Proc. 13th Midwest AI and Cognitive Science Society Conference, MAICS-2002, Chicago, pp. 161–168.

Building an Ontology for CIRCSIM-Tutor

Chung Hee Lee* Jai Hyun Seu** Martha W. Evens*

*Department of Computer Science
Illinois Institute of Technology, Chicago, IL, 60616
Phone: 312-674-1294
Fax: 312-567-5067
leechuniit@hotmail.com and evens@iit.edu

**Department of Computer Science
Inje University, Kimhae, Korea
Phone: 312-567-3851
Fax: 312-567-5067
jaiseu@cs.inje.ac.kr

Abstract

We have built an ontology to support our intelligent tutoring system, CIRCSIM-Tutor, designed to provide a backbone for the knowledge base, a reference for what kinds of things that the tutor can discuss with the student. The next step is to propagate this knowledge into case frames and logic forms for use in the system. In the process we are discovering gaps in the ontology and adding concepts to it. The difficulties involved in doing this by hand led us to investigate automatic approaches. We describe algorithms for machine learning of case frames and also for clustering the frame information to obtain an ontology.

Introduction

The focus of CIRCSIM-Tutor is on the baroreceptor reflex, the negative feedback system that controls blood pressure in the human body (Evens et al., 1993). This reflex is a source of confusion for many students, who need help in learning to solve problems. The tutor describes a perturbation of the cardiovascular system to the students and then asks them to predict the qualitative changes in the most important parameters (tell whether they will go up or down or stay the same). Then it launches a tutorial dialogue to help the student correct any mistakes or misconceptions.

* This research was partially supported by the Cognitive Science Program of the Office of Naval Research under Grant N00014-00-1-0660 to Stanford University and under Grant No. N00014-94-1-0338 to Illinois Institute of Technology. The content does not reflect the position or policy of the government and no official endorsement should be inferred.

An ontology can be conceptualized as a taxonomy (a big ISA tree) or as a semantic space. Its function is to define the relationships between the entities in the knowledge base and the language used by the tutor in asking questions and giving explanations. Ontologies have become increasingly popular with the spread of object-oriented programming, since the ontology defines the set of classes needed in an implementation and the hierarchical relationships between them, but they have been recognized as fundamental to knowledge-base construction for many years (Woods, 1975; Evens et al., 1980; Smith, 1985). They also turn out to be useful in information retrieval systems. (Evens et al., 1985; Nutter et al., 1990; Pretschner and Gauch, 1999).

Our current work on the ontology was sparked partly by the decision to move some of the knowledge from a frame system to a rule system in order to allow the expert tutors, our colleagues at Rush Medical College, Joel Michael and Allen Rovick, to make changes to the knowledge base more easily. We also need to create case frames for an expanded inventory of verbs as we collect and analyze more human tutoring sessions.

Our system will make use of this ontology on every level from surface language processing to inference. We plan to make the categories used in our logic forms, now chosen in an *ad hoc* manner, correspond to nodes in the ontology. The case frames used in parsing and text generation will describe their selectional restrictions in terms of the ontology, also.

This paper mentions some other work on ontologies, particularly by anthropologists, and describes the ontology for CIRCSIM-Tutor originally built by Glenn Mayer (1992). Then we look at some of the dialogues carried on by human tutors and explain why we need to expand this ontology. The next section discusses the new ontology. This is followed by a description of Dardaine's (1992) case frame tables and how the ontology enables us to use them more effectively. We then look at the logic forms that underlie the semantic analysis of student input and the generation of responses and show their relationship to the ontology. Finally, we describe our efforts to automate the process of acquiring case frames via machine learning and then clustering the fillers of the case frame slots to derive an ontology.

Related Work

Ontologies or ISA hierarchies have been particularly important in the understanding of sublanguages for special areas of expertise and the building of lexicons and grammars for those sublanguages (Grishman et al., 1986). Robert Amsler started his work on computational lexicology by building taxonomies of nouns and verbs based on dictionary data (Amsler, 1981). The same issues have become important in anthropology, especially in the development of ethnologies (Werner, 1978; Werner and Schoepfle, 1987). Werner used an ontology as the foundation for the research on the *Anatomical Atlas of the Navajo* (Werner et al. 1969/1981).

Glenn Mayer (Mayer et al., 1989) set out to build a knowledge base by parsing a chapter of cardiovascular physiology written by Joel Michael. In the process he developed an ontology for the cardiovascular system (1992). This ontology tells us that the aorta and the carotids are arteries and arteries are blood vessels. The vena cava is a vein and veins are blood vessels. Arterioles and capillaries are also blood vessels. The tutor assumes that the student knows these cardinal relationships and also knows that the heart consists of four chambers, the atria and the ventricles. Valves are also essential components of the heart, important to its role as a pump. All of these are then classified as CV body parts. There are a number of components of the nervous system involved in the control of blood pressure: adrenergic receptors, adrenergic fibers, alpha and beta receptors, sympathetic and parasympathetic and vagal nerves, all of which appear in the CV-NERVE part of the tree. There are a number of concepts from physics like volume and pressure and force and acceleration, compliance, afterload and preload, which are also part of this picture.

Our knowledge base involves a number of parameters or variables. Some are neurally controlled and some are not. Heart Rate and Total Peripheral Resistance and Cardiac Contractilty (also known as Inotropic State) are neural variables. Right Atrial Pressure (which is essentially the same as Central Venous Pressure), Stroke Volume, Cardiac Output are hemodynamic variables. So is Mean Arterial Pressure, which is also the variable controlled by the baroreceptor reflex and, therefore, the central concept in this particular world view.

Mayer's ontology also includes some body chemicals: epinephrine and norepinephrine, acetyl-choline, adenosine, calcium and sodium ions play important roles in the control of blood pressure by the baroreceptor reflex.

The New Ontologies

Analysis of human tutoring sessions has been the basis of the development of CIRCSIM-Tutor. If we look at the actual tutoring sessions we see three different contexts. First comes the context of cardiovascular physiology that is the basis for the tutoring sessions. Glenn Mayer's ontology, with some minor additions, handles this context very well. Second, there is the computer context of the tutoring sessions - the tutor tells the student to press the <enter> key or to use the mouse to move the cursor. The terms that we see fit into the tree in Figure 1. While it contains some ISA relationships, this tree is largely structured by meronymy (the part-whole relation).

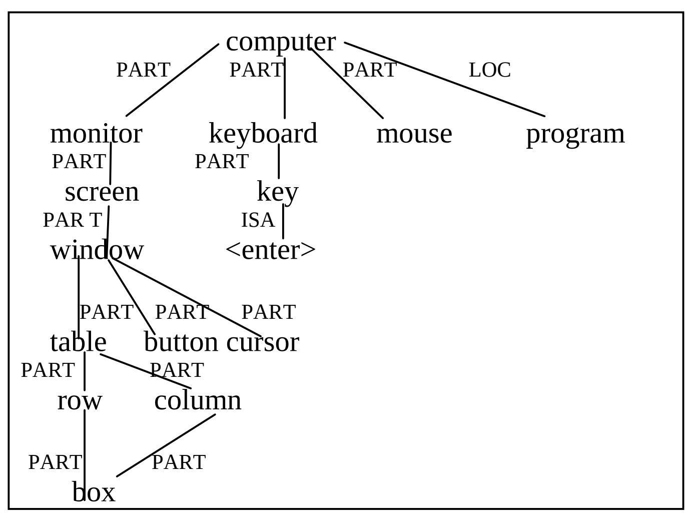

Figure 1. Ontology for the Computer Context

The third context that we see is even more abstract. The tutor is trying to guide the student to solve a problem defined by a procedure in the database. In this process the student and the tutor use verbs like "figure out" and "understand" with human subjects and concept nouns as objects.

We need to develop an ontology for the context of problems and procedures and analogies and solutions, of ideas and predictions, of questions and arguments. The evidence for the structure here is not as clear as in the other two ontologies but using *Webster's Seventh New Collegiate Dictionary* gives us the structure in Figure 2.

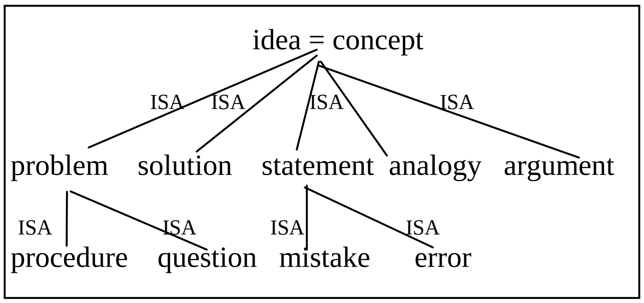

Figure 2. Ontology for the Problem Context

Case Frames and How We Use Them

Ten years ago Joanne Dardaine (1992) developed case frames for the verbs used in the first twenty-four human tutoring sessions. This effort was part of the process of building a large lexical database (Evens et al., 1991; Conlon et al., 1994). We are now building case frames for the new verbs in the fifty additional sessions that we have captured since that time.

The information in the table of case frames was designed to support the parsing and text generation needed to carry on a tutorial dialogue with a student. For each argument of the verb we include the syntactic role, the semantic role, the selectional restrictions, and also the occurrence information (an indication of whether that argument slot must be filled, or may be omitted, or may be omitted only if it is understood).

The syntactic roles include subject, object, indirect object, prepositional phrase, and a variety of clause types, based on Grimshaw (1990; Grimshaw and Jackendoff 1985). Our case frames like most other work on case, are based on the work of Fillmore (1968), but we have actually used the inventory of semantic roles developed by Allen (1987). We also include at least one example for each verb sense. A few of the new case frames that we have developed are shown in Table 2.

It is the selectional restrictions that interact with the ontology in important ways. If the parser is to make effective use of the case frame information, the selectional restrictions must be part of the ontology. If the student uses the word "combine," the system must try to figure out whether the student was talking about chemical reactions or about a human being confusing two ideas, the two senses represented in our example frames above. In order to make the distinction the parser must be able to discover that *bicarb* is a chemical. So we need to record that *bicarb* is an abbreviation for *sodium bicarbonate*, which is a *base*, which is a *chemical.*

Case frames also have an important role to play in text generation. Much of the tutoring dialogue involves changes in parameters and how one change causes the next one. The verbs increase and decrease mean both "go up/down" and "cause to go up/down." There are a number of ways to express these ideas besides "go up" and "increase" including "rise" and "fall" and "raise" (the causal form of "rise"). Which one we use depends partly on which arguments are available. It is the case frames that tell us how to map logic forms into syntactic choices. When Dardaine was bulding the case frames we were still working on a variety of approaches to parsing. So sometimes the selectional restrictions were phrased in terms of the ontology and sometimes they were not. We need to revise the old case frames so that they use terms in the ontology and make sure that these terms are used consistently in the new case frames. We may need to extend the ontology to make this feasible.

Logic Forms in the CIRCSIM-Tutor System

The ontology can be viewed as a definition of all the things that the tutor is prepared to discuss with the student. If we are to have an integrated and extensible knowledge base, it is essential that the slots in the logic forms be filled by elements of the ontology. The logic form (affect <var-list1> <var-list2>) underlies questions like “What variables determine CO?” or “What is the next variable affected?” and statements like “CO is determined by Heart Rate and Stroke Volume.” The logic form (mechanism <CV-mechanism> <var-list>) underlies questions like “By what mechanism is Heart Rate controlled?" and statements like "Heart Rate is neurally controlled." In these examples <var-list> is a list of one or more of the CV-parameters that dominate the discussion.

Automatic Extraction of Case Frames

The work of Hindle (1990) and Ravin (1990) convinced us that we should undertake an automatic approach to these tasks. Chung Hee Lee was inspired by the work of David Faure at the Laboratoire de Recherche en Informatique at CNRS in Paris Faure (Faure and Nédellec, 1998, 1999; Faure, Nédellec and Poibeau, 2000). Faure's ASIUM system uses machine learning to discover the subcategorization frames of verbs. These frames are like our case frames except that they do not include the column for the semantic roles shown in Table 2. Then the slots in the frames are clustered to build the ontology, using similarity measures from information retrieval.

We have had to make a number of changes in Faure's algorithms to compensate for the major difference in the data that we are using. The ASIUM system was designed to work on expository text. This text was carefully edited to contain complete and grammatical sentences. Our data consists of tutoring dialogues between professors at Rush Medical College and first year medical students, mostly questions and answers. The work described here is mainly based on the transcripts of twenty-five new dialogues collected in November, 1999. These dialogues are full of spelling errors and many of the student answers are consist of short phrases, not complete sentences. We often see noun phrases, without any verbs. Thus our system models mainly the language used by the tutor and it may turn out to be even more useful in generation than in parsing. The tutors are attempted to teach qualitative causal reasoning in physiology and the language is very domain specific. The most frequent verbs are mainly verbs of change, like *increase* and *decrease*.

Overview of Our Methods

We collect nouns and verbs (as shown in Table 3). Then we count their frequency. Our assumption is that the arguments of a particular verb are highly similar and this similarity also appears in the subcategorization frames. We sort the verb arguments by grammatical categories - subject object, indirect object, and prepositional phrase. The occurrence of the verbs *to be* and to *have* as auxiliaries as well as main verbs has caused us to postpone their analysis. The instantiated verb arguments are represented as:

<verb> <syntactic role | preposition: head word>*

There are two major differences between our work with verb arguments and Faure's work. He can assume his input text is correct and we cannot. Also, our goal is to build useful case frames for both input understanding and text generation, but Faure used it for understanding only. To achieve this goal we have to capture a compact representation of the selectional restrictions. We have adopted the similarity calculation used by Faure (Faure and Nédellec, 1998, 1999; Faure, Nédellec and Poibeau, 2000) This similarity calculation is useful not only for merging the arguments of two or more verbs, but also for splitting ambiguous arguments. Arguments to be clustered should have maximum overlap in order to form new classes. Thus arguments to be clustered must be strictly similar (distance is 0), while the distance between disjoint arguments without any words in common is the maximum value (distance is 1). The similarity is defined as the proportion of common head words in two arguments taking into account their frequency. Figure 3 shows the calculation of the similarity between clusters C1 and C2. Card(C1)and Card(C2) represent the number of head words in clusters C1 and C2. Ncomm is the number of different head words common to both C1 and C2. ψFC1 and. ψFC2

$$\mathrm{Dist}(C1, C2) = 1 - \left[\frac{\psi\ FC1 * Ncomm/Card(C1) + \psi FC2 * Ncomm/Card(C2)}{\psi\ f(word_i\ C1) \quad + \quad \psi\ f(word_i\ C2)} \right]$$

Figure 3. Distance Formula

are the sums of the frequencies of the head words of C1 and C2 And $word_i C1$ is the i-th head word of cluster C1 and also $f(word_i C1)$ its frequency. The weights Ncomm/Card(C1) and Ncomm/Card(C2) minimize the influence of the word frequencies by offsetting the attraction phenomenon of very frequent words among the verb arguments and increase clustering efficiency. For example, the similarity between the two clusters in Table 1 is calculated as follows.

C1: Decrease subject	C2: Fall subject
CVP 1	CVP 1
CO 4	CO 2
SV 3	SV 2
TPR 4	HR 1
MAP 3	
HEART 1	

Table 1. Verb Arguments

Ncomm/Card(C1) = 3/6 (3 of the 6 head words of C1 occur in C2 - cvp, co, sv), Ncomm/Card(C2) = 3/4, FC1 = 1 + 4 + 3 = 8, and FC2 = 1 + 2 + 2 = 5, thus the distance is

$$\frac{8 \times (3/6) + 5 \times (3/4)}{(1 + 4 + 3 + 4 + 3 + 1) + (1 + 2 + 2 + 1)}$$

This evaluates to 0.26, then the similarity is 1 - 0.26 = 0.74. Using this similarity measure, the arguments of the verbs are generalized so that new concepts are obtained. These concepts then replace their descendants in the selection restrictions in the verb case frames. Thus we can construct case frames automatically for reducing the garden paths in the parsing.process.

Next we modified the clustering method of Faure for learning ontologies and partitioned the verb arguments in terms of ontologies. ASIUM searches arguments using a bottom-up and best-first strategy, but we use a breadth-first and best-first search algorithm (Figure 4). This algorithm extracts a lot of taxonomic information given only a small number of common words in the clusters and also shows a multi-dimensional taxonomy in the domain. The supervised learning in this method is useful for controlling implausible similarities interactively between words as well as correcting clusters in the case of misspelled words. Sometimes new clusters have appropriate names. For example, the cluster {*heart rate*, *hr, co, bv, is*, *inotropic state*, *tpr*, *variable}* already suggests *variable* as the group name. This process can be used bottom-up to merge words and clusters to obtain a taxonomy. It can also be used top-down to break down the arguments in the case frames to represent the selectional restrictions in a compact form.

```
cluster-to-aggregate ← basic-words
new-cluster ← basic-words
repeat
  max-arg-length ← find-lonest-arg-length in cluster-to-aggregate
  for x = 1 to max-arg-length of cluster-to-aggregate
      for all cluster (Ci, Cj), Ci ∈ cluster-to-aggregate and C2 ∈ new-cluster
          candidate-cluster ← nil
          if dist (Ci, Cj) ← Threshold and Ci ≠ Cj
             then
             Cnew ← aggregate (Ci, Cj)
             candidate-cluster ← candidate-cluster ∪ Cnew
      endfor
      new-cluster ← validation of candidate-cluster
      cluster-to-aggregate ← cluster-to-aggregate ∪ new-cluster
   endfor
 until new-cluster = nil
```

Figure 4. Clustering Algorithm

Automatic Clustering to Obtain an Ontology

The verbs *increase, decrease,* and *change* all share the same subcategorization frames. We clustered their subjects and objects separately. The first result of clustering the subjects was a large cluster containing: *HR, Heart Rate, CO, IS, Inotropic State, TPR, variable, CVP. MAP, heart, SV*, and *CO.* Since all of these terms except heart are names of variables, we viewed this as a major success. When we examined the data to see if we could find out why *heart* was included, we discovered that the parser had misparsed a student sentence (from turn 36 of session K52):
"Okay, since the MAP is too high, in response, the body would want to make the heart less 'stretchy', and therefore decrease IS."
The parser decided that *heart* was the subject of *decrease*, and thus clustered *heart* with the variables that increase and decrease. We are considering weighting student input less than tutor sentences.

We tried raising the similarity threshold and looking at the resulting subclusters. We were hoping that we might be able to cluster the abbreviation HR with Heart Rate and IS with Inotropic State, but we did not succeed. We did get the cluster {CVP, CO, SV}. Since SV determines CO and CO determines CVP this seemed to make a lot of sense. Further refinement gave us separate clusters {SV, CO} and {CO, CVP} and then finally singleton clusters {CO}, {SV}, and {CVP}.

Although we did not mange to cluster IS and Inotropic State together as a pair, an experiment using a slightly different similarity measure produced the cluster {TPR, IS, Inotropic State, Variable}. This cluster seems to represent the group of neural variables (the variables controlled by the nervous system). Ordinarily, HR would form part of this group, but the tutoring sessions on which this work is based involved a pacemaker failure in which the HR was clamped and did not vary as neural variables ordinarily do. This result suggests that this approach to automatic ontology construction may well lead to multi-dimensional ontologies.

Conclusion

Summary

We have described the ontology that we have built for the CIRCSIM-Tutor knowledge base and the way that it is used in the parsing and generation of tutorial dialogues. In fact, the ontology is not one structure but three. One provides the structure for the knowledge of cardiovascular physiology that is the basis of the tutoring knowledge. Another describes the computer context of the tutoring sessions and supports the interaction in which the tutor helps the student get started using the system, the part of the dialogue in which the student learns how to "play the game." The third ontology supports another metalanguage, the discussion of where to attack the problems and how to solve them, the teaching of the problem-solving algorithm.

Current work on the ontology was sparked partly by the decision to move some of the knowledge from a frame system to a rule system in order to allow the expert tutors, Joel Michael and Allen Rovick, Professors of Physiology at Rush Medical College, to make changes to the knowledge base more easily.

Future Research

The ontology is fundamental to the CIRCSIM-Tutor lexicon and to the knowledge base as a whole. Future research on the lexicon involves adding more features to the entries. Other important features for verbs include the different kinds of sentential complements that each verb supports. These features also interact with the case frames. We also want to supply information about arguments for adjectives. This means establishing semantic classes for adjectives that determine the nouns that they can modify. We need to make sure that these classes correspond to nodes in the ontology or else modify the ontology structure to support them. Our adverb entries are definitely inadequate. The generator especially needs adverb placement information from the lexicon.

Our work on automatic extraction of case frames is only preliminary and there is much more work to be done to bring up to the level of the work that we have done by hand, but we believe that even if the process cannot be made fully automatic it will still improve our ability to move the tutor to new domains.

New work on the knowledge base is being carried forward by Jay Yusko and this ontology is being used to define the structure of the knowledge base for Version 4 of CIRCSIM-Tutor.

References

Allen, James. 1987. *Natural Language Understanding*. Benjamin/Cummings, Menlo Park, CA.

Amsler, Robert. 1981. A Taxonomy for English Nouns and Verbs. *Proceedings of the Annual Meeting of the Association for Computational Linguistics.* Stanford, CA. 133-138.

Conlon, Sumali, Joanne Dardaine, Agnes D'Souza, Martha Evens, Sherwood Haynes, Jong Sun Kim, Robert Strutz. 1994. The IIT Lexical Database: Dream and Reality. In Current Issues in Computational Linguistics: In Honour of Don Walker. Antonio Zampolli, Nicoletta Calzolari, and Martha. Palmer, eds. *Linguistica Computazionale*, Vol. IX. 201-225.

Dardaine, Joanne. 1992. Case Frames for a Lexical Database, *Proceedings of the Midwest Artificial Conference*, Starved Rock, IL, 102-106.

Evens, Martha, Bonnie Litowitz, Judith Markowitz, Raoul N. Smith, and Oswald Werner. 1980. *Lexical-Semantic Relations: A Comparative Survey.* Linguistic Research, Edmonton, Alberta.

Evens, Martha, James Vandendorpe, Yih-Chen Wang. 1985. Lexical-Semantic Relations in Information Retrieval, in S. Williams (ed.) *Humans and Machines: The Interface Through Language*, Ablex, Norwood, NJ, 73-100.

Evens, Martha, Joanne Dardaine, Y-F. Huang, Sun M. Li, Judith Markowitz, Frank Rinaldo, Margi Rinaldo, and Robert Strutz. 1991. For the Lexicon That Has Everything. *Proceedings of the Siglex Workshop,* Berkeley, CA, June, 1991. 179-187. An extended version appeared in James Pustejovsky and Sabine Bergler, eds., *Lexical Semantics and Knowledge Representation*, Springer-Verlag, Berlin. 219-233.

Evens, Martha, John Spitkovsky, Patrick Boyle, Joel A. Michael, and Allen A. Rovick. 1993. Synthesizing Tutorial Dialogues. *Proceedings of CogSci '93.* Boulder, Colorado, June. 137-142.

Faure, David, and Claire Nédellec. 1998. A Corpus-Based Conceptual Clustering Method for Verb Frames and Ontology Acquisition. *Proc. LREC Workshop on Adapting Lexical and Corpus Resources to Sublanguage Applications*, Granada, May. 5-12.

Faure, David and Claire Nédellec. 1999. Knowledge Acquisition of Predicate Argument Structures from Technical Texts Using Machine Learning; The System ASIUM. *Proc. Current Development of Knowledge Acquisition, EKAW.*

Faure, David, Claire Nédellec, and T. Poibeau. 2000. First Experiments of Using Semantic Knowledge Learned by ASIUM for Information Extraction task Using INTEX, *Ontology Learning Workshop ECAI-2000*, 7-12.

Fillmore, Charles J. 1968. The Case for Case, in Emmon Bach, and R. Harms (eds.), *Universals in Linguistic Theory*, Holt, Rinehart, & Winston, New York. 1-88.

Freedman, Reva K. (1996). *Interaction of Discourse Planning, Instructional Planning and Dialogue Management in an Interactive Tutoring System.* Doctoral dissertation, Northwestern University, Evanston, IL.

Grimshaw, Jane. 1990. *Verb Arguments*, MIT Press, Cambridge, MA, 1990.

Grimshaw, Jane, and Ray Jackendoff. 1985. *Report to the NSF on Grant IST-81-20403*, Department of Linguistics, Brandeis University.

Grishman, Ralph, Lynette Hirschman, and Ngo T. Nhan.1986. Discovery Procedures for Sublanguage Selectional Patterns. *Computational Linguistics*, 12(3) 205-215.

Hindle, Donald. 1990. Noun Classification from Predicate Argument Structures. *Proc. ACL* , 268-275.

Mayer, Glenn. 1992. *Creating a Structured Knowledge Base by Parsing Natural Language Text*. Ph.D. Dissertation, Department of Computer Science, Illinois Institute of Technology, Chicago.

Mayer, Glenn, Chiye Yamamoto, Martha Evens, and Joel A. Michael. 1989. Constructing a Knowledge Base from a Natural Language Text. *Proceedings of the 2nd Annual IEEE Symposium on Computer Based Medical Systems*, Minneapolis, 98-107.

Nutter, J. Terry, Edward A. Fox, and Martha Evens. 1990. Building a Lexicon from Machine-Readable Dictionaries for Improved Information Retrieval, *Literary and Linguistic Computing*, 5 (2) 129-138.

Pretschner, Alexander and Susan Gauch. 1999. Ontology Based Personalized Search. *IEEE Tools with Artificial Intelligence*, Chicago, IL. 391-398.

Ravin, Yael. 1990. Disambiguating and Interpreting Verb Definitions. *Proc ACL*, 260-267.

Smith, Raoul N. 1985. Conceptual Primitives in the English Lexicon, *Papers in Linguistics*, 18 (1) 99-137.

Werner, Oswald. 1978. The Synthetic Informant Model: On the Simulation of Large Lexical Semantic Fields. In M.D. Loflin and J. Silverberg, eds., *Discourse and Inference in Cognitive Anthropology.* Mouton, the Hague. 45-82.
Werner, Oswald and G. Mark Schoepfle. 1987. *Syste--matic Fieldwork.* Sage Publications, Newbury Park CA.
Werner, Oswald, Martha Austin, and Kenneth Y. Begishe. 1969/1981. *Anatomical Atlas of Navajo.* U.S. Department of Public Health, Washington. DC
Woods, William. 1975. What's in a Link? Foundations of Semantic Networks. Daniel Bobrow and Alan Collins, eds. *Representation and Understanding.*Academic Press, New York, NY.

Table 2. Some Sample Case Frames

Verb	Syntactic	Semantic	Selectional Restriction
anticipate	subject	experiencer	human, tutor
	that-clause	theme	action, fact
Ex. I was anticipating that IS was more important. K71-st-26-1			
arrange	subject	agent	human, tutor
	that-clause	subject	action
Ex. On two different heartbeats I arrange that the filling of the ventricle is exactly the same. K58			
belong	subject	theme	parameter
	advP	at-value	value
Ex. CO etc. can take any value that gets MAP back where it belongs. K70-tu-21-1			
combine1	subject	agent	human
	direct object	theme	entity
Ex. You have combined two very important but quite separate things.			
combine2	subject	theme	chemical
	with-NP	co-theme	chemical
Ex. Some of it combines with the bicarb to produce CO2 that is blown off. G1			
correlate	subject	theme	par.change
	with-NP	co-theme	par.change
Ex. A decrease in resistance would allow an increase in flow which usually seems to correlate with an increase in pressure. K59-st-39-1			

Table 3. Analyzed Words

Verbs (Including Phrasal Verbs)

GO UP, TAKE PLACE, AFFECT, CAUSE, CHANGE, CONTROL, CORRECT, DECREASE, DESCRIBE, DETERMINE, DILATE, END, ENTER, FALL, FINISH, FORGET, GET, GO, HAPPEN, HAVE, HOLD, INCREASE, INFLUENCE, IS, KNOW, LEAVE, LOOK, MALFUNCTION, OCCUR, PLAY, PREDICT, PRESS, PROVIDE, PUT, READ, REGULATE, REMAIN, REPRESENT, RESTORE, RETURN, SIGNAL, TALK, TELL, THINK, TYPE, UNCHANGE, UNDERSTAND, UNDO, VARY, WANT

Nouns (Including Phrasal Nouns)

ARTERIAL PRESSURE, BAROCEPTOR REFLEX, BLOOD VOLUME, DIRECT EFFECT, DIRECT RESPONSE, HEART RATE, INOTROPIC STATE, PREDICTION, NEURAL CHANGES, PACEMAKER MALFUNCTION, PHYSIOLOGICAL SYSTEM, RECEPTOR-MEDIATED REFLEXES, REFLEX EFFECTS, SOCIAL SECURITY NUMBER, STEADY STATE, SYMPATHETIC RESPONSE, VENOUS SYSATEM (note misspelling), VENOUS SYSTEM, +, - AFTERLOAD, ANSWER, BARORECEPTOR, BODY, BOOKLET, BRAIN, BV, CA++, CBVOLUME, CHANGE, CO, CVP, CVVOLUME, DECREASE, DETERMINANT, DIRECTION, DR, ENTER (the enter key), EVENT, FALL, HEART, HR, I, INCREASE, INPUT, IS, LIMIT, MAP, ME,, NAME, NORMAL, OK, PACEMAKER, PAGE, PERMEABILITY, PHASE, POST-TEST, PREDICTION, QUESTION, RECEPTOR, REFLEX, ROLE, RR, SEQUENCE, SS, SV, TABLE, TPR, TURN, VARIABLE, VESSEL, WE, XXX-XX-XXXX (coded social security number), XXXXXXX (coded name), YOU