Remko J.H. Scha

Logical Foundations for Question Answering
RIJKSUNIVERSITEIT TE GRONINGEN

LOGICAL FOUNDATIONS FOR QUESTION ANSWERING

PROEFSCHRIFT

ter verkrijging van het doctoraat in de Letteren
aan de Rijksuniversiteit te Groningen

op gezag van de Rector Magnificus Dr. L.J. Engels

in het openbaar te verdedigen

op donderdag 17 februari 1983

des namiddags te 4.00 uur

door Remko Jan Hendrik Scha

geboren te Eindhoven
Promotores:

Prof. Joyce Friedman
en
Prof. Frank Heny
Remko J.H. Scha

LOGICAL FOUNDATIONS FOR QUESTION ANSWERING

Contents.

Preface.

Chapter I. Methodological Preliminaries.

1. Introduction.
2. Artificial Intelligence.
4. Limitations of Predicate Calculus.
5. Conclusion.

Chapter II. Questions and Answers.

1. Introduction.
2. Characterizing the Content of a Question in Terms of its Propositional Answers.
   2.1. Hamblin: Questions as Sets of Possible Answers.
   2.2. Karttunen: Questions as Properties of True Answers.
   2.3. Answer-Propositions in Context.
   2.4. Groenendijk and Stokhof: Exhaustiveness.
   2.5. Problems with Rigid Designators.
3. Against the Primacy of Full-Sentence Answers.
5. The PHLIQA1 treatment: Questions and Answers as Describing Sets of Individuals.
   5.1. Introduction.
   5.2. Questions.
   5.3. Answers
   5.4. Indefinite Answers.
   5.5. Full-Sentence Answers.
6. Quantifying into Questions.
   6.1. Introduction.
   6.2. The "Compound Speech Act" Analysis
   6.3. A Proposal by Groenendijk and Stokhof.
   6.4. The PHLIQA1 Treatment.
7. The Pragmatics of Answering.
   7.1. Categories of Answers.
   7.2. Pragmatic Strategies.
8. Conclusion.
Chapter III. The PHLIQA1 Question Answering System.

1. Introduction.
2. External Requirements for a Question Answering System.
3. The Top Level Design of PHLIQA1.
5. The World Model Language.
6. The Data Base Language.
7. Translations.
8. The Control Structure of the PHLIQA1 Program.
9. The Behaviour of the PHLIQA1 System.

Chapter IV. Data Bases as Value Specifications.

1. Introduction.
2. Value Specifications.
3. Relational Data Bases Viewed as Value Specifications.
4. CODASYL Data Bases Viewed as Value Specifications.
5. Data Bases as Axiom Sets.

Chapter V. Translation Specifications: a Technique for Representing the Conceptual Information of a Question Answering System.

1. Conceptual Information: the Bridge Between Different Levels of Meaning Representation.
2. Conceptual Information in the Form of Translation Rules.
   2.1. Translation Rules.
   2.2. Type Constraints.
3. Translation Between Languages with Different Type Systems.
4. Identification Translations.
   4.1. The Problem of Compound Attributes.
   4.3. The Definition of Identification Translations.
5. Simplification Transformations.
6. Extending the Data Base Language.

Chapter VI. Alternative Knowledge Representation Techniques.

1. Introduction.
2. Rule schemes.
   2.1. Introduction.
   2.2. Rule Schemes for First-Order Languages.
2.3. An Example of the Use of Global Rule Schemes.
2.4. Theoretical and Practical Aspects of
      the Use of Global Rule Schemes.
3. Definitions Within One Language.
   3.1. Introduction.
   3.2. Translation and Evaluation.
   3.3. Interweaving Translation and Evaluation.
5. The Closed World Assumption.
6. Theorem Proving and Data Bases.
   6.1. Introduction.
   6.2. Reiter's proposal.
   6.3. The Exhaustiveness of the Answers in Reiter's System.
   6.4. The Closed World Assumption.

Chapter VII. Conclusion: Design Styles.

References.

Appendix A. Syntax and Semantics of the PHLIQA1 Languages.
1. Introduction.
2. The Type System of the PHLIQA1 Languages.
3. The Definition of the Expressions of a PHLIQA1 Language.
4. The Semantics of the PHLIQA1 Languages.
5. Semantic Anomaly.
6. Additions and Abbreviations.

Nederlandstalige Samenvatting.
Preface.

This thesis has a long history. It began in 1971 at Philips Electrologica in Apeldoorn when Herman Schweigmann suggested to me that I investigate the idea of developing a Natural Language Question Answering System with an ordinary data base as its knowledge base. In response to this suggestion, I developed the concept of a question answering program which would gradually transform the logical representation of a natural language query into a data base query by applying a series of "translation rules". Between 1972 and 1979 a system of this kind, called PHLIQA1, was actually developed at Philips Research Laboratories in Eindhoven.

Wim Bronnenberg, Harry Bunt, Jan Landsbergen, Piet Medema, Wijnand Schoenmakers, Eric van Utteren and myself were the participants in the first phase of this process which ended with the implementation of a Question-Answering System in 1975. Although the system incorporated many new ideas and displayed an interesting structure, its theoretical underpinnings were less than completely satisfactory. Occasionally, this would also show up in incorrect answers or otherwise undesirable responses.

Therefore, in the period from 1976-1978, important aspects of our approach were rethought. The system was redesigned at that time by Wim Bronnenberg, Jan Landsbergen, Wijnand Schoenmakers, Eric van Utteren and myself. In 1979 I wrote an elaborate description of the new program (Bronnenberg et al., 1980) while Bronnenberg, Landsbergen, Schoenmakers and Van Utteren implemented it (within a few months) and debugged it (within a week). The system was successfully demonstrated for an extensive period without displaying any unexpected behavior.

This thesis may be viewed as a theoretical complement to PHLIQA1. It expands on some important ideas underlying this system and compares them to possible alternatives. The issues focussed on here are some of those I feel particularly responsible for in the development of PHLIQA1: the over-all design and the semantics of questions, answers and data.

I wish to thank my former colleagues on the PHLIQA1 Project for everything which they have indirectly contributed to these pages. Particularly, I wish to thank Jan Landsbergen for years of stimulating cooperation, and to acknowledge his share in many aspects of the design of PHLIQA1. His crucial contributions to the development of the method of Translation Specifications, discussed in Chaper V, should be mentioned especially.

In writing this book, I have borrowed freely from some papers which I wrote some time ago, especially Bronnenberg et al. (1980) which forms the basis of Chapter III and Scha (1982) which is incorporated in large part in
Chapter V. The content of Chapter II was developed in a series of talks I have given at the Philosophy Departments of the Universities of Amsterdam, Groningen and Nijmegen.

I would like to thank both of my promoters, Professor Joyce Friedman and Professor Frank Heny, for their thorough and far-reaching comments on both content and form of my earlier drafts. If the present book is readable at all, that is largely due to their efforts and to the extremely active editorial assistance I have received from Dr. Livia Polanyi.