Chapter 1: Introduction

Computationally modeling and analyzing intelligence (CMAI) is the large issue toward which this thesis is directed. Situation theory (ST) provides the philosophical and logical basis for the approach of this work. There are many issues in CMAI. Many of these issues have in common various problems in knowledge representation (KR). This thesis provides an application of situation theory to some of these common knowledge representation problems, particularly in the area of reasoning about beliefs.

The central claim of this thesis is that situation theory is superior to classical logic as a foundation for knowledge representation in artificial intelligence. This claim is elaborated in the following four hypotheses:

1) A version of situation theory can be defined which has a characterizing logic (an "infon" logic) similar in form and expressivity to classical first order logic.

2) There is a semi-decision procedure for this new infon logic, and a theorem prover can be devised which implements it. Further, many of the techniques of automated theorem proving developed for classical logic can be applied to automated theorem proving in this new infon logic.

3) This new version of situation theory and the associated theorem prover is appropriate as a knowledge representation and reasoning system for theories of perception and belief.

4) Theories of perception and belief as defined by their embeddings in the new version of situation theory provide a better account of human reasoning than classical logic-based computational approaches to perception and belief.

Situation theory is a recently developed theory, primarily concerned with understanding information and meaning among people. It has seen only very limited application to computational problems in natural language understanding. Thus, one of the contributions of this thesis is to further develop the computational application of situation theory. Another contribution is a new set of problems for situation theory to address - what are the inference rules appropriate to reasoning in situation theory? Also, this thesis is intended to contribute toward a computational system which is as good at intelligent behavior as a typical person is, neither better nor worse.

Many of the arguments in this thesis for how to solve various problems in intelligence are based (informally, introspectively and anecdotally, not clinically or biologically) on how *people* solve those problems. Solutions are based on how people *do* achieve intelligent behavior, not on how they *might best* achieve intelligent behavior. In some fields (probability, psychology), this distinction is between *descriptive* theory (how people behave to achieve their goals) versus *normative* theory (how people *should* behave to achieve their goals, i.e. what is the optimal, correct approach). This thesis develops a *descriptive* theory.

A basic position of this thesis is that some form of the "knowledge representation hypothesis" (KRH) is correct. A version of this hypothesis is:

> Any artificial intelligent process consists of structural ingredients that serve at least two purposes.
> On the one hand, these structural ingredients represent a propositional account of the knowledge that the overall process exhibits.
> On the other hand, they also play a formal, causal, and essential role in its behavior.[1]

There is a detailed discussion of the relationship between this hypothesis and situation theory later in the thesis.

Also basic to this thesis is the idea that beliefs, meaning and information are key issues in the computational modeling and analysis of intelligence. *Belief* is used in this thesis rather than *knowledge*. Barwise suggests that belief and knowledge are different (if related) notions, neither more basic than the other.[2] *Knowledge* has something to do with the actuality of a proposition. A *belief* is a proposition of

[1] This hypothesis was first summarized by Brian C. Smith on p. 33 in [Smith 1982]. Smith's version is discussed in section 2 of this thesis.

[2] pp. 213-216 and pp. 265-272 in [Barwise&Perry 1983].

which an agent is aware and which that agent is prepared to act on *as though* it is true. The position of this thesis is that it is immaterial in the reasoning of an agent whether or not that agent's beliefs are *actually* true, since this information is unavailable to the agent.

Situation theory had its first major publication in 1983, *Situations and Attitudes* by Jon Barwise and John Perry[3]. Many of the elements of situation theory and situation semantics differ between situation theory papers. In some sense, there are many situation theories. However, there are certain things common to most of the approaches to situation theory. The central concept is the *situation*. A situation is a part of the real world. Thus, there are no "possible" situations (in the sense that there are "possible worlds" in the standard account of modal logic), a situation must be found in the real world. A situation *supports* a "way things are" or "state of affairs". These supported things are called *infons* in the current versions of ST. Examples of infons are the color of a rock and the existence of a happy mood. These are ways things can be, which a particular situation may support. Since a situation is a *part* of the world, it does not support the way *everything* in the world is. Thus, a situation is not the world.

A second central idea is that it is theoretically essential that everything is situated in the real world. Thus, to have a theory which can account for intelligent agents it is necessary to explicitly include in such a theory the possible effect of the agent's situation.

The 1983 situation theory presentation, *Situations and Attitudes*, has been superseded in various ways, yet it remains interesting in that it is still one of the largest single works in situation theory and semantics - among the various works on situation theory, it covers the most issues with a single formalism. The other major single work in situation theory is Keith Devlin's *Logic and Information* [Devlin 1991].

Organization of the Thesis

[3] [Barwise&Perry 1983]

The rest of this introduction provides the motivation for using situation theory for knowledge representation and reasoning in CMAI. Chapter 2 discusses the issues in knowledge representation and shows how situation theory is used in terms of these issues. Chapter 3 presents the new formulation of situation theory in detail. This chapter proves the first hypothesis of the thesis. Chapter 4 presents the basics of FELIX, the natural-deduction style theorem prover which reasons in both classical logic and also the “infon logic” of the new situation theory. This chapter proves the second hypothesis of the thesis. Chapter 5 presents new formalizations of theories of perception and belief which are grounded in the new situation theory. This chapter proves part of the third hypothesis of the thesis and provides arguments in support of the fourth hypothesis. Chapter 6 presents the extensions to FELIX which allow it to reason in these perception and belief logics. This chapter proves the other part of the third hypothesis of the thesis. These extensions, primarily the “multiple intensional context” mechanism, also allow FELIX to reason in the full new situation theory. Chapter 7 briefly discusses the execution of FELIX on a variety of test problems. Chapter 8 presents some conclusions drawn from the work presented in the preceding chapters.

There are three appendices for this thesis. The first appendix contains all of the proofs for theorems presented in Chapter 3 (the development of the new situation theory). The second appendix contains a small fragment of the source code for FELIX. This fragment defines the forward and backward reasons which FELIX uses in reasoning about the various logics (classical, infon, support, and belief). These reasons are discussed in Chapters 4 and 6. The third appendix contains many proofs found by FELIX, exactly as FELIX printed them.

Why Use Situation Theory in Knowledge Representation?

There are several reasons why applying situation theory to issues in knowledge representation is a substantial contribution toward CMAI. These reasons can be addressed in three groups - why knowledge representation issues matter, why situation theory matters, and why applying situation theory to knowledge representation matters.

This thesis accepts as given that the KRH is basically valid. As presented above, the KRH claims that the knowledge representation plays a “formal, causal, and essential” role in intelligence. Thus, it follows directly from the KRH that knowledge representation is central to intelligence. From this it is easy to claim that KR issues are important in computational accounts of intelligence.

The argument that ST is important to computationally modeling and analyzing intelligence is not as direct. Situation theory and situation semantics are addressed to problems of information and meaning - how can people mean, how can they possess and communicate information, what is meaning, what is information. It also addresses mental states such as knowledge and belief, and mental processes of inference and perception. The fundamental idea is that all of these issues must be understood in the light of the fact that the “reasoning agent” is situated in the world, and that the most basic concept in the analysis is the “situation” - some part of the world (generally, a part of the world “accessible” to the agent). In the introduction to *The Situation in Logic*, Barwise describes the goal of situation semantics (and situation theory[4]) by:

> “It now seems to me that the best way to understand what situation semantics is trying to do is to look at it as relaxing a certain simplifying assumption in the study of language and logic. The key insight, it seems to me, is that speech, writing, thought, and inference are *situated* activities. That is, they are activities carried out by intelligent, embodied, limited agents, agents situated in a rich environment, an environment that can be exploited in various ways. As such, these activities are always taken from an agent’s perspective within that environment, and they are about other portions, generally restricted portions, of that environment, portions to which the agent is somehow, directly or remotely, connected. And being activities, they have impact, they change the environment within which the agent operates.”[5]

The above quotation makes the intent of situation theory clearly and specifically relevant to modeling and analyzing intelligence.

The value of applying situation theory to knowledge representation is established by

[4] The terms situation theory and situation semantics are not well distinguished in the situation theory literature. Either term is sometimes used to refer to the entire discipline. I have chosen to use situation theory as the more general term, reserving situation semantics as a phrase to specifically describe situation theory as applied to natural language semantics.

[5] p. xiii in [Barwise 1988]

the following argument. The first claim is that a knowledge representation built from a philosophically coherent foundation is more broadly useful than one with an ad hoc basis. The notion of philosophical coherence intended here is that the many (relevant) philosophical problems which must be addressed are dealt with in a unified manner, that a limited number of mechanisms are invoked to handle a wide variety of issues instead of inventing a new mechanism for each hard problem. This criterion is related to Lakatos' idea of a mature scientific research program.[6]

In such a coherent foundation interactions between different elements of the theory are understood, because they are thoroughly planned and analyzed. The restricted number of such elements makes such analysis feasible. This is in contrast with an ad hoc approach where the interactions are not planned, and frequently not even understood. The multiplicity of mechanisms in an ad hoc theory contributes to the difficulty of analyzing their relationships.

ST is such a coherent philosophical foundation. The basic orientation to "situations", the centrality of information instead of truth, and the relation theory of meaning lead to solutions for a wide variety of problems. Among these problems are semantics of anaphora[7], semantics of noun phrase reference, self-referential statements, attitude reports (e.g. perception, belief, knowledge, and doubt), inconsistent beliefs, and common knowledge.

ST is a promising approach to computational KR because of its formal mathematical approach. Some philosophical work, while extremely interesting for CMAI, is not formalized in such a way as to make its use in mechanical reasoning feasible. Emmanuel Kant's "Critique of Pure Reason" is an example of this. ST is by intent a formal mathematical theory. Thus, it can be manipulated mechanically.

[6] The relationship, however, is not simple. "A mature science consists of research programmes in which not only novel facts but, in an important sense, also novel auxiliary theories, are anticipated; mature science – unlike pedestrian trial-and-error – has 'heuristic power'." from p. 88 in [Lakatos 1970]. The coherence I argue for in the philosophical foundation is both a demonstration of existing heuristic power (many problems already solved without resorting to ad hoc approaches) and a promise of more successful and useful predictions to come.

[7] Anaphora is the use of pronouns or pro-verbs to refer to a word or phrase which generally precedes the pronoun or pro-verb in an utterance.

This formal mathematical approach extends to the metatheory of ST. Originally, the metatheory was Kripke-Platek admissible set theory with ur-elements[8] (KPU set theory)[9]. The current work uses ZFC-/AFA set theory as originally described by [Aczel 1988]. From the computational vantage point, an interesting feature of the use of either of these theories as metatheories is that finitely-represented sets are adequate for ST. In the contemporary versions of ST, the ZFC-/AFA theory allows for the use of finite representations of certain kinds of infinite sets. These sets are called “hereditarily finite”.

Inference is a central issue in KR and CMAI. ST is explicitly, in part, about inference. This area of ST is not yet well developed, but some novel results have already been produced. Developing the application of ST to KR positions one to take advantage of ST’s results in inference, when they are sufficiently well developed.

This leads to a final point supporting the choice of ST as a foundation for KR. ST is a young subject under active development by a growing group of researchers. As improvements are made to ST, these improvements can be applied to KR. It is reasonable to assume such improvements are forthcoming.

[8] Ur-elements can be thought of as a base set of named constants or atoms. A set theory can be constructed which has no “constants” except for the empty set. Such a set theory has no ur-elements. Constructions of sets involving the empty set in an ur-element-less set theory can be mapped into a collection of ur-elements in such a way that for any set theory with ur-elements there exists an isomorphism with some ur-element-less set theory. Ur-elements are used to make the set theory more convenient to use, they do not alter its semantics.

[9] Note 4 (on p. 327) of p. 53 of [Barwise&Perry 1983] cites pp. 1-50 of [Barwise 1975] as a sufficient source for learning about KPU. They were doubtful in [Barwise&Perry 1983] about its appropriateness, but it was the most appropriate set theory of which they were aware at the time of writing that book.

Aczel 1988 *Non-well-founded Sets* by Peter Aczel, Stanford: CSLI, 1988.

Barwise 1975 *Admissible Sets and Structures* by Jon Barwise. New York: Springer-Verlag, 1975.

Barwise 1988 *The Situation in Logic* by Jon Barwise, Center for the Study of Language and Information: Stanford University, 1988.

Barwise&Perry 1983 *Situations and Attitudes* by Jon Barwise and John Perry, MIT Press: Cambridge, 1983.

Devlin 1991 *Logic and Information* by Keith Devlin, Cambridge University Press: for publication in 1991.

Lakatos 1970 "Falsification and the Methodology of Scientific Research Programmes" by Imre Lakatos, in *Criticism and the Growth of Knowledge*, Edited by Imre Lakatos and A. Musgrave. Amsterdam: North Holland, 1970. Reprinted in *Imre Lakatos: Philosophical Papers. Vol. 1: The Methodology of Scientific Research Programmes*, edited by John Worrall and Gregory Currie, Cambridge, England: Cambridge University Press, 1978.

Smith 1982 "Prologue to 'Reflection and Semantics in a Procedural Language'" by Brian Cantwell Smith, Technical Report MIT/LCS/TR-272 M.I.T., Cambridge, MA, 1982. Reprinted on pp. 32-40 in *Readings in Knowledge Representation* by Ronald J. Brachman and Hector Levesque (eds.), Los Altos: Morgan Kaufmann, 1985.