Integrating Abduction and Induction in Machine Learning

Raymond J. Mooney
Department of Computer Sciences
University of Texas
Austin, TX 78712-1188
mooney@cs.utexas.edu
March 3, 1997

1 Introduction

Abduction is the process of inferring cause from effect or constructing explanations for observed events and is required for tasks such as diagnosis and plan recognition. Induction is the process of inferring general rules from specific data and is the primary task of machine learning. An important issue is how these two reasoning processes can be integrated, or how abduction can aid machine learning and how machine learning can acquire abductive theories. My research group has explored these issues in the development of several machine learning systems over the last eight years. In particular, we have developed methods for using abduction to identify faults and suggest repairs for theory revision, and for inducing rule bases for abductive diagnosis. We believe that induction and abduction are two distinct reasoning tasks, but that each can be of direct service to the other in developing AI systems for solving real-world problems. Below I briefly review our work in these areas, focusing on the issue of how abduction and induction is integrated.\(^1\)

2 Abduction and Induction

Precise definitions for abduction and induction are still somewhat controversial. In order to be concrete, I will generally assume that abduction and induction are both defined in the following general logical manner.

- **Given**: Background knowledge \(B\) and observations (data) \(O\) both represented as sets of formulae in first-order predicate calculus where \(O\) is restricted to ground formulae.

- **Find**: An hypothesis \(H\) (also a set of logical formulae) such that \(B \cup H \not\vdash \perp\) and \(B \cup H \vdash O\).

\(^1\)Most of our group's publications are available in postscript on the world wide web at http://www.cs.utexas.edu/users/ml.
In abduction, $H$ is restricted to set of atomic ground formulae (called assumptions) and $B$ is generally quite large relative to $H$. On the other hand, in induction, $H$ generally consists of universally quantified Horn clauses (called a theory or knowledge base), and $B$ is relatively small and may even be empty. In both cases, following Occam's Razor, it is preferred that $H$ be kept as small and simple as possible.

Despite their problems, these definitions encompass a significant fraction of the existing research on abduction and induction, and the syntactic constraints on $H$ capture at least some of the intuitive distinctions between the two reasoning methods. In abduction, the hypothesis is a specific set of assumptions that explain the observations of a particular case; while in induction, the hypothesis is a general theory that explains the observations across a number of cases. The body of logical work on abduction, e.g. (Poole, 1973; Poole, 1989; Levesque, 1989; Ng & Mooney, 1991, 1992; Kakas, Kowalski, & Toni, 1993), generally fits this definition of abduction and several diagnostic models (Reiter, 1987; Peng & Reggia, 1990) can be shown to be equivalent or a special case of it (Poole, 1989; Ng, 1992). The work on inductive logic programming (ILP) (Muggleton, 1992; Lavrač & Džeroski, 1994) employs this definition of induction, and most machine learning work on induction can also be seen as fitting this paradigm (Michalski, 1983).

3 Abduction in Theory Revision

Theory revision or knowledge base refinement is the machine learning task of modifying an existing imperfect domain theory to be consistent with a set of data.

- **Given:** An initial theory $T$ and a set of positive examples $P$ and a set negative examples $N$ where $P$ and $N$ are restricted to ground formulae.

- **Find:** A "minimally revised" theory $T'$ such that $\forall p \in P : T' \vdash p$ and $\forall n \in N : T' \nvdash n$

Revising a theory may require both adding and removing clauses as well as adding or removing literals from existing clauses. Generally, the ideal goal is to make the minimal syntactic change to the existing theory (Mooney, 1995). Note that compared to the use of background knowledge in induction, theory revision requires modifying the existing knowledge rather than just adding clauses to it. Experimental results in a number of realistic applications has demonstrated that revising an existing imperfect knowledge base provided by an expert results in more accurate results than inducing a knowledge base from scratch (Ourston & Mooney, 1994; Towell & Shavlik, 1993).

Several theory refinement systems use abduction on individual examples to locate faults in a theory and suggest repairs (Ourston & Mooney, 1990, 1994; Wogulis & Pazzani, 1993; Baffes & Mooney, 1993, 1996; Brunk, 1996). Each of these systems use abduction in a slightly different way, but the following summarizes the basic approach. For each individual positive example that is not provable from the current theory, abduction is used to determine a set of assumptions that would allow it to be proved. These assumptions can then be used to make suggestions for modifying the theory. One potential repair is to learn a new rule for the assumed proposition so that it could be inferred from other known facts about the example. Another potential repair is to remove the assumed proposition from the list of antecedents of the rule in which it appears in the abductive explanation of the example. For example, consider the theory:
P(X) \Rightarrow R(X), Q(X).
Q(X) \Rightarrow S(X), T(X).

and the unprovable positive example:

P(a) \Rightarrow R(a), S(a), V(a).

Abduction would find that the assumption T(a) makes this positive example provable. Therefore, two possible revisions to the theory are to remove the literal T(X) from the second clause in the theory, or learn a new clause for T(X), such as T(X) \Rightarrow V(X). Another possible abductive assumption is Q(a), suggesting the possible revisions of removing Q(X) from the first clause or learning a new clause for Q(X) such as Q(X) \Rightarrow V(X) or Q(X) \Rightarrow S(X), V(X).

In order to find a small set of repairs that allow all of the positive examples to be proved, a greedy set covering algorithm can be used to select a small subset of the union of repair points suggested by the abductive explanations of individual positive examples, such that the resulting subset covers all of the examples. If simply deleting literals from a clause causes negative examples to be covered, inductive methods (e.g. ILP techniques like FOIL (Quinlan, 1990)) can be used to learn a new clause that is consistent with the negative examples.

As described in the references above, this general approach has proven quite successful at revising several real world knowledge bases in molecular biology, diagnosis, and student modeling for intelligent tutoring, significantly improving their accuracy on actual test data. These results demonstrate the utility of integrating abduction and induction for theory revision.

4 Induction of Abductive Knowledge Bases

Another important aspect of integrating abduction and induction is the learning of abductive theories. Induction of abductive theories can be viewed as a variant of induction where the provability relation \((\Rightarrow)\) is itself interpreted abductively. In other words, given the learned theory it must be possible to abductively infer the correct conclusion for each of the training examples.

We have previously developed a learning system, LAB (Thompson & Mooney, 1994; Thompson, 1993), for inducing an abductive knowledge base appropriate for the diagnostic reasoning model of Peng and Reggia (1990). LAB uses a fairly straightforward hill-climbing induction algorithm. At each iteration, it adds to the knowledge base the individual disorder \(\Rightarrow\) symptom rule that maximally increases accuracy over the training cases. The addition of rules terminate when the addition of any new rule fails to increase accuracy on the training data. Using real data for diagnosing brain damage due to stroke, this technique was shown to produce an abductive knowledge base that, according to one important evaluation metric, was more accurate than an expert-built abductive rule base and the deductive knowledge bases learned by several standard machine learning systems.

5 Conclusions

In conclusion, we believe our previous work on integrating abduction and induction has effectively demonstrated two important points: 1) Abductive reasoning is useful in inductively revising existing
knowledge bases to improve their accuracy and 2) Inductive learning can be used to acquire accurate abductive theories.

Acknowledgements

This research was partially supported by the National Science Foundation through grant IRI-9310819.

References


