

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Massey University Library
Thesis Copyright Form

Title of thesis:

- (1) (a) I give permission for my thesis to be made available to readers in Massey University Library under conditions determined by the Librarian.
- (b) I do not wish my thesis to be made available to readers without my written consent for ... months.
- (2) (a) I agree that my thesis, or a copy, may be sent to another institution under conditions determined by the Librarian.
- (b) I do not wish my thesis, or a copy, to be sent to another institution without my written consent for ... months.
- (3) (a) I agree that my thesis may be copied for Library use.
- (b) I do not wish my thesis to be copied for Library use for ... months.

Signed Ana Estela Antunes da Silva

Date 25.02.94

The copyright of this thesis belongs to the author. Readers must sign their name in the space below to show that they recognise this. They are asked to add their permanent address.

NAME AND ADDRESS

DATE

The AIM Methodology
Analysis, Interpretation And Modelling: A Methodology For The Role Of
The Knowledge Analyst In Knowledge Acquisition.

A thesis presented in partial fulfilment of the requirements for the degree
of Master of Science in Computer Science at Massey University.

Ana Estela Antunes da Silva
February, 1994.

Abstract

A methodology for the analysis, interpretation and modelling of knowledge is developed in this thesis - the AIM methodology. This methodology can be used by a knowledge engineer who has had no involvement with the elicitation process. The knowledge engineer who carries out the processes of the AIM methodology is called the knowledge analyst.

As a general description, the AIM methodology presents techniques together with their advantages and pitfalls in the processes of analysis, interpretation and modelling. The process of analysis consists of the refinement of the elicited knowledge. During this process, techniques such as the discourse analysis are used. The interpretation and modelling processes aim at representing the domain elements obtained during the analysis process. Models such as the model of expertise of the KADS methodology are used during the interpretation and modelling processes.

This research also aims at defining the activities involved in the knowledge acquisition phase in the development of knowledge-based systems. In this thesis, knowledge acquisition is considered as a set of processes which comprises not only the elicitation of knowledge, but also the refinement of the knowledge elicited through its analysis, interpretation and modelling.

Acknowledgements

Thanks are due to my supervisors, Mrs. Elizabeth Kemp and Mrs. Elisabeth Todd for their invaluable work and encouragement throughout this thesis.

Thanks to Mr. David Gray for the help with the data set used in this thesis.

Ministry of Foreign Affairs and Trade is thanked for the financial support of this study.

For all assistance, specially during my first year at Massey University, I thank Professor Mark Apperley.

Finally, thanks to my husband, Sila for his encouragement and patience.

Table of Contents

Abstract	ii
Acknowledgements	iii
Table of Contents	iv
List of Tables	viii
List of Figures	viii
Chapter 1: Introduction	1
Chapter 2: Literature Review	4
2.1 Introduction	4
2.2 Knowledge acquisition and elicitation	5
2.3 Elicitation techniques	8
2.3.1 Interviewing techniques	9
2.3.1.1 Structured interview	9
2.3.1.2 Semi-structured interview	10
2.3.1.3 Unstructured interviews	10
2.3.2 Protocol analysis	11
2.3.3 Construct elicitation	13
2.3.4 Card sorting	14
2.3.5 Laddering	15
2.4 Analysis techniques	16
2.4.1 Discourse analysis	17
2.4.2 Semantic networks	18
2.5 KADS methodology	20
2.5.1 The models in the KADS methodology	20
2.5.2 The model of expertise	22
2.5.3 The library of problem-solving activities - The interpretation models -	27
2.6 Other methodologies for knowledge-based systems	29
2.7 Conclusion	31

Chapter 3: Domain problem	32
3.1 Introduction	32
3.2 Dairy industry overview	32
3.3 Dairy production from pasture - An overview of New Zealand dairy systems	34
3.3.1 Problem domain	36
3.4 Research project: the development and evaluation of a knowledge-based system for dairy farmers	37
3.4.1 Project goals	37
3.4.2 Knowledge-based systems and dairy production	38
3.4.3 The selection of experts	39
3.4.4 The elicitation process - semi-structured interviews	39
3.5 Conclusion	40
 Chapter 4: Methodology	 41
4.1 The analysis process	43
4.2 The interpretation process	44
4.3 The modelling process	47
4.4 Conclusion	48
 Chapter 5: Case study and discussion	 50
5.1 Analysis	50
5.1.1 Analyze Transcripts process	50
5.1.1.1 The use of the discourse analysis technique	52
5.1.1.2 Example of application of the discourse analysis technique	52
5.1.1.3 Results of the use of the discourse analysis technique	55
5.1.1.4 Roles of a knowledge engineer	56
5.1.1.5 Practical considerations of the discourse analysis technique	57

5.1.2 Create Graphic Representation process	58
5.1.2.1 Examples of the use of semantic networks . .	58
5.1.2.2 Practical considerations of semantic networks	59
5.2 Interpretation	61
5.2.1 Select Template process	62
5.2.2 Create and Integrate Templates process	64
5.2.2.1 Create and integrate inference structure	64
5.2.2.2 Result of the integration of the templates . . .	65
5.2.2.2 Create task specifications	67
5.2.3 Practical considerations of the selection of templates	68
5.3 Modelling	69
5.3.1 Model Domain Knowledge process	69
5.3.2 Model Inference Knowledge process	70
5.3.2.2 Example of an instantiation of the inference structure	72
5.3.2.3 Practical considerations of the Model Inference Knowledge process	72
5.3.3 Model Task Knowledge process	74
5.3.3.1 Instantiation and refinement of the generic task template	74
5.3.3.2 Decomposition of the generic tasks	76
5.3.3.3 Practical considerations of the Model Task Knowledge process	78
5.4 Conclusion	78
 Chapter 6: Conclusion	 80
6.1 Results of the application of the AIM methodology	80
6.2 Future work	81
 References	 82

Appendix 1: Results of the discourse analysis process	90
A1.1 Domain Primitives	90
A1.1.1 List of concepts and attributes	90
A1.1.2 Relationships	91
A1.1.2.1 Short term period	91
A1.1.2.2 Long term period	110
A1.2 Task Model	118
A1.2.1 Functions obtained from the transcripts during the discourse analysis which led the construction of the Task Model	118
 Appendix 2: Domain Layer	 121
A2.1 Domain Primitives	121
A2.1.1 List of concepts and attributes	121
A2.1.2 Data dictionary	123
A2.1.3 Graphical representation of the summer-autumn management period	137
A2.1.4 Relationships	138
A2.2 Network of Primitives	141
 Appendix 3: Inference layer	 148
A3.1 Primitive inference actions and their input and output knowledge roles	149
A3.2 Concepts of the domain which were instantiated with the knowledge roles	151
A3.3 Relationships of the domain instantiated with the primitive inference actions	153
 Appendix 4: Task layer	 156
A4.1 Task decomposition	158

List of Tables

Chapter 3

Table 3.1: Summary of statistics for the New Zealand Dairy Industry . . .	34
---	----

List of Figures

Chapter 2

Figure 2.1. An Is_A semantic network	18
Figure 2.2. Levels of semantic networks	19
Figure 2.3: Intermediate models in the KADS methodology	21
Figure 2.4. The KADS four-layer model	25
Figure 2.5. Inference structure for hypothesis and testing	26
Figure 2.6. The KADS library of problem-solving activities	28

Chapter 3

Figure 3.1. The vertical structure of the New Zealand Dairy Industry . . .	33
Figure 3.2. Dairy farm events in relation to the pattern of pasture growth	35

Chapter 4

Figure 4.1. Data flow diagram of the AIM methodology	42
Figure 4.2. Techniques and representations of the Analyze Knowledge process.	43
Figure 4.3. Example of Network of Primitives	44
Figure 4.4. Techniques and representations of the Interpret Knowledge process	45
Figure 4.5. Inference structure of monitoring template	46
Figure 4.6. Task specification of monitoring template	47
Figure 4.7. Techniques and representations of the Model Knowledge process	48

Chapter 5

Figure 5.1. Task Model of the domain	55
Figure 5.2. High level network of primitives	59
Figure 5.3. Different points-of-view represented by semantic networks . .	60
Figure 5.4. Example of integration of the experts' networks	61
Figure 5.5. Inference structures	63
Figure 5.6. An initial inference structure for managing a dairy farm	65
Figure 5.7. Knowledge roles and primitive inference actions	67
Figure 5.8. Generic task specification of summer-autumn management of a dairy farm	68
Figure 5.9. Parts of the inference structure of the summer-autumn management period of a dairy farm	71
Figure 5.10. Final Inference structure of the summer-autumn management of a dairy farm	74
Figure 5.11. Decomposition of the generic task template	75
Figure 5.12. Example of decomposition of the generic tasks: Monitor Events and Diagnose Problems	77

Appendix 1

Figure A1.1. Task Model of the summer-autumn management of a dairy farm	118
--	-----

Appendix 2

Figure A2.1. High level network	142
Figure A2.2. LiveStock network - "is_part_of" relations	143
Figure A2.3. Other relations involving livestock	143
Figure A2.4. Mating information network - "Is_A" relations	144
Figure A2.5. Climate network - "is_part_of" relations	144
Figure A2.6. Cull cow network - "Is_A" relations	145
Figure A2.7. Grazing systems network	145
Figure A2.8. Soil condition network	146
Figure A2.9. Measurement methods network - "Is_A" relations	146
Figure A2.10 Feed supply network - "Is_A" relations	147

Figure A2.11 Average pasture cover network - "Is_part of" relations . .	147
---	-----

Appendix 3

Figure A3.1. Final inference structure of the summer-autumn management period	148
--	-----

Appendix 4

Figure A4.1. Generic task template of the summer-autumn management period	156
--	-----

Chapter 1: Introduction

Knowledge-based systems development has been the subject of many research studies for several years (Guida and Tasso, 1989; Bramer, 1987; Jackson, 1990). Knowledge-based systems can be distinguished from more conventional systems because they use representations of human knowledge instead of algorithmic approaches. They solve problems by heuristics or rules of thumb simulating human reasoning. They should give solutions to complex problems and/or give advice on difficult situations that an expert would do if facing the same circumstances.

One of the most important phases in the development of knowledge-based systems is the knowledge acquisition process. This is often described as a difficult and time consuming task (Motta *et al.*, 1989a; Aamodt, 1990; Althoff *et al.*, 1990). Factors such as the difference between the vocabulary used by experts and knowledge engineers, and the gap between the knowledge obtained from experts and the representations in the knowledge base are identified as some of these difficulties. Several methodologies for building knowledge-based systems have been considered in the literature (Motta *et al.*, 1989a; Karbach *et al.*, 1989; Motta *et al.*, 1989b; Wielinga *et al.*, 1991). Although none of them has been identified as the best development solution, the KADS methodology promises to succeed in several areas where other approaches have failed. Knowledge acquisition is one of these areas.

The methodology used in this research aims to position knowledge acquisition as a phase within the development of knowledge-based systems as well as to establish its main activities. This research considers knowledge acquisition as the phase which contains the processes of elicitation, analysis, interpretation and modelling of knowledge. Three of these processes are described in this methodology: analysis, interpretation and modelling. Together they comprise the AIM methodology.

The AIM methodology is mainly based on analysis techniques and the model of expertise of the KADS methodology. The research data set is also an important part of this study. The domain data set used to evaluate the AIM

methodology was obtained in the first phase of a research project at Massey University. The domain area concerns New Zealand dairy production during the summer-autumn period and involves the main decisions dairy farmers have to make to maximise their profits and the problems they face when making such decisions.

The AIM methodology also presents two different functions for the knowledge engineer. According to Hart (1992), to perform the role of a knowledge engineer a person should possess such characteristics as good communication skills, tact and diplomacy, logicity, technical knowledge and domain knowledge. Having a knowledge engineer with a background in the domain is a crucial point. It is very important that the expert does not think the knowledge engineer is too inexperienced or unclear about the domain during the elicitation process. However, it is difficult for the knowledge engineer to gain a good understanding of a completely new area in a short period of time and talk to the expert at the same level of knowledge. This has given rise to the question of who should be the knowledge engineer, a person with a technical background or a person with domain expertise.

To address this issue, this methodology considers two different roles for the knowledge engineer. The first role is responsible for the knowledge elicitation process and is directly involved with the expert. This person is called the knowledge elicitor. In this research, two participants of the dairy project played the role of the knowledge elicitor. The second role of knowledge engineer is called the knowledge analyst. This person is responsible for all the processes of the AIM methodology: analysis, interpretation and modelling. The knowledge analyst is not directly involved with the expert. The author of this thesis played the role of the knowledge analyst. Although considering the work of the knowledge elicitors as the input data set, the processes of this methodology are designed to be performed by the knowledge analyst. A similar view about different roles for the knowledge engineer is presented by Brule' & Blount (1989). The authors mention the advantage of working as a team which contains two knowledge engineers who would assist each other in the elicitation and analysis processes.

In the second chapter the supporting literature is described emphasising the importance of the knowledge acquisition phase and the different definitions given to this process. This chapter also focuses on the interchangeable way in which knowledge acquisition and knowledge elicitation are mentioned in the literature and explains the modelling approach of the KADS methodology. Chapter 3 describes the structure of the dairy project and its suitability for a knowledge-based system. In chapter 4, the AIM methodology developed in this research is explained. The chapter describes the structure of the AIM methodology and how the techniques and the model of expertise were used. In chapter five, the applicability of the AIM methodology is analyzed through a case study. A discussion of the results of the application of the methodology is also provided. Chapter six contains the conclusion drawn from the research and the future work that can be done.