Reprinted from *Proceedings of the Fifth International Joint Conference on Artificial Intelligence (IJCAI-77)*, August, 1977, p. 247

Truth Maintenance Systems For Problem Solving

Jon Doyle
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
545 Technology Square
Cambridge, Massachusetts 02139

This summarizes the full report (Jon Doyle, "Truth Maintenance Systems for Problem Solving," MIT AI Lab TR-419) which describes progress that has been made in the ability of a computer system to understand and reason about its own reasoning faculties. A new method for representing knowledge about beliefs has been developed. This representation, called a non-monotonic dependency system, extends and clarifies several similar previous representation forms for such knowledge, and has been employed in developing new strategies for representing assumptions, backtracking, and controlling problem solving systems.

A truth maintenance system is a combination of a representation for recording justifications for program beliefs and procedures for effecting any updating of beliefs necessary upon the addition of new information. Such a system can easily be used by processes for reasoning about the recorded program reasoning. In particular, processes for non-chronological, dependency-directed backtracking and hypothetical reasoning are particularly straightforward in implementation given the representations of a truth maintenance system.

The basic operation of a truth maintenance system is to attach a justification to a fact. A fact can be linked with any component of program knowledge which is to be connected with other components of program information. Typically, a fact might be connected with each assertion and rule in a data base, or might be attached, with differing meanings, to various subsystem structures. The truth maintenance system decides, on the basis of the justifications attached to facts, which beliefs in the truth of facts are supported by the recorded justifications.

A belief may be justified on the basis of several other beliefs, by the conditional proof on one belief relative to other beliefs, or by the lack of belief in some fact. The latter form of justification allows the consistent representation and maintenance of hypothetical assumptions.

Truth maintenance processing is required when new justifications change previously existing beliefs. In such cases, the status of all beliefs depending on the changed beliefs must be redetermined. From the justifications used in this judgement of beliefs, a number of dependencies between beliefs are determined, such as the set of beliefs depending on each particular belief or the beliefs upon which a particular belief depends.

Several useful processes are supported by the above functions and representations. It is a straightforward matter to interrogate the truth maintenance system representation for the basic material of explanations of beliefs. More sophisticated uses of the recorded justifications are in hypothetical reasoning, generalization, separation of levels of detail, and in dependency-directed backtracking.

Hypothetical reasoning is supported by the use of conditional proof justifications. These are justifications which support belief if a specified belief follows from a set of other beliefs. This capability is instrumental in summarizing discoveries in a manner independent of the hypotheses leading to their derivation.

The processes of generalization and separation of levels of detail are also supported by the mechanism of conditional proof. By using conditional proofs to remove dependence of beliefs on other beliefs, results can be justified independent of the particular quantities used in their computation, and results at one level of detail can be supported by reasons which are independent of results at lower levels of detail.

Dependency-directed backtracking is a powerful technique based on the representations of the truth maintenance system. This method employs the recorded dependencies to locate precisely those hypotheses relevant to the failure and uses the conditional proof mechanism to summarize the cause of the contradiction in terms of these hypotheses. Because the failure is summarized independent of the hypotheses causing the failure, future occurrences of the failure are avoided.

This research is supported by a Fannie and John Hertz Foundation graduate fellowship. This research was conducted at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643.