based on just input signals ("weak" conditions) in LEAP [254].

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to Professors Tom Mitchell and Lou Steinberg of Rutgers University for their fruitful guidances when he was at Rutgers as a visiting researcher. He would like to thank Dr. Katsuya Hakozaki, Dr. Masahiro Yamamoto and Mr. Nobuhiko Koike of NEC Corporation for their encouragement.

In Machine Learning: A Guide to Current Research, T.M. Mitchell, J.G. Carbonell, and R.S. Michalski (eds.) Kluwer Academic, 1986.

V-3

OVERVIEW OF THE ODYSSEUS LEARNING APPRENTICE

David C. Wilkins, William J. Clancey, and Bruce G. Buchanan

Stanford University, Department of Computer Science
Stanford, CA 94305

ABSTRACT

Human specialists employ impressive learning methods during their apprenticeship training period to augment their fledgling expertise. We describe an apprentice learning system under development that allows an expert system to use some of these same methods. These methods aid an expert system in transferring expertise to and from its knowledge base (i.e., in knowledge acquisition and intelligent tutoring).

Our approach to apprenticeship learning is embodied in a computer program, Odysseus, that watches the observable actions of a specialist. Justifications are created for each action of the specialist via a process of differential modeling between the specialist and the expert system. A learning opportunity occurs when no action justification is judged sufficiently plausible. This paper describes the three phases that Odysseus uses to learn via differential modeling: setting the stage for differential modeling by expanding the initial rule base and deriving rule justifications, detecting knowledge base differences by observing actions of a specialist and ranking proposed action justifications, and effecting knowledge base repair by rationalizing discrepancies and postulating new rules.

INTRODUCTION

An apprenticeship learning period is an important phase on the path to master expert status for human specialists.[1] During this phase, an apprentice specialist *learns by watching master specialists* and *learns by doing problem solving* under the supervision of master specialists. Our research investigates how to give an expert system the benefits of an

---

[1]By *specialist*, we mean a problem solver whose abilities are at the novice or master level, and who is either a human or an expert system.

diagnosis, these actions consist of all data requests made by a physician and the final diagnosis. Actions are rationalized by a process of *differential modeling* between the expert system and the specialist. Failure to find an adequate rationalization signals a possible deficiency in the expert system's domain or strategy knowledge. Using a taxonomy of deficiencies in conjunction with theoretical and experiential knowledge of the application domain, Odysseus automatically generates and tests conjectures to explain its inability to justify a specialist's action.

Odysseus is designed to work in conjunction with Heracles, an expert-system shell for solving heuristic classification problems, that was created by removing the medical knowledge from Neomycin [72]. Neomycin is a reorganization of the Mycin expert system that simulates the diagnostic process of medical experts, via a large body of abstract domain-independent strategy knowledge for hypothesis-directed reasoning. This strategy knowledge is used by Odysseus as a framework for detecting differences between the domain knowledge of a Heracles-based expert system and of a specialist. Odysseus has an abstract strategy language that allows comparison of the strategic behavior of the expert system and of a specialist [394].

ODYSSEUS' METHOD

Expanding Rule Base and Deriving Rule Justifications

There are two ways in which an existing expert system must be augmented before differential modeling of a human specialist can commence. First, the set of heuristic rules must be expanded via induction over past problem solving cases. The original set of rules is adequate for problem solving but, in our experience, is too impoverished to model the alternate problem solving behavior of other specialists in an apprentice context. Second, rules should be justified from first-principle knowledge or experience. Rule justifications allow a learning system to reason about the rules during the process of rationalizing discrepancies. The Leap learning apprentice for circuit design justifies rules in terms of circuit theory, a strong theory of the domain [254]. By contrast, only a weak theory generally underlies medical diagnosis, and Odysseus's justifications for rules

minimal rule generality (coverage), minimal rule specificity (discrimination), maximal rule collinearity (similarity), and maximal rule simplicity (number of conjunctions and disjunctions). The rule evaluator always gives preference to collinear forms of heuristic rules contained in the original rule base. The expanded rule set produced by the induction subsystem is necessarily incomplete; however, it bootstraps the differential modeling process that leads to its refinement. Later, we will discuss how the induction subsystem suggests missing rules to the repair subsystem during the process of rationalizing discrepancies.

Observing Actions and Detecting Knowledge Base Differences

Odysseus must decide whether an action of the specialist suggests a significant domain or strategic knowledge difference between the specialist and the expert system. For each observed action of the specialist, Odysseus generates an action justification set: $J(A) = (j_1, j_2, \ldots, j_n)$. An action justification structure, j_k, relates an action A to an abstract strategic goal G via a skeletal rule path, that is, $A \rightarrow R_1 \rightarrow R_2 \rightarrow \cdots \rightarrow G$. A typical goal might be the confirmation of a particular hypothesis. All skeletal rule paths beginning with A and leading to a goal are in the set $J(A)$; thus the set delimits the possible interpretations that can be attributed to the specialist's action. Using the original Neomycin rule base, the average size of $J(A)$ is 20 and the maximum size is approximately 400.

Action justification sets are posted on a blackboard, and a variety of knowledge sources (KSs) attach confirming and disconfirming evidence to individual action justifications. The more important KSs are as follows: The *Heracles simulator* KS processes the information obtained during the problem solving session and relates the current status of findings, hypotheses, and rules to individual action justifications. For example, if this KS believes that particular hypotheses have already been concluded, then it attaches negative evidence to all action justifications whose goal is to confirm one of these hypotheses. The *multiple interpretations* KS consists of heuristic rules that medical domain experts use to arbitrate between multiple interpretations. For instance, early in the consultation session with different action justifications confirming different hypotheses, the more general hypotheses are preferred. The *user model* KS records user