diagnostic problem solving. These are related to classification and planning tasks and are correspondingly used when

- generating and testing explanatory hypotheses: M-generate&test-hypotheses
- generating and testing plans for gathering needed information: M-generate&test-plan

We employ case-based retrieval in both generation of diagnostic hypotheses and generation of information gathering plans. Planning has not been the primary concern in this thesis, but we generate plans to illustrate our diagnostic approach as a whole. However, this thesis does not claim to make any significant contribution to the research on planning itself. Our main concern has been the classification stage, and its relation to the planning stage. Thus, planning is concerned solely as a part of the evaluation of educated guesses in the real world.

14.13 Future work

Our work constitutes a step in AI toward understanding and modelling contextual knowledge and its use in computational systems that support problem solving, in particular, diagnostic problem solving. The context notion has been commonly accepted as being important in adaptive behaviour. However, determined and serious attempts in the direction of a content theory of context within AI are scarce. Cognitive psychology is now in a position to give us some hints and descriptions which AI may use to form an operational theory of context. We investigate mechanisms for using contextual knowledge for improving the efficiency and the quality of computational systems. As such, we elaborate the connection between knowledge and ‘knowledge use’, including contextual knowledge.

We were not able to implement all of the parts of the system architecture we propose for diagnostic problem solving. The ‘inductive’ task (i.e., assessing the gathered information) has not been implemented at all. Consequently, a possible iterative application and a possible re-generation of hypotheses could not be demonstrated.

Our knowledge base was not big enough to properly test the context effects on the degree of improvement particularly in the efficiency of the system. In the future, larger case
bases, domains and context models are needed in order to observe the real consequences of integrating context into the system.

ConSID-Creek is currently unable to support medical practitioners in real use. The domain model needs to be improved, and the missing parts of the methods of the architecture should be implemented in order to assess the merit of the content theory of context proposed in this thesis.

Although learning has been an important aspect for diagnostic support systems as the mechanism for developing expertise, we could not explore learning in this thesis. In the future this will probably be a key issue for building more dynamic systems that can realistically support human diagnosticians.
References


[Barsalou 92] Barsalou, L. W., Cognitive psychology. Lawrence Erlbaum, 1992


[Haugeland 84] Haugeland, J., First among equals, in [Kintsch 84], pp 85-99


[Lester 91] Lester, J. C., Porter, B. W., Generating Context-sensitive explanations in interactive knowledge-based systems, AI 91-60 , University of Texas.

[Lockett 93] Lockett, A., Adaptsjon i Case-Basert resonnering, cand.scient (MSc) thesis, Dept. of Informatics, University of Trondheim, April 1995


[Öztürk 97b] Öztürk, P. An AI criterion for an account of inference: how to realize a task. Accepted in IJCAI’97 Workshop on Abduction and Induction in AI, Nagoya, Japan. 1997


[Thomson 86] Thomson, D. M. A model of cognition: Context effects and cognitive mechanisms. in European bulletin of cognitive psychology. vol. 6 No.2 April 1986


