Proceedings of the Twelfth National Conference on Artificial Intelligence, 1994, p. 1492.

Towards Situated Explanation

Raja Sooriamurthi and **David Leake**
Computer Science Department
Indiana University, Bloomington, IN 47405
{raja,leake}@cs.indiana.edu

In AI research on explanation, the mechanisms used to construct explanations have traditionally been neutral to the environment in which the explanations are sought. Our view is that the explanation process cannot be isolated from the situation in which it occurs. Without considering the intended use for an explanation, the explanation construction process cannot be properly focussed; without considering the situation the process cannot act effectively to gather corroborating information. The emphasis of this research is to view explanation as a means to an end and in this work the end is the successful functioning of the system requesting explanation. We develop a model of explanation as a situated, utility-based, hierarchical, goal-driven process.

Explanation is viewed as a memory process of adapting prior explanations of similar problems to fit the current situation. By using case-based reasoning we incorporate prior experience during the explanation process and also generate an initial set of hypotheses in a relatively inexpensive way. The system under development consists of a planner, acting in a simple simulated world, integrated with a case-based explainer. The over-arching goal of the explanation system is to aid the formulation and execution of plans. The occurrence of a planning or execution failure triggers explanation. Given the goals of the planner the explanation system has to decide when to explain, what to explain and how to explain in a dynamic environment. Related to earlier work on goal-directed diagnosis (Rymon 1993), a major issue being addressed in this research is how contextual changes influence the ongoing explanation process.

As an example of the issues to address consider the following scenario implemented in our system. The planner is trying to achieve the goal of catching a plane and generates two possible plans: driving to the airport or taking a taxi. Choosing the option of driving, the plan steps are executed in a simulated world and it is detected that the car does not start. In trying to determine why the car will not start our model of the ensuing explanation process involves:

- **Situated reasoning.** The explanation process responds to two types of situational changes:
  1. *Changes in the external environment:* For example, due to time constraints the explanation effort could be curtailed and the system might choose the alternative plan of taking a taxi. But if the system becomes aware of a delay in the flight it opportunistically uses the extra available time for continuing with the explanation effort.
  2. *Changes in the system goals:* The system will be able to produce explanations for three explanation purposes (Leake 1992): anomaly resolution, recovery and prevention. Changes in these explanation goals might produce explanations at levels of detail ranging from "something is wrong with the engine" to "drained-battery".
- **Knowledge planning.** An important facet of situated explanation is that all the information necessary to formulate an explanation might not be initially available to the system. The explainer has to determine what type of additional information is needed and how to procure it by planning to acquire knowledge (Hunter 1990). Prior cases help to package such knowledge planning actions that guide the information search.
- **Utility-based and hierarchical reasoning.** The utilities of the alternatives the explainer considers change with changes in the external environment. Our model of explanation allows the system to switch its focus to those alternatives that look most worthwhile in the current situation. The utilities of the alternatives also help to adjust the depth of the explainer's exploratory efforts. Case-based reasoning forms an effective mechanism to determine and update the situational utility of the alternative explanations.

References

Hunter, L. 1990. Planning to learn. In *12th Annual Conference of the Cognitive Science Society, Cambridge, Massachusetts*, 261–268. Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey.

Leake, D. B. 1992. *Evaluating Explanations: A content theory.* Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey.

Rymon, R. 1993. *Diagnostic Reasoning and Planning in Exploratory-Corrective Domains.* Ph.D. Dissertation, University of Pennsylvania.