1988(9)

An Expert/Novice Learning System with Application to Infectious Disease

Micheal Brody
Assistant Professor, Department of Science Education

Danny Kopec
Assistant Professor, Department of Computer Science

Larry Latour
Assistant Professor, Department of Computer Science

University of Maine,
Orono, Maine 04469-0122

Abstract

This paper describes a joint research project using expert system and cognitive mapping techniques to represent basic knowledge of sexually transmitted diseases in college populations. A prototype software system is described:

1) to aid experts and novices in the construction of a knowledge base for infectious diseases, and

2) to aid and advise the novice towards understanding the complex field of infectious diseases.

Paradigms are combined from three distinct fields of study: from cognitive science, to identify, analyze and organize the domain specific knowledge, from artificial intelligence to develop an automated tutoring system for construction of the knowledge base, and from software engineering to develop a well-designed modular tool set.

Subject Categories: expert systems, knowledge representation, knowledge bases, knowledge acquisition, cognitive science, problem solving, intelligent tutoring, intelligent interfaces, general software engineering.

EXPERT/NOVICE LEARNING SYSTEMS WITH APPLICATION TO INFECTIOUS DISEASES

Michael Brody, Danny Kopec and Larry Latour
University of Maine

The threat of AIDS is widely recognized as an important educational problem because no remedy has been found to effectively control its spead throughout society. Of particular interest is the college age population which has the highest incidence of AIDS in the United States. Although medical and health education professionals have made serious efforts to distinguish fact from myth, we are not aware of any intelligent software that has been developed to disseminate the critical knowledge relevant to this worldwide epidemic. Since sexually transmitted diseases (STD's) as a group and AIDS in particular are a personal problem which few people feel comfortable discussing. Use of the computer can provide an opportunity for learning about STD's while insuring anonymous interaction with an intelligent system.

The principle focus of the research described here is in three subdomains of AI, namely (1) knowledge representation, (2) learning, and (3) problem solving, with application to the identification, control and treatment of STD's. Fundamental AI issues which we address are:

1. **The effective extraction, representation and modification of expert/novice medical knowledge**

2. **The development of a computer system to facilitate growth and learning in the knowledge base**

3. **The problem solving techniques used by learners to apply this knowledge base within the context of real world needs for infectious disease control.**

Thus our overall goal is to use the knowledge and experience acquired through the study of the above issues to develop, evaluate and effectively employ an intelligent tutoring system to educate college age populations about STD's.

This paper describes prior work in artificial intelligence, concepts maps as a knowledge base respresentation, and software tools for the graphical representation and intelligent construction of concept maps. Finally we consider an intelligent tutoring system for STD's using the concept map knowledge base.

I Artificial Intelligence and Knowledge Representation

AI has achieved much in its brief history, especially in the development of cost-effective expert systems with diverse applications. Some examples are MYCIN for medical diagnosis and therapy with its generalization to EMYCIN, DENDRAL for the analysis of unknown chemical structures, and its extension to Meta-DENDRAL, and RI for VAX system configuration. However, as E.A. Feigenbaum has often stated (1983), "the "bottleneck" which remains, is the transfer of domain-specialist or expert knowledge to machine."

Considerable effort has been expended towards study in the area of cognitive skills, their acquisition, and their modelling with such systems as GPS (Newell and Simon, 1972), EPAM (Feigenbaum and Simon, 1962), HAM (Anderson and Bower, 1973) and ACT (Anderson, 1976), but the principle problems remain: (1) What is advice (expert knowledge) and how should it be represented, (2) How can advice be most effectively translated into executable procedures, and (3) How can previous erroneous behavior be modified and corrected.

It has been determined that there are essentially three stages of skill acquisition: encoding, proceduralization and composition. Yet the encoding, collating, organization, storage and retrieval of expert knowledge continuously perplexes those who study it, for it tends not to be constructed in a formal top-down manner, involves non-linear representational schemata, and is pattern-based. Experts seem to be able to perform *cognitive leaps* based upon acquired know-how or intuition. This enables them to draw upon *heuristics* which are not within the reach of novices, even when these novices attempt problem solving by text book methods. However, it is accepted that experts also tend to be poor in articulating precisely just how they are able to solve difficult problems within their domains of expertise.

II Concept Maps as Knowledge Representation

Research in knowledge representation has come from three distinct fields: AI, linquistics, and cognitive psychology. Work in the area of linquistics provides our work with the basic premise that knowledge is made up of concepts plus their relations to other concepts and that these relations can be reduced to a relatively small set.

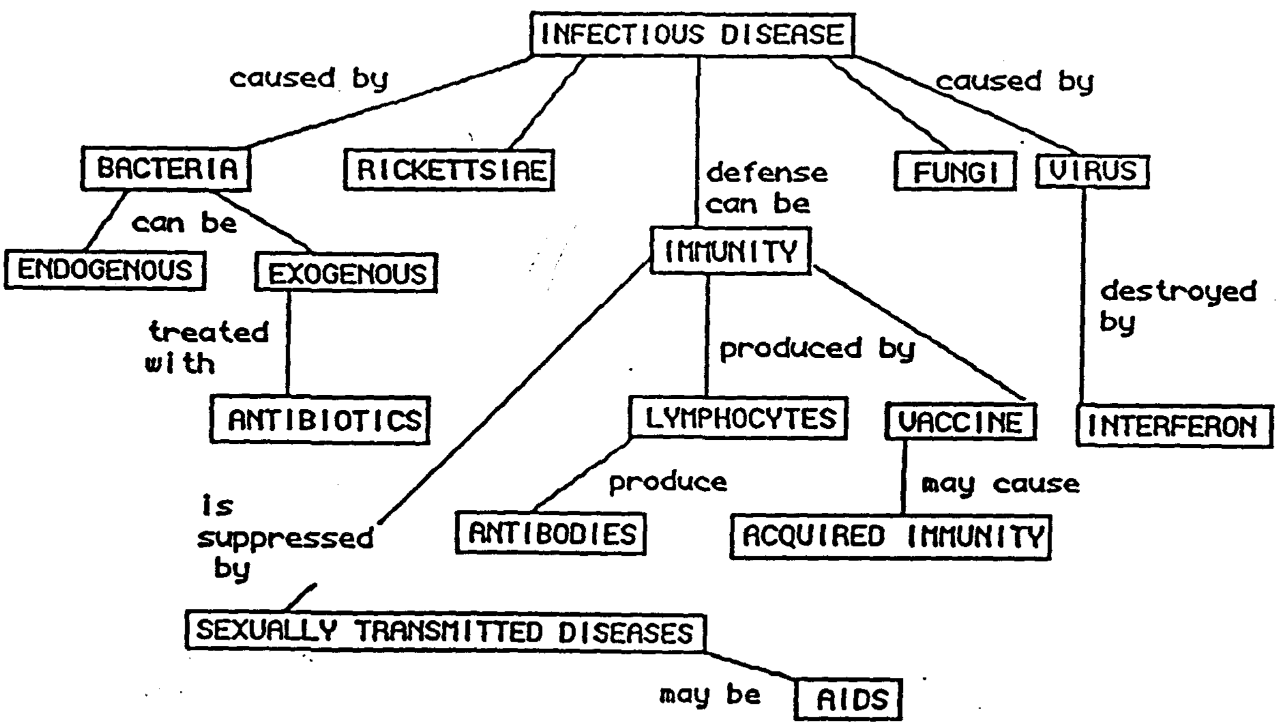
Artificial intelligence currently maintains as one of its central problems the representation of knowledge. One approach has been to combine small subsets of knowledge into larger complex networks connected by relations. The resulting representations of knowledge can be viewed as semantic networks: collections of concepts multiply-connected in complex networks.

Quite a few researchers in cognitive science have adopted the methods described above to help illustrate thinking and learning and have had students and researchers construct concept maps of their work in various disciplines.

We approach our study via a graphical knowledge representation methodology which is at present being successfully applied to learning in the science classroom (Novak and Gowin, 1984) and to informal adult learning situations (Brody, 1985). This approach is known as Concept Mapping and up to now has been somewhat independent of the AI domain intended here for its application. However, this technique's roots in cognitive psychology and learning suggest its suitability for the problem we are investigating.

Concept mapping provides a representational framework for subject/expert/novice knowledge in any domain. These concept maps can then be joined into rather complex semantic networks much as Collins and Loftus (1975) modified and stretched networks.

In this study concept maps are used to help extract expert and novice knowledge, identify key concepts involved in STD's, and represent them in easily understandable formats for both future teaching and learning in a computer-based intelligent tutoring system. They are intended to represent meaningful relationships between concepts in the form of propositions, i.e., two or more concept labels linked by words in a semantic unit. Concept maps work to make clear to both subjects and researchers the relatively small number of ideas they must focus on in any learning task. A concept map provides a visual road map showing the pathways we may take to connect meanings of concepts in propositions. Concept maps are also hierarchical; that is, the more general, more inclusive concepts are at the top of the map, with progressively more specific, less inclusive concepts arranged below them. Consider the concept map below:



This and similar concept maps have been and currently are being constructed by groups of experts in the medical field to collectively represent the knowledge base of experts' understanding of infectious diseases. In addition, through clinical interviews with students and the analysis of their comments, concept maps representing novice knowledge have been and are currently being constructed. By novice knowledge we mean knowledge which has partially misconceived and/or missing information. These two sources form the basis of a semantic network around which an intelligent tutoring system is being constructed. The net forms a map through which the user can approach a particular problem concerning STD's. Three issues central to this phase of our research are:

1. **differences between the subject specific maps produced by novices and those produced by experts**
2. **differences in the <u>techniques</u> of construction of such maps by experts and novices**
3. **differences in how experts and novices add to, delete or modify knowledge structures based on new information.**

III Automated Tools For Knowledge Base Construction

There are a number of ways in which computer automated tools can aid in the development of concept maps. These tools can be partitioned into two groups: **a. domain independent tools b. domain dependent tools** Further, each of these classes of tools can be partitioned into

1. "Standard" tools that help to construct and manipulate concept maps in user-directed ways, and
2. "Intelligent" tools which apply domain dependent and/or independent knowledge to aid the user in making decisions in the construction process.

We have in place a standard domain independent tool, SeeGraph (Latour & Johnson, 1987), a subsystem for storage, retrieval, and graphical representation of semantic networks. SeeGraph is discussed in the following section. In addition, we are in the process of developing the "Intelligent" tutoring system IntMap, for the construction of concept maps which employs the SeeGraph database. The system uses both domain independent knowledge about the structure of concept maps and domain dependent knowledge about the semantics of concept map knowledge to guide the user through the mapping process.

SeeGraph: A tool for the realization of Semantic Networks

Our software development group has constructed the prototype SEE, a Student's Educational Environment (Latour,1986) providing a set of tools to support investigations into the selection of a proper knowledge base structure. Incorporated in SEE is SeeGraph, a subsystem for storage, retrieval, and graphical representation of semantic networks. SeeGraph, simply put, is a network database with a graphical query language allowing users to add and delete nodes and arcs in multiple windows. To facilitate movement through the graph, operations for changing, magnifying, reducing, and more specifically focussing the user's current view are available. In addition, the user can create and manipulate text attributes of the nodes and arcs.

There are several conceptual advantages to representing semantic networks in a graphical fashion. The relation of one node to another through a connecting arc is unmistakable but limiting. SeeGraph gives the user a powerful spatial dimension in which to give meaning to the network at hand. Placement of objects in groups to show a sharing of common characteristics, use of node size to indicate hierarchy within groups, and thickening of arcs to show path precedence are all special effects which allow the user to give a feel for the logical representation of the graph and how to best use it as an interface to the underlying data objects.

SeeGraph views a semantic network as existing in an underlying universe which is for all intents and purposes unbounded. Each active user screen (many may be active simultaneously) is a window onto this universe. All SeeGraph operations are carried out through the use of the mouse and mouse activated menus.

A number of enhancements are currently at the specification stage for future SeeGraph versions, including: (1) an intelligent graphics interface to aid in the spatial placement and organization of nodes and arcs, (2) a 3-dimensional semantic network interface, and (3) formal tools to overlay domain dependent semantics on the domain independent graphs of SeeGraph. These include a parser to convert outlines into concept maps, and the aformentioned intelligent tutoring subsystem to aid the user in constructing domain specific semantic networks.

We have implemented a SeeGraph prototype in Ada on a network consisting of a Vax/780 connected to Vaxstation II and PC AT and XT workstations. In addition the prototype is being ported to SUN workstations and the same is being considered for the MAC II.

IV Design of Intelligent Tutoring System

The quality of learning is a pervasive issue in education and in particular in science education. Recent advances in this area have taken a particularly cognitive perspective (Carey, 1986) and many of the major findings have implications for education in general. This perspective begins with the principle that what learners already know about new information will influence their subsequent learning of that material (Champagne and Klopfer, 1984). The construction of the software system proposed in this project to represent knowledge bases of both experts and novices will be designed to allow the researchers to assess learners' prior knowledge and how that knowledge changes over time. Our study will entail human subject protocol analysis as well as interview techniques involving the computer system.

The process of restructuring knowledge has been investigated extensively in a number of areas including chess and physical science (Chi, Glasser and Ree, 1982) and it is our belief that this can help us understand learning in a new context. In particular we will be looking for exemplar novice-expert shifts in the restructuring of STD knowledge. The study and development of intelligent tutors has in recent years become a very active field of research (Anderson, 1987; Sleeman, 1982). This involves endeavors beyond the construction of expert systems and is predicated on models of cognition by Anderson (1982) and others which have led to a better understanding of skill acquisition and problem solving behavior. Our tutor, when completed, will lead the novice user through the relevant facts related to STD's to enable appropriate action.

Some examples of existing systems which serve as models of cognitive activity are Newell and Simons GPS (1972) and Anderson's ACT (1976). GPS employs means-ends analysis and difference tables to develop a *problem behavior graph* resulting in an explicit, operational, and sufficient model of human problem solving. The ACT system is able to perform a number of cognitive tasks. It includes short and long term memory components, as well as a programmable production system. Memory is modeled by a *propositional associative network* comprised of nodes representing concepts and arcs representing the relationships between them. It performs *knowledge compilation* whereby postulates are built into productions in the form of if-then rules which operate on the active components of the semantic network.

Our prototype system serves as a testbed for investigation into some of the above problem solving paradigms, as well as heuristic, rule based, goal-directed, non-linear, inductive and deductive problem solving. Its ability to cope with problems in STD detection, control and abetment will provide insight into plausible techniques for incorporating new information into existing databases. Our study of concept mapping Techniques and tools provides generalizable conclusions concerning how experts and novices learn about STDs.

Summary and Conclusions

The result of this investigation contribute to the design and implementation of an intelligent tutoring system to facilitate meaningful problem-solving. In particular, our software tools to construct concept maps of experts and novices provide an underlying knowledge base of STD's for the design of such a system. Our system is comprised of a collection of complex networks of concepts connected by relations; its performance will be carefully scrutinized in terms of its validity and reliability as a representation of expert/novice domain-specific knowledge.

Anderson, J.R., and Bower, G.H. (1973) Human Associative Memory Washington, D.C.: Winston.

Anderson, J.R. (1976) Language, Memory, and Thought. Hillsdale, N.J: Lawrence Erlbaum.

Anderson, J.R. (1982) Cognitive Skills and Their Acquisition Hillsdale, N.J.: Lawrence Erlbaum.

Anderson, J.R. (1982) Acquisition of Cognitive Skill. Psychological Review, vol. 8 No. 4, pp 369-406.

Anderson, J.R. (in press). Production systems, learning, and tutoring. In D. Klahr, P. Langley, and R.Neches (eds.), Self-modifying production Systems: Models of Learning and Development. Cambridge, MA: Bradford Books/MIT Press.

Brody, M.J. (1985) Transfer of Knowledge Within the Lake Ontario Sport Fishery, Oceans 85, Marine Technology Society, Washington, DC.

Carey S., Cognitive Science and Science Education. American Psychologist, vol. 41, no 10, pp 1123-1130.

Champagne, AB. and Klopfer, LE. (1984). Research in science education; the cognitive psychology perspective. In D. Holdzkom and PD Lutz (Eds) Research within Reach: Science Education, Washington DC: National Science Teachers Association.

Chi, M., Glasser, R., and Rees, E., (1982) Expertise in Problem Solving. In R. Sternberg (Ed), Advances in the Psychology of Human Intelligence, (vol. 1), Hillsdale, NJ: Lawrence Erlbaum.

Collins, A., and Loftus, E.(1975) A Spreading Activation Theory of Semantic Processing, Psychol Rev, 82, 407-28.

Feigenbaum, E.A. (1983) Knowledge engineering: the applied side. in Intelligent Systems: the unprecedented opportunity, eds. J.E. Hayes and D. Michie, Ellis Horwood, Chichester, England.

Hayes-Roth, B., and Hayes-Roth, F. (1978) Cognitive processes in planning. Rep. No. R-2366-ONR, Rand Corp., Santa Monica, Calif.

Hayes-Roth, B. (1980) Human planning processes. Rand Corp., Santa Monica, Calif.

Kopec, D. and Michie, D. (1983) Mismatch between machine representations and human concepts: dangers and remedies. Report to the EEC, subprogram FAST, Brussels, Belgium.

Latour, L. (1986) A programming environment for learning- SEE: a student's educational environment, Proceedings of the IEEE Conference on Ada Applications and Environments, Miami Beach, FL.

Latour, L. & B. Johnson (1987) SeeGraph: a description of a graphic based knowledge representation system, Internal Report, University of Maine Computer Science Dept. Orono, ME.

Newell, A. and Simon, H.A., (1972) Human Problem Solving, New Jersey: Prentice Hall.

Novak J., and D. Bob Gowin, (1984) Learning How to Learn. Cambridge University Press,

Cambridge, England.
Resnick, L.B. (1983) Mathematics and Science Learning: a New Conception. Science, 220, pp 477-8.
Sleeman, D.J. and Brown, J.S. (eds.,1982) Intelligent Tutoring Systems. London: Academic Press.