CS 786G *Cognitive Modeling*
Week 10 Discussion Paper
AI ain't ready yet, come back in a thousand years

Steve Woods
89800462

May 26, 1995

Smith (1991) is essentially a position paper regarding the CYC project and the positions of Lenat and Feigenbaum on representation and intelligence. In fact, it is more of an anti-position paper than a position paper insofar as the intent of the paper is more to bring attention to what Smith feels are the gaping wounds in the philosophy of Lenat and Feigenbaum and their quest for an intelligent (or at least knowledgeable) agent than to offer a detailed specification of an alternative. For the purpose of this short paper, we will refer to Lenat and Fegenbaum as L&F in regards to their position and work towards representing all of human knowledge in a single encyclopedia-like system that is capable of reasoning about this knowledge, and more.

Smith attacks L&F on essentially very straightforward grounds in this paper. First there is the *attack on generalization* in which Smith points out rather amazing assumptions are made, starting with some assumed general principles of intelligent agents that are essentially unproblematic at the level of observation, but which require rather involved realizations in implementation. Whereas L&F seem to regard this jump from general to realized trivially, the **middle ground** in which the actual details of these ideas must be well specified and detailed in notably missing. One such instance of the big jump is given below:

> Intelligent performance often requires the problem solver to fall back on increasingly general knowledge, and/or to analogize to specific knowledge from far-flung domains.

As one general conception about human reasoning this idea is without obvious fault, as a model of how to achieve this aspect of human reasoning it is perhaps incomplete.

Those familiar with current research in planning and problem solving of which (Woods 1991, Yang, Tenenberg & Woods 1996) are exemplars, use of hierarchical knowledge such as discussed in (Woods 1993, Woods 1994), or analogy as dealt with in (Gentner 1981), are well aware that all of these fields are far from "solved" or even from being well understood in the sense of allowing for implementations of human complexity in any foreseeable future. Seemingly ignoring this current state of research in these and other AI fields, L&F still manage to outline a research program in 3 definite stages that will (by the end of the century) result in a system with an *approximately complete* representation of human knowledge that additionally will be able to (one would suppose, intentionally) read and assimilate additionally information on its own. The third stage closely following has CYC (or whatever) undertaking its own program of research. These remarkable claims would seem to require correspondingly remarkable breakthroughs hitherto unreported across essentially the breadth of AI, but Smith points out that these are non-existent, or at the very least, undocumented. The remainder of Smith's paper is an attempt to get at the fundamental misunderstandings that he believes that L&F demonstrate through their approach and implicit assumptions, and to justify through examples their drastic oversimplifications of many areas of the study of intelligence.

Smith attempts to describe the fundamental deficiency of the work of L&F as the missing "middle realm". This middle realm is essentially the "how-to" for the broader approaches to intelligent behaviour which L&F describe so prolifically. As a starting point for discussion about foundational characteristics and behaviour of systems with some intelligent behaviour, Smith outlines 12 questions that are applicable to any potential reasoning system, and then suggests the corresponding approach of a pure logic derivation system, the L&F or CYC approach, and the approach that would be minimally required by an AI system of Smith's own understanding.

Rather than repeat these 12 here, it is sufficient for our purposes to observe that the main observable result from this analysis is to realize that the approach of L&F and that of a purely logical system (or inference system) are highly correlated. Essentially, they share characteristics that Smith identifies as fundamentally flawed or overly limiting for any "intelligent" system, including (but not limited to) the following:

- Representational content unsituated, implying that you can understand anything independently to how it is located in the circumstances of its occurrence. Any fact can be represented (and consequently "understood" independently of other related facts.

- Meaning independent of use, essentially ignoring opposing views that hold that the meanings of an individual part of a whole (say sentence) are as dependent on the whole as the whole is dependent on the individual for its meaning. The CYC view is strictly that an individual has well-defined and absolutely meaning in every usage and that meaning is the same in every usage. This is an excellent example of L&F's complete disregard for other related work in philosophy, anthropology and linguistics which argues similar points to no final conclusion.

- A single representational scheme is sufficient and adequate for the encoding of knowledge in an intelligent system. Much recent work has clearly demonstrated the utility of different representations for different tasks. Pattern recognition, vision and image manipulations, reasoning with analogy, hypothesis and logical entailment, expression by charts, diagrams and models, number systems and language use but some parts of the human cognitive toolset, where each known representation for these tasks has complexity and knowledge density differences, while each suiting a certain type of use. For instance, charts convey more spatial information quicker and certainly in a broader context than could be the case for a written expression of the chart information. Certainly it is conceivable that a single representation exists (thought this is far from certain, or even probably) that encompasses all of these, but it is certainly not the frame-based approach of CYC.

Smith's primary concerns (and consequently the focus of his alternative conception known as embedded computation), lies in the dis-connectedness or lack of a viewpoint inherent in the formalization (such as it is) given by L&F. This lack of situatedness, this detached view of thought and action, is precisely what L&F have in common with general purpose inferencing systems, and is in direct conflict with what Smith believes to be becessarily inherent in any intelligent system. Certainly the concept of absolute knowledge independent of an observer's context is dissipating in many fields with the rise of post-modernist thought. A specific instance in which this thinking has become rather formalized is physics where for example, observer context is an integral factor in measuring velocity or position, and in which the concepts of absolute or "observer-independent" velocity or position are meaningless.

In some sense one gets a feeling that CYCers expect intelligence to emerge from some large collection of "fundamental" approximation of human knowledge and application of heuristic search control strategies. Surely the additional issues of expression of intent and will somehow have been lost in the hurry to shuffle control and information.

From my point of view as a researcher in search and search control strategies in (necessarily) limited domains and applications, it is hard to believe that anyone could even conceive a search strategy that would be capable of integrating enough "common sense" so as to overcome the unbelievable mountain of potential inferences and relationships among human knowledge so as to give meaningful replies, or even any replies. Very simplified domains that only require limited types of knowlege cannot be addressed adequately. Hundreds of researchers are trying in these limited domains to integrate knowledge and search on even a limited scale[1]. Perhaps it is important that rather than a continual subdivision of AI into compartments, someone take a larger view, but it seems unlikely a necessary step that all current and past work be somehow be viewed as completed when in fact, the local-experts would disagree vehemently.

[1]One such example is constrained in the attempt to integrate hierarchical domain knowlege to search in constraint satisfaction problems (Woods & Yang 1994, Woods 1994) with interesting but necessarily limited results.

References

Gentner, D. (1981), The mechanisms of analogical learning, *in* S. Vosniadou & A. Ortony, eds, 'Similarlity and analogical reasoning', Cambridge University Press, chapter 7, pp. 199–241.

Smith, B. (1991), 'The owl and the electric encyclopedia', *Artificial Intelligence* **47**, 251–288.

Woods, S. (1993), A method of interactive recognition of spatially defined model deployment templates using abstraction, *in* H. Merklinger, M. Farooq, P. Roberge & R. Grodski, J.J.and Dobson, eds, 'Proceedings of the Knowledge -Based Systems and Robotics Workshop', Department of National Defence, Government of Canada, pp. 665–675.

Woods, S. (1994), Knowledge based model recognition in program understanding using constraint satisfaction, Cognitive Modeling termpaper submitted to Dr. Paul Thagard, University of Waterloo, Waterloo Ontario.

Woods, S. & Yang, Q. (1994), Knowledge based model recognition as constraint satisfaction, submitted for publication.

Woods, S. G. (1991), An implementation and evaluation of a hierarchical non-linear planner, Masters Thesis available as technical report CS-91-17, Computer Science Department, University of Waterloo.

Yang, Q., Tenenberg, J. & Woods, S. (1996), 'On the implementation and evaluation of Abtweak', *Compuational Intelligence.* to appear.