

# Towards a Unified Approach to Concept Learning

Pedro Domingos\*

Department of Information and Computer Science  
University of California, Irvine  
Irvine, California 92717, U.S.A.  
pedrod@ics.uci.edu  
<http://www.ics.uci.edu/~pedrod>

Rule induction (either directly or by means of decision trees) and case-based learning (forms of which are also known as instance-based, memory-based and nearest-neighbor learning) arguably constitute the two leading symbolic approaches to concept and classification learning. Rule-based methods discard the individual training examples, and remember only abstractions formed from them. At performance time, rules are applied by logical match (i.e., only rules whose preconditions are satisfied by an example are applied to it). Case-based methods explicitly memorize some or all of the examples; they avoid forming abstractions, and instead invest more effort at performance time in finding the most similar cases to the target one.

There has been much debate over which of these two approaches is preferable. While each one can be extended to fit the results originally presented as evidence for the other, it typically does so at the cost of a more complex, less parsimonious model. In classification applications, each approach has been observed to outperform the other in some, but not all, domains.

In recent years, multistrategy learning has become a major focus of research within machine learning. Its main insight is that a combination of learning paradigms is often preferable to any single one. However, a multistrategy learning system typically operates by calling the individual approaches as subprocedures from a control module of variable sophistication, and again this is not completely satisfactory from the point of view of parsimony.

In my thesis work I argue that rule induction and case-based learning have much more in common than a superficial examination reveals, and can be unified into a single, simple and coherent model of symbolic learning. The proposed unification rests on two key observations. One is that a case can be regarded as a maximally specific rule (i.e., a rule whose preconditions are satisfied by exactly one case). Therefore, no syntactic distinction need be made between the two. The second observation is that rules can be matched approximately, as cases are in a case-based classifier (i.e., a rule can match an example if it is the closest one to it according to some similarity-computing procedure, even if the example does not logically satisfy

all of the rule's preconditions). A rule's extension, like a case's, then becomes the set of examples that it is the most similar rule to, and thus there is also no necessary semantic distinction between a rule and a case.

The RISE algorithm is a practical, computationally efficient realization of this idea. RISE starts with a rule base that is simply the case base itself, and gradually generalizes each rule to cover neighboring cases, as long as this does not increase the rule base's error rate on the known cases. If no generalizations are performed, RISE acts as a pure case-based learner. If all cases are generalized and the resulting set of rules covers all regions of the instance space that have nonzero probability, it acts as a pure rule inducer. More generally, it will produce rules along a wide spectrum of generality; sometimes a rule that is logically satisfied by the target case will be applied, and in other cases an approximate match will be used. This unified model is more elegant and parsimonious than a subprocedure-style combination. Experiments with a large number of benchmark classification problems have also shown it to consistently outperform either of the component approaches alone, and lesion studies and experiments on artificial domains have confirmed that its power derives from its ability to simultaneously harness the strengths of both components (Domingos 1996).

The remaining proposed thesis work will focus on further elucidating what the bias of RISE is compared to that of its parent paradigms (and thus determining when it will be the more appropriate algorithm to use), on applying the ideas contained in RISE to problems like context dependency, feature selection, fragmentation avoidance, and data mining, and on bringing the use of domain knowledge into the RISE framework.

## References

Domingos, P. 1996. Unifying Instance-Based and Rule-Based Induction. *Machine Learning*. Forthcoming.

---

\*Partly supported by a PRAXIS XXI scholarship.