DECISION SUPPORT IN DEMENTIA CARE
Developing Systems for Interactive Reasoning

Helena Lindgren

Department of Computing Science
Umeå University
Till Moa och Josefin
Once I rose above the noise and confusion
Just to get a glimpse beyond this illusion
I was soaring ever higher, but I flew too high
Though my eyes could see I still was a blind man
Though my mind could think I still was a madman
I hear the voices when I’m dreamin’
I can hear them say

Carry on my wayward son
There’ll be peace when you are done
Lay your weary head to rest.

- Kerry Livgren
Abstract

There is a need to improve dementia care in Sweden. The main issues discussed are how to improve the competence of medical personnel and the quality of diagnosis and intervention. In this thesis the process of developing a decision-support system for the investigation of dementia is described, as one means to meet the need. The resulting prototype system DMSS (Dementia Management Support System) has been developed in cooperation with domain experts, and has been evaluated and redesigned in the process in an iterative development process. The process involves the assessment of evidence-based domain knowledge and its characteristics, the assessment of the procedural knowledge residing in clinical practice, and reasoning processes. Further, the terminology and main reasoning process integrated in the system have been validated. Qualitative methods have been used for these parts of the project for the purpose of assessing as many different aspects as possible, and for practical reasons due to the limited access to domain experts, patients and primary care physicians in the area. Triangulation of methods has been applied in order to validate results in the process. The development has been extended to also include prototypes for Japanese clinical environments.

Clinical investigation activities are complex processes, which are situated, emergent and directed by the individual need of the patient, but also restricted or enhanced by the available resources at different points and at different care levels in the process. For the purpose of creating a system which provides support throughout the investigation process, the domain knowledge and the clinical investigation process was analysed and formalised in a conceptual model of clinical activity, developed based on activity theory and case studies of patients. The need for methods for the transformation of informal results from field studies into formal knowledge and design is addressed by providing the framework, which integrates the conceptual model of clinical activity and a method for the assessment and transformation of the knowledge to be integrated in a decision-support system.

The model was used to identify actions and their characteristics suitable for formalisation in a decision-support system. Several sources of domain knowledge need to be integrated that express the knowledge differently, which increases the demands on a formalism for representation. The work towards formalising the diagnostic reasoning process in both typical and atypical patient’s cases is presented, where the evidence in ambiguous cases is valued within dif-
different frames of references in order to improve specificity. Different logical frameworks have been applied, evaluated and developed using case studies of patients. Two lines of work towards a dementia logic and flexible guideline representation is presented; the defeasible, non-monotonic approach where many-valued dictionaries are used in a context-based argumentation framework; and the monotonic approach of integrating reasoning in a fundamental view of transformations between logics, using general logics as generalised and categorical framework.
Demensvården i Sverige och i andra delar av världen har på olika sätt varit i fokus de senaste åren där man påtalat behovet att utveckla metoder och riktlinjer för hur vården ska bedrivas. Detta för att möta den växande andelen äldre människor som också utvecklar demenssjukdomar. Nationella projekt har drivits, företrädesvis i syfte att förbättra vård och omsorg av personer med demenssjukdom, men även för att förbättra diagnosticering och behandling. I denna avhandling beskrivs utvecklingen av det dator-baserade beslutsstödet för demensutredning, DMSS (Dementia Management and Support System), som syftar till att fungera som ett stöd för personer som arbetar med att diagnostisera och behandla personer med kognitiv sjukdom. Domänen valdes även på grund av dess komplicerade kunskapsinnehåll, där bland annat en spännvidd av olika typer av symptom, komplexa kliniska mätmetoder sett ur ett formaliseringsperspektiv, starkt teamorienterat arbetssätt, ställer krav på hur kunskap ska och är möjlig att formaliseras och integreras i ett beslutsstödsystem för att det ska bli användbart i kliniskt arbete. De olika studierna och delprojekten som beskrivs i avhandlingen syftar till att tillsammans skapa en grund för utveckling av ett kliniskt kognitivt verktyg som stödjer och utvecklar användarens kognitiva processer (lärande, beslutsfattande, resonemang, etc.), samtidigt som det stödjer utvecklingen av det kliniska arbetet vari systemet ingår. I detta arbete fokuseras demensutredning som applikationsområde.

Analyser har gjorts av den vidare användarkontexten, resonemangsprocesser, domän- och processkunskapen uttryckt i evidensbaserad litteratur och integrerad i klinisk praktik, terminologier samt formaliseringsregler som kan hantera domänkunskapens egenskaper och användarsituationens krav. Prototyper har utvecklats och utvärderats i en iterativ process i samarbete med domänexperten, för användande i klinisk praktik i Sverige och Japan. För dessa studier har kvalitativa metoder använts i syfte att fånga så många olika aspekter som möjligt angående formalisering och interaktion, samt av praktiska skäl då det funnits begränsad tillgång till expertanvändare och patienter. Triangulering av metod har tillämpats för att validera resultat.

Kliniska utredningsverksamheter är komplexa processer, som är situerade, emergenta och styrda av individens behov, men även begränsade eller möjliggjorda av tillgängliga resurser på olika vårdnivåer i vårdprocessen. Det behövs metoder och verktyg som kan användas vid utveckling av system som syftar till att stödja dessa verksamheter. Det finns exempel på metoder som utvecklats för transformation av informell klinisk kunskap till en formell struktur som kan implementeras i ett beslutsstödsystem, där verktyg har utvecklats primärt i syfte att


Hittills har i första hand kvalitativa aspekter och syften varit i fokus i de olika projekten. Därför behöver varje utvecklingslinje ytterligare utvecklas med kvantitativa mål. Utvidgade utvärderingsstudier pågår, som syftar till att undersöka fördelning mellan olika nivåer av komplexitet hos patienter och vilken typ av stöd som behövs för respektive. När systemet är integrerat i daglig verksamhet kan faktorer som hur användande av systemet påverkar användaren och verksamheten undersökas.
Preface

This thesis consists of three parts with an opening introduction and motivation of the work. Part I includes a description of the prototype system DMSS and user evaluations, Part II describes the assessment of knowledge for the integration in the system and Part III the process of formalising the knowledge. The first two parts are structured mainly as a monograph, with a summary of Paper I included, while the third part includes the Papers II-V concerning the formalisation of the domain knowledge with a short introduction.


---

1 Paper II is reprinted with permission of Springer Science and Business Media.
2 Paper III is reprinted with permission of IOS Press.
3 Paper IV is reprinted with permission of Editions EDK.
Related Publications

The following papers are based on different studies presented in Part I and Part II of the thesis.

6. Lindgren H. Conceptual Model as a Tool for Assessing Knowledge in Decision-Support Development. Accepted as poster presentation to MEDINFO’07.
Acknowledgements

I wish to express my sincere gratitude and many thanks to all the people who in different ways have helped me to complete this thesis.

I would especially like to thank my research and thesis advisor professor Patrik Eklund who has made this journey possible. Thanks for structure, guidance, visions and enthusiasm! I have learned plenty of creative ways of how to handle the joggling balls of life, cooperative work and research when time schedules simply do not seem to fit all the balls.

Further, I want to thank Sture Eriksson for the vast medical knowledge, experience and guidance you have shared during this work.

I also want to thank Gösta Bucht, Ewa Ställdal and Noriko Korube for providing means and contacts for extending the research scope and cooperations. I want to thank Noriko especially, for your enthusiasm and patience in introducing me and my research results to the world of Japan, and definitely not least, for introducing the world of Japan to me!

I want to thank all physicians, medical personnel and patients who have participated in studies, in Sweden and in Japan. I want to thank professor Tetsuro Miki especially, who has coordinated the studies in Japan and contributed with valuable comments on the Japanese version. I look forward continuing our cooperation!

Financial support has been provided by the European Union, the Vårdal and Knowledge foundations, Äldrecenterum Västerbotten, the Kempe foundations and the department of Computing Science.

Friends and colleagues at the department, I thank you for support and for being there in the fika-room whenever I want to be distracted from work.

To you who have provided light in the Latex-darkness, I want to express my deepest gratitude: Pedher, Stefan and Erik, Thanks! For not only spreading light, but also colours at the end of the tunnel, thanks again, Pedher! Katti, Kicki, Ulrika and all of you friends who have added fun and rest to this journey, many thanks! For adding peace and patience when mine has run out, thanks Erik. Thanks Stefan for your enthusiasm and encouraging response, patience and understanding during both good and bad days the past years. For helping me out when completing this thesis and reminding me to express creativity in every possible way, thank you Josefin! For reminding me to pursue dreams with our inherited stubbornness, thank you Moa! For making every day a good day, for keeping my mind off work and for being the best of friends, I cannot thank you enough.
Finally, and most important, I want to thank my family for your love, support and patience with me in intensive work periods. While I have been working with the knowledge of handling cognitive diseases in theory, some of you have entered the dusk of life and the beyond, living them and dying from them. Maddy, Georg and Einar, I thank you for wisdom and all good memories. I miss you tremendously.

Umeå, May 2007

Helena Lindgren
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>AMPS</td>
<td>Assessment of motor and process skills</td>
</tr>
<tr>
<td>Behave-AD</td>
<td>Behavioural pathology in Alzheimer’s disease</td>
</tr>
<tr>
<td>BPSD</td>
<td>Behavioural and psychological symptoms in dementia</td>
</tr>
<tr>
<td>CBD</td>
<td>Cortico-basal dementia</td>
</tr>
<tr>
<td>CHAT</td>
<td>Cultural-historical activity theory</td>
</tr>
<tr>
<td>CDR</td>
<td>Cognitive Drug Research project</td>
</tr>
<tr>
<td>CDSS</td>
<td>Clinical decision-support system</td>
</tr>
<tr>
<td>CL</td>
<td>Clinical guideline</td>
</tr>
<tr>
<td>CT</td>
<td>Computed tomography</td>
</tr>
<tr>
<td>DLB</td>
<td>Lewy body type of dementia</td>
</tr>
<tr>
<td>DMSS</td>
<td>Dementia Management and Support System</td>
</tr>
<tr>
<td>DMSS-SJ</td>
<td>DMSS - Swedish-Japanese version</td>
</tr>
<tr>
<td>DSM</td>
<td>Diagnostic and Statistical Manual of Mental Disorders</td>
</tr>
<tr>
<td>DSS</td>
<td>Decision-support system</td>
</tr>
<tr>
<td>EBM</td>
<td>Evidence-based medicine</td>
</tr>
<tr>
<td>ECG</td>
<td>Electro-cardiogram</td>
</tr>
<tr>
<td>EEG</td>
<td>Electro-encephalogram</td>
</tr>
<tr>
<td>EPR</td>
<td>Electronic patient record</td>
</tr>
<tr>
<td>FAST</td>
<td>Functional assessment scale</td>
</tr>
<tr>
<td>FTD</td>
<td>Frontotemporal degenerative dementia</td>
</tr>
<tr>
<td>GDS</td>
<td>Global deterioration scale</td>
</tr>
<tr>
<td>GP</td>
<td>General practitioner</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical user interface</td>
</tr>
<tr>
<td>HIS</td>
<td>Hachinski ischemic score</td>
</tr>
<tr>
<td>ICD</td>
<td>International classification of diseases</td>
</tr>
<tr>
<td>ICF</td>
<td>International classification of functioning, disability and health</td>
</tr>
<tr>
<td>KB</td>
<td>Knowledge base</td>
</tr>
<tr>
<td>MCI</td>
<td>Mild cognitive impairment</td>
</tr>
<tr>
<td>MMSE</td>
<td>Mini-mental state examination</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>NDM</td>
<td>Naturalistic decision-making</td>
</tr>
<tr>
<td>PCC</td>
<td>Primary care center</td>
</tr>
<tr>
<td>SPECT</td>
<td>Single photon emission computed tomography</td>
</tr>
<tr>
<td>UMLS</td>
<td>Unified medical language system</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>VaD</td>
<td>Vascular dementia</td>
</tr>
<tr>
<td>WAIS</td>
<td>Wechsler adult intelligence scale</td>
</tr>
<tr>
<td>ZPD</td>
<td>Zone of proximal development</td>
</tr>
</tbody>
</table>
7.2 Results

7.2.1 Compliance with Clinical Guidelines and EBM from Japanese Viewpoint 45
7.2.2 Compliance with Clinical Work Process 46
7.2.3 Compliance with Expert’s Assessment of Diagnosis 46
7.2.4 Purpose and Scope of a Japanese DMSS 48
7.2.5 Terminology 49
7.2.6 Design of Interactive Reasoning 50

7.3 Conclusions 51

8 Current State of DMSS and Future Work 52

PART II ASSESSMENT OF KNOWLEDGE FOR DECISION SUPPORT 53

9 How do Humans Reason and Make Decisions? 54

10 The Model of the Main Clinical Reasoning Process 59
10.1 Compliance with Physicians 60
10.2 Compliance with Guidelines 60
10.2.1 Step 1: Determine the existence of cognitive disease 61
10.2.2 Step 2: Determine the type of cognitive disease 62
10.2.3 Step 3: Determine the type of dementia 62
10.2.4 Step 4 and 5: Investigate grounds for and establish interventions 63
10.3 Discussion 63
10.4 Conclusions 64

11 Assessment of Domain Knowledge 66
11.1 Review of Cognitive Disorders in EBM and Clinical GLs 67
11.2 Identifying and Diagnosing Typical Cases 72
11.2.1 Step 1: Determine the existence of cognitive disease 72
11.2.2 Step 2: Determine the type of cognitive disease 73
11.2.3 Step 3: Determine the Type of Dementia 74
11.3 Identifying Atypical Cases 76
11.4 Summary and Conclusions 78

12 Validation of Terminology 80
### Contents

12.1 Relevant Classifications and Terminologies 81
   12.1.1 The Disease Classifications DSM-IV and ICD-9-CM 82
12.2 Concepts vs. Instances 84
12.3 Different but Related Semantic Types 85
12.4 Sign or Symptom 86
12.5 Characteristics of Symptoms 87
12.6 Implications for the Terminology of DMSS 88
12.7 Conclusions 89

13 Assessment of Clinical Activity - Summary of Paper I 92

**PART III FORMALISATION OF KNOWLEDGE** 95

14 Towards Dementia Logic 96
   14.1 Summary of Logics Used in Paper I-V 98
      14.1.1 Paper II, III: Argumentation Frameworks 100
      14.1.2 Paper IV: Basic Many-Valued Logic and NPL 103
      14.1.3 Paper I, V: General Logics and Institutions 105
   14.2 Knowledge Representation for DMSS - Summary 107

15 Summary and Future Work 109

**References** 112

Paper I

Paper II

Paper III

Paper IV

Paper V
CHAPTER 1
Objectives and Scope

Work is more than ever influenced by a constant change of conditions and terms that direct in what settings, how and by which means work may and is expected to be done. Perhaps medical care is one of the historically most predefined work environments, where roles of profession and division of labour have been comparatively well founded and defined, in spite of the fact that the medical domain knowledge has been and still is rapidly evolving. But also medical practice is subject to changes. By introducing for instance electronic medical records into medical practice a change is made of professionals’ duties, pattern of cooperation and work methods. This emphasizes special concerns in the development of new tools aimed to be used in clinical practice. The aim for the work described in the thesis, is to create a foundation for developing a successful, mediating tool that supports and extends the user’s cognitive processes (learning, reasoning, decision making, etc.) and supports the development of practice while functioning as a mediator of activity. The activity in focus is the clinical investigation of dementia. Such tools in this particular context are commonly denoted clinical decision-support systems.

A distinction is made in the thesis between domain and procedural knowledge, where domain knowledge refers to knowledge assessed by evidence-based medical studies concerning what is known, while procedural knowledge concerns such organisational knowledge as how things should be done, when, by whom and in cooperation with whom. Both types will be accounted for in the thesis, since it is argued that a successful decision-support system has to provide both. The foundation consists of an assessment of the domain knowledge as well as the knowledge residing in the work context. This is described as a process of transformation of informal knowledge into formal knowledge integrated in a decision-support system when possible and appropriate, supplemented with a design of the user interface which mediates the knowledge and activity in focus. The concept of interactive reasoning is introduced in the thesis for the purpose of capturing this component of a decision-support system, which is both mediating in terms of activity theory, and yet functions as an active part in a reasoning process.

Investigations have been made of 1) the wider perspective of the context in which the support system is aimed to be used, in this case the investigation of dementia, 2) cognitive processes used in medical practice, 3) the domain knowledge, and 4) formalisation languages that meet the actual use situation and the qualities of the domain knowledge optimally. The prototype system DMSS has
been developed and evaluated. A model for the assessment and transformation of knowledge for the integration in a decision-support system is being developed, which can be used as a tool for identifying and formalising evidence for reasoning processes supported by the system. The cultural-historical activity theory (CHAT) [Vygotsky, 1978; Engeström, 1987] was used as theoretical foundation for the user evaluations and for the assessment model. For the most recent work on formalisation, general logics is used as generalised and categorical framework [Meseguer, 1989].

The application area of cognitive diseases was chosen because of the need to improve the quality of dementia care, and because the domain is distinguished by its complicated nature with vaguely described features to take into consideration, the mixture of all possible types of symptoms; physical, neurological, cognitive, psychological, social and behavioural, as well as the incomplete and ambiguous domain knowledge. It is also a domain in which teamwork is highly beneficial. The domain was considered sufficient to provide a wide range of issues concerning knowledge representation, interaction design and computer-
supported cooperative work.

The thesis is divided into three parts, describing twelve studies and/or development/research projects (summarised in Figure 1.1). In Part I the prototype system DMSS (Dementia Management Support System) and development work, including user evaluation studies are described (Chapter 6, 7). An introduction will be given of the medical domain, and in particular, the situation for the person with dementia and the person’s family. A brief definition is provided of what constitutes a decision-support system in the thesis together with a review of decision-support systems developed for the domain.

Cognitive diseases affect the ability to use cognitive functions and perform activities. The same cognitive processes are used by clinicians in their decision making. A review is provided in Part II of how knowledge and clinical reasoning are described in literature and how reasoning is performed by experts and novices. Further, the studies made for the purpose of assessing procedural and domain knowledge are embraced in Part II. A model for the procedural knowledge in terms of diagnostic reasoning is introduced and evaluated (Chapter 10).

The domain knowledge, residing in evidence-based studies and clinical guidelines, was investigated and analysed in cooperation with a domain expert (Chapter 11). The results are described and provide a foundation for the analysis of formalisms in [Lindgren, 2005] and for the implementation of the knowledge in the prototype system.

Further, the existing ontologies (i.e., terminology models) developed for the domain of cognitive disorders has been investigated and compared with the terms used in the prototype system. This has been done with the purpose to validate the terminology used, and evaluate the utility of the ontologies for the decision-support system. The results are presented in Chapter 12.

In Chapter 13 a summary is provided of Paper I, including the results of the analyses of case studies of actual patients and of the domain knowledge, using activity theory as theoretical framework. The developed model of clinical activity was used to identify actions suitable for formalisation as an alternative view on clinical work processes to the ontological and workflow perspectives. The model incorporates both process and domain knowledge in a common structure that forms a foundation for the support to be integrated in the system.

A set of basic if-then rules were identified and implemented for the task of diagnosing typical cases, using 60 features among 150 symptoms defined in the system. These rules constitute the main stream of inference which is based on certain selected guidelines. It is expected to capture approximately 70% of the dementia cases. For evaluation purposes additional support was implemented in the system which allowed the physicians in a study to provide their view
of an argumentation mechanism for diagnostic reasoning in atypical cases of patients. The additional support mechanism is being developed on the basis of the analysis made of formalisms described in [Lindgren, 2005], which will make use of guidelines in which knowledge is augmented with more uncertainty. The difficulty in handling incomplete knowledge in practical intelligent systems has been described in literature [Berner, 1999; van Bemmel and Musen, 1999; Pearl, 1988]. The decision theories available to build knowledge systems upon have been basically either numerical, with the problems of assessing prior probabilities and making sense of numerical values representing non-numerical events in a complex context, or symbolic with difficulties in handling ambiguous information and making inferences valid (finding non-conflicting truths) in the same complex, imperfect context. The work towards a formalism suitable for the domain is summarised in **Part III**, as an introduction to Paper I-V and to future formalisation work. The ongoing formalisation work is grounded in the activity analysis in Paper I, using the model for identifying basic building blocks of logics as well as process structures for interactive reasoning. Thus, the scope of this thesis regarding formalisation of knowledge, is not to establish the theoretical foundations for a certain formal implementation technique, but to create a model which assures the integration of the different frames of interpretation of evidence in the process model of the activity in focus, which in this work constitutes the investigation of dementia.
PART I

DMSS - DECISION SUPPORT FOR THE INVESTIGATION OF DEMENTIA

Jag glömde sömnen i sängen
Jag glömde smaken på smörgåsen
Jag glömde piggheten i kaffet.
- Agnes Karbin

4 I forgot the sleep in my bed, I forgot the taste on/of the sandwich, I forgot the alertness in the coffee. Written by Agnes at the age of 9.
CHAPTER 2
Introduction to the Medical Domain

There is a need for tools at primary care centres (PCC) to report relevant information and to venture an initial diagnose among the class of diseases that the clinic treats. The group of individuals over the age of 70 was responsible for 40% of the visits to primary care centres in a study made in Linköping [Ólafsdóttir, 2001]. The prevalence of dementia in this group was 16% and other mental disorders represented 17%. The same study reported that only 25% of patients with dementia were detected by their general practitioner (GP). The majority of the patients that actually had dementia in the study had a psychiatric diagnosis in medical records (66%), mostly anxiety or sleep disorder.

However, the number of dementia patients each GP meets in daily practice is still too low to make the GP well trained in diagnostics and management of patients in the area of cognitive diseases. A GP meets approximately 1-4 new dementia cases each year. In addition, new treatment strategies are established in clinical routine directed towards cognitive deficiencies with behavioural and psychological symptoms in the presence of dementia (BPSD) [Finkel et al., 1996]. A large turnover of personnel also increases the need for tools for introduction of new personnel, especially in northern Sweden where the distances make it difficult to fill the need of physicians in primary care.

The care of persons with a dementing disease is costly for society. Approximately 40 billions Swedish Crowns are spent every year on health care and social services for persons with dementia in Sweden, even though the care is often done by relatives (estimated as 4-5 times the formal care) [Wimo and Jönsson, 2001]. Drug expenditure on patients with dementia is estimated low, comprising mainly acetycholinesteras inhibitors, antidepressants and neuroleptics. A study made in Spain shows that for that particular region as little as 1.6% of the total cost represented drugs. The main expense was social treatment and care, needed because of disability to perform daily activities and BPSD [Gutierrez Iglesias et al., 2000]. Studies have shown that by introducing drug treatment of Alzheimer’s disease there is a delay of institutionalisation for up to nine month with an overall saving of 17-30 % [Kelly et al., 1997]. Therefore it is of great interest to identify subjects in the early stages of dementia to be able to delay the development of severe cognitive impairment that causes disturbances in ability to perform activities of daily living [Knopman et al., 2001]. Evaluations are being made to determine which treatment options that may alter the rate of the syndrome Mild cognitive impairment (MCI) progression to dementia.
Cognitive diseases cause damages to the brain, which are manifested as cognitive disorders, neurological dysfunctions and in many cases behavioural and psychological symptoms as a secondary but difficult consequence. Often these diseases are degenerative, which means that they cause a progressive course with increasing difficulties in handling activities of daily life. When cognitive abilities are impaired, we usually think of memory impairment, since it is possibly the most common and the easiest to identify. But cognitive skills also include higher-level processes as abstract thinking, judgement, communication, problem solving, etc. Cognitive abilities are used in every activity performed in daily life, therefore social ability and ability to pursue an occupation will eventually be affected. In this process an increasing load is put on family and other people around the person who is suffering from a dementing disease, since the environment needs to compensate for the loss of ability. This means that when the disabled person does not remember to eat, or where to find the mail-box, or find his or her way around in the formerly familiar neighbourhoods and gets lost, starts irrational projects involving dangerous machines etc, other people must have the control and provide the comprehension of the world that the person with dementia lacks.

Beside cognitive impairment, BPSD such as restlessness, wandering, aggression, failure to recognise relatives and locations, are common among the persons with dementia. People around the person can experience a change of personality. The daily life has to be adjusted, often for all in the family in order to compensate for the cognitive impairment. At a point when the difficulties have grown beyond the family’s capacity to handle, even with home-care provided from society, a submission to a hospital or treatment centre is inevitable. This is always only a temporary solution and an individual may have to move back and forth during a period before the situation has been solved permanently. Then the new home usually becomes a household for a group of persons with dementia with care 24 hours a day or at a nursing home.

The local authorities are responsible for the daily care and make judgements about the kind and amount of support an individual needs. Based on these, daily treatment and care is planned and provided. There are certain personnel responsible for these judgements. These judgements are sometimes done on the basis of actual results of dementia investigations made by teams organised by the health care (the primary care and hospitals managed by county councils). These judgements can also be based on evidence contributed by occupational therapists, care personnel or nurses within the care provided by the local authorities. The health care is responsible for the medical care, including dementia investigations. The local authorities employ nurses with a special responsibility for seeing to that the medical care is provided. The physicians are responsible for diagnosing, although other personnel often contribute to the collection of evidence. In the
activity analysis presented in the thesis (Chapter 13), the focus is put on the care provided by the health care and especially on the process of investigating dementia.

In every day clinical practice the clinician uses different sources of evidence in the reasoning process of establishing a diagnosis. Literature reviews and evidence-based clinical guidelines may be used for this purpose, as well as expert consultations and individual trials. It is recommended that cognitive screening instruments are used to screen for dementia in individual subjects with suspected cognitive impairment. Screening tools are found to be useful to determine the degree of cognitive impairment. For the same reason certain structured informant interviews are recommended, as well as neuropsychologic batteries [Knopman et al., 2001]. The most widely used screening tool is the Mini-Mental State Examination (MMSE) [Folstein et al., 1975].

The presence of BPSD in an individual is best known by relatives or daily care personnel and are detected with informant interviews. Interventions are often directed towards behavioural disorders since these are usually difficult to handle for relatives or personnel in daily care and may lead to hospital admission. In the Linköping study [Olafsdóttir et al., 2001] more than half of the persons with dementia were diagnosed with a psychiatric diagnosis, mostly anxiety or sleep disorder. A possible explanation that was discussed was that the psychiatric symptoms, manifested as behavioural disorders overshadowed the cognitive disturbance.

When the world turns incomprehensible for the person with dementia, a lot of effort is put into understanding not understandable things, and for the people around into explaining over and over again who the people around are, where the person is and why he or she needs to do certain things. In the early stages this causes a lot of frustration and anger in the person when things he or she normally can do become impossible and he or she is experiencing an increasing dependence on other people. As a consequence, depression is common as well as aggressive behaviour, which is difficult for the family to handle. Because of the load dementing diseases cause the family, dementia is often called the disease of the family. Therefore, also the people around the person, who is in focus for investigation, are involved in the process in order to treat the person in a fruitful way. Just as the environment can diminish the effects of the dysfunctions in the person with dementia, the environment often plays a major role when symptoms of BPSD are increased.

The interest in the status of dementia care in Sweden is growing and we see extensive efforts in increasing the quality and amount of resources put into the care of the elderly population. Commissioned by the Swedish Government, after
pressure from the Dementia Association (the National Association for the Rights of the Demented), the National Board of Health and Welfare in cooperation with the Swedish Association of Local Authorities and Regions (SALAR\textsuperscript{5}), have been monitoring a national project providing support for teams in clinical practice in order to improve dementia care (reports can be found here: [The Swedish Association of Local Authorities and Regions, 2007]). At the beginning of 2004 experts from different professions in the domain came together with representatives from the relative’s association and formulated goals for the clinical teams to work towards. While today only 25% receive a correct dementia diagnosis in their first encounter with a general practitioner, a goal was set that at least 90% will receive a specific dementia diagnosis. Several of the other goals concerned cooperation and communication within teams and between teams at different levels of care, as well as providing information to relatives. An important issue that was extensively discussed at the expert meeting was how competence could be improved, especially for personnel working in daily care with persons with dementia. Another goal was that all persons in need for further investigation should be investigated using national clinical guidelines. The work on forming national guidelines was initiated 2006 [The National Board of Health and Welfare, 2007]. These needs to improve dementia care were also expressed by physicians in a study made to investigate the motivations for a decision-support system in the domain [Lindgren, 2001]. Considering the results of this study and these official national goals for improving dementia care, the current development of a decision-support system for the domain is seen as highly relevant and as one means to provide the domain knowledge to the personnel at the point of care. A decision-support system can be used as a tool for cooperation, consultation, continuing medical education and for the dissemination of clinical guidelines, in the process of examining and diagnosing illness of persons with suspected dementia.

\textsuperscript{5} SALAR is a joint federation since March 2007 between the Swedish Association of Local Authorities (SALA) and the Federation of Swedish County Councils (FCC).
Chapter 3
Clinical Decision-Support Systems - Definitions and Related Work

The term decision-support system (DSS) is widely used in the literature. Types of programs that fall into the category of clinical decision-support systems (CDSS) are

- Stand alone programs that interview a user about a clinical situation and give advice in return
- Programs that generate alerts automatically on the basis of clinical data
- Programs that enforce clinical practice guidelines.

The following definition is used in [van Bemmel and Musen, 1999].

...any piece of software that takes as input information about a clinical situation and that produces as output inferences that can assist practitioners in their decision making and that would be judged as "intelligent" by the program’s users.

In this work the term DSS is used to denote a system that can be used as a cognitive tool, which extends the user’s own reasoning ability among other cognitive functions. Thus, the purpose is not to replace the physician but to reinforce the decisions made by the physician. The effect the use of the system has on the user and on his or her work is also acknowledged, since a DSS does not only support decisions at hand, but also transforms the decision processes (Figure 3.1). Therefore, the concept of interactive reasoning is used to capture the level of complexity in the use situation.

The main reasons why clinical DSSs are needed, according to van Bemmel and Musen [van Bemmel and Musen, 1999], are 1) to reduce the amount and consequences of human errors, 2) to automate routine tasks, 3) to cope with information overload and 4) to disseminate clinical guidelines. An additional aspect is the possibility for the user to continue a medical education while using the DSS, which was highly valued in an evaluation made by Karlsson [Karlsson, 2001].

The success of a particular clinical DSS is measured with different techniques depending on which decision-support paradigm is used in the system. Karlsson identifies five different paradigms [Karlsson, 2001]. In Bayesian statistics the choice that generates the highest probability for success is the optimal choice while inductive methods and expert systems compare the accuracy of a result to
Figure 3.1 Components of a decision-support system aimed at supporting and developing complex work environments such as clinical practice.

An expert’s. The **protocol paradigm** includes guidelines and treatment policies and are evaluated by determining the conformity to these protocols. The **case-based reasoning paradigm** compares a particular case with other cases in a database in order to find a similar case. Since presented work in this thesis is mainly based on clinical guidelines as source of knowledge the approach taken in the work is protocol based, viewing the accuracy of the inferences produced by the system as dependent on 1) the domain knowledge, 2) the system’s compliance with process factors residing in the context of the system in general, and 3) specifically with the cognitive skills of the user.

Several attempts have been made to develop decision-support systems for the large domain of psychiatric diseases, of which the cognitive diseases represent a part. While the earliest goes back to the beginning of the 70’s, the DSP (**Diagnostic Decision Support System for Psychiatry**) is one of the most recent systems described in literature [do Amaral et al., 1996]. This system is based on the whole area-specific standard clinical guideline DSM-III-R [American Psychi-
The system makes use of a rule-based reasoning with probabilities and certainty factors [Shortliffe and Buchanan, 1975] attached to the rules, in combination with a heuristic categorical reasoning. Using only the rule-based part, the system suggested the correct diagnosis as the first hypothesis in 52.8% of the cases, compared to experts’ diagnosis of the same 53 patients, typical for the domain. Adding the heuristic part the system reached a level of 73.6% correct diagnosis. It has been shown that large decision-support systems that are aimed to cover a wide range of diagnoses tend to perform less satisfactory than systems with a narrow focus, which can reach performances with over 90% accuracy [van Bemmel and Musen, 1999].

The system Diagnostica developed by Gartner et al. [Gartner et al., 2000], uses the same scope, the psychiatric domain, but uses the more recent version of the DSM, DSM-IV [American Psychiatric Association, 1994], as knowledge source. The current version of the system makes use of a simple but clinically relevant abduction technique, with a three-valued logic and the well-founded semantics (WFS). The interface follows the guideline very closely, visualising and using the similar hierarchies of terms in tree-structures. The system will be extended with a possibility for the user to add preference orders among diagnoses as well as true differentials for diagnoses.

A different approach to give computerised support to clinical practice is the group of systems which are developed for the purpose of assessing the patient’s cognitive function. The Cognitive Drug Research project (CDR) is the most established and is used in clinical trials of evaluation of drugs [Wesnes et al., 2002]. These computerised cognitive screening tests measure time factors and errors and have shown to be reliable and useful for the purpose. The CDR is used for different groups of patients, not only for dementia patients.

In the domain of cognitive diseases the dementia diagnosis Alzheimer’s disease is the most common with a representation of up to 80% of dementia cases, depending on which study is referred. A prototype system that has been developed, which is aimed to capture cases of Alzheimer’s disease, is the rule-based system described by Herrero and colleagues [Herrero et al., 2002]. An extensive sample of features is taken into consideration in the system, specific for the Alzheimer’s disease. This system and the DSP system are aimed at covering the most common cases of patients in the domain. The DSP was also evaluated using typical cases, but was in spite of this not performing very well. The system was covering a wide range of psychiatric diagnoses which are manifested in a patient in similar ways, which made the features included in the reasoning process not as many as in other domains. However, the causes of the symptoms are fundamentally different. The problem with overlapping sets of characteristic features, as
well as cyclic diagnosis criteria and incomplete knowledge in the DSM guideline is addressed in the development of the system Diagnostica, however, techniques for handling these were only partially implemented into the system at the time for the article [Gartner et al., 2000].

A system has been developed in Great Britain for the dementia domain, which is used in some primary care centres integrated in the electronic medical record. It reminds the user of taking action if information added to the patient’s record evokes a suspicion of dementia [Iliffe et al., 2002]. The system provides questions for the user to consider in deciding whether he or she can diagnose the dementia or if the patient should be referred to specialists. An education material is included. This system was considered too general to be of use in clinical practice in the region of northern Sweden.

The purpose of studies presented in the thesis is to cover the subfield of psychiatric diagnoses that are classified as cognitive disorders, although other psychiatric disorders as well as systemic diseases are considered in a differential diagnosis process. A level of granularity will be aimed for that allows the capture of both typical cases and atypical cases, on the basis of a set of relevant guidelines, among which DSM-IV is one of the most widely used in clinical practice.
CHAPTER 4

Methods and Procedure

The decision-support system DMSS (Dementia Management Support System) is being developed as a joint project between the departments of Computing Science and Community Medicine and Rehabilitation at Umeå University. The project is partly funded by the European Union, the Vårdal and Knowledge foundations, department of Computing Science, Umeå university, Äldrecentrum Västerbotten and the Kempe foundations. The development process has been characterised by a close cooperation with medical domain experts throughout the process, both in the analysis of the domain knowledge, in the issues concerning formalisation techniques, and design of the interaction. Evaluations have been conducted in an iterative manner, focussing different issues involving different potential user categories depending on the purposes with the evaluations (Figure 4.1, Table 1).

In the development process methods and routines in different clinical practices have been analysed in order to identify the needs of practitioners at different levels of care [Lindgren, 2001]. As a result the prototype system contains screening tools for cognitive deficiencies, BPSD and functional dysfunctions that are used in the investigations. The domain knowledge used in the system is based on clinical guidelines and literature available in the area. The knowledge acquisition was done in cooperation with a domain expert. The procedural knowledge was assessed through interviews with professionals at different levels of care and case studies of actual patients in clinical practice with the system integrated.

Prototypes have been implemented using Visual Basic 6.0, while prototypes for mobile hand-held applications, as extensions to the system, are implemented using embedded Visual C++.

The prototype system has been developed in two main phases where the first phase was initiated by building the model of the reasoning process and a hierarchical structure of data which formed the base for the design of the graphical user interface (GUI). Initially the model and data structure were provided by the domain expert. The GUI was implemented and extended with a simplified knowledge base (KB) for diagnosis of the most common diseases, which was designed and implemented with if-then rules. This initial prototype (Prototype I) was evaluated by observing six general practitioners (GPs) using the system without patients present. The GPs were also interviewed. The evaluation was done in an iterative process where the data structure, KB and GUI were adjusted according to the GPs’ comments and failures revealed during the observations.

In phase two the domain knowledge was further analysed and formalised.
The KB concerning typical cases was updated and supplemented with less common diagnoses according to the results presented in Part II, together with support concerning interventions. The GUI and the data structure were adjusted according to the development of the KB. This work continued in cooperation with the domain expert. Additional support for atypical cases, sufficient for providing users a view of how such support can be implemented in a system, was integrated for evaluation purposes.

The second version of the prototype (Prototype II) was evaluated by observing six professionals using the system with current, and in their perspective, difficult cases of cognitive diseases. The professionals were three domain expert physicians (geriatricians), one training to become an expert and two persons from a team at a primary care centre (PCC), of which one was an experienced GP and the other person was a nurse responsible for dementia investigations at the PCC. None of these participated in the first evaluation. In this evaluation teamwork and consultation situations were captured.

In order to study the system integrated in clinical practice, case studies were
made where the system was used in the initial phase of investigation. The system used in the evaluation was an updated version based on the results in previous evaluation, here called Prototype III. Results of the evaluations of the prototypes I-III are summarised in Chapter 6.

A Japanese version of the prototype system, DMSS-SJ, was developed for evaluation and development purposes. DMSS-SJ was evaluated in a pilot study with experts of different specialities related to the dementia domain, who used the system on a wide range of different types of patient cases. The results of the evaluation are fed into the development of DMSS, as well as into the development of an extended Japanese decision-support system, which, in addition to the purposes of DMSS, aims at integrating decision support for managing persons with dementia within the Japanese health insurance system. The results are provided in Chapter 7. An overview of the studies is provided in Table 1.

<table>
<thead>
<tr>
<th>Prototype</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>DMSS-SJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary investigation focus</td>
<td>Diversity in needs, user individuality aspects</td>
<td>Teamwork, consultation, reasoning model, argumentation functionality</td>
<td>System’s compliance with investigation process</td>
<td>Reasoning model, terminology, design, diversity in patients’ cases, content</td>
</tr>
<tr>
<td>Number of users</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>User characteristics</td>
<td>GPs; geography, range in expertise</td>
<td>Experts, team; range of different expertise and work methods</td>
<td>Expert geriatrician</td>
<td>Experts, GPs; range of different medical expertise</td>
</tr>
<tr>
<td>Number of patients</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td>Patient characteristics</td>
<td>Current patients</td>
<td>Current complicated patients</td>
<td>Range of typical and atypical patients</td>
<td>Range of typical and atypical patients</td>
</tr>
<tr>
<td>Environment</td>
<td>PCCs in northern Sweden</td>
<td>Geriatric centre, PCC in northern Sweden</td>
<td>Geriatric centre in northern Sweden</td>
<td>Different clinics in a county of Japan</td>
</tr>
</tbody>
</table>

In parallel to the above mentioned, cognitive screening tools were implemented and integrated into the main system, as well as in the form of mobile extensions for hand-held computers. The mobile applications will not be covered in this thesis since they are not essential for fulfilling the purpose of the thesis. The description of the DMSS version of the system in the following chapter is focussed on the GUI and KB, and the interactive reasoning they mediate.
CHAPTER 5
System Description of DMSS

The Prototype system II presented in [Lindgren, 2005] is an extended and revised version of Prototype I, where the additional functionalities mainly concern diagnosis of less common diagnoses, intervention, screening-tools for cognitive deficiencies and the argumentative functionality, which were implemented for the purpose of presenting the functionality in an evaluation with users. The prototype system DMSS presented in this chapter is a revised and English version of Prototype II, III and DMSS-SJ based on the results from the evaluation studies presented in Chapters 6 and 7.

Figure 5.1 Main frame in the 3rd step in an atypical patient’s case.

The GUI is implemented based on the data structure described in the next section. The system consists of a main frame that guides the user through the reasoning process providing instructions and feed-back (Figure 5.1). The left
side of the frame contains the links to the different frames and functions as an overview of the content of the system. Basically, the functions at the top constitute containers of information organised by source or by tool, while the functions as the bottom constitute advanced reasoning support functionality. The same functions can be found in the menu. There is also a textbox showing information as the process proceeds. The main reasoning process is visualised on the right side of the main frame.

As supplement to the main frame, two frames are used for presenting the results of analyses; Profile (Figure 5.5) and Intervention (Figure 5.8), which are updated throughout the process. Other frames function as input and containers of data with categorisation of data depending on type and source in a similar structure as the medical records used in the region. These are Autoanamnèsis, Heteroanamnèsis (Figure 5.2), Status (Figure 5.3), Laboratory and radiology examinations (Figure 5.4) and Previous diseases (Figure 5.9). Additional frames consist of the integrated screening tools MMSE (Figure 5.6), FAST and the “behavioral pathology in Alzheimer’s Disease Scale” (Behave-AD) (Figure 5.7). Behave-AD is the only place in the system where it is possible to enter text, by request from professionals who are accustomed to add comments while using the tool.

When necessary data are missing, the frame where the data should be entered appears with the requested data highlighted in red (see Figure 5.2). The choice was made to let the regular frames appear instead of temporary frames containing only the missing data, in order to minimize the amount of different views which the user has to see, thereby minimizing the variation of cognitive input and support recognition.

5.1 Termination Model

The set of features specified in the system are formed as an hierarchical structure, which initially was provided by a domain expert. The structure formed a base for the structure of the GUI and has been adjusted in the development process. The terminology in the system was evaluated, both in the user evaluations and in the perspective of existing terminologies and classifications that was found relevant for the scope of the decision-support system. The latter evaluation is further described in Chapter 12, along with motivations for current terminology in the system. The vocabulary developed for the prototype is based upon the vocabulary in guidelines, medical terminologies and disease classifications such as ICD-9-CM, SNOMED-CT, ICF and domain experts’ use of medical vocabulary in the area. The classification of findings is based on source and type of manifestations.

Physicians use different sources when they obtain data; clinical interviews
with the patient, clinical interviews with relatives or treatment personnel, examinations, other professionals, etc. The source is important to take into account, in order to detect and understand ambiguities. This is currently made in medical records that are used in clinical practice. Therefore, the initial classification of signs in the project is made by source, e.g., the patient, the relatives, physical examination, etc.

Within the source categories autoanamnesis, heteroanamnesis and status there is a similar structure of the physical, behavioural and cognitive signs, but at different levels of granularity. The level of granularity depends on the presumed ability of the source to account for specifics of the impairment. A tool-tip function is used to explain concepts that were found less familiar among participants in evaluations. Another factor which determined the specificity of data in the system concerns the relevance of the data in the perspective of diagnosing cognitive disorders. Typically, cognitive symptoms are accounted for at the finest level in the system, while, for instance, the set of symptoms from the circulatory system.
are of coarse granularity.

Diseases are categorised according to the structure in the chapter of cognitive disorders in DSM-IV-TR [American Psychiatric Association, 1994] supplemented with less common dementia types. A “fish-eye view” is taken on diseases since basically all alternative systemic or organic causes have to be considered in the process, however, implementing core features for each and all is beyond the scope of this work. Therefore the finest granularity is used when analysing the core features immediately relevant for differential diagnosis of type of dementia, while other diseases are appearing in the system as reminders for the user to consider in the process without diagnostic criteria.

The value of the majority of the features in the system can be one of the set \{unknown, normal, abnormal\}. To capture the severity of the cognitive deficiencies an additional valuation is requested according to DSM-IV, which can be one of the set \{unknown, significant decline, mild decline\}. To capture the time aspect and other characteristics, the user describes the onset by choosing either of \{unknown, rapid (days to weeks), slow (more than 6 months)\}. Further,
for assessing the course of the decline the following can be registered: \{\textit{non-progressive, progressive, stepwise, fluctuating}\} (Figure 5.2). The valuation of the results of X-rays can be done with predefined descriptions among which the user can choose \{\textit{normal, Focal/vascular signs, frontotemporal signs, signs typical for Alzheimer’s disease, Other signs}\} (Figure 5.4). The valuations of presence, severity and importance of different x-ray results are subjects for adjustments as a result from recent evaluations, primarily as supplementary valuations in the perspective of alternative guidelines.

![Figure 5.4 Laboratory data and results from radiology examinations.](image)

The terminology of the system was adjusted based on user evaluations and on the ontology analysis and valuation presented in Chapter 12. To summarise recent adjustments, features and symptoms in the system were changed when possible from terms indicating dysfunction, to neutral terms. By constructing the basic vocabulary from neutral terms, misinterpretations can be avoided caused by different cultures and use contexts. The basic vocabulary is supplemented with scales for valuating instantiated features in a certain patient. The scales can be replaced or supplemented with other scales depending on routines in the local use context, a design which makes the system dynamic and adjustable to the development of medical knowledge.
Terms referring to syndromes and diseases are reformulated to be in accordance with terms found in the UMLS, primarily terms from SNOMED CT® and ICD 9 CM as far as found appropriate. Some original terms are kept that refer to a cluster of diseases or syndromes following the level of granularity in the design of the system. Further, there are terms which are used locally that is kept in the local design of the system, such as the term vätskestatus in Swedish, referring to a set of laboratory investigations (Body fluids is the temporary translation to English in Figure 5.4). The term is well established in Swedish clinical practice; therefore it is motivated to keep the term in the system. However, such concepts need to be carefully matched to concepts in other languages, when the system is translated.

Evaluations revealed that the vocabulary used in the system was highly context dependent, assuming context knowledge of Swedish practice. Therefore, terms and expressions have been clarified and extended to avoid misunderstandings and enhance a common interpretation of each term used in the system. In addition, the integration of international definitions of concepts drawn from mentioned terminologies and classifications in UMLS is planned.

5.2 Investigation Process

The “backbone” of the system is the clinical reasoning procedure, which is visualised on the right side of the main frame (Figure 5.1). It follows roughly five steps of which the first three concern diagnosis and the last two concern intervention. The model is described and evaluated in Chapter 10.

There are basically two ways to proceed in the use of the system. The analysis mechanism is used whenever the user wants the system to analyse the information regarding a certain patient. The user can choose to let the system function as a guide through the process, giving reminders and instruction of what to do next and feedback with suggestions of diagnosis and treatment actions by using primarily the analysis mechanism.

The user can also systematically investigate the patient’s resources in the way the user may be used to in clinical practice and enter all known information categorised by type and source before using the analysis mechanism. In this way the analysis mechanism can be used as a supplement to and confirmation of the user’s own analysis.

In evaluations, the experienced physicians tend to use the second method to a larger extent, especially in evaluations done with physicians in Japan.

The process is implemented with a control mechanism which require necessary data to be entered in the system before an analysis of the current step can be made and move the process on to the next step. This interactive guidance
component in the system is referred to as the process-level of the knowledge base in subsequent descriptions of work analyses and formalisation in the thesis.

After request from users in evaluations, the possibility to enter into the reasoning process when a state of dementia is known has been added, with a simple alert confirming that there is a state of dementia. This is done by using the re-analysis function in the main frame, when the preceding steps are not accomplished through the use of the system.

5.3 Diagnostic Reasoning Process

The first three steps of the reasoning process comply with the clinical guideline Diagnostic and Statistical Manual, 4th edition (DSM-IV) [American Psychiatric Association, 1994], which is a clinical guideline recommended in an evaluation done by the Quality Standards Subcommittee of the American Academy of Neurology [Knopman et al., 2001]. The guideline is supplemented with consensus guidelines for the diagnoses frontotemporal degenerative dementia (FTD), dementia of Lewy body type (DLB) [Neary et al., 1998; McKeith et al., 2000] and for Mild Cognitive Impairment (MCI) [Petersen et al., 2001]. A main stream of inference is implemented with if-then rules based on core sets of features drawn from these guidelines, which aims to detect and identify the part of cases of cognitive diseases where the aetiology fit the school-book descriptions. In these cases, which we distinguish as typical cases, the system suggests and motivates one single unambiguous diagnosis based on the data that has been entered into the system. In the process, the system provides reminders of what to consider, or requests for additional and necessary data that has to be entered. The level of firmness of a diagnosis which can be obtained in these cases we call level I, denoting the highest reliability of the three levels identified in the system (Table 2).

In order to provide the user the maximum support for committing to a decision in the complex domain of cognitive diseases, it was found necessary to provide not only one line of reasoning according to some clinical guideline, but to give the flexibility to adopt to the most applicable interpretations available in the context of a particular patient and the knowledge available. For the more complicated, atypical cases where several diagnoses fit, or where the known evidence is not enough to diagnose an obvious cognitive dysfunction in the context of these guidelines, the user may use extended support in the context of additional clinical guidelines. In the prototype a simple argumentation mechanism is implemented, sufficient for giving users the idea of how such support can be implemented, inspired by the framework of argumentation logics described by Fox and Parsons [Fox and Parsons, 1998]. The approach gives the possibility to keep tentative
diagnoses in the process even if two or more are contradictory.

It should be noted in this context that the proportion of unambiguous cases matching only one dementia diagnosis seems to be smaller than expected. Therefore, investigations is currently being made on proportions of patient’s cases of mixed aetiologies compared to cases of single aetiology and the corresponding proportion of typical vs. atypical cases defined by the system (pilot study in Chapter 7).

In the system the evidence concerning the patient is matched to arguments corresponding to a certain diagnosis according to guidelines and scales used in practice (Figure 5.5). The evidence is presented to the user interpreted into the numerical setting of MMSE, Behave-AD and Hachinski Ischemic Score (HIS), and the qualitative setting of the consensus criteria for FTD and DLB and the NINCDS AIRENS guideline for VaD. In these guidelines a certain mix of signs in a patient may generate a set of supportive or contradictory arguments for a certain diagnosis and a set of core arguments that are used to value the probability for the existence of a diagnosis according to some guideline or sets of guidelines.
The overall value of the core arguments may be either of probable, possible or unlikely in relation to the particular guideline used. The vocabulary used in the system consists of the terms used in respective guideline. The trustworthiness of a probable diagnosis depends on the specific guideline and the trustworthiness of other probable diagnoses. It is the user’s task at this point, to value the trustworthiness of different suggestions of diagnoses and to establish the differential diagnosis, since no aggregation function which compares the different hypotheses is implemented in the prototype system. The set of supportive or contradictory arguments and the valuation in the perspective of the guidelines are presented to the user, and contributes to the user’s evaluation. The user marks his or her choice among the set of alternative dementia diagnoses, an action which generates suggestions of interventions. This is only possible when necessary data is collected and in the case of an atypical patient. The user may want to re-evaluate the evidence if new evidence is collected, which can be done using the mechanism for re-evaluation in the main frame. In this case the decision made about diagnosis is withdrawn.

<table>
<thead>
<tr>
<th>Level</th>
<th>Qualities in suggested diagnosis/diagnoses</th>
<th>Inference techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>One single diagnosis based on sets of core features in DSM-IV and consensus GLs for MCI, FTD, DLB.</td>
<td>IF-THEN rules</td>
</tr>
<tr>
<td>II</td>
<td>More than one diagnosis based on sets of core features in DSM-IV and consensus GLs for MCI, FTD, DLB. Additional GLs is needed as supplement.</td>
<td>IF-THEN rules Argumentation</td>
</tr>
<tr>
<td>III</td>
<td>No diagnosis fit completely based on sets of core features in DSM-IV and consensus GLs for MCI, FTD, DLB. Additional GLs is needed as supplement.</td>
<td>IF-THEN rules Argumentation</td>
</tr>
</tbody>
</table>

A situation where more than one diagnosis is met by the criteria according to the main stream of inference, and the additional evaluation is needed, is referred to as level II in the valuation of reliability of suggested diagnoses. Additional information will be evaluated, weighted and presented at a level with more substance. In the case when the specified guideline criteria are not completely met (lacking some core feature), but findings indicate a state of dementia, then the argument aggregation technique is used to point towards a diagnosis. However, the reliability of tentative diagnoses in this situation is considered lower, here denoted as level III (Table 2).

Each step of the process of detecting typical cases is implemented in VB 6.0 using production rules. The vocabulary used in the programming code consists of medical terms used by a domain expert, which facilitates cooperation during de-
development and makes the code more accessible for a physician. For the reasoning process as a whole a categorical or diagnostic reasoning technique is used, where more general disorders are narrowed down to specific diagnoses, while at a component level a causal technique is used, for the purpose of confirming or rejecting hypotheses. Rules for each diagnosis or disorder are formed as one function for each condition, which can be used in a deductive way in the reasoning process. In practice, there is a matching of phenomena and collections of phenomena occurring in a patient, in a process where hypotheses are excluded and the set of possible diagnoses is narrowed down as much as possible, according to the initial set of guidelines. The knowledge base is currently being developed, a process presented in Paper I-V.

![Figure 5.6 MMSE - Mini-mental state examination [Folstein et al., 1975].](image)

To what extent screening-tools can be used as a direct source for evidence, has been discussed during the sessions with the domain expert. A tool such as the MMSE is not validated for other purposes than indicating an overall cognitive deficiency (Figure 5.6). It can for instance not be used to confirm a memory deficit without other evidence, in the process of diagnosing a cognitive disease.
5.4 Interventions

On the other hand, the tool Behave-AD was considered useful for detecting hallucinations and indicating the existence of depression, which can be manifested as secondary symptoms of dementia (Figure 5.7). The choice was made to reinforce the use of the tool in Swedish clinical practice by using it as the source in the system of the core feature visual hallucinations in DLB (this is not the case in the current Japanese version). Behave-AD is also used as the source for suggestions of interventions for BPSD, which will therefore be further described in the next section.

Figure 5.7 The behavioral pathology in Alzheimer’s Disease Scale - Behave-AD (brief translation from the Swedish version).

5.4 Interventions

There are interventions that aim to treat the disease corresponding to a diagnosis and interventions that are directed towards secondary consequences of the disease, like BPSD and the level of ability in activities of daily life [Eriksson et al., 2000; Doody et al., 2001]. The signs in a patient need to be interpreted considering the patient’s whole life context in order to establish fruitful interventions. For the interpretations there are different frames of reference (i.e., organic, social, personal, etc) that are used, and are represented in the system (Figure 5.8). A specific BPSD symptom may be caused by the degenerative process affecting the brain, but becomes unbearable in daily life because of, for instance, environmen-
tal factors or the patient’s coping abilities. This analysis is a continuing process intertwined with the diagnosing process, since some interventions may need to be actualised before the diagnosis is established. The fourth step in the reasoning process aims to investigate further the consequences of the cognitive decline and the existence and impact of BPSD. In the fifth and last step interventions that correspond to the diagnosis and regarding remaining secondary symptoms are established (Figure 5.8).

![Interventions](image)

Figure 5.8 Interventions.

In the process of establishing appropriate interventions, screening tools are used in clinical practice. One that is frequently used for detecting the existence and amount of BPSD is Behave-AD (*The behavioral pathology in Alzheimer’s Disease Scale*), which is one of the tools implemented in the system. The tool provides a profile that may be used for choices of treatment (Figure 5.7) as a product of the valuation of the presence of each symptom according to one of the distinct values 0, 1, 2 and 3. Some studies have been made regarding the usability of this tool for the purpose [Harwood et al., 1998], and one of the purposes for integrating the tool in the system is to gain further knowledge of to which extent
the profiles can be used in differential diagnosis, in the choice of treatment, as an overall evaluation of treatment and for follow-ups of patients.

5.5 Relation to Other Systems and Future Extensions

The electronic patient record system (EPR) used in the region of northern Sweden does not capture data at the level of granularity and specificity needed for the purpose of using the data for decision support or as a registry of patients for research purposes (Paper I). The current version of the decision-support system is built as a stand-alone system, while the possibilities of using data from the EPR in the reasoning process and sending reports to the EPR is investigated. The data structure implemented in the system is currently used to build a registry of patients for research purposes in the region. The possibility to integrate the decision-support system as an interface to the registry could significantly facilitate and motivate the gathering of clinical data for research purposes. The structure is also used for research concerning structured radiology reports in the domain.

Further, an investigation is currently being made of the data flow involving other systems, by using case studies of out patients investigated at a geriatric centre in northern Sweden. Preliminary results from the study include the pharmacy system and systems used at the radiology unit at the local hospital (Paper I).

The systems used by the municipality are separate from the systems in health care in the region. The database TILLIT is being built within a project governed by the municipality of Umeå in northern Sweden for the purpose of communicating patient’s data and facilitating cooperation between different levels of care. The possibility and motives for integrating the decision-support system into the municipality’s systems have been investigated, with the conclusion that the system should provide only the information needed, in a form useful as a base for the specific decisions that are made by municipality personnel. This information can be communicated through the TILLIT database. However, a project is planned which aims to further investigate the work situation of the municipality personnel in order to identify and develop tools needed in their work. A corresponding investigation is currently being conducted in Japan, where extensions to the main system will be identified for meeting the need for support in the management of persons with dementia.
5.6 Japanese Version DMSS-SJ

An analysis of the qualities of the medical domain knowledge expressed in clinical guidelines is needed [Lindgren, 2005] as well as a thorough work analysis of how the actual clinical work is performed in order to establish the requirements of a decision-support system. This is done in Paper I concerning the local praxis at a clinic in northern Sweden. However, since the system is being developed to meet an international need for assessment and support tools for treating the growing population of elderly people in different places, which has in common the growing amount of persons with cognitive deficiencies, there is a need for extending the scope of the system. Due to the urgent need to develop tools which can be used to facilitate the diagnosis and management of the increasing amount of persons with a dementing disease also in Japan, the limited version of DMSS named DMSS-SJ has been developed for research and development purposes.
A cooperation has established with a municipality in Japan and with local universities in the same area. The purpose is to develop a Japanese version of the decision-support system suitable for the Japanese welfare system and clinical practice. The Japanese system constitutes a sub-set of the Swedish version of Prototype II translated to Japanese. An initial demonstration of the system was given to a medical audience in Japan in June 2006 as an initiation of the project, and the system was evaluated in the beginning of 2007 with local physicians and patients cases. The results are described in Chapter 7. A second version is currently being used in clinical practice for evaluation purposes in Japan (Figure 5.10).
In the initial development phase, interviews were conducted with medical personnel at different levels of care, for the purpose of investigating the processes and methods used in the investigation of dementia in northern Sweden [Lindgren et al., 2002; Lindgren, 2003]. In addition, sources of conflicts which may cause the failure of a decision-support system for dementia investigation, were identified. These were current priorities concerning time spent on patients, economic aspects and attitudes in medical personnel and management. Positive aspects were the mainly positive attitude among the general practitioners towards tools which could facilitate communication within teams and with experts when consultation is needed. Also tools that can be used as an introduction for new personnel and for support in difficult cases of patients were desirable.

Further evaluations were conducted of Prototype I (primary care), Prototype II (expert evaluation) and case studies in clinical practice with Prototype III integrated in the investigations of patients. The results of these evaluations will be presented and discussed in this chapter. Firstly, an introduction is given to qualitative approaches to evaluating decision-support systems and to the theory and methods used in the evaluation studies.

6.1 Evaluation Paradigms

Work is more than ever influenced by a constant change of conditions and terms that direct in what settings, how and by which means work may and is expected to be done. Perhaps medical care is one of the historically most predefined work environments, where roles of profession and division of labour have been comparatively well founded and defined in each setting, in spite of the fact that the medical domain knowledge has been and still is rapidly evolving. But also medical practice is subject to changes. By introducing for instance electronic medical records into medical practice a change is made of professionals’ duties, pattern of cooperation and work methods. This emphasises special concerns in the development of new tools aimed to be used in clinical practice.

The fact that certain changes are bound to take place in the introduction of a new system may also be one of the reasons why up to 95% of clinical decision-support systems (CDSS) never reach clinical practice [Kaplan, 2001a]. There is a remarkably few studies where reasons for success or failure of CDSSs have
been studied. The focus of evaluations has been the capacity of the system in producing correct diagnoses compared to experts, or in conjunction with the user compared to the user without the system. Still why a system is not used in a certain setting, or why the user chooses not to comply with suggestions from the system, or why the system has little impact on the outcome of the treatment of patients, are phenomena we know very little about. However, the work of improving the quality of these systems and the quality of evaluations was initiated at the conference Medical Informatics in Europe’02 which resulted in a research database with evaluations made of clinical systems [Inventory of evaluation publications, 2006; Ammenwerth and de Keizer, 2005].

*Usability* and *utility* factors are often used in evaluations. To be measurable they have to correspond to a goal, i.e., user satisfaction compared to satisfaction of previous systems, or system’s functionality compared to required functionality defined in a system development context. In this process users and other people define their needs and procedure of work in a way they are not used to do, to people outside the work context that probably know little of what is really going on in clinical practice. To make needs and functionality understandable for the engineer, these aspects are defined in detailed quantifiable measurable terms that the potential users can agree upon. In this process, implicit factors that are more influential may be left out and that may become the cause of failure in a system implementation. The Scandinavian tradition of users actively participating in the development process is a method to meet some of these needs and to bridge the gap between the system engineering perspective on the system development and the clinical practice perspective [Bødker, 1996]. The work presented in this thesis follows the Scandinavian tradition of involving users throughout the development process, in Sweden as well as in the development of a Japanese version in Japan. In addition, a method for bridging the gap between the informal clinical reality and the formal structure in a decision-support system is being developed based on activity theory, which is presented in Paper I.

The *socio-technical approach* presented by Berg [Berg, 1999] among others emphasizes that the social aspects and technical aspects in a work environment are closely intertwined and cannot be considered in isolation. The work context is seen as a network where each unit is defined and relevant only in the context of other units, being technical as a stethoscope or social as roles, responsibilities etc. This perspective of medical care is similar to the *activity theoretical view* of activity systems [Engeström, 1992], where activity consists of a subject’s use of tools, while changing objects in a continuously changing context, a process that inevitably changes the subject as well. Both approaches point out the importance that every development process of an electronic clinical system should start in a thorough empirical examination of the work network or activity system where the application is aimed to be used, in order to identify issues and factors
essential in the unique work context in focus. This has to be done with the humble consideration that any description of work processes or clinical knowledge is a reconstruction of reality and may be out of date as soon as defined [Bødker, 1989; Berg and Goorman, 1999; Suchman, 1987]. In this work this perspective of a dynamic and changing use situation also includes the individual user, who develops skills and knowledge continuously.

Kaplan suggests a framework sprung from a social interactionist theory for evaluation of CDSSs. The framework consists of four specified foci: need for, or system supporting communication, system’s impact on care, on user’s sense of control and system’s status in the context, the 4Cs [Kaplan, 2001b]. The view of people learning in interaction with others, is an important consideration that theories of social interactionism emphasize, as well as cultural-historical activity theory. Vygotsky has contributed with the concept of zone of proximal development, ZPD to identify activities that a person can accomplish and learn to master only in interaction with more skilful peers [Vygotsky, 1978; Wells, 1999]. ZPD is one of the concepts in activity theory that in combination with concepts as breakdowns, focus shift, transformations between activity levels etc, create an understanding of what is happening when the motives for an activity is challenged.

Another approach where focus is put on the context of what the actual motive is in the use situation is proposed by Boisen et al. [Boisen et al., 2003]. In addition to usability and utility, they suggest the notion of “copability” to shift focus from the ideal use of the system to the real use in daily practice. Copability denotes the user’s ability to handle and alter the object with the system as mediating tool. Implicitly the ability can be different depending on which mediating tool that is used, and the ability mirrors the success of the tool. For instance, a person may need to live with a disease as diabetes, a situation which may become easier with a tool that tracks insulin levels and gives immediate feedback of the results of drug intake. This ability is proposed to be measured by which coping-strategies are used, where problem-oriented strategies are counted as more successful than emotion-oriented strategies. What is then the difference between copability, usability and utility except for the factors measured? According to Boisen et al. the notion of usability is often limited to the interface between the user and the system, while utility puts focus on the interface between the system and the outcome of the task, how the system meets the requirements in order to accomplish the task. To take into account a larger scope of context, for instance the scope of the whole target activity, which includes meaningfulness and motivation issues for the user, we need other terminology. Activity theory provides a framework that takes the activity as the minimal unit of analysis. Every activity is directed towards a person’s motive, which corresponds to a need, a need that
the person may not be aware of. The motive on the other hand is known and when a motive is frustrated, the person often gets upset and behaves the most unpredictable [Kaptelinin, 1996]. In order to understand behaviour in a use situation, it is important to identify goals for lower level actions that contribute to the fulfilment of the larger perspective of the motive, as well as the motive. In activity theory terms as breakdowns, conflicts, shift of focus and transformations have been used to illustrate what is happening when a motive is frustrated.

To view the ability to handle a difficult situation in a context of learning and development is also interesting. What does the person really need from the significant other that knows more, in order to transfer the difficult activity, which for the moment lies in his or her zone of proximal development ZPD, out of the ZPD and into the set of activities which the person can do independently? Is it plain knowledge, or support to change coping-strategy? Or does the more skilful peer function as a compensation for a not so successful coping-strategy? Within social work and rehabilitation the notion of “conscious use of self” is frequently used to illuminate the role the therapist has in the interaction with a client. By using self in a conscious way in the interaction, the therapist allows and supports the client to develop skills in a particular context of purposeful activity. In order to become the extension of the client in this development process, the therapist must be able to identify the activity in relation to the ZPD of the client. The active role the therapist has in the activity is similar to the role of the teacher in a pedagogic context, or the researcher in developmental work research, where his or her actions also become objects of data collection and critical analysis in the process [Engeström et al., 2003]. In the introduction of a novice into the domain of cognitive diseases and clinical practice, the expert functions as the more skilful peer. The role of a decision-support system in this context has the potential of being a supplement to the expert in the development of knowledge and skills in the novice, which is one aspect to consider in the introduction of a decision-support system into clinical practice.

The present support system’s position in relation to the user’s ZPD was evaluated in the initial studies, and also the integrated reasoning model in conjunction with the limited domain knowledge communicated through the interface. The more active role was taken by the observer in the initial evaluations, in order to investigate how the use of the system was internalized in terms of activity theory. In order to increase the reliability of the results in the qualitative studies, triangulation of methods is used. Further, in order to gain as many different aspects as possible on the system, the different studies involve different professionals, with different amount of experience, and focus different use situations. Cultural-historical activity theory has been used in analyses of data and in motivating the re-design of the system, based on the evaluations.
6.2 Motives for Evaluation

The purpose of the evaluation of Prototype I was to receive comments and material for the design of the system from GPs with different amount of experience, from different parts of the county of Västerbotten.

The purpose of the evaluation of Prototype II was 1) to validate the model of the reasoning process in a clinical setting, 2) to receive comments on the argumentation functionality in the system from experts, 3) to investigate how the system can be used in a consultation situation, and 4) to investigate how the system can be used in a teamwork situation at a primary care centre. Other aspects which were highlighted and observed in the sessions were also included.

The evaluation of Prototype III was focussing process factors such as the system’s compliance with the reasoning process of the physician and with the logistic process of investigating dementia in clinical practice.

6.3 Methods and Material

Prototype I was evaluated at three occasions with totally six GPs, from three primary care centres, participating in the evaluation. The amount of experience in diagnosing dementia ranged from a few occasions to more than 10 years within primary care. A qualitative approach was taken since the purpose was to obtain as many and different perspectives on issues concerning the system in use as possible. The GPs were observed using the system and interviewed. In the analysis of data each GP was treated as a case in which data was analysed in the framework of activity theory.

The evaluation study of Prototype II was designed as a qualitative case study with the aim of collecting as many different comments on the system as possible, in a setting as close to a real setting of clinical practice as possible. Therefore the physicians were instructed to use current patients in their use of the system, but without the presence of the patient. Consequently, the patients’ cases used were of different kinds; two cases of suspected AD, two cases of suspected or diagnosed VaD and a case of mixed aetiology with no diagnosis yet established.

Six professionals participated in the evaluation; the three available domain expert physicians in the region who had not been involved in the development of the system, one training to become an expert and two persons from a team at one of the primary care centres (PCC) which cooperates with the expert team, of whom one was an experienced GP and one was the nurse responsible for dementia investigations at the PCC. The professionals were observed while using the system with current and in their perspective difficult cases of cognitive diseases and were interviewed. Open questions were used in the interviews. Three of the
6.4 Results

The aim for the evaluations of Prototype I was not only to discover reasons for breakdown situations, but also to investigate which potential the system had to function as a tool for mediation of activity as well as for skill acquisition. The information provided in the breakdown situations were of different kinds, depending on which focus the user had for the activity in that situation, e.g., the diagnostic reasoning process or the prototype system. The approach was found to be useful, as several aspects of the system were highlighted through the framework and modified in the development process. Some of the results were that the graphical user interface was supplemented with explanations of medical terms, which may not be familiar to GPs, some words used in the system were changed to become more comprehensible, and the reasoning model was found not immediately familiar to the persons participating in the evaluations, but also not a subject for objections. Therefore the conclusion was drawn that they agreed to proceed according to the model.

The system allowed and gave support for differences in work methods among the users but put constraints on ad hoc approaches to the investigation of a patient’s resources.

The analysis of observations and interviews resulted in four categories of issues that will be presented; results regarding the reasoning procedure and diagnostic approaches, support for diagnosis and support for teamwork and consultation, and issues concerning learning and development, which will be discussed in the next section, based on results from the evaluation with GPs.
6.4.1 Design for Learning and Development

A continuing medical education was considered one of the most important contributions a CDSS could make to clinical practice [Lindgren et al., 2002]. Providing evidence-based knowledge in the domain in a readable form is one way of supporting learning in a user. The system also gives the user tools or information about tools for collecting data, for instance scales for functional abilities and screening tools for cognitive deficiencies.

Another aspect is the possibility to provide models for reasoning in the decision making process. This has been done by implementing the model for clinical reasoning that was developed in co-operation with domain experts for the purpose of diagnosing cognitive diseases and establishing intervention (evaluated in Chapter 10). The model contains medical frameworks for the interpretations of information in a patient’s case. These models of interpretation are used to determine appropriate intervention in the process. For a physician in primary care who is unfamiliar with dementia investigations, the integration and mediation of these models through the use of a system facilitates the introduction into the domain and into clinical practice. In the use of the system the user may eventually internalize the whole reasoning process or parts of it, as well as factual knowledge. For an experienced GP that already has internalized models for the reasoning processes in uncomplicated and more common cases, the system can provide support in difficult and unusual cases. The user should in both cases be able to choose the amount of support he or she wants from the system.

To be able to learn, the user may have to be made aware of his or her reasoning strategies, i.e., transform operations to an action or activity level. This is done when the system does not give an expected response and a breakdown occurs. In Activity theory this is viewed as a positive thing as it provokes development in the user [Bødker, 1989]. In the evaluation that was made, breakdowns mainly occurred when the system responded with unknown information or the information asked for by the system was of higher granularity than the user was acquainted with. A shift of focus is made in this situation. Instead of working with the patient as the object and with the CDSS as mediating tool, what caused the breakdown will become the object of the activity. In the evaluation, the user’s internalized knowledge or lack of knowledge became the object together with the knowledge expressed by the system.

However, breakdowns may as well occur that put features of the CDSS into focus. If the system is not user-friendly then the user will have to spend valuable time considering the CDSS as the object instead of the illness of the patient. This shift of focus might be acceptable in a short initial phase when the system is introduced into clinical practice but will most likely in the long run cause a
failure in the implementation of the system. In a situation of a breakdown the system should be able to provide additional information to make the user able to proceed with the new knowledge within a reasonable time. This way the system prevents to put the user in his or her ZPD where help from another person is needed.

6.4.2 Reasoning Procedure and Methodology

There were differences in how the individual domain experts approached a task of diagnosing. They valued sources of information differently, they used different tools for detecting deficiencies and they were differently careful in collecting evidence before expressing hypotheses. In spite of the differences in methodology they all agreed that the line of reasoning in the system was in accordance with their own reasoning procedure: “…this is how we work”. One noted that often the process had proceeded a few steps when the patient is referred to the expert team. The patients in this case came from a primary care centre where the team had received education in and a continuing support for dementia investigation from the expert team.

In the case-studies, the analysis-mechanism caused breakdowns and was found unintuitive, which can be improved by making the system analyse the material each time new data is entered. One of the critical features which were directing the differential-diagnosis, was a comparison of the signs obtained in radiology examinations. The issue whether to integrate the possibility to weight the results from such examinations needs to be further discussed and analysed.

The necessary features fluctuating cognitive ability and exposure to toxic substances, were missing in four of the five patient’s cases, which stopped the system to guide further in the reasoning process. This fact was not well highlighted in the GUI of the system, which caused breakdowns. This indicates that the missing features need to be exposed in a more clear way than it is done in the evaluated version. Further, whether these features constitute a hinder for other physicians in Swedish clinical practice as well, need to be further evaluated.

In the case where non experts were unfamiliar with certain screening tools for detecting deficiencies the issue was raised of whether a more ad hoc method of collecting certain data could be allowed. The system suggests screening tools and presupposes the use of certain tools for the purpose of gaining reliable data which is used in the investigation process.

6.4.3 Diagnostic Support

The system was by all physicians viewed as useful for detecting the easier cases in primary care centres so that unnecessary referrals could be avoided. Another need that was expressed by all physicians was additional support for the detection
of early cases of dementia. The system provided support for MCI but lacked support in the cases with a more diffuse aetiology without a clear episodic memory deficit or judgment deficit.

The stream of inference that complied with the guideline DSM-IV may be enough to detect easier cases of dementia, however, when facing complicated cases the expert physicians looked for and expressed a need for additional support. The argumentation mechanism in the system was viewed as useful and as a solution, since it was possible to see the evidence in perspective of different guidelines and other tools for interpretation. Even in the case where the system did not provide a clear suggestion of a diagnosis because of the insufficient knowledge in the domain, the presentation of arguments was considered helpful in the reasoning process.

It was considered important to illuminate the problem of mixed aetiologies where the diagnosis may be for instance, “Alzheimer’s disease with high degree of vascular elements”, where interventions have to be directed towards both diagnoses. This was accomplished by the system in a satisfactory way.

In the evaluation sessions the expert physicians used the inferences in the system to confirm their own hypotheses and conclusions in difficult cases of patients. In the most difficult case where the physicians at that point could not diagnose, the system confirmed that none of the implemented guidelines could provide a suggestion of a single reliable diagnosis when the information was analysed.

6.4.4 Support for Teamwork and Consultation
Dementia care involves several professions at different levels of care, where each level represents a distinct context. Further, each profession contributes with knowledge that is essential to forming the holistic view of a patient that is needed in the management process. The knowledge has to be communicated to be formed and in this process there is a need for cognitive tools that support cooperation. For a GP, the evaluated system can serve as a basis for consultation with an expert physician in a case where a patient otherwise may have to be referred to an expert team. In this situation the system contributes to moving the dementia investigation as activity from a place beyond the ZPD of the physician, into his or her ZPD. Through the use of the system the activity becomes possible for the user to perform in cooperation with the expert. The system has also the potentials of serving as a basis for determining the amount and type of home care needed from the municipality services.

A key issue when developing a CDSS for practitioners at different levels of care is to adjust the terminology to prevent that the different granularity of concepts becomes a hinder for communication [Karlsson, 2001]. Instead the
system should provide a common language for communication between different levels, by for example giving explanations of concepts when needed. In this way, the language remains a tool that mediates activity instead of becoming the focus of activity. The different levels of granularity used in different clinical settings were one aspect taken into consideration in the evaluation, which caused some adjustments in the user interface.

The possibility to keep the information collected in a system instead of having the information about a patient spread over different types of records during the investigation process was considered an advantage. Both to use the system as a checklist and use it as a common ground in the consultations within the team and with personnel from the expert team were considered useful properties of the system. The investigation process is often a highly cooperative process where different professionals contribute with their special competence and knowledge. In this perspective, the integration of additional sources of information that are used in the investigation process were suggested by the physicians. One example is the tool “Assessment of Motor and Process Skills” (AMPS), which is developed and used by occupational therapists.

6.5 Discussion and Conclusions

The implemented process in a CDSS, which aims to support an investigation process in a case of cognitive disease, should be in line with the reasoning processes used by the domain experts in the area. The number of expert physicians in the domain of cognitive diseases is limited in the health care of northern Sweden. Three participated in the evaluation of Prototype II, out of the five expert physicians who were working clinically in the region at the time of the evaluation. Of the two physicians not participating in the evaluation, one was involved in the development of the system and the other expert was on leave during the evaluation period. The evaluation showed that in spite of differences in routines and how to approach the task of diagnosing, they all agreed that the line of reasoning in the system was in accordance with their own way of reasoning. To investigate how significant the methodological differences actually are and what impact they may have on the use of the system, a field study needs to be made with observations of the clinicians using the system in their daily work.

To which extent the methodology of the experts differs from the methodology used by GPs depends on the extent the GP refers suspected dementia cases. In the cases when the GPs run the investigation process, certain tools are used which are accessible and manageable for GPs in their work context while referrals are often used to get access to investigation methods managed in a county hospital, e.g., certain X-ray techniques, screening-methods for detecting cognitive deficiencies.
or BPSD, etc. In these cases the investigation process is often taken over by an expert team. Some of these tools may as well be used in the primary care centres to improve the reliability of diagnoses established in primary care and prevent referrals of less complicated cases. These tools can be disseminated by the system. Other tools are expensive or seldom used and are only used in difficult cases, which make them more appropriate to be used by experts who have access to extra resources in a hospital.

Another need for support, which was expressed by all physicians, was additional support for the detection of early cases of dementia. There is currently poor support for this in published guidelines and evidence based research (see analysis in Part II). Providing a signal of cognitive deficiencies early that may become followed up with further investigations is one way of framing the early cases of dementia.

Since the number of physicians participating in the evaluations was low, these results represent solely indications of that the design directions taken in the development of the prototype, such as whether the model of clinical reasoning and the argumentation mechanism, can be considered appropriate. Concerning design of details in the graphical user interface modifications have being made according to the evaluations during the development process. Further evaluations are needed when the system integrates full support for atypical cases and can be used in clinical practice. Initially, an extended case study is planned with the system integrated in the process as in Evaluation study III, however, involving more patients and physicians. The purpose is to validate the results of the initial case study and to investigate the outcome of the system.
CHAPTER 7
Pilot Evaluation of Japanese Version
DMSS-SJ

When focus is set on the terminology, issues such as the distinction between facts, judgements, qualities of evidence, what is normal vs. abnormal, levels of dysfunctions, distinction between acts of decision and acts of data collection, sources of evidence, etc become essential to clarify in order to formalise the knowledge correctly. An analysis of the qualities of the medical domain knowledge expressed in clinical guidelines is needed [Lindgren, 2005] as well as a thorough work analysis of how the actual clinical work is performed. This is done concerning the local praxis at a clinic in northern Sweden (Paper I). However, since the system is being developed to meet an international need for assessment and support tools for treating the growing population of elderly people in different places, which has in common the growing amount of persons with cognitive deficiencies, there is a need for extending the scope of the system. Due to the urgent need to develop tools which can be used to facilitate the diagnosis and management of the increasing amount of persons with a dementing disease also in Japan, the limited version of DMSS named DMSS-SJ has been developed for research and development purposes.

In cooperation with the IT-company DKK and Ehime University in Japan, the need for adjustments of the system to the Japanese society is currently being investigated. This will be accomplished through iterative evaluations of the prototype DMSS-SJ concerning system’s compliance with clinical practice in Japan, compliance with clinical guidelines and compliance with decisions made by experts in Japan including:
- scope and purpose of the system,
- design of GUI (interaction), and
- content (integrated terminology, evidence, clinical guidelines, clinical screening-tools, reasoning mechanisms, etc).

The work presented here is the results of an initial pilot evaluation conducted in Japan.
7.1 Methods and Material

The study design was qualitative, with the aim of capturing as many aspects as possible in interviews and in sessions where physicians are observed using the system with patient’s cases. The physicians were asked to prepare up to 5 patient’s cases with suspected or established cognitive disease with information about the patient available on paper before the sessions. The prototype system DMSS-SJ was introduced (“hands-on”) at an open session by Umeå University to available medical experts at different departments of Ehime University hospital and one primary care clinic. In the initial session one patient’s case (case 1 in Table 4) was analysed while aspects of the system was discussed. After that, individual sessions were held with physicians using the system with their patient’s cases. Of the seven physicians who participated in the study, five participated in the initial session. Further, a few of these had used the system for a short period on their own, in ongoing investigations of patients, or had tried the system with completed cases.

**Table 3.** Demographic characteristics of the seven physicians participating in the study, and the clinical diagnoses of the 21 patient’s cases. The cases marked in bold are typical cases and for those marked in italic support is not included in the design of the prototype.

<table>
<thead>
<tr>
<th>Physician</th>
<th>Type of Clinic</th>
<th>Pat 1</th>
<th>Pat 2</th>
<th>Pat 3</th>
<th>Pat 4</th>
<th>Pat 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>University hospital, Dept Neuropsychiatry</td>
<td>C-J Disease</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>B</td>
<td>University hospital, Dept Neuropsychiatry</td>
<td>VaD DLB FTD VaD</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C</td>
<td>University hospital, Dept Pathology</td>
<td>AD Chr 14 CBD</td>
<td>AD</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>D</td>
<td>University hospital, Dept Geriatrics</td>
<td>DLB VaD AD VaD MCI</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>E</td>
<td>PCC (patients in group home)</td>
<td>AD AD</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>F</td>
<td>Psychiatry clinic</td>
<td>CBD Pick’s disease CBD</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G</td>
<td>PCC/private hospital (patients in group home)</td>
<td>VaD AD Parkinson dementia</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

All the physicians participating in the study had profound experience in diagnosing and treating patients with dementia diseases and several have contributed with research in the domain. In their current work environments they were focussing different problem areas: two were primary care physicians, both also
monitoring group homes for persons with dementia and one managing a small hospital, two were experts in neuropsychiatry at a specialist clinic, thus experts on cognitive dysfunctions, one was an expert in Geriatric medicine, monitoring a geriatric clinic, one was an expert on BPSD in dementia, working in a psychiatric clinic and one was currently researching on pathology in dementia. Therefore, their patient’s cases represent both typical patients as well as extremely rare cases of dementia. The proportion was 5 cases of 3 different rare diagnoses for which support was not integrated in the DMSS-SJ system; 10 typical cases and 6 atypical cases in the sense of [Lindgren, 2005]. One of these 6 atypical cases was an extremely rare type of AD (less than 10 known cases in Japan). Table 3 shows the demographics of participating physicians and the clinical diagnoses of their respective patients.

The physicians had prepared 1-5 familiar patient’s cases each. Some patients had been in their care for years during the progress of the disease, which were the case of the PCC physicians who were monitoring group homes.

The sessions were recorded on video and analysed. Notes were taken and the patient’s cases in the DMSS-SJ system were saved for analysis. The observers took an active role in the sessions and interviewed the physician during the use of the system. In sessions with two physicians additional persons were participating, functioning as interpreters. In some sessions additional physicians were participating, giving comments on features in the system during the sessions.

7.2 Results

The results are preliminary, since the study is a pilot study, based on a limited amount of physicians and patient’s cases. However, the results indicates important aspects and amendments which need to be addressed and handled before proceeding further with more extensive evaluations with the system integrated in clinical practice in Japan.

7.2.1 Compliance with Clinical Guidelines and EBM from Japanese Viewpoint

The evaluation showed no differences in how the main diagnostic procedure is processed when compared to international clinical guidelines and evaluations of the diagnostic procedure in Sweden [Lindgren, 2005].

There are differences in how certain features and concepts are used in clinical practice in the different countries, which lead to a clarification of these in the Japanese version in order to prevent misunderstandings.

A handbook in dementia diagnosis and management, written also for educational purposes (in Japanese) [Miki, 2005] was briefly reviewed and compared to
the content of the system. The guidelines and clinical tools reviewed in the book were mainly the same as those integrated in the original DMSS system, and of which the DMSS-SJ system presents a sub-set. However, the main guideline used in clinical practice for the assessment of VaD among the physicians participating in the study is the NINCDS AIRDEN criteria, which is not integrated in the set of guidelines integrated in the system for identifying typical cases. However, the criteria is integrated for the assessment in atypical cases. A change of initial criteria for assessing VaD should therefore be done in future Japanese versions.

7.2.2 Compliance with Clinical Work Process

One of the two primary care physicians who participated in the study refers the persons with suspected dementia to the geriatric department at the university hospital for diagnosis, since they do not have access to X-ray equipment at the PCC. At the geriatric centre and at one PCC clinic they schedule an initial meeting with the patient early one day. The patient stays at the clinic for examinations such as X-rays and laboratory work ups during the day and meets the physician again later the same day. If the physician has received all results (which is usually the case), then the patient receives a diagnosis, and treatment begins. An initial follow-up is scheduled two weeks later and there are continuous follow-ups until the patient is considered finished in their perspective. During the investigation period, the patient is in focus for interventions such as day care and rehabilitation by other health care professionals.

Two of the physicians had access to the system during a period of one month before the evaluation sessions were held. The physicians used the content of the system printed on paper as a checklist in the encounters with new patients, and one explained that he was verifying his diagnoses by using the system after the patient’s data was collected.

7.2.3 Compliance with Expert’s Assessment of Diagnosis

The proportion of rare cases was high in the limited amount of patient’s cases in this pilot study. In the cases that were clinically diagnosed with the diseases Creuzfeldt-Jacob’s disease (C-J disease), Pick’s disease and Cortico-based dementia (CBD), the system could not assess the correct diagnosis since guidelines for these were not implemented in the system. Instead, the system suggested in two of the CBD cases a typical AD and in the third case an atypical FTD. This case was misdiagnosed during 10 years as an untypical AD, until a neuropsychiatry expert excluded the AD diagnosis. The correct diagnosis was obtained by autopsy. The system suggested the C-J disease case to be a typical VaD, while in the case of Pick’s disease the system left no suggestion since the evidence did not match the criteria of any of the guidelines integrated in the system. Instead
further investigation was recommended.

**TABLE 4.** Overview of patients participating in the study. Which response in form of a suggestion of diagnosis that was given by the system is shown, and by what strategy the system handled the diagnosis: as a typical or atypical, and whether support for the diagnosis was implemented in the prototype.

<table>
<thead>
<tr>
<th>Case</th>
<th>Clinical Diagnosis</th>
<th>Suggested diagnosis</th>
<th>Typ</th>
<th>Atyp</th>
<th>Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Creutzfeldt-Jacob disease</td>
<td>VaD</td>
<td>X</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>VaD (vascular dementia)</td>
<td>AD (Alzheimer’s dis.)</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>DLB (Lewy body dementia)</td>
<td>probable DLB</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>FTD (fronto-temporal dementia)</td>
<td>FTD/AD</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>VaD</td>
<td>VaD</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>AD chromosome 14</td>
<td>FTD/AD</td>
<td>X</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>untypical AD (clin) CBD (autop)</td>
<td>FTD/AD</td>
<td>X</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>AD</td>
<td>AD</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>DLB</td>
<td>DLB</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>VaD</td>
<td>VaD</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>AD</td>
<td>AD</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>VaD</td>
<td>VaD</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>AD</td>
<td>FTD/AD</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>AD</td>
<td>FTD/AD</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>CBD (cortico-basal dementia)</td>
<td>AD</td>
<td>X</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Pick’s disease</td>
<td>Info. not sufficient</td>
<td>X</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>CBD</td>
<td>AD</td>
<td>X</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>VaD</td>
<td>VaD</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>AD</td>
<td>FTD/AD</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Parkinson dementia</td>
<td>Parkinson dementia</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>MCI (Mild cognitive impairment)</td>
<td>MCI</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Part from these cases for which the system did not integrate the clinical guidelines, the system’s suggestions of typical diagnoses correlated with the clinical diagnoses in 9 of the 10 cases. This was not in the case of Patient 2 due to the different clinical guidelines that was used for diagnosis. In the system DSM-IV is used for the assessment of VaD in typical cases, while the clinical differential diagnosis in this particular patient’s case was based on the more strict guideline NINCDS AIRDENS, thus providing the clinical diagnosis AD instead of VaD, which the system suggests. In the remaining six cases the system did not suggest one single diagnosis. Instead the system presented a summary of critical features which supports each diagnosis for which clinical guidelines were implemented in the system.

In one case the system interpreted the case to be atypical since the evidence supported a probable DLB diagnosis, which means that one of the core features
was missing. The evidence is however enough support for the physician for establishing the clinical diagnosis DLB.

The dominant situation (5 out of 6) was that the evidence supported both AD and FTD, however, not the FTD criteria completely so that it could be judged to be a typical FTD. One of these cases was an example of a rare type of AD with early onset (less than 10 known cases in Japan) and three other were patients who were currently living in group homes. These four cases were cases in which the disease had been proceeding further than in the other cases, a conclusion which was supported by the score they had on the Hasegawa scale (below 10 points). In these cases more symptoms of a various kind had been developed, which in distribution resembles FTD. Whether this is an indication of that the system is more suitable for diagnosis early in the development of the disease, or that the system actually identify complicated cases which have been misdiagnosed in earlier stages, needs to be further investigated. In one of these cases, the physician explicitly was going to use the output of the system to reconsider the clinical AD diagnosis. A summary of the patient’s cases is provided in Table 4.

A need to be able to re-consider features and diagnoses in a patient’s case with the support of the system was expressed by the physician who had been using the system during some weeks. In complex cases the physician wants to be able to change the valuation of the importance of certain features, such as x-ray results, in the process.

The amount of data that is possible to enter in the system was considered too much for a general practitioner to enter in daily practice, and too time-consuming. The way of using the system by the physicians was to systematically enter all known features by frame before using the analysis function. The alternative, to enter only the critical data when the system requires data, was considered confusing in the way it is designed in the system.

7.2.4 Purpose and Scope of a Japanese DMSS
The need expressed by representatives from primary care is mainly support for

- Diagnosis,
- Assess the level of progress and severity in an individual patient,
- How to care for the patient – Interventions, primarily other than drug treatment,
- Determine the level of aid the patient need in relation to the care provided by Japanese health care system.

It is more important to aid the assessment of severity and how to care for the patient than to aid diagnosis, in the perspective of the primary care physician
who refers the majority of his patients for diagnosis. Three of the physicians who receive referred patients for diagnosis expressed that they use, or would like to use the system for verifying their clinical diagnoses, primarily in the difficult cases, and as a checklist for clinical investigations. A few of these also viewed the content of the system to be too rich for the primary care and too time consuming to enter all information.

The screening tools most commonly used in clinical practice were the Hasegawa test, MMSE and FAST. Further, a clinical test called “NM Skill system” was used for assessing the difficulties in daily life. A common need expressed by all physicians is tools for assessing BPSD and for support on how to handle the care for persons with BPSD. The Behave-AD was perceived as an interesting option for the assessment. These tools will be integrated in future versions of the Japanese system. Further, the integration of a module for handling the valuation in the perspective of the Japanese health insurance system requires a careful analysis of the specific domains of analysis and which information can be transformed between valuation frameworks. This particular extension module of the decision-support system will also be a subject in future evaluations with the aim of integrating the support.

To summarise these design aspects, the usefulness and purpose of the system differs, mainly depending on whether the local routines at a certain clinic include diagnosis or not. Therefore, the system should distinguish these processes 1-4 listed above, in order to provide the support the local practice needs to integrate. Further, the design of the system should also be clearer about what information is needed for the different purposes.

7.2.5 Terminology

The terminology in the Swedish system has currently been evaluated towards international clinical terminologies and classifications. Terms, as well as definitions, have been identified in these frameworks. These are in English and there are none yet translated into Japanese. Therefore, the Swedish terms in the original system is currently being translated into English, which will constitute the common language for different future versions of DMSS.

A critical difference between the Swedish/English language and the Japanese language concerns how negations are treated in common language, when “yes” means “yes” as an answer to a question formulated in a particular way. Several confusions and discussions were caused by the different ways of expression. As a result negated expressions are replaced with neutral terms such as “Cognitive functions” instead of “Affected cognitive functions”. The valuation of the terms is limited to the checkbox with alternatives, in which the context-dependent terms “yes” and “no” are replaced with “normal” and “not normal”, except when related
to direct questions. By using neutral terms, the content of checkboxes can easily be replaced in the development, if there would arise reasons to provide the user a more detailed scale for assessment of severity, for instance. The approach is similar as the approach used in the International Classification of Function, Ability and Health (ICF) [WHO, 2001], where neutral terms are used for ability and scales provide information if and to which extent the ability is decreased.

Other expressions which appeared as unclear to the Japanese users were mainly context-dependent expressions typical for the Swedish clinical practice and medical society. These are currently being clarified and reformulated in the English and Japanese versions.

Further, some medical expressions were used differently by the Swedish and Japanese clinicians. The differences lie in what is included in for example “focal neurological signs”, which includes Apraxia in Japan, however, not in Sweden. In order to make the data useful for research purposes, such distinctions need to be clarified and implemented in a way that the inferences become valid and data is the same, regardless if it is collected in Japan or in Sweden.

The physician who was the most fluent in English among the physicians expressed that he would prefer an English version in order to use the original English terms and in this way avoid misunderstandings. However, since the medical education in Japan is conducted in Japanese with Japanese literature with only a few exceptions, a Japanese system is considered vital.

7.2.6 Design of Interactive Reasoning

All physicians systematically entered all information available about a patient, before using the analysis function. In one case, the interactivity of the system reminded the physician about necessary evidence that was missing. However, the features were missing only because the physician slipped and missed the object in the interface when entering the information. The way of interacting with the system by letting the system guide the gathering of only the necessary information was demonstrated to six of the physicians, who tested the method in one patient’s case each. The physician who was the most familiar with the system, perceived this method as very confusing. The other physicians did not comment on the procedure, other than they confirmed that it would be good to have a selection of necessary information in order to make the interaction and the procedure faster.

The distinction of the different sources was difficult to internalise by most of the physicians, especially to distinguish the patient’s own view. The initial reaction was that it seemed to be multiples of the same features. A second reaction was that some thought that it could be difficult for relatives to distinguish the symptoms at the level of detail presented in the system, however, not for
7.3 Conclusions

An evaluation is made of the initial prototype version of a Japanese dementia management support system (DMSS-SJ) based on a Swedish decision-support system under development for supporting the investigation of dementia. The results are preliminary, since the study is a pilot study, based on a limited amount of physicians and patient’s cases. However, the results indicates important aspects and amendments which need to be addressed and handled before proceeding further with more extensive evaluations with the system integrated in clinical practice in Japan. The results include differences in terminology, usage of clinical guidelines and screening-tools, and a difference in how they perceive the interactive reasoning compared to the results from Swedish evaluations. The results indicate also similarities, primarily in reasoning procedure, the logistics of the investigation of dementia and the major purposes of a decision-support and management system for dementia care.
CHAPTER 8
Current State of DMSS and Future Work

The prototype systems DMSS and DMSS-SJ are currently used in controlled settings in Sweden and Japan for evaluation and development purposes. The prototype systems integrate international clinical guidelines, terminologies and a reasoning procedure, which is validated towards clinical guidelines and clinical practitioners in Sweden and Japan. Further, each system integrates local process knowledge, in the northern of Sweden and a municipality in Japan, respectively.

Qualities in the domain and process knowledge have motivated the ongoing development of formalisation techniques for the purpose of handling incomplete, ambiguous and imperfect domain knowledge and integrating process knowledge within the same formal framework.

The development and implementation of the conceptual model of the knowledge base presented in Paper I is the major task, which will generate a re-design of the GUI and interaction as well as rulebases in accordance with local practices. In this work terminologies and further formalisation of guidelines are central, as well as a creative solution for implementing the interactive reasoning.

To this point in the development process, qualitative aspects and goals have been in focus, rather than quantitative goals in order to assure the creation of a foundation as rich as possible for the content and design of the system. Therefore, each line of investigation needs further development before the system can be fully integrated into clinical practice. This includes quantitative studies on the effects of distinguishing between typical and atypical patient’s cases (currently being done in a study in Japan), quantitative studies on specifics of the interaction with the system and qualitative studies focussing the effects on the user’s development of knowledge and skills. Further investigation of the information transformations in the investigation process is needed in order to investigate details concerning the system’s relation to other systems in the user environment.

Extensions will be developed, which integrates support for care at different levels, as well as decision support for other care personnel than physicians.

The development of DMSS and deployment in clinical practice will continue within the nearest future as a research and development project in cooperation with industry, local universities and health providers in Sweden, Japan and Chorea, governed by Umeå University and LifeScIntel AB.
PART II

ASSESSMENT OF KNOWLEDGE FOR DECISION SUPPORT

In this part some aspects of human decision making and acquisition of knowledge will be described that are of particular interest in the formalisation of the domain knowledge. The traditional views of humans’ ways of handling incomplete knowledge in decision making, which have been highly influential on the design and development of decision-support systems, are presented in Chapter 9. A review of how expertise is manifested in the medical domain is given and compared to how novices perform diagnostic reasoning. Further, a model of the clinical reasoning process as performed in dementia investigation is introduced and evaluated in Chapter 10. Based on the model, the domain knowledge of cognitive disorders as presented in evidence-based medicine and in clinical guidelines is reviewed and summarised in Chapter 11. Further, the terminology used in the system was validated and the results are summarised in Chapter 12.

The work of assessing the clinical work context is summarised in Chapter 13 as an introduction to Paper I, where a conceptual model of clinical activity is introduced and used for the transformation of informal knowledge in EBM together with knowledge residing in the clinical work practice, into a semi-formal structure. The purpose is to use the structure for assessing and formalising the knowledge suitable for formalisation, and distinguishing the knowledge which needs to be provided by a careful design of the interaction, promoting interactive reasoning.
CHAPTER 9
How do Humans Reason and Make Decisions?

The cognitive processes humans use in reasoning and decision making have been studied, in particular within neuropsychological, social psychological and cognitive scientific studies. There are two main approaches dominant in these studies. The first is the traditional experimental setting, which has been the most frequently acknowledged in computer related issues. The focus has often been to investigate isolated processes in order to compare humans’ performance and strategies to formal frameworks implemented into computer systems. We find here studies which present results of humans’ inadequate reasoning with numerically represented probabilities and how decisions are affected by how linguistic information is presented to the decision maker. The main critique to these studies is the poor ecological validity of studies done in a controlled laboratory setting.

The second approach investigates the influence of the environment on decisions, and larger contexts of decision making are framed in social or organisational theories [Nardi, 1996; Engeström, 1992]. Naturalistic decision-making (NDM), represented by the work of Patel among others [Patel et al., 2002], is a term used for research focussing complex decision situations in particular. NDM has emerged as a research area with complex, rapidly changing work environments in focus, where decisions are considered impossible to be studied separated from their context. This is highly relevant in medical situations where decisions are commonly continuously made in a teamwork setting often under stress and time pressure.

In psychology and cognitive science, studies have been made of humans’ strategies for decision making and judgement on the basis of incomplete knowledge. There are three normative viewpoints used when humans’ methods of handling ignorance is to be explained [Smithson, 1989]; the knowledge seeker, the certainty maximizer and the statistician orientation. Together they form the modern western ideal approach for humans in dealing with ignorance, where the first step is to maximally reduce ignorance by seeking full knowledge ignoring nothing that is relevant. Secondly, gain maximal control by exposure to and responding to fully informative, certain, controllable stimuli or environment. Lastly, when ignorance is further irreducible, treat uncertainty probabilistically and ignore other types of ignorance and finally, select the alternative which maximizes the expected utility. The statistician approach is concerned with criteria for
rationality in thought and behaviour and makes use of Bayesian probability and the maximization of expected utility. A crucial assumption in this perspective is the independence of uncertainty and utility, which should make the assessment of subjective probabilities value-free. These theories according to Smithson assume a

\[ \text{well-adjusted, fast-learning, adaptively responsive, rational individual who follows their prescriptions.} \]

Focus of debates and research is largely descriptions of deviations from these prescriptions, since the common human decision maker seems to be subject to biases and deficient in applying probability theory to decision problems. We will review a few of these aspects in the following section, starting with the statistical viewpoint.

Studies have shown that humans assign different and overlapping numerical estimates of probability to linguistic expressions like possible, often, rare, etc. Usually recommendations based on these studies are to use numerical expressions rather than linguistic expressions, or to minimize the set of expressions. However, these studies do not indicate whether the linguistic terms change with time and context, or if the terms are vague and/or non-specific [Smithson, 1989]. A study made by Karlsson shows that medical terms are highly context-dependent where a consensus of meaning is created locally in a work environment [Karlsson, 2001]. The studies also do not indicate whether the terms with large overlaps and variance actually represent some other nonprobabilistic kind of uncertainty as the notion of possibility defined by Zadeh [Zadeh, 1978].

Zimmer [Zimmer, 1983] claims that humans process information for the purpose of making predictions more in the way of putting forward arguments than making estimates of parameters. This would imply that the humans participating in these studies were using cognitive processes requiring more effort than their usual mode.

Over- or underconfidence in the truth of one’s predictions and over- or underconfidence in one’s ability to distinguish true predictions from false, have been subject of research. Studies show a general tendency towards overconfidence in experts, which increases with task difficulty.

People tend to neglect information of disconfirming or excluding nature, a phenomenon addressed in research as the confirmation bias. We also tend to dismiss information that does not confirm already held beliefs. The confirmation bias refers to selective attention in accessible information as well as selective information seeking, interpreting and testing. These biases are manifested in clinical reasoning, which stand in contrast with the classical decision-theoretical paradigm as an ideal.
Extensive investigations have been made for the purpose to describe how clinicians perform the complex task of clinical diagnosis [Evans and Patel, 1989], [Patel et al., 2002]. One of the strategies found is that working hypotheses are generated early in an abductive manner, when only a few facts have been collected, through some kind of pattern recognition. In pattern recognition a physician recognises a particular sample of data as signifying a particular patient state, which is an ability that is developed with experience. Then the hypotheses are confirmed or rejected, while new information is gathered and new hypotheses are generated. Studies made by Patel et al. show that if an expert includes the correct hypothesis in the initial set of hypotheses, the remaining reasoning tends to aim at confirming the hypothesis instead of generating new hypotheses [Patel et al., 2002]. A less experienced physician has more difficulty in evaluating hypotheses and therefore more difficulty in eliminating incorrect hypotheses. The less experienced also tend to continue generating new hypotheses, even after the correct diagnosis is produced. Thus, what seems to distinguish the expert from the novice is the knowledge of what not to do. The experts participating in these studies may have been guilty of a confirmation bias, which, if it is the case, is shown to be highly productive in many situations, in these studies as well as in others.

Human clinical reasoning is categorised by Kassirer into three abstract categories: **probabilistic**, **causal** and **diagnostic** [Kassirer, 1989]. Formal models for each type have been developed. Humans are not good statisticians when they perform probabilistic reasoning as described in previous chapter.

Causal (backward) reasoning is based upon cause-and-effect relations between clinical variables and can be viewed as the reverse of diagnostic (forward) inferences where the reasoning proceeds from consequences to causes (Figure 9.1). The major difference between the approaches is the amount of data that has to be accounted for. To be able to perform diagnostic inferences the physician
has to handle a wide range of information collected in each case and among the
more or less relevant information extract the important features that supports a
diagnosis. When the information is ambiguous, or too confusing, the approach
of causal reasoning is a way to come to a conclusion by narrowing down the
factors to work with. It is also a way of identifying additional information that is
needed in the process. Humans tend to rely more on causal data than diagnostic
data when making judgement under uncertainty. Novices use causal reasoning to
larger extent than experts. Experts on the other hand tend to use causal reasoning
only when they face problems outside the scope of their expertise, or facing very
unusual cases, or when they try to explain their reasoning to others [Evans and
Patel, 1989]. Consequently, the novice supervised by an expert may be given the
causal version of the expert’s reasoning, which may reinforce the novice’s own
causal reasoning rather than his or her diagnostic reasoning.

It is shown that diagnostic reasoning is connected to very high levels of ac-
ccuracy in diagnosis [Patel and Groen, 1991]. Human diagnostic reasoning is
sometimes seen as following rules similar to IF-THEN production rules where
certain conditions demand appropriate actions. Studies have shown that diagnos-
tic reasoning is more dynamic than solely production rules or pattern recogni-
tion, since the development of expertise is not connected to an increased quality
of representations [Magnani, 1992]. Typically, diagnostic reasoning is used to
generate explanations, until loose ends appear. These are tied up using causal
reasoning. The amount of generated loose ends decreases with experience. In
the cases where pure diagnostic reasoning is used, cases which involved experts
in the study, the reasoning also produces the correct diagnosis. It appears there-
fore appropriate to encourage the use of diagnostic reasoning as done in medical
education and training, also in the use of a decision-support system.

There are also differences between novices and experts in ways used to tackle
daily clinical practice. Novices tend to use single purposes defined by official
norms, single-layered surface relations and detailed actions. Experts handle sev-
eral conflicting purposes while using official as well as unofficial norms and also
personal judgement in the process. Other aspects of expertness are ability to
handle complex multi-layered social relations and use high-level actions. This is
for example shown in interactions with patients where data is collected. While
the novice directs the conversation following an implicit or explicit manual or
protocol, the expert is able to follow the patient’s choice of subjects and still
gather the information needed [Beuscart-Zéphir et al., 2001]. A consequence
that has been shown is that experts tend to be more reluctant to use computer
applications that put constraints on the activity and reduce flexibility.

Clinical practice is extensively augmented with the implicit situated knowl-
edge shared in a specific clinical situation that needs not to be expressed because of the mutual understanding among team members. Such knowledge typically consists of how terms and language are used, how situations should be handled, who is responsible of what, priorities, etc [Berg and Goorman, 1999; Karlsson, 2001].

In clinical practice experts put more effort on the representation of the problem by “sizing up” the situation rather than generating decision options, a process called situation assessment [Patel et al., 2002]. Novices on the other hand tend to jump to generating hypotheses and implementing solutions, sometimes based on conflicting data, in situations where experts defer their decision until they have more evidence, while meeting the immediate need of the patient.

To summarise, expert strategies commonly involve a range of heuristics and are associated in research with high levels of accuracy. Experts represent their decision problem in a way where comprehensive patterns emerge from the data in order to be able to handle the mixture and quantity of data. However, the heuristics used are typically viewed as undesirable in classical decision research. The aspects summarised in the review are taken into consideration in the design of DSMSS in order to meet the needs of both expert and novice users, and in particular, providing support in the process of developing skills and knowledge in novice users.

In the following chapters the evidence-based knowledge of the domain of dementia will be analysed from different view-points and supplemented with heuristics used by domain experts. Further, in Chapter 13 a summary of the results of case studies of patients in clinical practice are presented (Paper I), which reveals indications of the compliance of a physician’s reasoning process with the reasoning process implemented in the system as well as environmental factors.
Chapter 10  
The Model of the Main Clinical Reasoning Process

A model for the clinical reasoning process was formulated by a domain expert and implemented in the prototype system for investigation of cognitive diseases [Lindgren et al., 2002]. It guides a diagnostic reasoning process, which keeps a wide perspective initially with differential diagnoses, a wide range of features, and narrows down the focus while differential diagnoses are excluded and general syndromes guide way towards specific diagnoses. The purpose of the model is to give less experienced physicians a guide into the diagnostic reasoning process in order to improve their diagnostic skills. The model is augmented with tasks of investigation, valuation, etc.

The model consists of the following five steps:

1. determine the existence of cognitive disease,
2. determine the type of cognitive disease,
3. if a state of dementia, determine the type of dementia,
4. interpret findings by using models of reference and establish intervention accordingly.
5. evaluate and modify interventions.

The model was evaluated at two different occasions, first with general practitioners with different amount of experience testing the system. One of the purposes was to find out whether the experts’ view of the process was intuitive and useful for general practitioners. The main purpose of the second evaluation was to validate the model with other domain experts.

In order to not only evaluate the model in the local setting of dementia care in northern Sweden, the model was put into the wider context of evidence-based medicine. This was done by validating the model in the perspective of clinical guidelines and the guideline DSM-IV in particular (see the next chapter for description), for the purpose of investigating the compliance of the model with the reasoning expressed in guidelines.
10.1 Compliance with Physicians

Seven general practitioners at different primary care centres in the region of northern Sweden, with different amount of experience, and the three domain experts in the region that were available at the time of evaluation participated in the evaluations. They were observed and interviewed during and after their use of the system.

The primary care physicians except one used the system with the first three steps of the model integrated concerning diagnosis [Lindgren, 2001]. In the initial evaluation, the naming and the use of the analysis mechanism, which moves the process to the next step, were viewed unclear. A poor initial description of how the system works and for a physician non-informative name for the analysis mechanism were considered as the main reasons for the confusion, which was diminished in the following evaluations after modifications of the user interface. Another reason may have been the fact that when questioned after their use, none of the primary care physicians were familiar with the diagnostic steps. However, none questioned the procedure during their use of the system, which indicates that they did not oppose to the model of the process. It should be noted that there were large differences in routines among the general practitioners participating in the evaluations, in how the investigations of cognitive diseases were done.

There were also differences in how the individual domain experts approached a task of diagnosing. They valued sources of information differently, they used different tools for detecting deficiencies and they were not equally careful in collecting evidence before expressing hypotheses. In spite of the differences in methodology they all agreed that the line of reasoning in the system was in accordance with their own reasoning procedure: “...This is how we work”.

One physician noted that the process often had proceeded a few steps when the patient is referred to the expert team. The patients in this case came from a primary care centre where the team had received education in and a continuing support for dementia investigation from the expert team. The possibility to enter the reasoning process in the system when the first steps already have been accomplished without the system, was considered important in the case where the needed support concerns dementia diagnosis or intervention.

10.2 Compliance with Guidelines

The overriding purpose of clinical guidelines is to improve the quality of care for patients. Clinical guidelines are developed and validated by groups consisting of experts in the specific area. Since most good health care depends on multidisciplinary teams, the implementation of clinical guidelines should be planned from this perspective [Feder et al., 1999]. Therefore it is desirable that the multidisciplinary group participates in the development in the way that they plan the local
10.2 Compliance with Guidelines

settings and adapt the guideline, as well as the evaluation within the local setting.

Clinical guidelines (GLs) should be based on evidence-based medicine (EBM), presented and referred to in the guideline. They often have a broader scope than systematic reviews and provide an overview of the management of a condition or the use of an intervention. Clinical guidelines can be used by individual clinicians for continuing professional education [Feder et al., 1999; Karlsson, 2001]. Clinicians may also use guidelines in their day to day practice to answer patient related questions. Evidence-based clinical guidelines can be used as evidence in daily practice, along with other types of relevant evidence like systematic reviews, individual trials and expert advice. Guidelines should be seen as only one strategy to find evidence since a guideline cannot cover all the uncertainty in clinical practice [Feder et al., 1999].

The guidelines used in the prototype system are the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) [American Psychiatric Association, 1994], the consensus guideline of the Lund and Manchester groups for frontotemporal dementia (FTD) [Neary et al., 1998] and the consensus guideline for Lewy Body dementia (DLB) [McKeith et al., 1996]. DSM-IV is used as the basis since it contains evaluated criteria for the state of dementia and is viewed as useful in clinical practice. The chapter Delirium, Dementia and Amnestic and Other Cognitive Disorders in DSM-IV is analysed in the following comparison with the model of clinical reasoning process. A thorough account for the content in the guidelines will be given in next chapter.

10.2.1 Step 1: Determine the existence of cognitive disease

The first phase of the diagnostic process starts when it is expressed by the patient or relatives that “Something is wrong, because...”. Usually a situation is described where the cognitive ability does not hold in a particular practical situation. In order to determine the existence of a cognitive disease anamnesis is obtained from relatives and a screening for cognitive dysfunctions is made with certain tools. In the model memory deficit and judgement deficit is used as a measure of cognitive dysfunction and as criteria for cognitive disease. If present, further investigations need to be done.

In DSM-IV memory impairment is distinguished as the first feature mentioned in the criteria for each type of dementia diagnosis and for amnestic disorders, but for delirium and mild cognitive impairment memory dysfunction is not a necessary condition, rather an example of one symptom among others.

Sometimes judgement is included in executive cognitive functions. However, judgement deficit is not explicitly mentioned in DSM-IV, but was considered as an essential indicator that further investigation has to be done, even if a memory deficit is not a problem for the patient. A judgement deficit may affect the safety
of people around the patient since, for example, driving ability is commonly diminished. Therefore, this feature is explicitly accounted for in the DMSS.

This step in the process is implicitly stated in DSM-IV as an indicator of dementia. At this point the recommendations in other guidelines are related to which methods that should be used in the investigation process.

10.2.2 Step 2: Determine the type of cognitive disease

The second phase in the model is a differentiation diagnostic procedure to eliminate alternative hypothetical causes of the cognitive decline, other than dementia. To accomplish this, further investigations are needed. Firstly, cognitive functions other than memory are investigated. The severity of the cognitive decline, exposure of toxic substances (including drugs), onset and progress are examined. At this stage there is enough evidence to give the appropriate cognitive disease diagnosis among those defined in DSM-IV, unless the cognitive decline is caused by something else. Therefore X-rays, laboratory tests and physical examination are also made in order to rule out other causes.

Just as in the previous step, this differentiation among causes is integrated in the criteria for each dementia diagnosis in DSM-IV and is not stated as a separate step. However, the differential diagnosis accomplished by this step is based on a common description for all dementia criteria in DSM-IV. Therefore the model is considered consistent with the guideline.

One deviation from the DSM-criteria at this step is the heuristic clause of judgement deficit which implies a state of dementia, in spite of intact semantic memory. This is considered as a deviation of factual knowledge rather than procedural knowledge, therefore the issue will be further discussed in the next section.

10.2.3 Step 3: Determine the type of dementia

When step 1 and 2 are accomplished and a state of dementia is established in a patient, the reasoning process moves on to determine which type of dementia the patient is suffering from. The guideline DSM-IV is preferred in clinical practice because it contains a level of granularity that is not too high for clinical use [O’Brien et al., 2000]. However, since it lacks some of the dementia types, other guidelines have to be considered as supplement.

In the guideline DSM-IV only the two most common types of dementia, Alzheimer’s disease (AD) and vascular dementia (VaD), are specified with diagnostic criteria. Dementias due to other general medical conditions are listed without criteria but with a reference to the causing disease (e.g., Parkinson’s disease, Huntington’s disease, Pick’s disease, HIV, head trauma, hypothyroidism, etc). Dementia of Lewy body type (DLB) and frontotemporal degeneration (FTD) are
not described in the guideline DSM-IV.

The model is in accordance with the synthesis of the guidelines DSM-IV, consensus criteria for FTD and DLB in diagnosing dementia.

10.2.4 Step 4 and 5: Investigate grounds for and establish interventions

At this stage, the presence and amount of BPSD (behavioural and psychological symptoms in dementia) and ability to function in daily life need to be investigated, in order to determine appropriate interventions. DSM-IV provides a coding system for diagnoses in which it is possible to code complicated states with BPSD symptoms, but it provides no suggestions of tools for investigations. Reviews of evidence-based medicine such as the “practice parameters” concerning diagnosis and management of dementia [Petersen et al., 2001; Knopman et al., 2001; Doody et al., 2001; Eriksson et al., 2000], give account for cognitive screening tools, functional tests, etc that a physician can use in order to investigate deficiencies. They also provide recommendations for interventions. Taking these into account, the model contains the necessary actions needed to complete the investigation process based on evidence-based medicine.

10.3 Discussion

The medical domain of cognitive diseases is complex with several imprecisely described features to consider in a diagnostic process. The diagnostic reasoning process performed by an expert and the model of reasoning implemented in the prototype system, proceeds from general information and syndromes toward more specific information and diagnoses. It can be viewed as a continuous evaluation of features, focussing on qualitatively different aspects of the deficits, as type or quantity of cognitive features, time aspects, severity of cognitive dysfunctions and qualities like fluctuation and descriptions of progress. Some of these are viewed in figure 10.1 in which the complexity is increasing for each step in the process.

In this evaluation process an expert uses his or her observation skill to distinguish for example fluctuating cognitive ability from disorientations caused by a stressful environment, often through some kind of rules of thumb, or heuristics. The use of heuristics can be questioned, since they are considered as one cause of failures in human reasoning. Research shows that expert physicians develop heuristic methods with increased skill and expertise, and perform better than novices. Novices tend to rely more on the ideal procedure of decision making, with extensive information gathering and creation of hypotheses, trying to understand the underlying pathophysiology of the patient’s problem by mapping from a causal reasoning procedure. In the diagnostic process the expert physician makes
use of observation skills in combination with heuristics to capture abnormal and in some cases urgent symptoms, which often direct the reasoning process towards a diagnosis faster than would be the case if only the clinical guidelines were used in a solely diagnostic reasoning process. In some cases these heuristics may even be the primary reason why the physician defer the decision making when there is contradicting or atypical evidence, cases in which the novice tend to jump to conclusions and implement solutions on inconsistent grounds [Patel et al., 2002]. If heuristics used by an expert can become explicit and motivated, these can also be used by, and become integrated in a non expert’s reasoning. The few heuristics identified in this work so far can be motivated to a certain extent by EBM, but the larger emphasis put on these features is primarily motivated by practical and safety reasons.

10.4 Conclusions

A model of the clinical reasoning process in the domain of cognitive disease investigation has been evaluated and found in accordance with domain experts’ view of their work process, as well as the implicit process in the clinical guideline DSM-IV and other guidelines. In this process heuristic rules used by experts have been identified.

The reasoning process implemented in the prototype system was considered by the participating expert physicians accurate and in line with their own diagnostic reasoning. The possibility to enter the reasoning process in the system when the first steps already have been accomplished without the system, was
considered important when support was needed for differential diagnosis or intervention.

The primary care physicians were not familiar with the process as manifested in the prototype system, but were not reluctant in using it.

The three first steps of the model concerning diagnosis comply with the chapter used in the guideline DSM-IV. The first step of the model can be viewed as determining whether the patient may suffer from any of the diseases in the chapter with cognitive disorders or not, the second step narrows down the inferences to which subgroup in the chapter that is the most likely, i.e., which cognitive disease. In case of a state of dementia, the third step focuses on the subgroup of dementia types. In this phase other guidelines have to be considered as well. Other guidelines have to be integrated in order to accomplish the third, fourth and fifth step since there is no support for diagnosing all dementing diseases or for intervention in DSM-IV.
CHAPTER 11
Assessment of Domain Knowledge Concerning Cognitive Disease Diagnosis

In this chapter a background and status are given of the domain knowledge concerning cognitive disorders. The different types of dementia are of special interest, since these are the most common and affect a large portion of the elderly population. The purpose of this section is to give an introduction to the domain with the perspective of formalising the domain knowledge concerning diagnosis in a logic that can handle the incompleteness, ambiguity and uncertainty in the domain knowledge. An initial analysis is presented of to what extent the knowledge represented in the guidelines may become represented in logic.

Further, the analysis constitutes a basis for the formalisation of the knowledge which is implemented in the prototype system using production rules. In the formalisation process, such things as the connective words and the structure of the guidelines have been used in order to capture critical knowledge since there is no explicit logic in the guidelines. The formalisation has been made in co-operation with a domain expert. Where different interpretations have been made, the physician has determined which one will be used. The physician has also prioritised the guidelines.

There are two levels of decisions made in the diagnostic process that have to be distinguished. A guideline provides knowledge in natural language with extensive use of imprecise terms. The physician has to evaluate the symptoms manifested in the patient, determine what is abnormal and correlate the findings with linguistic expressions given in guidelines like “fluctuating cognitive ability” and “significant decline from previous level”, etc. Once these decisions about how to interpret the findings are made, the next level can be reached in which the diagnostic process is in focus [Lindgren, 2006].

It can be noted in this context that guidelines are evaluated by comparing the conclusions drawn by users using and interpreting the content of the guidelines in certain patient’s cases (inter-rater reliability). Poor inter-rater reliability means for example that different users may value differently what constitutes a “significant semantic memory deficit”, and come to different conclusions in the presence of certain evidence in a patient.

The results will be structured according to the first three steps of the reasoning model concerning diagnosis described in previous chapter. Firstly, an overview of the domain knowledge will be given.
11.1 Review of Cognitive Disorders in Evidence-Based Medicine and Clinical Guidelines

The primary types of dementia are degenerative cognitive diseases that are caused by organic processes affecting the brain. The main common consequence of dementing diseases is a diminishing ability to use cognitive functions in activities of daily life. This is usually noticeable initially in activities where memory is essential. Cognitive functions are used in every kind of activity. A lack of ability is the most noticeable in higher level activities where previous knowledge is needed, where new knowledge needs to be integrated and in activities where it is necessary to act according to knowledge by organizing, planning, making decisions, etc. Often these kind of cognitive abilities are central in a work environment. A key definition in DSM-IV of a cognitive decline correlating to a dementing disease, is that the decrease of cognitive ability has to

...generate a significant impairment in social or occupational functioning and represent a significant decline from a previous level of functioning.
[American Psychiatric Association, 1994]

Therefore the dementia diagnosis is dependent on the individual’s pattern of activity and the requirements these activities put on cognitive ability, and not as much the extent of the cognitive dysfunction per se.

The “gold” (completely reliable) method to detect degenerative diseases affecting the brain is autopsy, a method by which substances are removed from the brain to be examined. Since the method is likely to cause brain damages it is not used for the purpose of differentiating dementia diagnoses in clinical practice. Studies have been made where patients have been followed during the progress of a dementing disease with an autopsy made post-mortem. Present features in the patient and diagnoses established during treatment periods have been compared to the results of the autopsy. Autopsy is a reliable method for all dementia types except for vascular dementia for which there is no current gold standard for the pathological diagnosis.

The amount of different symptoms in populations have been examined and based on these correlations clinical guidelines have been developed by groups of experts in the domain. The evidence-based studies are often ambiguous and incomplete and are therefore valued in these consensus groups of experts in order to find a common view of the existing domain knowledge, in order to make it applicable in clinical practice. The results of evidence-based studies are presented as amounts of occurrences within a statistical framework. Clinical practice on the other hand deals with patient cases in which the features are manifested or not at a certain point of time. In the guidelines the domain knowledge often is formulated in the form of diagnostic criteria for different diseases. The
criteria can point more or less firmly toward a diagnosis. In some guidelines the uncertainty aspects inherited from evidence-based studies are formulated in words like possible, probable and unlikely, in others a set of core features has to be present and determines whether a diagnosis can be set. These two types of representations correspond to the two different streams of reasoning described in the following sections.

There are few guidelines that deal with the coexistence of diseases. Dementing diseases that affect the functions of the brain often cause psychiatric and behavioural symptoms like depression or paranoia (BPSD) in addition to cognitive symptoms. BPSD may be accounted for in some guidelines in a differential diagnosis phase or as additional features. In clinical practice these symptoms are often the ones that cause the problems in daily life for the patient as well as the people around the patient and need to be treated as soon as possible, regardless the biological cause. The cause may though have an important impact on the choice of treatment. Other cases of coexisting diseases are typically where there are vascular features in addition to features typical for other dementing diseases.

The Diagnostic and Statistical Manual of Mental Disorders, 3rd edition (DSM-III-R) and 4th edition (DSM-IV) [American Psychiatric Association, 1994], produced by the American Psychiatric Association is a guideline for dementia diagnosis among other psychiatric diagnoses. DSM uses definitions for mental disorders, including diagnostic criteria and is compatible with the chapter on mental disorders in International Classification of Diseases (ICD-10), only DSM is more detailed.

Other diagnostic criteria used for dementia diagnosis are the Hachinski Ischemic Score (HIS), NINCDS AIRENS and the California criteria for VaD [Knopman et al., 2001], and the NINCDS ADRDA for AD. The NINCDS-ADRDA was evaluated by Lowenstein et al. [Loewenstein et al., 2001], and was found not detailed enough in its criteria

\[\text{However, a more parsimonious model specifying memory, verbal abilities, visuospatial skills, executive function, and higher as well as lower functional activities of daily living fit the data better than the NINCDS-ADRDA model.}\]

The guideline is shown to have good sensitivity but modest specificity, providing not much guidance for differential diagnosis. The Quality Standards Subcommittee of the American Academy of Neurology has evaluated the current criteria for the diagnosis of dementia according to new evidence [Knopman et al., 2001]. The conclusion was made that the definition of dementia used in the DSM-III-R [American Psychiatric Association, 1987; American Psychiatric
11.1 Review of Cognitive Disorders in EBM and Clinical GLs

Association, 1994], and DSM-IV has good to very good reliability for diagnosing dementia and was recommended to be used routinely

*The essential feature of Dementia is impairment in short- and long-term memory, associated with impairment in abstract thinking, impaired judgement, other disturbances of higher cortical function, or personality change. The disturbance is severe enough to interfere significantly with work or usual social activities or relationships with others. The diagnosis of Dementia is not made if these symptoms occur... in Delirium... (DSM-III)*

Epidemiological studies have shown that by using the diagnostic criteria for dementia three groups of subjects are yielded: those with dementia, those without dementia that have normal cognitive abilities and finally subjects with cognitive impairment (usually memory) that cannot be classified as normal or suffering from dementia. Literature uses the term *mild cognitive impairment* (MCI) in these cases and the state is viewed as a transitional stage between normal ageing and very mild Alzheimer's Disease (AD). For these subjects there is an increased risk for developing AD [Petersen et al., 2001].

The most common dementia diagnosis is Alzheimer's disease (AD) which represents around 80% of the dementia diagnoses. AD is recognised by its slow onset, progressive nature and is more devastating when onset comes early in life (presenile before the age of 60), although early onset is rare. In 2% of the presenile cases Alzheimer's disease is transmitted as a strong autosomal dominant gene with strong penetrance. Most of the early onset cases is caused by mutations of the chromosome 14, of which most cases occur before the age of 50. Semantic memory dysfunction is typical for AD.

Dementia with Lewy bodies (DLB) accounts for from 15% of dementia cases at autopsy and up to 20% in some studies, which makes DLB the second most common degenerative dementia [McKeith et al., 1996; Greicius et al., 2002]. This particular type of dementia has a different course prognosis and treatment response compared to other dementia types but is often mistaken for vascular dementia (VaD) or AD. This is easily done since an additional vascular pathology may be present in as much as 31% of the cases of DLB [McKeith et al., 2000]. In an investigation made by McKeith et al. [McKeith et al., 1994], 35% of DLB received scores satisfying VaD using the Hachinski Ischemic Score (HIS). Between 15-50% met the criteria for AD when current diagnostic criteria were used (the DSM-III-R [American Psychiatric Association, 1987] identified 35% as AD). An explanation for these kinds of mistakes was the attempt to find one diagnosis that fit all symptoms

*...the incorrect clinical diagnosis was partly a consequence of attempting*
to find a unitary explanation for all symptoms, rather than opting for a ‘mixed pathology’ category. [McKeith et al., 2000]

An overlap of pathology (results from autopsy) has been reported where 15% of DLB cases have severe AD pathology, 55% have some and only 30% have no more AD pathology than age matched controls.

Diagnostic criteria for DLB are formulated as follows

*Fluctuation in cognitive function, persistent well-formed visual hallucinations, and spontaneous motor features of parkinsonism are core features with diagnostic significance in discriminating DLB from AD and other dementias.*

However, the presence of the core features in an individual is also not reliable. Visual hallucinations are absent in 46-61% of the cases and parkinsonism never occurs in up to 20% of the cases. Sensitivity to neuroleptic drugs was noted in 33% of patients with Lewy body dementia [Gnanalingham et al., 1997]. A revision of these criteria is being made and the most recent investigation made by Serby and Samuels [Serby and Samuels, 2001] claims that fluctuating cognition is not as reliable as previous considered with a prevalence of only 30%. Dopaminergic drugs were associated with the presence of visual hallucinations. The consensus criteria have poor sensitivity (22%-75%) which has aroused the suggestion that REM sleep disorder and depression could increase specificity. The specificity is 79%-100% depending on which study is referred [Greicius et al., 2002]. Recent studies have shown memory deficit in DLB similar with those of AD, part from that episodic memory is worse in AD. DLB patients have substantially greater impairment in attention, working memory and visuo-perceptual ability. A suggested reason for the low specificity is the difficulties in identifying fluctuating cognition, as shown by two inter-rater reliability studies where the agreement between experts is barely better than chance. Registration of qualitative performance characteristics (i.e., inattention, distraction, mental set establishment and shifting, confabulation, perseveration, intrusions) in initial test situations is shown to increase the differentiation of DLB from AD.

There are also difficulties in correctly diagnosing cerebro-vascular pathology in dementia. Recent research provides a perspective that distinguishes between some or any vascular lesions and pure vascular pathology, rather than simply considering vascular dementia as present or absent. Pure vascular pathology accounted for dementia in 9-10% while some vascular pathology exists in as much as 29-41% of the dementia cases [Holmes et al., 1999; Lim et al., 1999].

The Quality Standards Subcommittee of the American Academy of Neurology (AAN) has evaluated current criteria for dementia diagnosis and investigated
key issues in the management of dementia to make treatment recommendations based on evidence-based literature. The results were collected in Practice parameters [Doody et al., 2001; Knopman et al., 2001; Petersen et al., 2001]. Some of the results are presented as guidelines, other as optional and some as standard procedures.

The diagnostic criteria for Alzheimer’s disease (AD) presented in the DSM-III-R have sufficient reliability and validity and should therefore be used according to the recommendation [Knopman et al., 2001]. The criteria for vascular dementia (VaD), dementia with Lewy bodies (DLB) and frontotemporal dementia (FTD) have imperfect reliability and validity but may be of use in clinical practice. The committee suggests that the necessary criterion of memory disorder should be excluded from the definition of dementia so that VaD, DLB and FTD could be better integrated in the domain of dementia, since memory disorders are not necessarily a part of the initial phase of these disorders.

The Hachinski Ischemic Score (HIS) as modified by Rosen, was considered by the committee the most suitable diagnostic criteria for VaD, while lacking neuro-imaging criteria and since the DSM-III-R, NINCDS AIRENS and California criteria had low sensibility.

The committee describes the development towards earlier detection of cognitive impairment and the following necessity to better distinguish between MCI, VaD, AD, and other forms of dementia in an early stage. To accomplish this there has to be a modification of current definitions of dementia along with further evaluation of different diagnostic parameters. Current diagnostic criteria are good at detecting pathology per se but not at detecting pure pathology [Holmes et al., 1999]. In addition, none of the criteria performed well for mixed cases in evaluations. This may be one of the reasons why GPs are reluctant to give specific dementia diagnoses and instead use the general coding of being a state of dementia.

To summarise, the main problems of the domain knowledge are the general weak and ambiguous statistical knowledge of occurrences of features, the lack of means to capture early stages of dementia, especially the cases of unusual dementia types, the blurriness in how to handle coexisting diseases in diagnosis, and finally, how managing differential diagnosis. Some of these problems will be addressed in the following sections.
11.2 Identifying and Diagnosing Typical Cases

A main stream of reasoning is identified, which is considered sufficient to capture a majority of dementia cases in a diagnostic process. The chapter concerning cognitive diseases in DSM-IV [American Psychiatric Association, 1994] and core features defined in consensus guidelines for MCI [Petersen et al., 2001], FTD [Neary et al., 1998] and DLB [McKeith et al., 1996] are used as a synthesis of guidelines for identifying typical cases. The cases of patients where the synthesis guideline support one unambiguous suggestion of a diagnosis we define as typical cases.

The model of the diagnostic reasoning process described in previous chapter constitutes the line of reasoning by which the different tasks are accomplished.

11.2.1 Step 1: Determine the existence of cognitive disease

In order to determine the existence of a cognitive disease anamnesis is obtained from relatives and a screening for cognitive dysfunctions is made with tools like the MMSE [Folstein et al., 1975]. In the model memory deficit and judgement deficit is used as a measure of cognitive dysfunction and as criteria for cognitive disease. If present, further investigations need to be done.

Judgement deficit is not mentioned in the context of cognitive disorders in DSM-IV, but was considered by the expert as an essential indicator of that further investigation has to be done, even if a memory deficit is not a problem for the patient. Therefore judgement deficit is integrated as one sufficient criterion for a cognitive disorder.

Memory impairment is defined in DSM-IV for each type of dementia diagnosis and for amnestic disorders as

\[ \ldots \text{impaired ability to learn new information or to recall previously learned information} \]

In this work the choice was made to make the different types of memory dysfunctions explicit for research purposes, following the distinction made in [Tulving, 1972]. Therefore episodic, semantic and short term memory have to be affected in order to meet the criteria for memory impairment. For the diagnoses MCI and Delirium the type of the memory dysfunction is not explicitly stated in DSM-IV. Therefore, the terminology in the system uses finer granularity than the guideline in determining which cases will be considered as a cognitive disorder.

In clinical practice a distinction is made of sources of evidence. Status means the current medical investigator and heteroanamnesis means that the data is obtained from relatives, care personnel or other sources.
11.2 Identifying and Diagnosing Typical Cases

11.2.2 Step 2: Determine the type of cognitive disease

The second phase is a differentiation diagnostic procedure to eliminate alternative hypothetic causes of the cognitive decline, other than dementia. At this stage a wide range of diagnoses need to be covered. In guidelines there are indications for which differential diagnoses should be investigated more carefully. The closest related cognitive diseases are those that are the most accounted for in evidence-based studies as well as in clinical guidelines. These are delirium, depression, mild cognitive impairment (MCI) and amnesia.

In addition other diseases that temporarily may affect the central nervous system have to be considered, like brain tumours, CVS, hydrocephalus, Parkinson’s disease, subdural haematoma, etc. Some of these diseases may also cause a state of dementia, which will be accounted for in the next step of the reasoning.

The clinical gold standard reference criteria for dementia include the dementia criteria of DSM-III-R [American Psychiatric Association, 1987], DSM-IV [American Psychiatric Association, 1994], ICD-9, ICD-10 (International Classification of Diseases) or NINCDS ADRDA [O’Brien et al., 2000]. Criteria for dementia in the guideline DSM-IV are integrated in the description of criteria for the different types of dementia. There has to be an impairment in semantic memory together with one or more of the following symptoms; apraxia, aphasia, agnosia and/or disturbance in executive functioning. The cognitive decline has to be severe enough to cause significant impairment in social or occupational activities and put the patient at a lower level of functioning. If the decline is not affecting social or occupational function to the extent described, a state of mild cognitive impairment (MCI) can be considered as a diagnosis. MCI is not included in the guideline DSM-IV but is described in a practice parameter [Petersen et al., 2001]. Between 6-25% of cases of MCI develop dementia each year. The diagnosis of MCI is therefore important in the effort to detect early stages of dementia. Since memory deficit is a key feature in diagnosing dementia and MCI according to the guidelines, cases of less common dementia types are not detected in which memory deficit does not necessarily appear in early stages of the dementia progress.

Another criterion for dementia is that the cognitive deficits cannot be judged to appear exclusively during the course of a Delirium. In the criteria for Delirium in DSM-IV there are specific cognitive features that should be affected, which are consciousness and that cognitive ability may fluctuate during the course of day. In addition, the onset should be sudden, the cognitive decline should not be better accounted for by a dementia and there should be evidence for the symptoms to be a direct consequence from a general medical condition or intoxication. The temporal parameter in the condition of sudden onset is specified in the guideline
as “usually hours to days”. This condition can be compared to the condition of gradual onset for Alzheimer’s disease, which in other guidelines is specified so as that the cognitive decline should have been present more than six months.

The term general medical condition is used in the guideline as a shorthand term for communication purposes, whereas a more specific terminology is expected to be used in clinical practice. For convenience we use the same notion here, whereas in the system this has to be taken into consideration and be adjusted accordingly to meet the purpose. The additional condition of judgement deficit in the rule for dementia is not in compliance with the clinical definition of a state of dementia. One possibility to avoid controversial adjustment of the definition could be to implement a separate conclusion in the form

JudgementDeficit $\rightarrow$ possibleDementia,

which would indicate that further investigation needs to be done. However, we chose to use judgement deficit as a sufficient criteria for strongly suspected dementia and include such cases in further investigations.

### 11.2.3 Step 3: Determine the Type of Dementia

When step 1 and 2 are accomplished and a state of dementia is established in a patient, the reasoning process moves on to determine which type of dementia the patient is suffering from. The available guidelines have been evaluated regarding interrater reliability, sensitivity, specificity and likelihood ratios [O’Brien et al., 2000; Qizilbash et al., 2002]. Criteria for Alzheimer’s disease (AD) were found to be suitable for certain types of research where typical cases are desired, but unsatisfactory for clinical practice. To obtain best possible results a combination of guidelines was suggested.

The guideline DSM-IV is preferred in clinical practice because it contains a level of granularity that is not too high for clinical use [O’Brien et al., 2000]. But since it lacks some of the dementia types, other guidelines have to be considered as supplement.

In the guideline DSM-IV the two most common types of dementia are described with criteria which are Alzheimer’s disease and vascular dementia (VaD). Dementias due to other general medical conditions are listed without criteria but with a reference to the causing disease (e.g., Parkinson’s disease, Huntington’s disease, Pick’s disease, HIV, head trauma, hypothyroidism, etc). Dementia of Lewy body type (DLB) and frontotemporal degeneration (FTD) are not described in DSM-IV chapter in focus.

In addition to the criteria for a state of dementia described earlier, the course is characterised for Alzheimer’s disease by a gradual onset and a continuing
11.2 Identifying and Diagnosing Typical Cases

decline of cognitive functions. The implicit temporal parameter in these conditions is specified in other guidelines in the way that the cognitive disorders should have been present more than six months. In order to establish a diagnosis of Alzheimer’s disease all other medical conditions that may cause the dementia have to be excluded, according to DSM-IV. This condition is removed in this work in order to provide the opportunity to derive more than one diagnosis, since the coexistence is possible and this is a critical situation where atypical cases of patients can be identified. In this case the inferences of the main stream can provide no more diagnostic guidance to the user, since additional grounds for analysis will have to be included, which will be discussed in the next section.

For vascular dementia there have to be focal neurological signs and symptoms or laboratory evidence indicating a cerebrovascular disease related to the disturbance, in addition to the dementia criteria, according to DSM-IV. The criteria are more liberal than in earlier guideline DSM-III in which both criterion above have to be met, in addition to a stepwise deteriorating course and patchy distribution of deficits.

The co-occurrence of criteria for both AD and VaD in the same patient is common. In these cases it may not be possible to determine which disease causes the cognitive decline and it is likely that both affect cognitive ability. DSM-IV recommends that both diagnoses may be assessed in these cases. Other formulations used are variations of “Alzheimer’s disease with high amount of vascular elements”.

The frontotemporal degenerative dementia (FTD) has a different course compared to AD and VaD. Memory deficit or typical cognitive dysfunctions as apraxia or aphasia may appear later in the course of the disease. Instead a change in personality is initially seen with social dysfunction, judgement deficit, behavioural disinhibition and emotional blunting among others. In the consensus criteria for diagnosis of FTD developed by the Lund and Manchester Groups [Neary et al., 1998] features are listed and valued as core, supportive, exclusionary and relatively exclusionary. Each of these features is described in the guideline. For instance, “emotional blunting” refers to

…an inappropriate emotional shallowness with unconcern and loss of emotional warmth, empathy, and sympathy, and an indifferent to others.

The condition “other neurological symptoms” refers to signs such as myoclonus, cerebellar ataxia, corticospinal weakness and choreoathetosis, which all indicate damages on other parts of the brain than the frontotemporal region.

It can be noted that the description for onset and progress is the same as in AD which makes the cognitive, neurological and behavioural features the significant
elements in a differential diagnosis process. Therefore some of the exclusion features are designed to exclude AD.

Consensus criteria for diagnosis have been developed for the dementia of Lewy body type [McKeith et al., 1996]. There has to be evidence for a profound cognitive decline that significantly decreases social and occupational function. The course differs from AD and VaD in that a predominant memory deficit may not necessarily occur in the early stages. Three core features were defined where two or more are essential for a diagnosis of probable DLB and one is essential for possible DLB. The core features are 1) “fluctuating cognition with pronounced variation in alertness and attention”; 2) “recurrent well-formed and detailed visual hallucinations”; and 3) “spontaneous motor features of parkinsonism”, also denoted as “extrapyramidal” symptoms. Fluctuating cognition is also seen in VaD and delirium.

The memory requirement in the definition of dementia will cause early stages of DLB and FTD to be missed. The heuristic rule added to the previous two steps regarding judgement deficit makes it possible for the system to capture early cases of FTD.

11.3 Identifying Atypical Cases

In previous section typical cases of patients were identified as those for which it is possible to suggest one single diagnosis, based on the knowledge in the guideline DSM-IV and core features in the consensus guidelines for DLB and FTD. Other cases will be defined as atypical, which are those where more than one diagnosis fit the patient’s symptoms or none fit although there is an obvious cognitive disease. A clinical decision-support system in the domain needs to be able to provide additional support that can handle the uncertain descriptions in guidelines.

In the consensus criteria for diagnosis of FTD defined by the Manchester and Lund Groups [Neary et al., 1998] features are listed and valued as core, supportive, exclusionary and relatively exclusionary. Core features are cardinal to the clinical syndrome. All core diagnostic features must be present and all exclusive features must be absent to fulfil the criteria for diagnosis. Supportive features are not necessary for diagnosis but “their presence adds substantial weight to the clinical diagnosis” [Neary et al., 1998]. Supportive features may be present only in a period of the disease or not at all in a person, but are characteristic and have high diagnostic specificity when present. The relatively exclusionary features do not firmly exclude a diagnosis of FTD but should generate caution against the diagnosis. A history of alcoholism indicates that the abuse is the alternative cause of the frontal lobe syndrome, but the alcohol abuse may also be a secondary
manifestation of the behavioural disturbance caused by FTD.

Regarding investigations specified in the guideline forming some of the exclusive features, it is declared that they should be absent “when the relevant information is available”. The investigations are brain imaging (CT and MRI) and laboratory tests of metabolic and inflammatory disorders (MS, AIDS, syphilis etc). These are therefore here regarded as relatively exclusionary.

The guideline NINCDS ADRDA specifies criteria for a possible and probable diagnosis of AD [O’Brien et al., 2000]. In addition there are features listed that supports a probable AD and features making the diagnosis more unlikely.

VaD is treated in a similar way in the guideline NINCDS AIRENS criteria. This criteria for a probable VaD is stricter than in DSM-IV, which criteria corresponds to a possible VaD in the NINCDS AIRENS criteria. The features that are making VaD unlikely or uncertain are early onset and progressive course, which are criteria for differentiating between VaD and AD.

Another tool used for diagnosing VaD is the Hachinski Ischemic Score (HIS) which defines 13 features attached with weights of 1 or 2. The present features are summarised, where a sum over 6 indicates VaD.

Apart from sensitivity to neuroleptica, the core features for DLB are supplemented with features which are secondary manifestations of the core features, for instance gait disturbances and frequent falls. Vascular signs and history of epilepsy are described as contradictory features for DLB.

In a prospective validation study of the consensus criteria for DLB, 50 cases were compared after autopsy, using the DLB consensus criteria, NINCDS ADRDA and NINCDS AIRENS [McKeith et al., 2000]. 35 probable diagnoses of DLB, AD and VaD where accurate and three were inaccurate. Of the 30 cases which met criteria for possible AD, 21 cases were DLB and two were VaD at autopsy. Of the eight possible cases of DLB six were AD and one supranuclear palsy. Based on these numbers the definitions of possible diagnoses in the criteria seem highly unreliable. However, since not all cases meet criteria for a probable diagnosis they have to be used. In the study 31% of DLB cases had additional vascular pathology (insufficient to alone cause dementia) which caused a misdiagnosis in two cases.

In the guidelines contradictory or exclusive features are typically defined for the purpose of differential diagnosis. Still, the guidelines are not performing well in mixed cases. The features in the different guidelines are not equally important and cannot be formally compared in an easy way.
11.4 Summary and Conclusions

The domain knowledge concerning dementia diagnosis was investigated. Focus has been on the most common dementia diagnoses: AD, VaD, DLB and FTD. There are diagnostic criteria for these which allow formalisation using simple if-then rules, provided that the interpretation of the collected evidence into the linguistic terms for features used in the guidelines, is done by the physician. In what way screening-tools may be used in the assessment of evidence, contributing to formal inferences in a system, remains to be analysed.

The level of complexity in the evidence and in the manifestation of the disease, is mainly determined by how widespread the deficit is within the group of a certain type of symptoms, how severe the symptoms are and time aspects on the manifestations in an individual. Thus, the temporal as well as qualitative and quantitative aspects in a certain feature contribute to the distinction between diagnoses. Regarding the core features, these aspects are in this thesis and in the prototype system mainly captured by expressions representing the values present, absent or unknown.

The inter-rater reliability of some of the clinical guidelines has been investigated with satisfying results so far only for the dementia criteria. This means that two physician do not necessarily interpret findings in a patient the same way, or map the findings the same way to criteria in a guideline. Some of the values may be obtained through investigation instruments, which are validated and reliable. Examples of instruments are neuropsychological screening tests, or the test for detection of motor and process dysfunctions shown in activity (AMPS), which is a diagnostic instrument used by occupational therapists. One example of a feature which is poorly defined is that of fluctuating cognitive ability. Several suggestions have been proposed to make the distinction more reliable; one is by using cognitive screening tests to detect differences in ability from one time point to another. Thus, the reliability of data supplied to a decision-support system can be increased by requiring that certain screening tools are used in the investigation. Since evidence is obtained from different sources in a patient’s case, the source implicitly indicates a level of reliability of the data. Data obtained from an expert (neuropsychologist, occupational therapist, geriatrician, etc) is typically of finer granularity, and of higher reliability, than datum obtained from the patient for instance, which has to be taken into account when conflicting evidence is obtained.

Another aspect of granularity concerns differential diagnosis, where the set of diagnoses accounted for in the reasoning process are represented with different specificity concerning the diagnostic evidence. The guideline DSM-IV for instance, uses the complementary diagnosis code “dementia caused by other
general medical condition” for the less known types. Higher granularity was considered essential to provide in supporting clinical practice. Therefore a “fish-eye view” is taken in this work where the domain of cognitive disorders is in focus as presented in the corresponding chapter in the clinical guideline DSM-IV. A finer granularity than provided by DSM-IV is required for the diagnosis of different types of dementia. The distinction of other causes of a cognitive disorder than those specified in the chapter of DSM-IV, however, is represented at the lowest level without specific diagnostic criteria. It was considered sufficient to direct the user’s attention to alternative explanations of the cognitive disorder without providing diagnostic criteria for each since this was judged being beyond the scope of the domain of cognitive disorders and of this work.

Two main levels of complexity in patients’ cases have been identified in relation to the clinical guidelines, where typical cases are defined as being in accordance with the guideline DSM-IV and sets of core features in consensus guidelines for FTD, DLB and MCI. Inference rules in propositional logic have been defined for diagnosis according to these guidelines and implemented in the decision-support system described in the following part. The rules are diagnostic (or evidential), which means that the reasoning moves from evidence to causes. They can be interpreted as follows: the presence of a certain set of features increases the level of belief in the corresponding syndrome or diagnosis with the highest possible measure, provided there is only one suggestion of diagnosis in this context.

Atypical cases can be categorised into cases where more than one diagnosis fit according to these guidelines, or none fit completely although there is evidence indicating a cognitive disorder. In order to distinguish the atypical cases, the requirement of excluded other diagnoses was removed from the core criteria for Alzheimer’s disease. A formalism that allows comparison of possible diagnoses in a decision process is needed in the atypical cases.

Sets of supporting and contradictory features have been identified according to the mentioned guidelines as well as other guidelines containing terms which express more uncertainty.
CHAPTER 12
Validation of Terminology

The earlier versions of the decision-support system integrate a terminology that has evolved locally from both clinical practice in Northern Sweden and from the most commonly used clinical guidelines in the area [Lindgren, 2005]. In the integration of additional clinical guidelines and the integration of support in atypical patient’s cases, further analysis of the terminology in the system as well as in the clinical domain is required in order to facilitate the merging of several clinical guidelines in the domain into the decision-support system, facilitating the development of the knowledge in the system and make the system available to other clinical contexts. A long-term purpose is to make the outcome of the system useful for research purposes, which puts additional demands on the validity and quality of the evidence used in the system. In the development of the system the terminology of the system has also shown to be crucial for the system to be seamlessly integrated at different levels of care [Lindgren, 2005].

Ontologies (i.e. terminology models) are used to structure reality as it is known by, for instance, a medical domain into general classes of entities (i.e. substances, qualities and processes) and their relations according to a theory of the domain. The theoretical definition of a particular entity (universal) is separated from its instances in the real world. Ontologies developed for serving as formal reference ontologies, generally do not contain the level of detail required for classifying disease instances and reasoning about them clinically. In contrast, systems for decision support contain vocabularies designed for special purposes and mechanisms to reason about instances, but allow in general not reasoning about general classifications of diseases [Burgun et al., 2005]. The use of ontologies in decision-support systems is increasing, [Shahar et al., 2003] is one example. There are extensive efforts in creating common medical ontologies for clinical practice, where the Unified Medical Language System (UMLS) is an example of an effort to unify terminologies and classifications in the domain [Humphreys et al., 1998; U.S. National Libraries of Medicine NLM, 2007].

Since there is no agreement on a common medical terminology in the domain of cognitive diseases, the vocabulary used in the prototype system has evolved from the local clinical praxis in northern Sweden and from the guidelines that are used. However, the international efforts to gain a consensus of a terminology for health care have been fruitful in that now the WHO:s International classification of functioning, disability and health (ICF) [WHO, 2001] is recommended by the Swedish National Health Ministry to be used in Swedish health care, as a
supplement to the International classification of diseases (ICD-9-CM) [WHO, 1992]. Work has been done to identify sub-sets (core sets) of concepts relevant for different medical domains, in order to facilitate the use of the ICF in clinical practice [Cieza and others., 2004a; Cieza and others., 2004b]. However, so far we have seen no such effort for the domain of cognitive diseases.

The clinical guidelines give account for which evidence is essential in the investigation of cognitive disorders in a particular patient’s case, which are typically signs and symptoms and characteristics of findings. The evidence is integrated in the decision-support system and used in the reasoning process towards diagnosis, a process supported by the system. In this Chapter we will focus on the representation of the evidence, which constitutes the outcome of the sub-actions in the investigation process. Aspects concerning ontology and decision-support will be discussed in the context of existing ontologies in the domain, aspects such as the distinction between concepts and instances, different yet related semantic types and characteristics of symptoms. The terminology used in the system is analyzed and compared to existing terminologies and classifications in order to validate and develop the terminology. A secondary purpose is to evaluate the utility of existing terminologies and classifications for the purpose of decision support in the domain of cognitive disorders. In Paper I a patient-oriented view is taken in order to identify the relevant concepts and form an example of a simple patient ontology based on ICF, a terminology of functioning, ability and health. In this chapter motivations will be given of choices of terminologies and classifications for future work. Firstly, a summary will be given of the ontologies that were used in this work.

### 12.1 Relevant Classifications and Terminologies

There are classifications of diseases and syndromes which present its concepts at a coarse level (i.e. ICF, IDC-9-CM, DSM-IV), compared to terminologies which build terms from low-level building blocks (i.e. Read Codes, SNOMED CT®, FMA). In the decision-support system we use concepts of a wide range of granularity, in order to both capture distinct signs, which can be mapped to a concept in a terminology, and use the findings for clinical reasoning, i.e. navigating among classifications of diseases according to clinical guidelines. Therefore, not only one of the existing terminologies or classifications is sufficient for our purposes. We need for instance, both disease and disability classifications, as well as terminologies covering biological and organic properties, for capturing evidence in the different levels of the reasoning activity.

We investigated the relevant clinical guidelines, which are developed for the domain of cognitive disorders, or the wider domain of mental disorders, in order to find ontologies for the domain. Of the clinical guidelines the Diagnostic and
Statistical Manual for Mental Disorders 4th edition (DSM-IV) [American Psychiatric Association, 1994] is the most widely spread classification of psychiatric disorders and the most analysed concerning vocabulary. DSM-IV contains references to the guideline ICD-9-CM of which codes are used for diagnosis. WHO:’s International classification of functioning, disability and health (ICF) was added as a classification not focussing on diagnoses but on disability [WHO, 2001]. ICF is also recommended as mentioned earlier by the Swedish National Health Ministry to be used as a supplement to the ICD-9-CM [WHO, 1992].

Part from ICF we used the terminology SNOMED CT® [SNOMED® International, 2007] for capturing distinct cognitive features. SNOMED CT® contains terms from the domain of cognitive disorders of high granularity and specificity. SNOMED CT® is the result of merging the terminology Read Codes/Clinical Terms Version 3 (Read CTV3), which seeks to construct concepts from primitive building blocks governed by validation rules, with the SNOMED (Systematised Nomenclature of Medicine).

Finally, we supplemented the search for relevant terminologies and classifications by using the Unified Medical Language System (UMLS) [U.S. National Libraries of Medicine NLM, 2007]. UMLS is developed for the purpose of unifying concepts and relations drawn from different medical classifications and terminologies into a unified medical ontology. The Metathesaurus is the concept-structure of UMLS containing mappings to terminologies and classifications such as the ICD-9-CM, DSM-IV, SNOMED CT® among others.

12.1.1 The Disease Classifications DSM-IV and ICD-9-CM

The clinical guideline DSM-IV [American Psychiatric Association, 1994] is developed for the domain of mental disorders with a 5-axis structure to support a holistic evaluation of a patient’s functioning. The axes represent the mental dysfunctions in focus (axis I), developmental conditions and personality disorders (axis II), other medical conditions than psychiatric, which may affect the current condition (conditions listed in ICD-9-CM, axis III), social functioning and impact of symptoms (axis IV), and a global assessment of functioning (axis V). For the purpose of investigating cognitive disorders the diagnostic criteria in the chapter Delirium, Demential, and Amnestic and Other Cognitive Disorders was used for the formalisation of the domain knowledge in the support system supplemented with consensus criteria for the diagnoses Lewy Body Dementia (DLB) and frontotemporal dementia (FTD) since diagnostic criteria for these are not included in DSM-IV.

For our purposes, the axis focussing the current condition (the chapter mentioned) and the axis taking other conditions into account (using concepts from ICD-9-CM) are useful. The other axes represent knowledge about the patient
that is in the region of Northern Sweden assessed by other means than DSM-IV, which use optimal granularity for the particular domain of investigating dementing diseases.

In DMSS the categorisation of syndromes and diseases follows the structure in DSM-IV and ICD-9-CM with the notion of other medical condition as a cause of dysfunction. However, since the level of granularity in DSM-IV was considered too coarse in guiding dementia differential diagnosis, other clinical guidelines were integrated in the phase of differentiating between types of dementia [Lindgren, 2005]. In addition, some of the concepts in DSM-IV, such as dementia not otherwise specified, are also of epistemic nature in contrast to ontology theory. Such concepts exist for the purpose of viewing the state of the knowledge concerning instances and not the knowledge about reality at the abstract level [Bittner and Smith, 2003]. This particular concept corresponds to the term state of dementia used in the system. However, the purpose of the term in the system is not to be used as a final diagnostic term as done in practice (which is also seen as bad practice), since the purpose of DMSS is to assess knowledge of the type of dementia present in the patient.

The coarse level in DSM-IV is also shown in that there are no relations among concepts in DSM-IV except at the level of diagnosis. Further comparison of the categorisation of symptoms has to be made from descriptions in natural language. In [Dunne and Chute, 1999] a study is presented which has parsed the content of DSM-IV in order to extract the clinical significant signs, symptoms, findings and conditions that are present. The result is a “latent terminology” implicitly represented in the guideline consisting of terms of high granularity and specificity. It was found that the terms supplemented the concepts already available in UMLS Metathesaurus. Whether this means that the terms were sufficiently represented in the UMLS or that they could enhance the UMLS is not clear, therefore we continued investigating these sources for our purposes.

In the perspective of diagnostic inference, it is interesting to note that more than half of the concepts were duplicated for different conditions and only a minority of concepts occur only once when viewing the whole DSM-IV [Dunne and Chute, 1999]. This indicates that the set of psychiatric signs and symptoms is limited and that there is a particular combination that defines a disease or syndrome. The diagnostic criteria in the DSM-chapter in focus also share several concepts, for instance, implicitly defined as the syndrome dementia.

DSM-IV provides diagnostic criteria in the form of complex predicates, in order to enabling disease instances to be linked to diseases or syndromes. Such predicates are usually not represented in ontologies [Burgun et al., 2005]. DSM-IV also makes use of references to normalities in the form of examples instead
Chapter 12 Validation of Terminology

of formal definitions. In the case of a milder onset period of a disease such as a schizophrenic prodrome, which precedes the main disease, this condition is classified as a subclass to schizophrenia in the terminology SNOMED CT®. However, in an ontology it is expected to be represented as a temporal part of schizophrenia [Burgun et al., 2005]. A similar example is the state of a mild cognitive impairment (MCI) which in many cases can be viewed as an early stage of a dementing disease, however, distinguished conceptually and in disease classifications from the syndrome dementia.

12.2 Concepts vs. Instances

One key issue in connecting disease instances to concepts (or universals), is that there has to be a clear distinction of what the instance is and what the universal truth about the phenomenon is and their relations. The user adds into the decision-support system manifestations found in a particular patient that are used as clinical evidence in the reasoning process. These manifestations are instances of concepts but are also related to the causing disease, which in turn is an instance of a disease concept. Thus, there are relations on the concept level and between instances, but also between instances and concepts (Figure 12.2). In the system, all information about the patient that is entered into the system is assumed to be instances in the reasoning in a particular patient’s case. However, the basis of the reasoning is the medical knowledge, represented as universal truths about the world as it is known to this point. The instances and their relations are therefore matched to the knowledge structures with the purpose of connecting manifestations to disease instances. However, in the domain of cognitive disorders the ground truth or gold criteria is typically evidence obtained through autopsy of the brain, which leaves full knowledge about the existence of a certain disease instance to post mortem. Consequently, the instances which can be evident in a patient can only be matched to abstract concepts at the level of determining the causing dementia disease. Syndromes on the other hand, which are defined by sets of manifestations as observable entities in a patient, can be determined at the instance level in clinical practice. The usage in DSM-IV of the concept Dementia of Alzheimer type, distinguished from the concept Alzheimer’s disease (ICD-9-CM), corresponds well to this view of manifestations directly related to a syndrome and indirectly to the causing disease (Figure 12.2). However, in UMLS the two concepts are synonyms.

The weak link between assessable evidence and the gold truth about the distinct dementia types in evidence-based medicine is partly solved in the clinical guidelines by the usage of terms which mirror the statistical evidence-based domain knowledge, such as possible or probable, in the assessment of diagnosis.
12.3 Different but Related Semantic Types

Similarly, and based on the guidelines, DMSS provides valuations to what degree the evidence at the concept level supports the manifestation of a certain disease in the particular patient. Although the conceptual valuations are provided by the system, it is the user’s task to commit to an interpretation of the evidence and establish a diagnosis, i.e., judge what disease instance is present in the patient.

**Table 5.** Definitions of some central semantic types in UMLS.

<table>
<thead>
<tr>
<th>Semantic type</th>
<th>Semantic sub-type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finding</td>
<td>Sign or Symptom</td>
<td>An observable manifestation of a disease or condition based on clinical judgement, or a manifestation of a disease or condition which is experienced by the patient and reported as a subjective observation.</td>
</tr>
<tr>
<td></td>
<td>Laboratory or test result</td>
<td>The outcome of a specific test to measure an attribute or to determine the presence, absence, or degree of a condition.</td>
</tr>
<tr>
<td>Health Care Activity</td>
<td>Laboratory Procedure</td>
<td>A procedure, method, or technique used to determine the composition, quantity, or concentration of a specimen, and which is carried out in a clinical laboratory. Included here are procedures which measure the times and rates of reactions.</td>
</tr>
<tr>
<td></td>
<td>Diagnostic Procedure</td>
<td>A procedure, method, or technique used to determine the nature or identity of a disease or disorder. This excludes procedures which are primarily carried out on specimens in a laboratory.</td>
</tr>
<tr>
<td></td>
<td>Therapeutic/Preventive Procedure</td>
<td>A procedure, method, or technique designed to prevent a disease or a disorder, or to improve physical function, or used in the process of treating a disease or injury.</td>
</tr>
</tbody>
</table>

12.3 Different but Related Semantic Types

There is a difference between the outcome of a laboratory or diagnostic procedure, which is a factual value (laboratory or test result), and the procedure itself (for definitions used in UMLS, see Table 5). Additionally, there is a distinction between the test result and how it can be valued as a sign or symptom, which, in turn, can be valued as indicating a disease or syndrome. These differences are due to the different levels of activity in the process. DMSS typically provides support at a higher and more complex level of activity. The type of finding that is the known result of a laboratory examination is to be valued normal or abnormal when entering the information into the system. By not using the actual results, the user needs to value the results in the context of investigating dementia. In DMSS the laboratory investigations are used to receive indications of alternative causes for the cognitive decline, i.e. other diseases or syndromes. Therefore, the user’s valuation when entering the information is to determine 1) the disease...
or syndrome the test result points at, and 2) the influence this condition has on the cognitive decline. Similarly, the actual results from neuropsychological tests (with the semantic type diagnostic procedure) are not integrated into the reasoning process, only the interpretations of these, as a part of a global valuation related to different cognitive dysfunctions. The interpretations of radiology examinations are entered by choosing among a few typicality descriptions of findings related to cognitive disorders. Thus, in DMSS several examinations, which are classified as diagnostic procedure in UMLS (e.g. EEG, ECG, MRI, SPECT, etc), are categorised as laboratory examinations with the information at a higher level of interpretation than a laboratory or test result in the terminology of UMLS. It can be useful to also provide support at the lower level of activity, such as for the interpretations of these values in the context of cognitive diseases (further discussed in Paper I). When DMSS is integrated with the EPR, the results of laboratory measurements can easily be integrated and used in the process.

12.4 Sign or Symptom

The definition of a sign or symptom in UMLS includes entities sprung from objective observations (examinations made by healthcare professional) and subjective complaints from the patient. However, there is no distinction between these sources in the structure of entities (Table 5). In DMSS these sources of evidence are separated in order to value these as different entities in the reasoning process. In the use of ICF the sources should be accounted for, however, only in a general way and not for each piece of evidence.

The majority of the symptoms and signs in DMSS can be directly mapped to the corresponding terms found in UMLS (i.e. SNOMED CT, Read Codes). However, for some distinctive clinical terms for certain cognitive deficits there are more than one concept in UMLS which are close, but none are exactly the same as in DMSS. This may be a consequence from different tradition of categorisation of cognitive deficits in research, but also different traditions in how these findings are denoted in clinical practice and in patient records. In the local context of the support system the categorisation of memory types follows [Tulving, 1970; Tulving, 1972]. There are distinct concepts for these memory types as subcategories in the branch of semantic types for normal mental functions. However, no corresponding terms for pathologic memory function can be found in UMLS. Instead, the concept Memory Impairment is used partly with the definition of a dysfunction of episodic memory: diminished or inability to recall past events. Dysfunction of this particular memory type (episodic memory) is not described other than with descriptions such as Amnesia for remote events as a subcategory. The additional definition for Memory Impairment: inability to remember or recall
bits of information includes deficits of the memory types semantic memory and short term memory (subcategory Amnesia for recent events). In DMSS these memory types are distinguished although the criteria in DSM-IV specify memory impairment according to the definition in UMLS: impaired ability to learn new information or to recall previously learned information [American Psychiatric Association, 1994]. The distinction is made in DMSS for educational and research purposes.

ICF, in contrast to the terminologies integrated in UMLS, uses a neutral language without distinction between normal and pathological functions. For example, the term memory function is used as the universal concept, and the term is assigned a value in a particular patient's case according to a scale which contains values from no dysfunction to complete dysfunction and the neutral qualifiers not specified and not applicable. In DMSS the default value unknown is used for evidence, until the presence or absence of dysfunction is determined. Consequently, the ICF scales correspond well to the scales used in the system.

12.5 Characteristics of Symptoms

In the reasoning process the user needs to value the amount and severity of entities. There is some support for such valuations in UMLS, provided in the form of the scale able-difficulties-unable related to cognitive abilities. Since we use two grades of difficulties in order to distinguish between mild cognitive impairment and a state of dementia, this particular scale is not useful for our purposes. However, there are some cognitive screening (assessment) scales integrated into UMLS such as the Global deterioration scale finding for assessment of primary degenerative dementia (GDS) with codes for the different levels of cognitive decline, and the Mini-mental state examination (MMSE) [Reisberg et al., 1982; Folstein et al., 1975], however, without content. Both are of the semantic type Intellectual Product and the MMSE also Diagnostic Procedure in UMLS.

ICF is structured with neutral concepts, which are to be combined with a qualifier which assesses the level of dysfunction for each phenomenon. The scales of qualifiers use linguistic terms and range typically from no impairment, mild, moderate, severe, to complete impairment, with the qualifiers not specified and not applicable for completion. The qualifiers mild and moderate may correspond to the levels 3 and 4 in GDS and be useful for differentiating between mild cognitive impairment and dementia in the reasoning process.

Other characteristics of the symptoms such as descriptions of the time-related factors onset, progress and fluctuating cognitive disability are defined as qualifiers in SNOMED CT®.
Chapter 12 Validation of Terminology

12.6 Implications for the Terminology of DMSS

The majority of terms used in DMSS for common signs and symptoms can be directly mapped to concepts provided by SNOMED CT®. Similarly, a majority of the terms used for diagnoses are found in the classification of diseases DSM-IV and ICD-9-CM. In the cases when terms used in the system do not correspond directly to a concept integrated in any of the terminologies in UMLS, the purpose of the particular term typically is to guide the reasoning process, indicating levels or characteristics of a finding. ICF supports the valuation of findings, and distinguishes between sources of evidence, both constituting important functionalities of a decision-support system for the domain. ICF also use neutral terms in that not only dysfunctions can be identified, but also resources and positive effects of environment on ability. ICF allows identification of dysfunctions without necessarily connect the dysfunctions with a diagnosis. This corresponds well to the second dimension of the reasoning process described in Paper I. ICF is structured in a more systemic way, in contrast to earlier more hierarchal terminologies for dysfunctions such as ICIDH [WHO, 2001], with interactions between components such as body functions, body structures, activities and participation and environmental factors (Figure 12.1). Further, work has been done in identifying rules for how health-status measurements can be linked by health professionals in a reliable way from a various clinical screening tools to ICF as a common classification [Cieza and others., 2002]. This is in accordance with the formalisation work presented in Paper I. Based on these aspects we expect ICF

Figure 12.1 A systemic view of categories acknowledged by ICF. The classification does not provide definitions of the relations between the categories.
12.7 Conclusions

The terminology of the prototype system DMSS was validated in the perspective of existing classifications and terminologies in the domain. The information used in the investigation process and in the system was found highly diverse, in that information can be traced to a wide range of classification and terminologies, of which none is sufficient on its own for capturing evidence in a patient’s case. In addition, the existing terminologies and classifications were found insufficient, when the reasoning process is to be captured in a uniform way, together with domain knowledge and evidence concerning the patient. Ontologies are one suggested solution to handle diverse data [Pisanelli et al., 2000] and formalise guidelines [Kumar et al., 2006]. However, ontologies have limitations in that they are often multiple for a particular domain, typically seen in the medical domain. They also require a choice of logic for the implementation, of which several are used. In Paper I an approach is presented where both process and domain knowledge are integrated in the same theoretical framework, for the purpose of structuring the content of the system for formalisation and design. Based on the analyses described in this chapter and in Paper I, and based on the user evaluations presented in Chapters 6 and 7, a synthesis terminology model can evolve which is drawn from appropriate clinical knowledge sources and different clinical practices. This way the aim of avoiding misunderstandings and supporting a common interpretation of each term used in the system, independently of use context and independently of logics used in the formalisation can be met.

to be useful for structuring the terminology of DMSS concerning the evidence in a patient’s case. However, if ICF is to be used as the basic structure for function and dysfunction, it needs to be supplemented with other terminologies and classifications that account for pathology, i.e., a tool with higher granularity for the assessment of causes for dysfunctions. To what extent ICF needs to be extended and supplemented needs further investigation. Figure 12.2 provides a simple example of a patient’s case in the perspective of dementia diagnosis and in the structure of UMLS. In Figure 12.3 the example is extended with ICF as the interface between the patient and the domain knowledge structures, following the knowledge model developed in Paper I.
Chapter 12 Validation of Terminology

Figure 12.2 Patient example of assessing the semantics of a diagnosis in the UMLS framework.
12.7 Conclusions

Figure 12.3 Extended patient example of assessing the semantics of diagnosis in a combined framework integrating both UMLS and ICF frameworks. Each ICF code comes equipped with one or two values indicating level of dysfunction.
Chapter 13
Assessment of Clinical Activity - Summary of Paper I

Clinical investigation activities are complex processes, which are situated, emergent and directed by the individual need of the patient, but also restricted or enhanced by the available resources at different points in the process. For the purpose of creating a system which provides support throughout the investigation process, i.e. functioning as a cognitive tool for the user, the clinical investigation process was assessed and formalised. The main purpose for the work presented in Paper I was to use an activity analysis as a means for identifying actions which are suitable for formalisation in a decision-support system and which actions to implement in the form of interactive reasoning. This was done using cultural-historical activity theory as framework for activity analysis. The analysis is based on analyses of the domain knowledge (Chapter 11) and qualitative case studies of investigations of actual patients. A secondary aim was to investigate to what extent the theory is useful for our purposes. The third aim for this work was to provide an introduction to formalisation from the perspective of clinical actions.

As a result a conceptual model for clinical investigation activity in the domain was created. The model is used to frame the investigation process as an alternative method to the formal ontology and workflow approaches, in order to identify relevant actions to formalise for decision-support in the system. An overview is presented in Figure 13.1. Further, General logics is used as formal categorical framework in exemplifications. The model is used as a bridge between the informal view on clinical actions presented in Part II of the thesis to the formal implementation of reasoning in Part III. By using activity theory the dynamic and situated character of purposeful activity is captured, and General logics provides means to integrate complex knowledge and reason over different logics. The combination of these frameworks can be used in the form of the semi-structured clinical activity model, as a framework for knowledge acquisition in decision-support development. The adapted activity theory is found well suitable for the purpose in the semi-formal version presented in this work, providing a base for the formalisation of the knowledge exemplified in Paper V.

A decision-support system aims at supporting a reasoning process performed over time as the basic structure of the clinical activity in focus, with tasks executed depending on decisions made and available information in the current context. The conceptual framework that is presented takes the complexity of reality into account, such as the different levels of granularity of actions, the
Activity System Model
Patient Model
Assessment and change of patient’s life situation
Activity Model

Figure 13.1 Interpretation of investigation activity based on Engeström’s model of activity system and ICF.

situated character of clinical actions, the role of contexts in the execution of actions and the possibility to assign tasks to teams and not only to single actors. The model provides the distinction between three levels of activity; 1) the main process constituting the main activity, 2) sub-processes supporting the main process, and 3) actions which create data and manage the logistics of the investigation activity at the least complex level. The model consists of an activity system model, a patient’s model and an activity model integrating the former two in the execution. As a result, a design of the knowledge base of a future version of DMSS is formed, consisting of three vital components; the Patient MAP functioning as an ontology providing the patient data, the basic clinical guideline-based knowledge base (GL-KB), and the meta-knowledge base (Meta-KB) directing the decision-support and interactive reasoning, using the former two structures.

The model captures the data and their sorts and qualities, the tools that are used in assessing the data, the actions of transforming the data and contextual factors such as who is responsible for a certain action. A key issue is how the semi-formal structure can be transformed into a formal structure for representa-
tion. A critical feature is the formalisation of the content of scales and values provided through the use of different tools in the process. In order to create the framework for the formalisation, we use category theory as a meta language for the identification of the fundamental building blocks, independent of the logic language used in the implementation of the knowledge. An introduction to the formalisation work is presented in Paper I. However, the focus in Paper I has been on providing a method for the transformation of the informal knowledge into the semi-formal model, and further, into the formal framework of category theory and institutions. Therefore, a deeper formal analysis of to what extent the model is useful for the integration of the knowledge into the framework remains to be investigated. An initial formal analysis is provided in Paper V, where results from the activity analysis are applied in a case study for the purpose of demonstrating the case when a splitting of entailment systems can be done. To this work, the model provides a specification at a coarse level, with the terminology of category theory introduced, which facilitates the transformation. To what extent the structure is useful for assessing the components of different logics is subject for future work.
PART III

FORMALISATION OF KNOWLEDGE

In this part an introduction is provided to the work on formalising the domain knowledge. In Paper I the work analysis generates implications for the formalisation and provides an introduction to logic from a general perspective of clinical actions. Paper II-V represents case studies in which different logical frameworks are applied and evaluated. The following papers are summarised:


CHAPTER 14
Towards Dementia Logic

A top-down approach has been used in the work on formalising the domain knowledge, which is grounded in clinical practice, through thorough work and domain knowledge analyses without early assumptions on which formalisation techniques that are going to be used (Part II). By establishing the requirements on such qualities as expressivity of the formalisation, based on the qualities of the knowledge that is aimed to be integrated and communicated to a user, appropriate formalisation techniques can be developed and used.

The domain knowledge has been analysed in an iterative refinement process where firstly a basic set of propositional rules was identified that captures the typical cases, i.e., school-book cases. It was estimated that approximately 80% of all dementia cases could be captured by this knowledge base [Lindgren, 2005]. However, recent case studies indicates that a larger part than estimated actually are atypical, complicated cases which require a more elaborative knowledge base that can handle and provide support in these patient’s cases. However, the domain knowledge expresses a high degree of uncertainty and is partly incomplete and ambiguous or imprecise. Consequently, inconsistencies are inevitable and need to be handled in the formalisation process. This has lead the formalisation work into alternative approaches in order to integrate support for also atypical patient’s cases. While the knowledge base in the prototype systems that has been evaluated, has been represented using production rules supporting primarily the typical cases, the support for atypical cases has been mainly been implemented as a dialogue with the user for evaluation purposes, providing evidence interpreted within different sets of guidelines, leaving the final decisions to the user.

A decision-support system, aimed at providing tools and guidance, needs to integrate a formalisation that supports a mapping between different scales and frames of interpretation (i.e. clinical guidelines) to accomplish a refinement of evidence and to support reasoning in difficult cases [Lindgren, 2005]. Since the different guidelines and qualifier-tools express the request for information about qualities of evidence differently, the expressivity of the logic language or languages chosen for the representation of the domain knowledge is crucial. A desired property of the formalisation is also that it allows the integration of new tools depending on local policies, or as an effect of the development of the domain knowledge.

Evidence-based approaches in medicine are well rooted in quantitative research methods and provide statistical tools to establish validity of formulas.
by hypothesis testing. However, this typically provides confidence concerning isolated independent decisions or choices. Validity of guidelines, i.e. sets of decision or choice rules, such as appearing in differential diagnosis frameworks, however, cannot be managed by conglomerated hypothesis testing. Rules interact, are usually cascaded, and purely statistical analysis of rules become unfeasible.

Developments in medical decision support still struggle with the leap from statistically rooted evidence-based choices to logically well-founded decision making with respect to sets of rules representing guidelines. In a knowledge and reasoning scenario, sets of rules are always syntactically represented in some logic, where inference mechanisms provide tools by which conclusions are drawn. In a logical framework, language constructions and its semantics need to fulfil correctness and completeness criteria. Furthermore, a set of rules formulated in a particular logic needs additionally to be validated, in the end by motivations provided in evidence-based medicine.

Statistics and logic do indeed meet, even if meeting points and computational mechanisms therein are still rather unclear. Recent decades, however, present several candidate solutions to “statistical logic”. Key issues are e.g. conditionalities, where we already know that purely Bayesian approaches cannot be logically formulated. However, turning probabilities to possibilities, knowledge to beliefs, certainties to many-valuedness, and so forth, opens up a wide range of candidate logic machineries subject to investigations concerning suitability for medical decision support frameworks [Zadeh, 1978; Dubois and Prade, 1993; Benferhat et al., 1997; Fox and McBurney, 2002; Parsons and Hunter, 1998].

The papers II-V included in the thesis are such examples of case studies in which the application domain of diagnosing dementia is used to investigate the usefulness of different formalism candidates for the purpose of decision support in the domain. Paper II evaluates a probabilistic argumentation technique where changes in support are used in the reasoning process. The approach was found insufficient, therefore an argumentation framework is extended in Paper III using qualitative measures drawn from clinical guidelines for valuing support for different hypotheses in a reasoning process, integrating the notion of contexts for valuation as different frames of reference. Paper IV contributes to investigating dementia diagnosis in a many-valued framework integrating values drawn from the unit interval based on expert assessment, while Paper IV extends the example by applying general logics as a theoretical framework for accomplishing transformations between logics for the purpose of refining evidence.

A combination of techniques is required to meet the complexity of the domain, enabling flexible guideline representations and dynamic support in decision-making. Therefore, recent work has used General logics as a common formal,
categorical framework for identifying the basic building blocks of logics. For this purpose categorical structures such as institutions can be used to define the relations between different logic languages in order to assure valid transformations between languages in the reasoning process [Goguen and Burstall, 1984; Meseguer, 1989]. To accomplish this, structure-preserving functions are used, which can be seen as an implementation of a morphism in terms of category theory. The purpose is to assure that the internal structure of a logic is preserved in the transformations. This line of work is currently in progress. Therefore, the summary in this chapter and the Papers I and V can be seen as an introduction to future work.

The second line of investigations that have been conducted has been addressing the reasoning process as a process executed over time, with new, sometimes conflicting data continuously added to the reasoning process, which need to be handled. Further, when using clinical guidelines as source of knowledge, which express the knowledge in a qualitative way using expressions in natural language without clear associations to quantitative measures, techniques that have the potential of providing valuations with a larger adherence to the knowledge source than numerical based techniques become highly interesting. This line of work includes investigations of the defeasible formalisation approaches, and argumentation techniques in particular ([Lindgren, 2005], Paper II, III).

A summary is given in the next section of the work done in applying techniques for associating weights to evidence, rules, or sets of rules. Mainly, the summary will provide an overview of the papers included in the thesis. The chapter concludes with a synthesis sketch of the different approaches in the perspective of the structure of the knowledge base evolved in Paper I and in the framework of category theory. The sketch constitutes the foundation and outline for future work.

14.1 Summary of Logics Used in Paper I-V

Traditionally, a logic $L_{\Sigma}$ over a signature $\Sigma = (S, \Omega)$, where $S$ is the set of sorts and $\Omega$ is the set of operators producing terms, consists of a set $L_{\Sigma}$ of formulas and a satisfaction relation $\models \subseteq Alg(\Sigma) \times L$, where $Alg(\Sigma)$ is the set of all algebras over the signature $\Sigma$. We frequently write $\Phi \models \varphi$, $\Phi \subseteq L$, to mean that for all $A \in Alg(\Sigma)$ we have $A \models \varphi$. $\Phi$ becomes a theory if $\Phi \models \varphi \Rightarrow \varphi \in \Phi$. In this situation, satisfaction transforms to being $\models \subseteq PL \times L$, where $P$ is the powerset functor.

Logic calculus involves inference rules with proof derivation being the relation $\vdash \subseteq PL \times L$. A logic system can thus be seen given by $\mathcal{L} = (L, \models, \vdash)$.

When the signature $\Sigma$ is given, the set of terms is defined traditionally. There
14.1 Summary of Logics Used in Paper I-V

is a choice to be made whether the constants and operators, for the formation of well-formed formulas over terms, themselves form terms over its signature, or if well-formed formulas are treated separately in a metalanguage. We use the latter standpoint in Paper III, while in current work, using the framework of General logics in Paper V, we use the former standpoint.

Propositional logic $\mathcal{L}^p = (\mathcal{L}, \models)$ can be viewed as a situation in form of a one-sorted signature where $\Omega$ consists of constants, $\neg$ as a unary operator, and $\wedge$ as a binary operator, with disjunction $\vee$ and implication $\rightarrow$ as the usual shorthand forms based on $\neg$ and $\wedge$. Note that we may interpret formulas in $\Phi$ to be true formulas. Thus we could equivalently say $(p, \text{true})$ is in $\Phi$ whenever $p$ is in $\Phi$. Similarly, we would have $(q, \text{false})$ in $\Phi$, whenever $\neg q$ is in $\Phi$. We then make truth values in the semantic domain more visible. This is useful when we extend to many-valuedness, for the purpose of integrating medical scales with sets of different values.

Two-valued logic can be extended to involve uncertainties and many-valuedness in several ways. Most commonly, extensions involve only truth values, and term sets remain as two-valued crisp sets. Full extensions of logic calculi involving generalized powersets of terms is still, however, largely undeveloped. Extension involving truth values moves to use various lattices of truth values. Requirement concerning properties of these lattices vary from logic to logic. Intuitionistic style logic involves Heyting algebras, where many-valued logic generally speaking uses completely distributive lattices. Many-valued logic specified as “fuzzy logic” typically uses the unit interval of truth values, in particular for engineering applications.

In Paper III sets of truth values (or dictionaries) are used based on qualitative measures expressed in clinical guidelines without foundational concerns of the lattice structures of the dictionaries. Ordering among elements in dictionaries are discussed mostly from a pragmatic point of view and in relation to the context-based model of argumentation presented in the paper. Further, in Paper III only the syntactic aspects of argumentation is addressed, leaving the semantic properties to future work. However, the paper illuminates the complexity of values when different sets of values are required in the same reasoning process. It is argued that a combination of dictionaries should be used for local computations of levels of support for a particular hypothesis, for the hypothesis to be comparable with other hypotheses. In the local computations for each hypothesis, dictionaries can be used that correspond closely to clinical guidelines, which may capture the level of significance a particular feature holds in the context of a certain set of clinical guidelines concerning a hypothesis. This motivates the extension of the argumentation framework to include the notion of context.
to distinguish localities, and further, to capture the different localities in terms of category theory and establish sound transformations between the contexts, which could possibly require different logics for formalisation (Paper I, V). Firstly, an introduction to the argumentation notation used in Paper III is given.

14.1.1 Paper II, III: Argumentation Frameworks

The notion of defeasibility and defeasible reasoning was well-known and used in philosophy of law around the time when Toulmin defined a conceptual model for argumentation [Toulmin, 1958]. In the last decades an important cross-breeding from different disciplines (philosophy, computer science, law, decision making) has developed especially the theoretical foundations for argumentation. For a thorough analysis of the evolvement of argumentation approaches within AI see the review by Chesñevar and co-workers [Chesñevar et al., 2000]. The main idea of argumentation is to structure reasoning in the way that rules that support a conclusion can be defeated when new information arise. By chaining these defeasible inferences in order to reach a conclusion, arguments are created, consisting of certain components according to Toulmin’s model. Arguments differ from proofs in the way that they are defeasible, tend to leave some premises implicit, tend to use vague, open textured terms and be “open world”. Another difference is that all possible arguments do not necessarily have to be considered in argumentation, only those which have been actualised, which is applicable and used in a law context for instance [Chesñevar and Simari, 2002]. A key issue in argumentation frameworks is how to define the relations between arguments, in order to create preferences and reasons for defeat among arguments. One of the most influential contributions to computational argumentation is the work of Dung [Dung, 1995], who defines a theory of argumentation within first-order logic. A central notion is the acceptability of arguments and he proves that argumentation can be seen as a special case of logic programming with negation as failure. He also shows that most of the major non-monotonic reasoning approaches are special forms of this theory of argumentation. In [Lindgren, 2005] and in Paper II and III two argumentation approaches are applied and evaluated in the perspective of dementia diagnosis [Parsons, 1998; Parsons, 2004; Fox and Parsons, 1998]. The following is an introduction to the structure developed in Paper III as a result of the evaluations.

Let $L_{mvL}$ be some many-valued extension of a propositional logic $L^\pi$ with respect to a dictionary $D_0$ of (truth) values. In the presentation in Paper III of the argumentation logic $L^{AL}$, propositions in the base logic, i.e. some many-valued extension of propositional logic, are embedded into clauses in $L^{AL}$ which are of form $(i : l : d)$, where $i$ is its name (or index), $l$ is a well-formed formula, i.e. $l \in L_{mvL}$, and $d \in D_1$. Note that $D_1$ must be distinguishable from $D_0$. A set
of such clauses is called a database (of clauses). We write $I_\Delta$ for the name (or index) set related to a database $\Delta$, i.e. $I_\Delta = \{i \mid (i : l : d) \in \Delta\}$.

An argument $a$ (for a well-formed formula $p$) is a triple $(p, G, d)$, where $p \in L^{AL}$, $G \subseteq I_\Delta$, and $d \in D_2$. The set $G$ represents the set of supporting clauses for the proposition, $p$. Again note that $D_2$ must be distinguishable from both $D_1$ and $D_0$. In Paper II we considered the case $D_2 = D_0$. The set of arguments for $L^{AL}$ is denoted $\text{Arg}(L^{AL})$.

As pointed out, we are not considering formal semantic theories in Paper III. However, as there is an intuitive understanding of diagnosis contexts, we provide some basic ideas concerning formal treatment of such contexts in the paper. It is clear that contexts are to be based on validation mappings. Moreover, we consider contexts of the form $v : \text{Arg}(L^{AL}) \to D_1$.

Note that we obviously need suitable transformations between $D_0$, $D_1$ and $D_2$.

$\vdash_{AL}$ is defined and used to reason about changes in beliefs in the framework. In the building process when the rules are used, values are handled and combined, in order to reach a value of validity of a proposition. Every distinct argument with the value $d$ concerning $p$ has to be taken into account and combined in an aggregation process. A number of different arguments for a certain claim have to be mapped into a single measure, which is a process called flattening. The flattening function $\text{flat}^A$ in the sense of [Fox and Parsons, 1998] maps a set of possibly conflicting arguments $A_p$ for a proposition $p$ to an overall measure of validity $d$ in the proposition, i.e.:

$$\text{flat}^A : A_p \mapsto (p, d)$$

where $d$ is some combination of values in $D_2$.

In order to capture the clinical reasoning process, which is performed within different contexts of interpretation, or frames of reference, the argumentation logic framework can be seen extended with the notion of context given respectively by $D_0$, $D_1$ and $D_2 \subseteq D$ together with corresponding flattening functions (Paper III). A context-based argumentation logic framework for a knowledge domain will then consist of the triple $(A, D, F)$ where $F$ is a set of flattening functions that maps sets of arguments $A \subseteq \text{Arg}(L^{AL})$ to sets of values $D_i \subseteq D$, corresponding to the different frames of understanding used in the reasoning process. In this scenario, the flattening function $\text{flat}^A$ maps a set of arguments $A_p$ for a proposition $p$ to an overall measure of validity $d$ in the proposition in...
the context \( \mathbf{v} \), i.e.;

\[
flat^A : A_p \mapsto (p, d)
\]

where \( d \) is some combination of values in \( D \), and \( \mathbf{v} \) is the local context of interpretation, or validation, of the evidence.

The semantics applied in Paper II is limited to values concerning changes in confidence in a proposition, which was found insufficient for our purposes. Therefore, the semantic domains in the example in Paper III are given by

\[
D = \{\text{excluded, unlikely, contradictory,}
\]

\[
\text{uncertain, supportive, core,}
\]

\[
\text{possible}^-, \text{possible}, \text{possible}^+,
\]

\[
\text{possible}^+, \text{probable}^-, \text{probable},
\]

\[
\text{probable}^-, \text{probable}^+, \text{confirmed}
\]

as an alternative to the approach taken to argumentation in Paper II. However, the argumentation framework is general in that different dictionaries can be integrated containing modalities, ordered sets, the unit interval, possibilities, etc. A conditional uncertainty over \( L^{AL} \), or \( L \) for short, is a mapping

\[
\tau_{L^{AL}}^{cond} : L^{AL} \times L^{AL} \times L^{AL} \to [0, 1]
\]

where we write \( \tau_{L^{AL}}^{cond}(a \mid b, X) \) instead of \( \tau_{L^{AL}}^{cond}(a, b, X) \). Clearly, \( \tau_{L^{AL}}^{cond} \) should fulfil suitable properties ([Fox and Parsons, 1998]). For describing conditional uncertainty we actually do not need to fix our semantic view concerning \( \tau_{L^{AL}}^{cond} \) neither in form of probabilities or as possibilities, or as something else.

Semantics of clauses could be defined e.g. as \( (i : a \rightarrow b : d) \) being true if and only if

\[
\tau_{L^{AL}}^{cond}(b \mid a, X) \geq \tau_{L^{AL}}^{cond}(b \mid \neg a, X)
\]

for all terms \( X \) over the signature for which \( (i : X \rightarrow b : d) \) is true for any \( d \in D_i \).


14.1 Summary of Logics Used in Paper I-V

14.1.2 Paper IV: Basic Many-Valued Logic and NPL

In Paper IV Hájek’s basic many-valued logic (BL) [Hájek, 1998] is used as extension from two-valued (propositional) logic. The basic many-valued logic, establishing soundness and completeness over 1-tautologies, i.e., absolute truth of formulas, involves two conjunctions, where (strong) conjunction is semantically specified by a selected t-norm. The basic many-valued logic framework nicely incorporates Łukasiewicz and Gödel logic, simply by a particular choice of a t-norm. The set of axioms for the basic many-valued logic needs for Łukasiewicz and Gödel logics be extended with one axiom.

Basic many-valued logic (BL) \( L_{BL} = (L_{BL}, \models_{BL}) \) is a propositional language, syntactically with propositional variables, a propositional constant \( \bot \) (false), and additionally with binary connectives \& (strong conjunction) and → (implication) as operators in \( \Omega \). If \( \varphi \) and \( \psi \) are formulas in \( \Phi \) then so are \( \varphi \& \psi \) and \( \varphi \rightarrow \psi \). Semantically, formulas are mapped to truth values in the unit interval \([0, 1]\), thus extending to many-valuedness saying \((p, i)\) is in \( \Phi \) whenever the truth-value \( p \) is in \( \Phi \) to the degree \( i \). Strong conjunction is by evaluation mappings \( e \) semantically defined by a given t-norm \( * \), according to \( e(\varphi \& \psi) = e(\varphi) * e(\psi) \). Further, \( e(\bot) = 0 \) and \( e(\varphi \leftarrow \psi) = \max\{x \mid e(\varphi) * x \leq e(\psi)\} \). A weaker conjunction is defined as \( \varphi \land \psi = \varphi \& (\varphi \rightarrow \psi) \), and negation is according to \( \neg \varphi = \varphi \rightarrow \bot \).

The restriction to 1-tautologies is rather strong from application point of view. In [Eklund and Klawonn, 1992], graded formulas are allowed and used, including more flexible ways to incorporate connectives than approaches like [Pavelka, 1979] does. Further, and in order to enable methodology such as parameter estimation within evidence-based medicine, we may use only continuous logical connectives. This type of many-valued logic has been shown to be close to neural network like structures. In fact, we can even transform these neural propositional logic programs into corresponding generalized neural networks, where synaptic functions need not only be weighted sums but rather correspond to connectives of the logic. These transformations then enables us to use efficient learning algorithms involving parameters represented by certainty values of formulas [Eklund and Klawonn, 1992].

This extension to gradation of formulas can be further generalized to many-valued propositional logics involving a set of binary connectives rather than just one for (strong) conjunction. In the Neural Propositional Logic (NPL) approach [Eklund and Klawonn, 1992] the language consists of the same propositional constants as in [Pavelka, 1979], the unary logical connective \( \neg \), the binary connective \( \rightarrow \), and a finite set \( B \) of binary connectives that are continuous, associative, and monotone increasing in both arguments. Continuity and monotonicity...
are required when NPL programs are transformed to neural-like networks where truth values are optimized with gradient descent techniques [Eklund and Klawonn, 1992]. The semantics of negation is $e(\neg \varphi) = 1 - e(\varphi)$ and for implication we define $e(\varphi \rightarrow \psi) = \min\{e(\psi) - e(\varphi) + 1, 1\}$. This corresponds to the situation in BL where $*$ is the Łukasiewicz $t$-norm.

The analysis in the case study presented in Paper IV uses NPL with its flexibility to include various types of connectives. The graded formulas included in the case study produce diagnoses of different reliability, depending on the significance of different evidence used in the inferences. The grades, provided by a medical expert, are based on subjective experiences of diagnosis and can not be accessed in an explicit way from the domain knowledge. Therefore, the results can only be used as an exemplification of the internal structure of the logics used in the example. The same difficulties arise in any application of a numerical approach to formalising clinical guidelines, where the assessment of measures such as probabilities is required. Assessing numbers is a problem because the domain knowledge typically is text-based and mediates qualitative assessments of evidence, interpreted for clinical practice from studies in evidence-based medicine. The statistical results from these studies can vary and are often incomplete. However, what makes NPL an interesting approach is that it have the potentials to be of use to construct graded formulas based on a larger set of patient’s cases. In this way, the assessment of grades becomes less ad-hoc and more reliable than using subjective interpretations. Thus, the approach has the benefits of using both patient data sets and clinical guidelines as source and grounds for the formalisation, which will be taken into consideration in future work when the usage of DMSS has generated a database of patient data that can be used for developing the domain knowledge.

Other numerical approaches, or approaches where numbers are associated as weights, were reviewed and evaluated in case studies in [Lindgren, 2005]. The conclusion was drawn that these approaches were of limited use, since the assessment of weights was required from experts, which is a less satisfactory solution, as discussed in [Lindgren, 2005], and because the expressivity of the approaches were too limited for our purposes. In order to provide support in the difficult patient’s cases, the different types of support expressed in clinical guidelines need to be captured, i.e., by distinguishing the support generated by core features for a particular hypothesis, and the support generated by supportive features as defined in clinical guidelines. While a reliable numerical foundation lacks for the assessment of the different types within the same formalism, a combination of techniques is required to capture the multi-dimensionality in the knowledge. In order to capture the different techniques and the reasoning process within the same formal framework, in order to assure valid transformations between logics.
14.1 Summary of Logics Used in Paper I-V

When needed, we use category theory as a metalanguage for the formalisation and general logics, which includes the necessary logical structures for the purpose. In the following subsection a brief introduction is given to Paper I and V and the motives of using the framework.

14.1.3 Paper I, V: General Logics and Institutions

A category consists of objects (entities and their structures), morphisms, which are the functions between the objects, compositions of morphisms, and for each object an identity morphism. Functors are morphisms between categories, which map both the object and the morphisms between the categories. A category can have other categories as its objects, which will be the case in the following introduction to the formalisation.

To capture the evidence in a patient’s case, suitable underlying categories are used with knowledge represented within the categorical logic framework, in turn being based on selected underlying categories. The basic building block in the reasoning process is the entity and its characteristics. The entity is a part of a set of knowledge, corresponding to a model of the knowledge at hand at a certain point in the investigation process. In our case, the model corresponds to the Patient MAP introduced in Paper I. The transformation of the knowledge (i.e. action), may constitute a mapping of the entity into a new set of knowledge by morphisms defined for the particular category.

In atypical patient’s cases the transformation of knowledge may require mappings between different underlying categories for evaluating the evidence in another context of integrated tools, using other morphisms. This is the case when the evidence needs to be interpreted using different logics. Therefore, we consider the category of logics (LOG) which includes the different logics needed for the representation of the evidence and reasoning at different phases of the investigation process. The objects in this category are logics, defined as 5-tuples $L = (\text{Sign}, \text{Sen}, \text{Mod}, \models, \vdash)$, each consisting of the category \text{Sign}, the $\models$ relation defining the semantic relations between the set of sentences constructed by the \text{Sen} functor, operating on the \text{Sign} category, and the models constructed by the \text{Mod} functor, and the syntactic relation $\vdash$ defining the entailment relations. A logic consists of an institution and an entailment system as defined in the original work of Goguen, Burstall and Meseguer [Goguen and Burstall, 1984; Meseguer, 1989], summarised and exemplified by Helgesson [Helgesson, 2007].

The morphisms that are part of the category LOG, which are used for mappings between logics are defined in [Meseguer, 1989]. \text{Sign} is a category which has signatures as previously defined in this section as objects. Mappings can be formalised on different levels within the LOG structure, i.e., between different signatures within the same logic or between logics, depending on the knowledge...
The model of clinical activity developed in Paper I is used to structure reasoning and investigation processes and their content. The results concerning differential diagnosis of cognitive diseases were used in a case study in Paper V, as an example of mappings between entailment systems for refinement of clinical evidence. Instead of aiming for integrating knowledge within one single logic with the associated limitations discussed in the thesis, a dynamic structure, which allows flexible knowledge representation, is the subject for Paper V. Flexible in the sense that a formalisation can be suitable for different levels of care, requiring knowledge of different granularity and reliability (i.e., many-valuedness of truth), according to different clinical guidelines, but also knowledge of different character, suitable for different professionals at different points in time. In Paper

Figure 14.1 Example of mapping between two categories of patient data in the Patient MAP: Function (ICF) vs. Dysfunction (pathology) (DSM-IV). The functor maps the whole structure including the objects and their relations. The semantics concern the first step of the main reasoning process: DEO cognitive disorder in the context of the guideline DSM-IV.
V, a case study is provided in which an entailment system is split into different entailment systems, based on different signatures of interest. The necessary theoretical framework is established, including a formal description of a mapping between the entailment systems.

### 14.2 Knowledge Representation for DMSS - Summary

This section provides a synthesis of the different approaches to formalisation in the perspective of the structure of the knowledge base evolved in Paper I.

As described in Paper I, a knowledge base needs to integrate three dimensions of knowledge; 1) the basic rulebase grounded in clinical guidelines, 2) the meta-rulebase which handles the reasoning process and the ambiguities in the domain knowledge, the analyses using different tools, and directs the interactive reasoning; and 3) the Patient MAP which structures the information concerning a particular patient and the related evidence (Figure 14.2). Due to the fundamentally different characteristics of these functionalities and their knowledge structures, different formalisation approaches are needed. Differences lie primarily in the need to reuse information for follow-ups in daily practice, research, and developing the domain knowledge. While the basic rulebase may be built in a modular structure, by different sound and validated logic languages, with emphasis on entailment for reasoning about particular patient’s cases, the Patient MAP mediates the substance and validation of evidence in the perspective of sets of patients’ cases. Further, to provide means to adjust the support for local clinical practice, the meta-rulebase needs to be flexible and adhere to differences in treatment protocols and policies at different work environments. While the major part of the basic rulebase and Patient MAP can be built on international domain knowledge, the meta-rulebase needs to be adjusted and adjustable to local and individual routines. In this work, a creative design of the user interface is needed, which should mediate the knowledge and functionality of the knowledge base through mechanisms of interactive reasoning, suitable for different professionals.

By defining a context of interpretation, the rules can be validated within this context, in order to assure sound reasoning. The different contexts are exemplified as sets of guidelines (GLset) represented using different suitable logics in the framework of general logics in Figure 14.2. However, non-monotonicity may arise when the results of interpretations in different contexts need to be combined and compared. Some arise when integrating different clinical guidelines. Since the framework of general logics aims to integrate primarily techniques that are monotone, in order to assure valid transformations, the inconsistencies needs to be acknowledged and handled. These can be handled by the meta-rulebase through an approach similar to the argumentation technique presented in Pa-
Further, there are several logics that are being developed, suitable for implementing ontologies, which would also be appropriate for formalising the Patient MAP. Establishing which formalisms to be used is the next step in the formalisation process, and to implement and validate each module to patient sets. Further development of the formal and categorical framework will be needed, for instance, investigating and developing structures for accomplishing transformations of entities to sets of entities.

Figure 14.2 Structure of the three dimensions of knowledge base concerning main reasoning process.
CHAPTER 15

Summary and Future Work

In this thesis the work on developing a foundation for a dynamic decision-support system is described, which takes the development-potentials in the user into account, as well as the development of care processes. This includes organisational factors as well as characteristics of the domain knowledge that impose constraints on formalisation methods. Further, the decision-support system for the domain of investigation of cognitive diseases that is being developed is presented, including methods for development and formalisation.

The prototype system for decision-support in the domain (DMSS) has been developed in an iterative process and evaluated in different user settings, as one means to improve dementia care. The work context and the evidence-based domain knowledge have been investigated. The system is built on the data structure and reasoning model initially provided by the geriatrician who has functioned as the user representative and domain expert in the development process. The structure and reasoning model have been developed and validated, and the formalisation of the domain knowledge has been evolving during the development process.

The evaluations of terminology resulted in a re-formulation of symptoms presented in the GUI in neutral terms, in a similar way as in ICF, leaving valuations of terms to scales integrated in the system. This was found necessary in order to prevent misinterpretation of terms due to ambiguities when the system is implemented in different cultural environments. In addition, it was found of importance to use terms and concepts defined in international ontologies in order to make the system useful as a tool for research and for disseminating and enhancing a common medical terminology.

Currently, a case study is being done for the purpose of investigating the relation to other systems in clinical practice, by studying actions made throughout investigations of dementia concerning particular patients with the system integrated. Initial results are presented in this thesis concerning the assessment of the user context, and the valuation of the role of the decision-support system. Further, the conceptual model of clinical activity based on the cultural-historical activity theory is presented, which can be used as a tool for decision-support development for clinical practice. The method is being developed and evaluated, especially regarding its usefulness in the transformation of informal knowledge into formal evidence for implementation. The case studies are being extended with additional patient’s cases and professionals in order to validate the prelim-
In the knowledge acquisition process certain dimensions have been identified, which have major impact on representation and the quality of the outcome of inferences made by the system. These are the level of granularity of evidence and inferences, the level of complexity in the characterization of evidence and the level of ambiguity in the suggested diagnoses provided by the system. The reliability of suggested diagnoses increases by specificity of the evidence, which in turn may depend on the ability of the source to account for specificity. This is taken into consideration in the design of the presented system. As a result of these aspects two categories of patients were defined; typical patients, who comply with the clinical guidelines DSM-IV and consensus criteria for DLB, VaD and MCI, without ambiguity, and atypical patients, who present mixed aetiology or incomplete evidence for diagnosis. For typical cases the inferences implemented with if-then rules according to these guidelines are sufficient to provide a suggestion of a diagnosis, while in the atypical cases additional clinical guidelines are contributing to the decision support. The inferences used for support in these cases need to perform defeasible reasoning, since the domain knowledge is incomplete and produces contradictory hypotheses of diagnosis. The work presented in this thesis shows that it is possible to integrate the domain knowledge, also when it consists of partially conflicting clinical guidelines, into a decision-support system. This can be accomplished by formalising the domain knowledge integrated in a process of diagnostic reasoning, taking the qualities of different clinical guidelines into account and providing the user an interpretation of the evidence in a patient’s case within different contexts. Thus, the purpose of the decision-support system is not to deliver the one and only reliable suggestion of a diagnosis or none at all, but to present evidence in a form that makes the evidence useful for interactive decision making in clinical practice. To accomplish this, a structure of the knowledge base includes different structures for 1) the patient data structure, 2) a basic knowledge base and 3) a meta-knowledge base governing interactive reasoning and its content and tools. The structure and its content is generated by the work analysis, and will be used in further development and evaluations.

A combination of techniques is considered the most applicable solution for formalisation of the domain knowledge of cognitive disorders, in order to provide different levels of support depending on the complexity of the patient’s case. For typical cases classical first-order logic is sufficient, while in cases of ambiguity and mixed aetiology, i.e., atypical cases, techniques which allow defeasible reasoning and multi-valuedness have been investigated. In the search for suitable techniques, a few techniques were found particularly interesting for the purpose, which have been presented and applied to a subset of the knowledge
domain [Lindgren and Eklund, 2005b; Lindgren and Eklund, 2005a; Eklund and Lindgren, 2006; Eklund et al., 2007]. The technique which will be used for the atypical cases must allow features to have different levels of significance related to different diagnoses and allow multi-diagnosis. The techniques that make use of more than one intermediate level of positive support (which can be attached to a rule or a set of rules) and that allow working with several contradicting hypotheses in the reasoning process within different contexts, comply with clinical guidelines in the domain and are therefore of interest.

The first steps have been taken in the formalisation of the domain knowledge, and the theoretical work will continue in the context of a foundational view of transformations between logics. The medical and theoretical motivations for this work and some formal indications concerning the transformations are presented in the thesis. The development work is continuing with the representation of the evidence within the different contexts, validation of the knowledge base and the integration into clinical practice.

A Japanese version is being developed in cooperation with a municipality in Japan. Extensions for supporting municipality personnel in the care of dementia residences is being developed in cooperation with municipalities in the county of Västerbotten and in Japan. When the system is fully accessible in daily practice further evaluations will be done, which will show what impact the contextual issues discussed in the thesis have on the use of the system. The impact the system may have on the users in the form of a cognitive tool for development of skill and knowledge, will also be investigated when the system is integrated in clinical practice.
References


Helgesson, R. (2007). A categorical approach to logics and logic homomorphisms. Ummad 07.676, Department of Computing Science, Umeå University,


Wells, G. (1999). The zone of proximal development and its implications for


