Odyssey: A Program to Access Medical Knowledge

ChinLi Fan, Michael J. Lincoln, Peter J. Haug, Charles W. Turner, and Homer R. Warner

Department of Medical Informatics, School of Medicine, University of Utah

Abstract

Odyssey is part of a software package of computing tools designed to aid the third year medical students at the University of Utah while they are on the Internal Medicine clerkship. The Odyssey program is designed to facilitate student access to select knowledge bases of high clinical relevance. Access to the Odyssey program is from Macintosh computers located on the medical wards. The knowledge bases available to the students at the present time include a literature reference database and the HELP system's collection of decision-making frames.

1. Introduction

Odyssey is a program developed to assist in teaching third year student clerks on the Internal Medicine service at the University of Utah. The project began when the medical staff identified several educational problems. The first problem is that students fail to read the medical literature. In this they differ very little from many physicians already in practice. Williamson surveyed U.S. primary practitioners and found that 81% thought that finding good literature took an inordinate effort. Of these, 65% felt it was impractical to try to keep up with the medical literature.[1] Students are frustrated by the convoluted language structures one must learn to search on-line databases unassisted and the time demands of leaving the ward to use the library. And when they do search the medical literature, the retrieved articles are often irrelevant or of poor quality.[2]

The second problem was that third year clerks had poor decision making skills. Most students learn decision making informally by acting like apprentices. Formal teaching and literature about decision making are rarely available. Many groups have been interested in teaching decision making using various computer-assisted Decision Support Technologies (DSTs). Unfortunately, many DSTs are only used as research tools or have fallen into disuse, even in the DST developer's institution.[3] When available, their use may require extensive computer expertise. Another mechanism is needed to teach students about decisions.

Finally, the patient care area is remote from the library. To to promote maximum learning efficiency, patient care experiences must be immediately reinforced with reading and reflection. But students must extricate themselves from the wards and travel to a storehouse of knowledge, the library. Critical patient decisions are often made long before students have a chance to learn. Unless the student can "strike when the iron is hot," the impact of even the best library or knowledge base is minimal. Students fail to use learning resources properly because they cannot access them at critically relevant times.

Computer programs such as Health Evaluation through Logical Processing (HELP)[4] and Quick Medical Reference (QMR)[5] are efforts to provide powerful solutions to these problems. They succeed, in part. A student can enter a patient's findings and produce a list of possible diagnoses. But neither program permits students to enter a disease or pathophysiologic concept and access information about that disease or concept. Students who wish to do this must manually search multiple knowledge bases, including the medical literature and textbooks. It is often impossible to do this in a timely, relevant, and efficient manner.

2. The Odyssey program

Odyssey is a computer program that runs on Macintosh Plus computers located on the medicine wards at the University of Utah. It is available at all hours to serve as a database manager, allowing students access to several knowledge bases. The search terms entered by the student can be natural, imprecise words. A powerful, transparent system known as "relations" uses these nonspecific terms to perform powerful searches of several knowledge bases, such as literature references and HELP decisions.

The literature knowledge base consists of approximately 1500 journal articles and monographs relevant to Internal Medicine. Experts selected these references because they were interesting, methodologically sound, and clinically relevant to third year students. The HELP knowledge base is downloaded by Odyssey's communication routines from the hospital Tandem mainframe computer. Odyssey allows students faced with immediate clinical problems to examine relevant literature and HELP decisions.

3. Relations

The program allows the creation of special structures we call "relations." Relations connect key words in a hierarchical fashion to allow users access to very specific medical concepts. An example of a relation is pneumonia. "Pneumonia" can be identified using the key words "lung" and "infection." This is because the parent relation called "infection" has several children, including "pneumonia," "bronchitis," "sinusitis," etc. As relations are traversed from the top to bottom of the tree the concepts expressed go from the general to the specific.

0195-4210/0000/0488$01.00 © 1987 SCAMC, Inc.

Imagine a patient admitted at 3:00 A.M. with fever, cough, and shortness of breath. The student considers a diagnosis of pneumonia. He vaguely remembers some lectures about pneumonias from the second year of medical school. Sitting down at the Macintosh, he opens the Odyssey program. In a search window, he enters the key words "lung" and "infection." The search window (see Figure 1) contains three

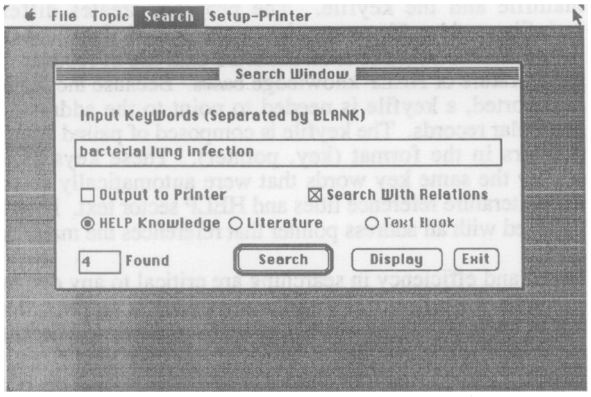

Figure 1. Search Window

Macintosh buttons, labeled "HELP decisions," "Literature," and "Textbook." The student clicks the "Literature" button to search the literature knowledge base and then clicks the "Search" button. Two seconds later a second window is displayed, listing highly relevant references to articles, with attached abstracts. He notes that some concern viral and parasitic lung infections, so he decides to narrow the search. Going back to the search window, he adds "bacteria" as a third key word and repeats the search. Now fewer references are retrieved, all dealing with bacterial lung infections (see Figure 2). He now clicks a "Retrieval" button beside two of the most promising references. This prints a hard copy of the abstracts. He then returns to the search window and clicks on the "HELP Decisions" button and the "Search" button. Odyssey repeats the search with the same key words, but this time extracts relevant titles for HELP frames that contain decisions about bacterial lung infections (see Figure 3). After reviewing the titles the student can now display any HELP frame and review the disease findings used by the frame and the logic by which it makes decisions.

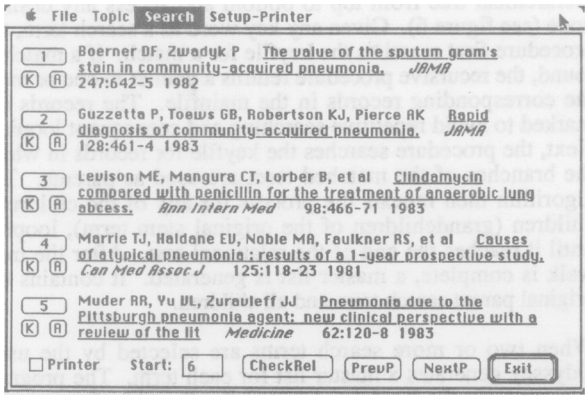

Figure 2. Literature references

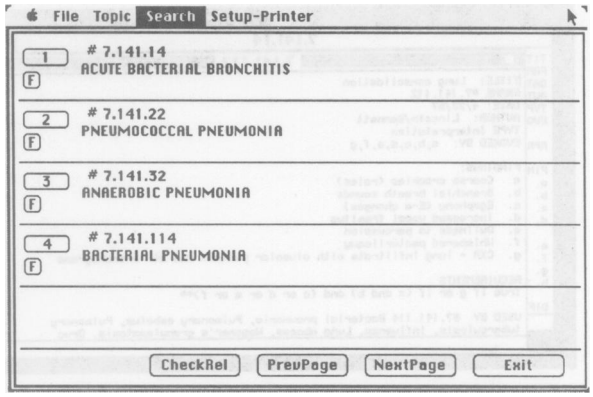

Figure 3. HELP knowledge frames

The student is about to use one of the most powerful features of the Odyssey system, the ability to access HELP decision frames. HELP frames come in two styles, "clusters" and "sectors". Clusters are decision frames that usually describe pathophysiologic concepts. Sectors are decision frames that usually describe disease entities. Clusters act as subroutines that may be called by other clusters or sectors. For example, the cluster "Pulmonary venous hypertension" searches the patient record for findings of dyspnea, orthopnea, rales, edema on chest radiograph, elevated wedge pressure, etc. The presence or absence of these findings is scored and a diagnostic decision is made regarding the presence or absence of the concept "pulmonary venous hypertension." The existence of "pulmonary venous hypertension" is used as a diagnostic subroutine by other clusters ("Congestive heart failure") or sectors ("Hypertensive cardiomyopathy"). Clusters save time writing HELP frames and express useful teaching concepts.

The advantage to the student is that Odyssey provides a mechanism to translate the programming language used in the logical structure of a HELP frame (sector or cluster) into plain English. The student can examine this natural language translation of the frame's decision logic in an attempt to learn how the medical expert that wrote the frame made decisions about cases similar to the student's. In this frame translation, imbedded clusters are marked by a special character "@" (see Figures 4 & 5). The student can select such a line of text with

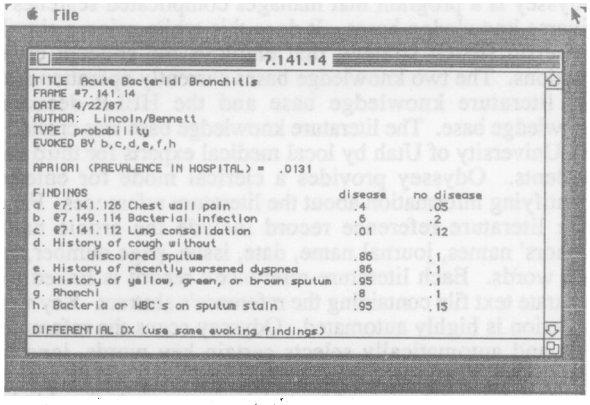

Figure 4. A frame with three imbedded clusters

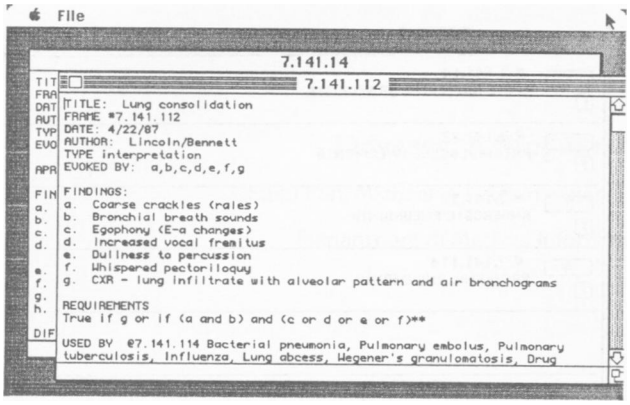

Figure 5. An imbedded cluster

the mouse, activating a hypertext mechanism that superimposes a second window containing a translation of the cluster. This cluster translation may contain other nested clusters that can also be opened. The student can toggle between these different windows, comparing the structure and function of the nested frames.

Relations provide several advantages to the Odyssey system. The relation structure can be built by the medical expert separately from the construction of the actual knowledge bases. When the knowledge bases are updated, the relations need not necessarily change. The power of relations allows users to search knowledge bases using common English words to express complicated concepts. Odyssey can gives feedback on searches within seconds, listing the reference titles and abstracts. The user can use this feedback to modify search terms until the desired knowledge is found.

4. Program Implementation

The Odyssey program is written in Aztec C and implemented on a Macintosh Plus computer system. This system uses a Motorola 68000 processor, 1 megabyte of main memory, 800 kilobyte internal floppy disk drive, 20 megabyte external hard disk, and Imagewriter II printer. Odyssey downloads HELP frame text through a communications link with the Tandem computer on which HELP resides. After downloading frames, Odyssey can create the key word list automatically from the HELP frame titles.

Odyssey is a program that manages complicated searches of diverse knowledge bases. It does this while using a simple, powerful search strategy that employs the concept called relations. The two knowledge bases currently operational are the literature knowledge base and the HELP decisions knowledge base. The literature knowledge base was created at the University of Utah by local medical experts for third year students. Odyssey provides a clerical mode for entering identifying information about the literature references. Fields in a literature reference record include the article name, authors' names, journal name, date, issue, page number, and key words. Each literature reference's record is linked to a separate text file containing the reference's abstract. Key word selection is highly automated. Odyssey scans the reference's title and automatically selects certain key words, ignoring participles, conjunctions, and other irrelevant terms. For an article entitled "Therapy of bacterial pneumonia in the elderly," the program would supply the key words "therapy," "bacterial," "pneumonia," and "elderly." If desired, these key words can be edited. The HELP knowledge base is created by downloading HELP frame titles from the Tandem computer. Sector text and sector text modifiers (HELP data dictionary terms describing the frame) automatically supply key words that may also be edited if desired. The HELP frame title and associated key words are stored as text fields within a HELP frame record structure.

Two file formats are used in the Odyssey program, the mainfile and the keyfile. The program creates different mainfiles and keyfiles every time a new knowledge database is selected. The mainfile is an unsorted data file of records from the literature or HELP knowledge bases. Because the mainfile is unsorted, a keyfile is needed to point to the addresses of particular records. The keyfile is composed of paired keys and pointers in the format (key, pointer). These keys are the simply the same key words that were automatically selected from literature reference titles and HELP sector text. Each key is paired with an address pointer that references the mainfile.

Speed and efficiency in searching are critical to any database. The maximum speed of an Odyssey search is on the order of $\log_2 N$ (where N is the number of items in the array) because a binary search algorithm is used. In certain instances, the search strategy may be even faster because the keyfile is ordered alphabetically. Before the binary search algorithm is employed, an algorithm we call "address segmentation" is used to select one of 26 alphabetical search ranges. For instance, a key term "aneurysm" would only be searched for under the "A's." After an alphabetic range is selected, the binary algorithm searches for an exact match with the key word. When a match is found, a linear search algorithm comes into play to examine the keyfile in either side of the exact match, looking for closely related terms. As an example, "aneurysmal" and "aneurysms" would cluster near "aneurysm" in the keyfile.

The relation file created by the medical expert is also referenced by a mainfile and keyfile. The relation mainfile is an unsorted file containing many relation records. The text fields of a relation record contain strings naming the trunk (parent) and branches (children) of a relation. A single relation record can contain up to 5 synonyms for the trunk and 25 branches. As an example, the record holding information about the relation "infection" contains the synonyms "infectious," "infected," and "germs." It also contains the branches "viral," "bacterial," "parasitic," etc. Branches may be terminators or full-fledged parent relations with their own children. A relational keyfile is needed to locate mainfile records. Odyssey uses a recursive procedure to traverse a hierarchical tree from top to bottom and access any desired node (see figure 6). Given any key word as a search term, the procedure first searches the keyfile for a match. If a match is found, the recursive procedure returns a pointer to one or more the corresponding records in the mainfile. The records are marked to avoid repetitive searching and consequent looping. Next, the procedure searches the keyfile for records in which the branches of the matched term are used as parents. The algorithm then repeats the process for any of the children's children (grandchildren of the original stem term), looping until it reaches the end of the relational tree. After the tree-walk is complete, a master list is generated. It contains the original parent search term and all children.

When two or more search terms are selected by the user, Odyssey generates a master list for each term. The program then employs the Boolean operators "and" and "or" to start the search. First, all children within a master list are "or'd" with the stem term. Odyssey then "ands" each of the master lists to

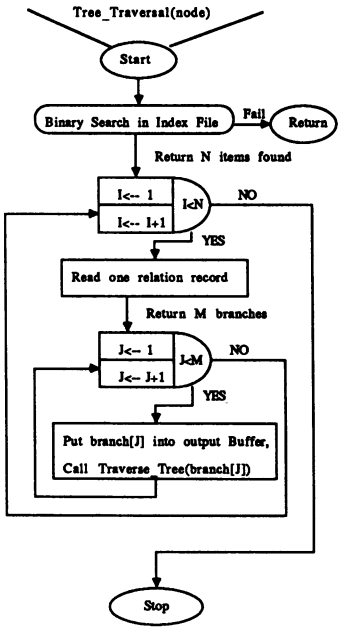

Figure 6. Tree_traversal flowchart

narrow the search specifications. The resulting specifications are used to search the knowledge bases. A example involves a search on the key words "heart" and "infection." "Heart" is the parent of "endocardium," "myocardium," "pericardium," "valves," etc. "Infection" is the parent of "bacterial," "viral," "parasitic," etc. These branches (children) are treated as alternatives (Boolean 'or') to the stem key word during the search. Odyssey employs the Boolean operator 'and' to compare the stem key word "heart" (and children) with the stem key word "infection" (and children). This search will retrieve literature references to subacute bacterial endocarditis, viral cardiomyopathies, and a few other heart infections. Because a relations exists that shows infection as a child of inflammation, a broader search on "heart" and "inflammation," will find these original references and some new references discussing alcoholic myocarditis and other non-infectious heart inflammations.

Help frames contain detailed information on the sensitivity, specificity, evoking strength and frequency of various findings as easily understandable text documents. They also contain the prior probabilities and scoring algorithms used to make each of these HELP frame decisions. The text format allows medical experts to write decision frames in an easily understood syntax and also facilitates debugging of the logic. A translator program is designed to make the process of coding the text frames into programming language for the Tandem semi-automatic and interactive with PTXT, the system dictionary. These textual HELP frames are also useful because the students can access them from within Odyssey using powerful hypertext features. The creator of the text version of a HELP frame must identify any item in a frame which is derived from a cluster frame by using bold type. When the student selects a boldface string while examining a HELP frame, a pointer is passed to Odyssey, which opens the cluster HELP frame and displays the resulting text file in another window overlying the original window.

5. Summary

Odyssey is part of a larger effort at the University of Utah School of Medicine to augment student learning. The Department of Medical Informatics is studying the effect of Odyssey and several other computer-assisted instructional aids on learning in the third year of medical school under a grant from the National Library of Medicine. The University of Utah School of Medicine admits 100 students to each class. In the third year, these students are randomly allocated to various hospitals in Salt Lake, including the University Hospital, LDS Hospital, and the Veteran's Hospital. We are allowing students at these hospitals access to a variety of tools on the Macintosh, including the Odyssey program. A team of education experts and statisticians will study the effect of these computer tools on grades, performance on standardized exams, decision-making ability, staff evaluations of student performance.

This work was supported in part by a grant # 1 R01 LM 04604-01 to the University of Utah from the National Library of Medicine.

References

1 Warner HR, Williamson JW, Haug PD, Lincoln ML, Turner CW, Whitman N, et al. Access to Knowledge Through Models, National Library of Medicine Grant # 1RG1 LM 04604-01

2 Williamson JW, Goldschmidt PG, Colton T. The quality of medical literature: an analysis of validation assessments. In: Bailar JC, Mosteller F, eds, Medical uses of statistics, Waltham, Massachusetts: New England Journal of Medicine Books, 1986.

3 Lundsgaarde HP. Evaluating medical expert systems (in press)

4 Pryor TA, Gardner RM, Gardner PD, Warner HR. The HELP system. J Med Syst. 7: 87-101, 1983

5 Miller RA, Masarie FE, Myers JD. Quick medical reference (QMR) for diagnostic assistance. MD Computing. 3(5): 24-48, 1986

6 Horowitz GL, Jackson JD, Bleich HL. PaperChase: self-service bibliographic retrieval. JAMA; 250(18): 2494-2499, 1983.

7 Warner HR, Wooley FR, Kane RL. Computer assisted instruction for teaching clinical decision-making. Comput. Biomed. Res 7: 564-574, 1974.