Abduction, Induction and Completion Policies (Research Note)

Nicolas Lachiche[1] and **Pierre Marquis**[2]

Abstract.
In this paper, we are mainly concerned with abduction and confirmatory induction. While both reasoning models can be related to deduction w.r.t. a completed knowledge base, we claim that they do not basically rely on the same completion policies. Intuitively, induction requires individuals of the domain of interest to be circumscribed: every unknown individual is supposed to behave like the known ones. No additional assumption about the domain of interest is required. Contrastingly, abduction requires properties of such individuals to be circumscribed: explanations are supposed to be among the known ones.

1 Introduction

Many researches have been devoted so far to abduction and induction. Abduction is generally defined as inference to the best explanation. Thus, reasoning from a set of symptoms to their underlying causes (diseases) is often viewed as a straightforward abductive task. Inductive reasoning aims at deriving general laws from observations. For instance, generating classification rules from classified examples is basically an inductive activity.

While they look quite different at a first glance, abduction and induction can be considered closely related from several points of view. Particularly, both models have been pioneered by Aristotle, and further investigated by many logicians (e.g. C.S. Pierce) for centuries.

- From a logical point of view, both models are not truth-preserving. Thus, every conclusion derived by abduction or induction can be questioned in the general case (this is why such a conclusion is often called a *hypothesis*); accordingly, abduction and induction are non-monotonic. For each reasoning model, various bias (also called preference criteria) have been pointed out in order to characterize the best conclusions which can be drawn. For instance, consistency w.r.t. some background knowledge is often required both by abduction and induction.
- From the point of view of applications, abduction and induction have been used as basic reasoning components in many AI fields, including diagnosis, machine learning, speech recognition, etc.

In many works, inductive generalizations are required to have an explanatory power: every inductive hypothesis must explain why the corresponding observations are classified as they are classified. In this situation, the difference between abduction and induction is not so clear: From an epistemic point of view, both reasoning models aim at deriving best explanations. Accordingly, abduction and explanatory induction are closely related from a computational point of view: both can be considered as a form of "reversed deduction".

Despite many common features, our basic claim is that abduction and induction are different in essence: they do not aim at deriving conclusions with the same epistemic status. By investigating how similar they are, we should be able to see how different they are, and to have a better understanding of both reasoning models.

In the sequel, we consider abduction and induction from a logical point of view. Following Helft and Flach (among others), we first recall that induction does not reduce to "reversed deduction". Accordingly, we mainly focus on confirmatory induction in this paper. While abduction and confirmatory induction can be considered as deduction w.r.t. a completed knowledge base, we show that they do not basically rely on the same completion policies. Thus, after recalling how Clark's completion can be used both for abduction and induction, we present a model for confirmatory induction; in contrast to Clark's completion, the completion policy on which this model is based does not apply to abduction in a simple way. Intuitively, completion policies for confirmatory induction require individuals to be circumscribed, only: every unknown individual is supposed to behave like the known ones. No additional assumption about the domain of interest is basically required for induction. Contrastingly, completion policies for abduction require properties to be circumscribed: explanations are supposed to be among the known ones.

2 Formal Preliminaries

Throughout this paper, we shall consider function free first-order logic as a representation language. Formally, let L consist of a set L_P of predicate symbols (with the equality symbol $=$) and a finite and non-empty set L_C of constant symbols. Constant symbols are intended to represent individuals of the domain of interest, while predicate symbols allow us to represent their properties. Let X be an infinite set of variables. Let $FORM_L$ be the first-order logic language of formulas built in the usual way using standard connectives, the set L of non-logical symbols and the set X of variables. As usual, every finite set of formulas will be identified with the conjunction of the formulas in the set. Let us simply recall here that an atom is a *positive literal* and a negated atom is a *negative literal*. A formula is *ground* when no variable occurs in it. A formula is *closed* when no free variable occurs in it. A *clause* is a finite disjunction of literals the variables of which are implicitly universally-quantified.

Let $BOOL \triangleq \{T, F\}$ be the set of truth values *true* and *false*. An *interpretation* I (on $FORM_L$) is a pair $\langle D_I, M_I \rangle$, where D_I is a non-empty domain and M_I is a total function that associates an element of D_I to each constant symbol of L_C and a total n-ary function from

[1] CRIN-CNRS and INRIA-Lorraine, Bâtiment LORIA - B.P. 239, F-54506 Vandoeuvre-lès-Nancy Cedex, FRANCE
[2] CRIL / Université d'Artois, rue de l'Université, S.P. 16, F-62307 Lens Cedex, FRANCE

ECAI 96. *12th European Conference on Artificial Intelligence*
Edited by W. Wahlster
Published in 1996 by John Wiley & Sons, Ltd.

D_I^n to $BOOL$ to each *n*-ary predicate symbol of L_P. As usual with first-order logic languages including equality, the predicate symbol $=$ is interpreted as the identity relation over D_I for every interpretation I. An *assignment* v for an interpretation I is a total function from X to D_I. Let D_I^X denote the set of all the assignments on D_I. The *semantics* of a formula A in an interpretation I is noted $[\![A]\!](I)$; it is a total function from D_I^X to $BOOL$. The semantics of a formula A in an interpretation I given an assignment v for I belongs to $BOOL$ and is noted $[\![A]\!](I)(v)$: it is defined in the usual compositional way. We say that an interpretation I given an assignment v (for I) is a *model* of a formula A iff $[\![A]\!](I)(v) = T$; A is then said to be *satisfied in* I. When a formula A has a model, it is said to be *consistent*. Otherwise, it is said to be *inconsistent*. When every model of A given any assignment v is a model of B given v, B is said a logical consequence of A, noted $A \models B$. When $[\![A]\!](I)(v) = [\![B]\!](I)(v)$ for every interpretation I given any assignment v for I, then A and B are said *equivalent*; this will be noted $A \equiv B$.

3 Induction is Not Only "Reversed Deduction"

While induction is often considered as a form of reversed deduction, it does not reduce to it. For instance, Helft [6] and Flach [3] gave another account for induction, referred to as the nonmonotonic semantics for induction [12], or *confirmatory induction* [4]. In such approaches, generalizing a given set of examples does not consist in deriving hypotheses which explain observations, where explanation is captured by the logical entailment relation. Indeed, in many situations, pointing out the properties that all examples share leads to an interesting form of generalization, and such an inference does not come down to explanatory reasoning.

For instance, if all what I know is that Tweety is a bird, Tweety flies, Superman flies, Garfield is a cat and every cat is not a bird, then provided that every unknown individual behaves like the known ones (i.e. every individual possesses every property shared by both Tweety, Superman and Garfield), I can assume that Woodstock flies whenever I know that Woodstock is a bird. I can also draw the same conclusion with every other bird. More formally,

$$H \triangleq \forall X(bird(X) \Rightarrow flies(X))$$

can be considered as a generalization of

$$E \triangleq \{bird(Tweety),\ flies(Tweety),\ flies(Superman),\ cat(Garfield),\ \forall X(cat(X) \Rightarrow \neg bird(X))\}.$$

While H is not a logical consequence of the examples E, it is satisfied by every known individual.

Clearly enough, E is not a logical consequence of H, hence H cannot be considered an explanation of E (at least, with a deductive meaning): Assuming that every bird flies does not explain at all why Tweety is a bird, why Superman flies, why Garfield is a cat or why every cat is not a bird. Thus, deriving H from E can be considered as induction, but not as explanation.

In contrast to the "explanatory view" of induction, this form of generalization is based on the assumption that every unknown individual behaves like the known ones. Such an assumption is often referred to as an "induction principle" or a "hypothesis space bias". Whenever this principle is correct, the inductive hypotheses derived thanks to such kind of confirmatory induction are satisfied.

Pioneered by Hempel [7], this principle (sometimes called "similarity hypothesis") is the basis on which several proposals for the notion of inductive generalization, including those of Helft [5], [6] and Marquis [11], are based. This principle can be formally translated in the metalanguage of the logic, the language of which is used to represent both examples and generalizations. For instance, if the representation language is first-order logic, we can represent this principle by the second order formula:

$$\forall H((\forall X \in D_E, H(X)) \Rightarrow \forall X, H(X)).$$

where D_E is the set of all constant symbols occurring in E.

Adhering to the "similarity hypothesis" allows one to avoid many problems that occur when the standard explanatory view is considered. Particularly, there are no mutually inconsistent generalizations of a given set of examples; moreover, the fact that there is no more specific generalization in first-order logic in the general case does not create any difficulty in this framework. Finally, there is no general/specific dilemma: in the standard view of generalization, the more general a generalization is, the less plausible (but the more useful) it is. In a similarity-based framework, all the properties (and particularly the most general ones) that are shared by every individual are of interest. Another strong point of the "similarity hypothesis" is that it is not new from a practical point of view: actually, many generalization algorithms, including the well-known ID3 classifiers [13], rely on it (even if this is not stated formally in the seminal papers on ID3).

4 Abduction and Induction w.r.t. Clark's Completion

Interestingly, both abduction and confirmatory induction can be considered as deduction in a completed knowledge base. At the inference level, both reasoning models consist in deriving a set of hypotheses from a set of examples (and a background theory). Hypotheses are characterized intensionally by some completion principles (e.g. by formulas in the metalanguage of the logic used to represent both the examples and the hypotheses) and the problem consists in deriving them extensionally (i.e. in the language of the logic).

From this general point of view, abduction and induction appear closely related and, expectingly, we can benefit from the computational frameworks that have been developed for one of them to emulate the other one (*mutatis mutandis*). For instance, Clark's completion [1] can be used with profit for abduction and induction. Particularly, in the model for confirmatory induction proposed by De Raedt and Bruynooghe [14], Clark's completion plays a central role. Indeed, a clause c is a hypothesis for E if no counter-example of c can be logically deduced from $Comp(E)$, where $Comp(E)$ denotes Clark's completion of E.

For illustration, let us consider the example:

$$E \triangleq \{bird(Tweety),\ flies(Tweety),\ flies(Superman),\ cat(Garfield)\}.$$

On this example, we have:

$$\begin{aligned} Comp(E) \triangleq \{&bird(Tweety),\ flies(Tweety),\\ &flies(Superman),\ cat(Garfield),\\ &\forall X, (bird(X) \Rightarrow (X = Tweety)),\\ &\forall X, (cat(X) \Rightarrow (X = Garfield)),\\ &\forall X, (flies(X) \Rightarrow\\ &((X = Tweety) \vee (X = Superman)))\}.\end{aligned}$$

As $\exists X/bird(X) \wedge \neg flies(X)$ is not satisfied by $Comp(E)$, its negation $\forall X,\ \neg bird(X) \vee flies(X)$ is considered confirmed by E.

Clark's completion can also be used for abductive reasoning [2], [8], [9]. For instance, given a background theory

$$Th \triangleq \{\forall X\ (bird(X) \Rightarrow flies(X))\},$$

$flies(Woodstock)$ has only one non trivial abductive explanation, namely $bird(Woodstock)$. Adhering to the abductive completion principle, we shall assume that $bird(Woodstock)$ holds. Clearly enough, $bird(Woodstock)$ can be deduced from $flies(Woodstock)$ and Th completed with the formula $\forall X\ (flies(X) \Rightarrow bird(X))$ (Clark's completion of Th).

5 A Model for Confirmatory Induction

In this section, we present a model for confirmatory induction. This model can be viewed as deduction w.r.t. a completed theory. It basically relies on direct interpretation of the "similarity hypothesis": known individuals are supposed to be good representatives of all individuals. Our model is basic: In contrast to many other proposals for induction, no additional assumption about the domain of interest is made. Particularly, no predicate is completed in this approach. Accordingly, the completion policy on which this model is based cannot be used for abduction in a simple way.

Our model for induction is closely related to Hempel's definition of confirmation, which is based on the "similarity hypothesis" discussed before. Roughly, a set of examples E confirms a hypothesis H if H is satisfied by every individual of E.

More precisely, Hempel's proposal for the concept of confirmation is [7]:

> E confirms H if and only if E would entail H if the universe were reduced to the set of individuals appearing in E.

Stepping back to the example given in Section 1, "every bird flies" is confirmed by "Tweety is a bird", "Tweety flies", "Superman flies", "Garfield is a cat" and "every cat is not a bird" because every known bird (Tweety) flies.

In order to define the concept of confirmation in a formal way, we need the notion of *expansion of a formula over a domain* [10].

Definition 5.1 *The expansion of a first-order formula Φ over domain D, noted $exp_D(\Phi)$, is inductively defined as follows:*

$$\begin{aligned}
exp_D(\phi) &\triangleq \phi,\ if\ \phi\ is\ ground\\
exp_D(\neg\phi) &\triangleq \neg exp_D(\phi)\\
exp_D(\phi \wedge \psi) &\triangleq exp_D(\phi) \wedge exp_D(\psi)\\
exp_D(\phi \vee \psi) &\triangleq exp_D(\phi) \vee exp_D(\psi)\\
exp_D(\phi \Rightarrow \psi) &\triangleq exp_D(\phi) \Rightarrow exp_D(\psi)\\
exp_D(\forall x\phi) &\triangleq \bigwedge_{d\in D} exp_D(\phi[x \leftarrow d])\\
exp_D(\exists x\phi) &\triangleq \bigvee_{d\in D} exp_D(\phi[x \leftarrow d])
\end{aligned}$$

where $\phi[x \leftarrow d]$ is the formula which results from replacing in Φ every free occurrence of x by d.

As a particular domain, we shall consider the *Herbrand domain* of E, i.e. the set D_E of all constant symbols occurring in E.

For instance, the expansion of the formula

$$H \triangleq \forall X(bird(X) \Rightarrow flies(X))$$

over the Herbrand domain of

$$\begin{aligned}E \triangleq\ &\{bird(Tweety),\ flies(Tweety),\ flies(Superman),\\ &cat(Garfield),\ \forall X(cat(X) \Rightarrow \neg bird(X))\}\end{aligned}$$

is the ground formula

$$\begin{aligned}&(bird(Tweety) \Rightarrow flies(Tweety))\\ &\wedge (bird(Superman) \Rightarrow flies(Superman))\\ &\wedge (bird(Garfield) \Rightarrow flies(Garfield)).\end{aligned}$$

The expansion of a formula Φ without equality over D_E is a ground formula which can be considered as equivalent to Φ, provided that every constant symbol occurring in Φ also occurs in E, when Herbrand models over D_E are considered, only. Equivalently, we may take into account the logical axioms of equality, the domain-closure assumption (the domain elements are all named by constants occurring in E) and the unique-name assumption (different constants of E refer to different domain elements) as implicit axioms in the logic.

On this ground, we can give a formal definition for the notion of confirmation:

Definition 5.2 *Let E and H be two closed first-order formulas. E confirms H, noted $E \mid\!\sim H$, if and only if $E \models exp_{D_E}(H)$ where $exp_{D_E}(H)$ is the expansion of the formula H over the Herbrand domain D_E of E.*

Accordingly,

$$H \triangleq \forall X(bird(X) \Rightarrow flies(X))$$

is confirmed by

$$\begin{aligned}E \triangleq\ &\{bird(Tweety),\ flies(Tweety),\ flies(Superman),\\ &cat(Garfield),\ \forall X(cat(X) \Rightarrow \neg bird(X))\}\end{aligned}$$

since the formula

$$\begin{aligned}&(bird(Tweety) \Rightarrow flies(Tweety))\\ &\wedge (bird(Superman) \Rightarrow flies(Superman))\\ &\wedge (bird(Garfield) \Rightarrow flies(Garfield))\end{aligned}$$

is a logical consequence of E.

We are now ready to define our notion of generalization. Intuitively, a generalization of E is a most general formula which is confirmed by E but is not implied by E. For simplicity, we restrict such formulas to clauses, only.

Definition 5.3 *Let E be a closed first-order formula and H a clause which does not contain any occurrence of the equality symbol. H is a generalization of E if and only if*
- $E \mid\!\sim H$,
- $E \not\models H$,
- *If H' is a clause such that $E \mid\!\sim H'$, $E \not\models H'$ and $H' \models H$, then $H \models H'$.*

Thus, $\forall X(bird(X) \Rightarrow flies(X))$ is a generalization of

$$\begin{aligned}E \triangleq\ &\{bird(Tweety),\ flies(Tweety),\ flies(Superman),\\ &cat(Garfield),\ \forall X(cat(X) \Rightarrow \neg bird(X))\}\end{aligned}$$

Thus, while confirmation is a necessary condition for induction in this framework, it is not a sufficient condition for it. Particularly, every formula from E cannot be considered as a generalization of E (intuitively, it does not convey any new information w.r.t. E).

Requiring that every generalization does not contain any occurrence of the equality symbol allows us to prevent from considering the domain-closure axiom (every individual belongs to D_E) as a generalization. Indeed, the very purpose of induction is to point out general laws that apply to unknown individuals. If the domain-closure axiom were considered as an acceptable generalization, the existence of unknown individuals would be denied; inductive reasoning would be of no use in this situation. Actually, our objective is to derive general laws that would be true if the domain of interest were reduced to the known individuals. The domain-closure axiom is used as a technical medium to derive such laws by forcing the domain to be closed; however, it must not be taken into account: we do not believe in it when induction is concerned.

In our framework, a generalization is a most general clause which is not a logical consequence of E but is a logical consequence of E together with the equality axioms, the domain-closure assumption and the unique name assumptions. Thus, on the running example, H is a logical consequence of E and $\forall X((X = Tweety) \vee (X = Superman) \vee (X = Garfield)$, and $(Tweety \neq Superman) \wedge (Superman \neq Garfield) \wedge (Tweety \neq Garfield)$. Accordingly, our notion of generalization can be viewed as deduction w.r.t. a completed theory. The completion principles can be represented as the domain-closure axiom plus the unique name assumptions in a first-order language with equality. While such a completion policy embodies a circumscriptive flavor, individuals are circumscribed, only (no property is restricted here). This contrasts with the great majority of completion policies introduced so far (mainly, Reiter's closed-world assumption and its generalizations, Clark's predicate completion and various forms of circumscriptions à la McCarthy) and used to define both abduction and induction.

Clearly enough, our model is strongly related to those of Helft [5], [6], Marquis [11], and De Raedt and Bruynooghe [14]. However, in contrast to these approaches, we do not adhere to any additional completion principle here (like Clark's completion, closed-world assumption, or, from a model-theoretic point of view, inference w.r.t. minimal model(s)). Particularly, positive facts and negative ones play the same role in this framework, while this was not the case in the approaches mentioned above.

6 Discussion

While abduction and induction can be considered as deduction w.r.t. a completed knowledge base, the completion principles on which both reasoning models rely are seemingly different: on the one hand, every unknown individual is supposed to share the same properties than the known ones (so we focus on individuals of the domain of interest), and nothing else is required; on the other hand, the explanation of a given property is supposed to be among the known ones (and we focus on the properties of the domain of interest). Expansion is a direct interpretation of the "similarity hypothesis". Clark's completion is a looser interpretation; instead of restricting individuals to the known ones, it leads us to assume that all individuals satisfying a predicate p are those who are already known to satisfy p, i.e. all other individuals do not satisfy p. Thus, the completion achieved thanks to Clark's technique is more than what confirmatory induction basically needs.

While both completion principles are circumscriptive in some sense, they lead to hypotheses that are significantly different from an epistemic point of view. Thus, an abductive hypothesis is a factual piece of knowledge used to explain a given observation (and possibly nothing else); contrastingly, an inductive generalization is a general law, which extends to the unknown individuals some of the properties shared by the known ones (and does not necessarily explain something).

Another significant difference between abduction and induction in this framework lie in the status of completion formulas. While completion formulas can be considered as part of the knowledge base when abduction is concerned, they must be considered only as convenient technical media for induction; particularly, they must not be incorporated in the knowledge base once generalizations have been derived. For instance, completing a knowledge base with $\forall X, (flies(X) \Rightarrow bird(X))$ means that being a bird is the only non trivial cause we know for flying; as far as abduction is concerned, this formula must be added to the knowledge base when it is generated because we believe it at that time (this does not question the possibility to disbelieve it in the future). On the contrary, a completion formula such as $\forall X, (bird(X) \Rightarrow (X = Tweety))$ is only used for technical needs during the derivation of inductive hypotheses and must be rejected then such hypotheses are generated.

REFERENCES

[1] Keith L. Clark, *Logic and Databases*, chapter Negation as Failure, 293–322, H. Gallaire and J. Minker, plenum press edn., 1978.

[2] L. Console, D. Theseider-Dupré, and P. Torasso, 'On the relationship between abduction and deduction', *Journal of Logic and Computation*, **1**(5), 661–690, (1991).

[3] Peter Flach, *Inductive Logic Programming*, chapter A framework for Inductive Logic Programming, 193–210, S. Muggleton, academic press edn., 1992.

[4] Peter A. Flach, 'Rationality postulates for induction', in *Proceedings of the Sixth Conference on Theoretical Aspects of Rationality and Knowledge (TARK 1996)*, ed., Yoav Shoham, pp. 267–281, (1996).

[5] Nicolas Helft, *L'Induction en Intelligence Artificielle : Théorie et Algorithmes*, Thèse de doctorat, Université d'Aix-Mareille II, 14 septembre 1988.

[6] Nicolas Helft, 'Induction as nonmonotonic inference', in *Proceedings of 1^{st} International Conference on Principles of Knowledge Representation and Reasoning*, pp. 149–156, Toronto, (1989).

[7] C. G. Hempel, 'A purely syntactical definition of confirmation', *Journal of Symbolic Logic*, **8**, 122–143, (1943).

[8] K. Inoue, 'Linear resolution for consequence-finding', *Artificial Intelligence*, **56**, 301–353, (1992).

[9] K. Konolige, 'Abduction versus closure in causal theories', *Artificial Intelligence*, **53**, 255–272, (1992).

[10] René Lalement, *Logique, réduction, résolution*, chapter IV.3. Le théorème de Herbrand, 163–168, études et recherches en informatique, Masson, 1990.

[11] Pierre Marquis, 'Building up inductive generalizations from facts', in *Proceedings of 10^{th} European Conference on Artificial Intelligence (ECAI)*, pp. 446–450. Wiley & Sons, (1992).

[12] Stephen Muggleton and Luc De Raedt, 'Inductive logic programming: Theory and methods', *Journal of Logic Programming*, **19**, 629–679, (1994).

[13] J. Ross Quinlan, *C4.5 : Programs for machine learning*, Series in machine learning, Morgan Kaufmann, 1993.

[14] Luc De Raedt and Maurice Bruynooghe, 'A theory of clausal discovery', in *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence (IJCAI'93)*, pp. 1058–1063, (1993).