In Proceedings, Workshop on Models of Complex Human Learning, Cornell University, June 27-29, 1990, p. 32.

Automated Knowledge Acquisition Using Apprenticeship Learning Techniques

David C. Wilkins
Department of Computer Science
University of Illinois
Urbana, IL 61801

Apprenticeship is the most effective means for human problem solvers to learn domain-specific problem-solving knowledge in knowledge-intensive domains. This observation provides motivation to give apprenticeship learning abilities to knowledge-based expert systems. The paradigmatic example of an apprenticeship period is medical training. Our research investigated apprenticeship in a medical domain.

The described research illustrates how an explicit representation of the strategy knowledge for a general problem class, such as diagnosis, provides a basis for learning the domain-level knowledge that is specific to a particular domain, such as medicine, in an apprenticeship setting. Our approach uses a given body of strategy knowledge that is assumed to be complete and correct, and the goal is to learn domain-specific knowledge. This contrasts with learning programs such as LEX and LP where the domain-specific knowledge (e.g., integration formulas) is completely given at the start, and the goal is to learn strategy knowledge (e.g., preconditions of operators) (Mitchell, 1983). Two sources of power of the Odysseus approach are the method of completing failed explanations and the use of a confirmation theory to evaluate domain-knowledge changes.

Apprenticeship learning involves the construction of explanations, but is different from explanation based learning as formulated in EBG (Mitchell, 1988) and EBL (DeJong, 1986); it is also different from explanation based learning in LEAP (Mitchell, 1989), even though LEAP also focuses on the problem of improving a knowledge-based expert system. In EBG, EBL, and LEAP, the domain theory is capable of explaining a training instance and learning occurs by generalizing an explanation of the training instance. In contrast, in our apprenticeship research, a learning opportunity occurs when the domain theory, which is the domain knowledge base, is incapable of producing an explanation of a training instance. The domain theory is incomplete or erroneous, and all learning occurs by making an improvement to this domain theory.

Our approach is also in contrast to the traditional empirical induction from examples method of refining a knowledge base for an expert system for heuristic classification problems. However, with respect to the learning of certain types of heuristic rule knowledge, empirical induction over examples plays a significant role in our work. In these cases, an apprenticeship approach can be viewed as a new method of biasing selection of which knowledge is learned by empirical induction.

An apprenticeship learning approach, such as described in this talk, is perhaps the best possible bias for automatic creation of large 'use-independent' knowledge bases for expert systems. We desire to create knowledge bases that will support the multifaceted dimensions of expertise exhibited by some human experts, dimensions such as diagnosis, design, teaching, learning, explanation, and critiquing the behavior of another expert.

Wilkins, D.C., Knowledge Base Refinement Using Apprenticeship Learning Techniques, *Proceedings of the Seventh National Conference on Artificial Intelligence*, August, 1988, 646-651.

Wilkins, D. C., "Apprenticeship Learning Techniques for Knowledge Based Systems," Report STAN-CS-88-1242 and KSL-88-14, Department of Computer Science, Stanford University, December 1988, 153 pp.

Wilkins, D. C., "Automated Knowledge Acquisition Using Apprenticeship Learning Techniques." in *Machine Learning: An Artificial Intelligence Approach, Volume III*, Y. Kadratoff and R. Michalski (eds.), Los Altos, CA: Morgan Kaufmann, 1989, in press.

Buchanan, B. G. and Wilkins, D. C. (eds.), *Readings in Knowledge Acquisition*, Menlo Park, CA: Morgan Kaufmann, to appear.