Machine Learning, 14, 137-138 (1994)

Guest Editorial

This special issue focuses on a broad theme in learning: evaluating and changing the representation. The issue was motivated by the workshop on Evaluating and Changing Representation held at the International Joint Conference on Artificial Intelligence in Sydney, 1991.

A quick look through the machine learning literature shows that the problem of representation is a reoccurring theme. Michalski first coined the term *constructive induction* (Michalski, 1983) to describe the situation where new terms are introduced into the representation. Several topical workshops since then include the workshop on Change of Representation and Inductive Bias in 1988 (Benjamin, 1990) and the workshop on Representational Issues in machine learning held at Cornell in 1989.

The influence of representation for the success of machine learing systems is well known. The representation language for a learning application is formed by the vocabulary (signature) as well as by the restrictions on the chosen formalism (e.g., number of literals). In practical applications the design of the representation for input, available domain knowledge, expert interaction, and output can be time-consuming and critical for success. Research here encompasses, empirically and theoretically, the evaluation of representations, their change, and the construction of new terms.

The importance of representation is well known throughout the artificial intelligence community. Representation interacts strongly with the tractibility of inference (Brachman & Levesque, 1984), the comprehensibility of knowledge in knowledge acquisition, and the learnability of concepts (Haussler, 1988). It is clear that knowledge representation has strong fundamental links with machine learning that have not yet been fully explored (Morik, 1989).

A change of representation can be used for several reasons. If performance losses are observed—for instance, by means of standard cross-validation techniques—it may be appropriate to revise the representation by reducing the number of basic features. On the other hand, if a concept cannot be determined, new terms in the vocabulary extend the hypothesis space. Changing representation can also produce cleaner and more comprehensible knowledge for inspection or validation by a domain expert. For instance, conceptual clustering and unsupervised learning can be viewed as the task of introducing a single new term into the language (the simplest change in representation). From a theoretical perspective, allowing change in representation is just the same as introducing a broader class of hypotheses. Therefore, various theoretical analyses such as PAC or Bayesian readily apply to changing representations. For instance, existing frameworks that introduce new terms, such as hidden Markov models (Rabiner & Juang, 1986), grammar induction, and unsupervised learning, have all been analyzed using fairly standard theoretical techniques.

In the context of analytic learning, representation change allows a much broader space of possibilities, but each sharing the common goal of improving the performance of inference—for instance, the utility measures now used in speed-up learning. Macro operators can be developed to improve planning systems (Korf, 1985), and different representations of the state space can lead to more efficient search.

Techniques presented in this special issue cover a range of the issues we have just discussed. Wnek and Michalski (this issue) introduce new features on the basis of learned rule sets. The rules with the highest predictivity and the highest number of covered positive examples become a new feature, which is then used to reformulate the examples for further learning.

The demand of a new term is determined on the basis of an evaluation of the representation language. In the approach of Stefan Wrobel (this issue), the evaluation refers to a problem-solver. A new term is introduced in order to distinguish between successful rule applications and those that lead to a contradiction.

Kietz and Morik (this issue) use a representation formalism that is a standard in the field of knowledge representation. The basic formalism allows learning in polynomial time. It is enhanced by constructive induction only if a new term is needed to distinguish between two disjoint classes.

Two approaches in the framework of logic programming are presented by Rouveirol (this issue) and Puget (this issue). Rouveirol enhances the inversion of resolution. Puget uses Clark's completion to learn complementary concepts from finite failures.

Katharina Morik, Univ. Dortmund, Germany
Francesco Bergadano, Univ. Torino, Italy
Wray Buntine, NASA Ames Research Center, USA

References

Benjamin, P. (1990). *Change of representation and inductive bias*. Boston, MA: Kluwer Academic.

Brachman, R., & Levesque, H. (1984). The tractability of subsumption in frame-based description languages. *Fourth National Conference on Artificial Intelligence*, (pp. 34–37), Austin, Texas.

Haussler, D. (1988). Quantifying inductive bias: AI learning algorithms and Valiant's learning framework. *Artificial Intelligence*, *36*(2), 177–222.

Korf, R.E. (1985). Macro-operators: A weak method for learning. *Artificial Intelligence*, *26*(1), 35–77.

Michalski, R. (1983). A theory and methodology of inductive learning. *Artificial Intelligence*, *20*(2), 111–161.

Morik, K. (1989). *Knowledge representation and organization in machine learning*. Springer: Lecture Notes in Artificial Intelligence Series.

Rabiner, L., & Juang, B. (1986). An introduction to hidden Markov models. *IEEE ASSP Magazine*, January, 4–16.