Inductive reasoning is sometimes deductive

position paper

Éric GRÉGOIRE and **Lakhdar SAÏS** [1]

Abstract. In this position paper, some differences between inductive reasoning and abduction are investigated. From an epistemological point of view, abductive reasoning is related to the non-deductive inference of explanations (which are often required to be basic facts) that, together with some background theory, allow a specific observation to be deduced or nonmonotonically inferred. In the machine learning field, an important family of inductive reasonings consists in inferring rules that generalize a series of examples. In this paper, the focus is laid on a distinction between both forms of reasoning that is traced back to one possible nature of generalization. More precisely, this nature together with properties linking deduction and induction are used to show that generalization from examples is sometimes performed just using deduction, although the resulting process is correctly interpreted as inductive reasoning. On the contrary, most authors would not classify the logically corresponding forms of deductive explanation as abductive ones.

1 INTRODUCTION

From an epistemological point of view, inductive reasoning as it is often defined in the machine learning field and abduction are clearly two different processes. It is often claimed that both are non-truth-preserving forms of inference that differ in their respective goals. On the one hand, abduction is used to infer an explanation for some specific observed situation. The inference should be non-deductive; with the help of some background theory, the explanation, which often takes the form of basic facts, allows the specific situation to be deduced or nonmonotonically inferred. On the other hand, inductive reasoning in the machine learning field is often related to the process of generalizing from examples. Accordingly, two basic differences at the ontological level are generally acknowledged between both forms of reasoning. When inductive reasoning (resp. abduction) is under consideration, rules (resp. facts) are thus inferred that explain a series of examples (resp. a specific situation). In this paper, the focus is rather laid on a distinction between both forms of reasoning that is traced back to the various possible natures of generalization. More precisely, one such specific nature together with properties linking deduction and induction are used to show that generalization from examples is sometimes performed just using deduction, although the resulting process is correctly interpreted as inductive reasoning. On the contrary, most authors would not classify the logically corresponding forms of deductive explanation as abductive ones.

Accordingly, we shall focus on the different possible natures of generalization from examples, taking the formal relations between logical induction and deduction into account. In this respect, we take position in what appears to be an unsettled debate in the machine learning community, where several interpretations of the nature of generalization seem to conflict one with the other. Our contribution can be summarized as follows. First, we differentiate between three seemingly different approaches to generalization from examples: the inductive, the deductive, and the set-theoretic ones, respectively. Following Flach's contribution [7], we acknowledge that the set-theoretic account of generalization can be reinterpreted into logical terms. But, we also claim that there is not a unique correct logic-based approach to the paradigm of generalization: on the contrary, both the seemingly opposite deductive and inductive approaches can be accepted. Moreover, we recall that deduction and induction are not just simple opposite logical processes underlying two definitions for generalization, but obey properties so that, under some precise circumstances, one process is used to perform the other one. Actually, we show that these properties allow one to connect the various approaches to generalization and to shed light on the crossed dualities between, on the one hand, induction and deduction, and, on the other hand, generalization and specialization.

2 VARIOUS APPROACHES TO GENERALIZATION FROM EXAMPLES: AN INTUITIVE ACCOUNT

Actually, many symbolic machine learning techniques are concerned with the inference of useful generalizations of some basic examples or facts. Let us describe the standard logic-based approach. For example, from the fact that Clyde the elephant has a trunk and from a similar observation about other elephants, we might want to generalize and infer that all elephants have a trunk. From a logical point of view, such generalizations are *induced* from the given examples or facts. Since induction and deduction are converse logical operations, generalizations can be defined as formulas that allow one to deduce and thus explain basic examples or facts. At this point, the links between generalization, specialization, deduction and induction might thus seem straightforward: generalizations are assertions that are obtained using induction from facts and examples; backwards, they specialize into these examples or facts using deduction. Accordingly, this logic-based approach to the definition of generalization (and the result-

[1] CRIL, Université d'Artois, rue de l'Université, S.P. 16, F-62307 LENS Cedex, FRANCE, e-mail: {*gregoire,sais*}*@lens.lifl.fr*.

ing generalizations themselves) will be called inductive. However, this issue is more intricate and the newcomer to the machine learning field often gets confused by several clashing intuitions about the logical nature of generalization. Let us describe these alternative intuitions and how they conflict with the above inductive approach.

A first apparent contradiction to the inductive approach is that many machine learning algorithms that are known to perform inductive generalizations actually compute formulas that are *deductive* (vs. inductive) consequences of the examples. For example, in their simplest logical setting, inductive machine learning algorithms like AQ [10] or ID3 [5] can be interpreted as yielding generalizations that, actually, are deductive consequences of the examples. In the following, we show why and to which extent inductive generalizations can be obtained using deduction.

Actually, several authors even define generalizations as deductive consequences of the examples (e.g. [4, 8, 9]). Let us refer to this opposite logic-based approach as the *deductive* one. In the deductive approach, 'red' is interpreted as more general than 'red and circle' since the former piece of information can be deduced from the latter one. The underlying idea is that 'red' encodes less information than 'red and circle' and is thus more general since it is a less constrained description and allows for more ways to complete the information that it conveys. The inductive and deductive approaches are clearly converse ones that translate opposite intuitions. From a model-theoretic point of view, a more general formula should exhibit less logical models according to the inductive approach; whereas it should exhibit more models according to the deductive approach. Actually, we claim that both logic-based approaches to generalization have opposite but sound motivations, so that they can be both defended. Deciding which is the correct one cannot be addressed without making the epistemological (vs. logical) nature of generalization more precise. Moreover, we shall show that when the logical framework is extended into a nonmonotonic one, it becomes even harder to relate the inference relationship of the logic to the notion of generalization in the general case. Anyway, the deductive approach to generalization is often used implicitly within an inductive reasoning schema. On the contrary, most authors would not defend that this formal schema is a form of abduction when it is reinterpreted as a deductive approach to explanation.

This issue is even more intricate because of a third intuition about generalization that involves set-theoretic considerations. This last intuition concerns open formulas that can be interpreted as concepts, i.e. formulas containing free variables, like $French(x)$ denoting the set of French people. Intuitively, a concept whose extension is larger should be considered as being more general. For example, the set of European people is larger than the set of French people. Accordingly, $European(x)$ is more general that $French(x)$ when these formulas are interpreted in the real world. Note that it should be possible to deduce that a French person is a European one. Thus, at the intuitive level, this set-theoretic approach resembles the deductive one and conflicts with the inductive one, nothwithstanding the fact that a formal comparison between the approaches is not straightforward since the set-theoretic one does not apply in a direct manner to the propositional case.

In order to discuss this matter in more details, we need some logical background.

3 LOGICAL PRELIMINARIES

In order to represent examples and facts, domain knowledge (also called background theory) and learned formulas, many symbolic machine learning techniques make use of logic-oriented languages that can be reduced to fragments of the language L of standard first-order logic. In the following, we suppose that the set of individual constants of L is both finite and not empty and that there is no other functions allowed in L.

The well-formed formulas (thereafter, formulas or wffs) of L are built as usual. The symbols $\neg$, $\wedge$, $\vee$ and $\Rightarrow$ represent the negation, conjunction, disjunction and the material implication connectors, respectively. A set of wffs will be identified with the conjunctive wff whose conjuncts are the elements of the set. For example, given a wff F and a set A of wffs, we can speak of the wff $A \wedge F$.

Interpretations, satisfiability and models are defined in the usual way for closed wffs of L [2]. Let us stress that this standard semantics is often extended to wffs containing free variables [6, 3]; however free variables are then interpreted as (roughly) universally quantified ones. This differs from the intuition we have about open formulas like $French(x)$, which denotes the *set* of French people. Accordingly, we shall consider closed wffs only, except when we express set-theoretic concepts. How these concepts could be interpreted inside this logical framework will be the object of further comments. As we shall restrict ourselves to Herbrand interpretations only, the symbol $\models$ will actually represent the standard first-order entailment relation according to this Herbrand semantics (it will thus concern closed formulas, only). Let us also stress that wffs and sets of wffs will be considered from a semantical point of view: wffs (and sets of wffs) are equivalent iff they exhibit the same models.

4 DEDUCTION vs. INDUCTION

Let us now address the central issue of this paper, i.e. *how can we explain that inductive generalizations can sometimes be obtained using deduction, whereas this does not apply to corresponding abductive schemas?*

First, let us give a standard (purely logical) definition of inductive generalization. In the following, D and F will be two consistent wffs of L representing the domain knowledge and the basic examples or facts to be generalized, respectively. Most often, F will be the disjunction of these examples or facts. It is generally required that $D \models F$ does not hold.

Definition 4.1 *A wff G is an inductive generalization of F w.r.t. D if and only if $D \wedge G \models F$*

Obviously enough, deduction and induction are not identical processes; they coincide mainly in the trivial cases involving reflexivity of $\models$, i.e. let $A \in L : A \models A$ and $D \wedge A \models A$. However, a simple way to perform induction through deduction has hardly been exploited in the machine learning literature. It is based on the contraposition property of $\models$, i.e. let $A, B \in L : A \models B$ if and only if $\neg B \models \neg A$. Accordingly, we derive the following alternative (purely deductive) definition of inductive generalization that is equivalent to Definition 1.

Theorem 4.1 *A wff G is an inductive generalization of F w.r.t. D if and only if $D \wedge \neg F \models \neg G$*

Accordingly, the inductive generalizations of some given examples can be derived in a purely deductive way from the examples and domain knowledge. They consist of the negated deductive consequences of the conjunction of the domain knowledge with the negated examples. Such a property has been envisaged by [1]. (Obviously, the main difference from deductive theorem proving is that we are here to discover G whereas when theorem proving is under consideration, we are often just required to prove that a given G is derivable.) Moreover, it does not answer the question addressed here. Indeed, algorithms like AQ or ID3 are not based on the use of contraposition. Actually, the key relationship that is often used in an implicit manner to perform inductive generalization using deduction is the following one.

Let *Concept* be a new 0-ary predicate and D, F and G be wffs of L.

Theorem 4.2 *if $F \models G$ then $G \Rightarrow Concept \models F \Rightarrow Concept$ if $D \wedge F \models G$ then $D \wedge (G \Rightarrow Concept) \models F \Rightarrow Concept$*

This theorem is best understood as follows.

Inferring the inductive generalizations of a wff F consists in inducing wffs G from F. However, when the formulas are interpreted as sufficient conditions for the satisfaction of a concept, i.e. when F implicitly denotes $F \Rightarrow Concept$, it **can** consist in deducing G from F.

This result plays a key role in our argumentation. It thus explains that the induced description of a concept (i.e. its sufficient conditions) can sometimes be obtained using deduction. In their simplest logical setting, algorithms like AQ or ID3 that compute inductive generalizations using deduction can be interpreted in these terms.

Let us stress that the logical rules of Theorem 2 do not involve any epistemological contents by themselves but are just purely formal results. In theory, we could thus apply them to abduction. However, in the context of diagnosis or explanation searching, whenever the explanation or diagnosis G can be deduced from the observed situation F, most authors would not agree that it is abductive. In most situations, some authors would even be reluctant to say that G is as a real explanation or diagnosis for F, especially if G has to take the form of facts (vs. rules in inductive generalization). Moreover, it seems hard to reinterpret Theorem 2 in the context of explanation or diagnosis while obeing the intuition behind these tasks (in particular, by what should *Concept* be replaced?).

5 A FEW ADDITIONAL COMMENTS ABOUT THE CONFLICTING APPROACHES TO GENERALIZATION

Actually, Theorem 2 and the transformation principle that it conveys shed some light on the complexity of the actual relations between the set-theoretic and the two logical definitions of generalization. Indeed, in addition to the existence of three approaches to generalization that can be defended because of their specific acceptable motivations, it allows the inductive one to be implemented thanks to deductive techniques (thus using implementation techniques pertaining to the deductive approach) and allows the set-theoretic one to be reinterpreted in terms of the inductive one, as we shall see. Accordingly, the transformation principle formalized in Theorem 2 has played several roles in the literature. First, it is used quite implicitly by [11] and [13] to differentiate between concepts and concepts recognition functions, the latter authors introducing the so-called I-implication technique to avoid interpreting free variables. In defending the inductive approach to generalization, [12] refers to such a relation in an informal manner, claiming that when we assert that a wff G that can be deduced from a wff F is more general than F, then we are probably interpreting F and G as sufficient conditions of a common concept. We disagree with Niblett in the sense that there is some viable direct intuitions that justify interpreting G as more general than F. Finally, [7] used a transformation similar to the one involved in Theorem 2 (i.e., replace the concept $F(x)$ by $\forall x(F(x) \Rightarrow Concept(x)))$ to relate the inductive and set-theoretic approaches to generalization, in an attempt at showing that these approaches are two sides of the same coin.

Actually, it seems to us that, to some extent, the set-theoretic approach is more directly connected to the deductive one. When set-theoretic concepts are under consideration, the nature of generalization seems uncontroversial: the extension of a more general concept includes the extension of a less general one. For instance, the set of European people includes the set of French people. The set-theoretic approach resembles thus the deductive one since we should be able to infer that any French person is a European one, provided some adequate background knowledge is available. Also, asserting that (the set-theoretic concept) $European(x)$ is more general than (the set-theoretic concept) $French(x)$ can be translated in our logical framework as $\models \forall x(French(x) \Rightarrow European(x))$, which in turn allows $\forall x French(x) \models \forall x European(x)$ to be derived. Obviously, when we must deal with wffs including both open subformulas representing set-theoretic concepts and other closed subformulas (e.g. the background knowledge and the examples), we cannot simply use Thayse's semantics [6, 3] or bind the free variables to obtain a global closed wff that can be interpreted under our Herbrand semantics. Indeed, no real distinction would be observed between universally quantified variables and free variables of concepts that would be bound. We claim that this is a *technical* default pertaining to standard logic that should not obscure the possible deductive interpretation of the set-theoretic approach. On the other hand, it is sometimes claimed that the logical counterpart of the set-theoretic approach is the inductive one. To this end, [7] translates the concept $F(x)$ into $\forall x(F(x) \Rightarrow Concept(x))$ (to some extent, also used by [13]). Although this translation schema allows closed wffs and transformed set-theoretic concepts to be handled within the same logical framework, it should be noted that it translates a specific epistemological view of the generalization task. In that respect, when the issue of defining the universal logical nature of generalization is under consideration, it introduces a side-effect bias towards the inductive approach.

Actually, we claim that deciding which one amongst the two opposite logical definitions of generalization is 'correct' is a domain(or task)-dependent issue that requires the epistemological (vs. logical) nature of the form of generalization that we want to model to be made precise. Let us elaborate a little more to defend this view. Let $A, B \in L$ and $A \models B$, which wff is more general: A or B? We have already illustrated in the introduction that diverging intuitions can guide this choice towards one of the formula vs. the other one. We claim that this choice does not depend on the intrinsic nature of the $\models$ relationship but on the examples and task at hand, as illustrated before.

Moreover, when the logical framework is extended into a nonmonotonic one, it becomes even harder to relate the inference relationship of the logic to the notion of generalization in the general case. Let $\mid\!\sim$ be a nonmonotonic inference relationship. The argument favouring the inductive approach that rests on the idea that a more general information should bring more information does not hold anymore. Indeed, in the schema $D \wedge A \mid\!\sim B$, B is not necessarily semantically entailed by $D \wedge A$. Formula B can bring some additional information with respect to $D \wedge A$, and conversely. Also, the link between the set-theoretic and the deductive approaches vanishes. Indeed, let $\models \forall x(B(x) \Rightarrow A(x))$ be a logical translation of the set-theoretic assertion: the concept $A(x)$ is more general than the concept $B(x)$. The corresponding nonmonotonic assertion $\mid\!\sim \forall x(B(x) \Rightarrow A(x))$ does not necessarily express that the extension of B is included (with some possible exceptions) in the extension of A. It can sometimes translate the opposite idea. To illustrate this, replace $B(x)$ by $Mammal(x)$ and $A(x)$ by $Terrestrian - mammal(x)$; clearly, in this example we want the extension of the latter concept to be included in the extension of the former one.

ACKNOWLEDGEMENTS

This work has been supported by the Ganymède II project of the Contrat de Plan Etat/Nord–Pas-de-Calais and has benefited from discussions at CRIL with B. Bessant, P. Marquis and E. Mephu. We thank B. Mazure for his help in formatting the final version of paper.

REFERENCES

[1] Morgan C.G. Hypothesis generation by machine. *Artificial Intelligence*, 1971.

[2] Chang C.L. and Lee R.C.T. Symbolic logic and mechanical theorem proving. *Academic Press, New-York*, 1973.

[3] Thayse A. (ed.). From standard logic to logic programming. *J. Wiley, London*, 1988.

[4] Nicolas J. Ally: un système logique pour la généralisation en apprentissage, (thèse d'université). *Université de Rennes 1*, 1987.

[5] Quinlan J.R. Induction of decision trees. *Machine Learning*, 1986.

[6] Genesereth M.R. and Nilsson N.J. *Logical Foundations of Artificial Intelligence*. Morgan Kaufmann, 1988.

[7] Flach P.A. Generality revisited. In *Proceedings of ECAI'92 Workshop on Logical Approaches to Machine Learning.*

[8] Besnard Ph. and Grégoire E. About deductive generalization. In *Proceedings of JELIA'92.*

[9] Besnard Ph. and Grégoire E. Deductive generalization in a default logic setting. In *Proceedings of NIL'91.*

[10] Michalski R.S. Learning flexible concepts: conceptual clustering. In *Machine Learning: an AI Approach III.*

[11] Michalski R.S. A theory and methodology of inductive learning. In *Machine Learning: an AI Approach.*

[12] Niblett T. A study of generalisation in logic programs. In *Proceedings of European Working Session on Learning ESWL'88.*

[13] Kodratoff Y. and J.-G. Ganascia. Improving the generalization step in learning. In *Machine Learning II, Michalski R.S. et al. (eds.).*