One of the important lessons learned in the construction of the DENDRAL programs and other domain specific problem solving programs is that a large amount of domain specific knowledge is necessary to guide the problem solving process effectively. Programs having a high problem solving performance in a specific task area have incorporated ad hoc approaches to knowledge representation. Programs used for experiments in general problem solving have used representations which do not lend themselves to knowledge bases of the size used by performance programs. Research is just starting to appear which organizes the knowledge used for problem solving as a knowledge base requiring management.

The basic requirement for a knowledge base is that knowledge can be appropriately accessed during problem solving. The comprehensive management of a knowledge based problem solving system, however, includes (1) the capability for the acquisition of knowledge from an expert in the task area and (2) the capability to explain the problem solving process. A workable approach to knowledge acquisition involves accepting knowledge from the user expressed in a form which is conceptually close to the way he thinks about it. Management of the acquisition process requires knowing how to integrate new knowledge into an existing knowledge base.

Current techniques involve directing the acquisition process so that questions to the user are well focussed and knowledge is acquired in a series of small manageable steps. Explanation, which embodies the reverse translations as acquisition, may be used for several purposes: (1) maintaining the trust and credibility of the user when the system acts in the role of consultant, (2) providing user/system interaction during a problem solving session, (3) informing a novice of the relevant domain knowledge for solving a particular problem, and (4) providing a part of a knowledge base debugging tool. Research into providing capabilities for knowledge acquisition and explanation have not been a major part of previous problem solving systems. The benefits of a system incorporating such capabilities include an ease and flexibility for exploring new representations and strategies and the possible extension of a problem solver to more than one task area.

Current research on the management of knowledge bases draws on several other areas of computer science research. For example, the development of extensible programming languages was an effort to minimize the conceptual distance for a programmer who had to express his ideas in terms of abstract programming symbols. A programming language is defined in terms of data structures, its operations, and its control structures. Extensibility is the language capability which allows a programmer to define extend the language by creating new types of these structures for his own convenience. It relies on the technique of allowing new entities to be defined in terms of a set of basic primitives. It was generally thought that extensibility in the programming language would result in clean and efficient programs and that these programs would be much easier to write. For the purposes of knowledge base research, the important lesson from this work was that the amount of knowledge for a user to mold the nature of a system to fit his requirements had been seriously underestimated. The system remained too ignorant to provide much help.
Another area of research relevant to knowledge base research is data base research. Knowledge base and data base researchers are currently attempting to define the differences between their respective fields. There are certain obvious differences. It is not unusual to find government or commercial data bases of more than one billion characters. This is roughly a thousand times larger than any knowledge base used in artificial intelligence. With a huge data base, researchers must be concerned with efficient retrieval of information. The information is generally used as input to separate programs performing specific tasks such as report generation, payroll, or a display of information for a user. The data base contains limited knowledge about itself and its uses.

As knowledge base researchers have moved to separate data from code, they have tried to create systems which reflected the dense interconnections necessary for problem solving. Thus, corresponding to the data structures, operations, and control structures of an extensible languages are the objects, operations, and strategies involved in problem solving. The objects are the entities that can be manipulated, the operations define what can be done to the objects, and the strategies control the problem solving process efficiently. In these terms, data base research has only been concerned with the objects.

Much of the progress in knowledge base research has resulted from the use of meta-knowledge, that is, knowledge about knowledge as a means towards creating new capabilities. Two forms of meta-knowledge are distinguished - schemata and derived models. Schemata are models which express the structure and variability of knowledge in terms of its components. Rigorous adherence to the philosophy of having a knowledge base organized by schemata implies that each object, operation, or strategy has a schema which describes its structure. The usual expectation is that each schema is used for many entities and every new entity in the system is acquired by a standard program which uses the schema to guide the acquisition process. Derived models summarize the instances of knowledge that the system has acquired so far. They may be used to create expectations beyond what is explicit in the schemata for new knowledge based on examples drawn from the knowledge base.