Some comments on rule induction

IVAN BRATKO
Kardelj University
Josef Stefan Institute, Ljubljana, Yugoslavia

DONALD MICHIE
The Turing Institute, Glasgow, UK

Abstract

Computer induction is discussed in the light of concerns recently expressed by Bloomfield about the use of such algorithms in expert systems construction.

1 Introduction

In "Capturing Expertise by Rule Induction" (this issue) B.P. Bloomfield expresses doubts concerning computer induction as an aid to the development of knowledge-based programs. He seems to imply that even though it may not be easy for a domain expert to generate rules exclusively by hand, nevertheless hand crafting somehow ensures more dependable or otherwise more satisfactory results.

Although the title of his article speaks of "rule induction", Bloomfield confines his concerns for the most part to the properties of one particular package for rule induction, an early entry-level PC product known as Expert-Ease (McLaren, 1984). This package's induction algorithm is misleadingly equated by Bloomfield with Quinlan's ID3. Expert-Ease actually embodies the ACLS algorithm (Paterson, Niblett and Shapiro, 1982) which extends ID3 into the numerical domain using a method due to Andrew Blake of Edinburgh University. In any case, it may be useful to offer comments on some of these concerns, based on our own experience with Expert-Ease and Expert-Ease-like systems.

2 Concerns and comments

Concern: Bloomfield is unhappy with the proposition that experts can provide suitable sets of example decisions for rule induction. "And if experts cannot fully explicate the basis of their expertise how can we assume that they are in any better position to provide an adequate training set?"

Comment: This is not an assumption, but a regular empirical finding. The phenomenon was first documented in 1976 by Chilausky, Jacobsen and Michalski. In their experiment the expert's hand-crafted rules gave 83% accurate diagnosis on a test sample of soy bean data, whereas a rule-base induced from an expert-supplied training set gave over 99% accuracy. The phenomenon of the "subarticulate" expert has been discussed elsewhere (Michie 1986b, 1986c).

Concern: Bloomfield makes the point that without an adequate set of attributes for characterizing states of the problem, an induction system such as Expert-Ease cannot be expected to derive a complete solution.

Comment: Nor, alas, can any other method for knowledge capture, such as "dialogue acquisition". But inductive tools do seem to help the expert to identify more precisely where attributes are missing, and also to discard redundant attributes which may wrongly have been thought significant. Even when the task is complicated by the presence of random noise in the data, the inductive systems C4 (Quinlan, 1986a, 1986b; Quinlan et al., 1986) and Assistant (Bratko and Kononenko, 1986) reliably detect insignificant attributes. If in a medical case the expert has assumed "age" to be a relevant variable for a given diagnosis, application of such an induction system to sample decision data may reveal that it is not.

Provided only that the training set is not too small, the induction system will discard or include such an attribute depending on the attribute's actual discrimination power. Quinlan (1984) analyses how the size of the training set affects the accuracy of induced rules.

Concern: Bloomfield gives an example of the use of Expert-Ease on a problem of building a recognizer of those points in an X–Y plot which lie on a 45-degree straight line passing through the origin.

Bloomfield makes a very natural comment to the general effect that the rule seems weak and ad hoc to anyone who has spotted the more elegant rule which is constructable as soon as one supplies a logical attribute which gives "true" if $x = y$ and "false" otherwise.

Comment: This is no more than a repetition of the familiar fact that a rule cannot exceed the power of the description language provided, which of course applies as well to hand-crafted rules as to inductively generated rules.

3 Concluding remarks

Since Bloomfield does not mention the criterion which perhaps matters most, i.e. experience with applying induction to practical problems, let us end with some findings from our own practice. Rule induction, as any other known method of knowledge-acquisition, is not without difficulties. However, when used interactively by a knowledge engineer and a domain expert in the knowledge elicitation process, it has produced in many cases far better results than dialogue techniques, judged in terms of speed of acquisition, economy and accuracy. Examples of industrial applications can be found in Michie (1986a, 1986b) and in Slocombe et al. (1986), medical applications in Bratko and Kononenko (1986) and in Quinlan et al. (1986), and agricultural in Michalski and Chilausky (1981).

More advanced rule induction tools than Expert-Ease also allow provision of highly articulate run-time explanation facilities, as demonstrated academically by Shapiro and Michie (1986) and in a commercial context by A-Razzak and colleagues (1986).

References

A-Razzak, M., Hassan, T. and Ahmad, A. (1986) *ExTran 7.2 User Manual*, Glasgow: Intelligent Terminals Ltd.

Bratko, I. and Kononenko, I. (1986) Learning diagnostic rules from incomplete and noisy data. *Proc. UNICOM Seminar on The Scope of AI in Statistics* (ed. B. Phelps). To be published in book form by Technical Press.

Chilausky R.L., Jacobson, B. and Michalski, R.S. (1976). An application of variable-valued logic to inductive learning of plant disease diagnostic rules. *Proc. VI Internat. Symp. on Multi-Variable Logic*, Utah.

McLaren, R. (1984) *Expert-Ease User Manual*, Glasgow: Intelligent Terminals Ltd.

Michalski, R.S. and Chilausky, R.L. (1980) Learning by being told and learning by examples: an experimental comparison of the two methods of knowledge acquisition in the context of developing an expert system for soybean disease diagnosis. *Int. J. of Policy Analysis and Inf. Systems*, 4, 125–161.

Michie, D. (1986a) Current developments in expert systems. *Proc. 2nd Austr. Conf. on Applications of Expert Systems*, Sydney, 163–182. Also to appear in book form: Applications of Expert Systems (Ed. J.R. Quinlan), Wokingham: Addison-Wesley.

Michie, D. (1986b) The superarticulacy phenomenon in the context of software manufacture. *Proc. Roy. Soc.* (*A*) 405, 185–212.

Michie, D. (1986c) Towards a knowledge accelerator. *Advances in Computer Chess 4* (Ed. D. Beal), Oxford: Pergamon Press; also reproduced as Chapter 21 in *On Machine Intelligence* (2nd Edn), Chichester: Ellis Horwood, 1986.

Paterson, A., Niblett, T.B. and Shapiro A, (1982) *ACLS User Manual*, Glasgow: Intelligent Terminals Ltd.

Quinlan, J.R. (1984) Learning efficient classification procedures and their application to chess and games. *Machine Learning and Artificial Intelligence Approach* (Eds. R.S. Michalski, J.G. Carbonell and T. M. Mitchell, Palo Alto, Ca: Tioga, 463–482.

Quinlan, J.R. (1986a) Induction of Decision Trees. *Machine Learning*, 1 81–106.

Quinlan, J.R. (1986b) Simplifying decision trees. *Working Paper*, Cambridge, Mass: MIT AI Laboratory. To appear in *Internat. J. Man-Mach. Studies.*

Quinlan, J.R., Compton, P.J., Horn, K.A. and Lazarus, L. (1986) Inductive knowledge acquisition: a case study. *Proc. 2nd Austr. Conf. on Applications of Expert Systems*, Sydney, 183–204. Also to appear in book form: *Applications of Expert Systems* (Ed. J.R. Quinlan), Wokingham: Addison-Wesley.

Shapiro, A.D. and Michie, D. (1986) A self-commenting facility for inductively synthesised endgame expertise. *Advances in Computer Chess 4*, 147–165.

Slocombe, S., Moore, K.D.M., and Zelouf, M. (1986), Engineering expert system applications. Delivered at the British Computer Society Annual Conference, Brighton. Typescript available from K.D.M. Moore, BP Ltd, Britannic House, London.

Declaration of interest

One of the authors (D.M.) is a Director of Intelligent Terminals Ltd, which developed and supplies commercially the copyright product Expert-Ease.