# A Case-Based Approach to Knowledge Navigation*

Kristian J. Hammond
The University of Chicago
Department of Computer Science
Chicago, 1L 60637
hammond<Ocs .uchicago. edu

Robin Burke
The University of Chicago
Department of Computer Science
Chicago, IL 60637
burke@cs .uchicago .edu

Steven L. Lytinen
Depaul University
Department of Computer Science
Chicago, IL 60617
lytinen@Blance. cs . depaul. edu

## Abstract

The AI Lab at Chicago has begun development of a new set of software agents designed to manage the flood of data colloquially called the "information superhighway". Our approach takes its lead from case-based technology [Riesbeck and Schank, 1989; Hammond, 1989; Kolodner, 1993] in that we are building systems that emphasize the use of examples over explicit queries or questions for communicating with the user.

## 1 Introduction

One of the important prospects for the national information infrastructure [Gore, 1993] is the wide-spread availability of on-line data and information services on a large scale. However, existing means of information access will not scale up to a network of this size. We believe that access to appropriate information, contacting the right source at the right time, is the most significant obstacle to making a large-scale information infrastructure work.

We have been examining two kinds of systems aimed at this problem: preference browsers (called FIND ME systems) that help users navigate through information spaces and either find or construct responses to fit their needs and information retrieval engines (called FAQ FINDER systems) that use a natural language question-based interface to distributed information sources, specifically files organized as question/answer pairs such as FAQ files.

Two features distinguish these systems. The first is that they are derived from the case-based ideas of using *retrieve* and *adapt* as the core problem solving model. The second is that they use existing archives and databases as resources to be "mined" on demand rather than as fodder for batch processors that learn new concepts or construct new knowledge bases independently of a user. This too draws from the case-based philosophy of waiting until a problem arises to solve it. In this video, we will demonstrate both FIND ME and FAQ FINDER and examine how the techniques of case-based reasoning transfer over to the task of information retrieval.

*This work was supported in part by the Office of Naval Research under contract N0014-85-K-010

## 2 Find Me Systems

One important class of information access problem is *prrference-based browsing,* moving around in an information space, guided by a user's personal preferences. Tasks of this type include shopping for consumer items, selecting videos or television shows, and selecting personnel out, of a large pool of candidates. It is usually not possible for a user to articulate all of the factors that might contribute to a decision in such a domain. This makes these domains poor candidates for standard feature-based database search, which requires that users create a query that encodes the desired properties they seek. However, it is much easier for a user to respond to an example if given one. In making personnel decisions, it is easier to evaluate an example, to say, "This person would be OK, but I'd prefer someone more experienced," than it is to decide what kind of specific query would get the desired result. In effect, people cannot "Say what they like." but arc able to "Know it when they see it."

We have designed an approach to this problem that allows users to see and respond to examples within different parts of the search space as a means of moving within it [Hammond *et a*/., 1995]. The choice of where to go next is dependent on users' evaluation of the example they are currently examining. A consumer might say "I like this car, but I really need more luggage space." Such feedback tells the system in what direction to push the search next. Ideally, an interaction with a FIND ME system would be like a conversation with a helpful clerk,

a conversation that is always mediated through reference to a particular object: the example currently under consideration. Because these systems extract features from examples, they allow users to navigate through a space of of possibilities, finding what they want, without ever having to build explicit queries. Retrieval depends instead, on a human/computer "conversation" about specific examples that the systems offer up to the user.

### 2.1 The central ideas

One of the core ideas of the FIND ME systems is to develop user interfaces that are metaphorically linked to known artifacts. Just as the Macintosh environment makes use of the desk-top metaphor, we have constructed interfaces that are analogs to existing artifacts that aid search within a domain.

The second major insight that we bring to knowledge navigation is that we can use the differences between a presented example and a user's target to formulate a description of the target itself which is then used to access a knowledge base of further examples, an iterative process of exploring the space of examples.

Finally, we have designed the overall FIND ME notion with an eye to avoiding hierarchical search of a space. To assist the user's search, FIND ME systems have active *clerks* that constantly view a user's current preferences and suggest alternative tracks through the space that reflect the clerk's unique goal organization and retrieval strategies.

## 3 Faq Finder

We are also developing a class of systems, called FAQ FINDER systems, that use a natural language question-based interface to access distributed text information sources, specifically text files organized as question/answer pairs such as FAQ files. In using these systems, a user enters a question in natural language and the system attempts to find an information source that answers the question, and then find the closest matching question/answer pair. These systems combine three technologies: statistically based IR. engines, syntactic natural language analysis and semantic networks. In particular, they combine the SMART information retrieval system [Buckley, 1985], a natural language parser that uses on-line word sense information, and a semantic net derived from Princeton's WORDNET.

The power of our approach rises out of two features: We are using knowledge sources that have already been designed to answer the commonly asked questions in a domain and as such are more highly organized than free text. We do not need our systems to comprehend the queries they receive. They only have to identify the files that are relevant to the query and then match against the segments of text that are used to organize the files themselves (e.g. questions, section headings, key words, etc.).

The most natural kind of interface to a database of answers is the question, stated in natural language. While the general problem of understanding questions stated in natural language remains open, we believe that the simpler task of matching questions to corresponding answers is feasible and practical.

As it stands, the FAQ FINDER project is an automated question-answering system that uses the files of "Frequently-Asked Questions" (FAQs) associated with many USENET newsgroups [Hammond *ti al.,* 1995]. These files are cornpendiums of the accumulated wisdom of the newsgroup on topics that are of frequent interest. FAQ FINDER takes a user's query on any topic, attempts to find the FAQ file most likely to yield an answer, searches within that file for similar questions, and returns the given answers.

While the FAQ FINDER program is in some sense a browser, our actual goal is to free users from doing any browsing at all. Rather than forcing users to traverse a knowledge space, we are providing a system that will do this for them. A similar approach can be applied to the querying of other large textual information sources, such as procedure and documentation manuals.

## References


[Buckley, 1985] C Buckley. Implementation of the smart information retrieval system. Technical Report 85-686, Cornell University, 1985.

[Gore, 1993] Al Gore. The national information infrastructure: Agenda for action. Technical report, United States Executive Office of the Vice-President, 1993.

[Hammond *et al.,* 1995] Kristian Hammond, Robin Burke, Charles Martin, and Steven Lytinen. Faq finder: A case-based approach to knowledge navigation. In *The Proceedings of the 1995 IEEE Conference on Artificial Intelligence Applications.* IEEE, IEEE Press, February 1995.

[Hammond, 1989] Kristian Hammond. *Case-Based Planning: Viewing Planning as a Memory Task,* volume 1 of *Perspectives in Artificial Intelligence.* Academic Press, San Diego, CA, 1989.

[Kolodner, 1993] Janet Kolodner. *Case-Based Reasoning.* Morgan Kaufmann Publishers, Inc., San Mateo, California, May 1993.

[Riesbeck and Schank, 1989] Christopher K. Riesbeck and Roger C. Schank. *Inside Case-Based Reasoning.* Lawrence Erlbaum Associates, Hillsdale, NJ, 1989.