# A Unified Framework for Abductive and Inductive Reasoning in Philosophy and AI

**Atocha Aliseda***

## 1 INTRODUCTION

In the history of philosophical thought abduction and induction have been frequently confused, either by regarding one as an instance of the other or by not defining the precise scope of each notion. The former confusion may be considered terminological, while the latter refers to the nature of the reasoning processes themselves.

In order to see abductive and inductive reasoning in better focus, we suggest to fix the terminology and identify main features. Our proposal puts forward a general notion of explanatory reasoning with a number of parameters that can be varied to get different kinds of explanatory reasoning. The inferential parameter sets the logical relationship between explananda and explanandum, which may be deductive or statistical. Additional parameters are triggers and outcomes. Triggers are the preconditions on the background theory $\Theta$ and the observation $\varphi$ in need of explanation. $\varphi$ may be a novel phenomenon, or it may actually puzzle the theory $\Theta$. Outcomes are the explanations themselves, outputs of processes for computing explanations. These may have various forms: facts, rules or even theories.

On the basis of these parameters, we construct a taxonomy for different types of explanatory reasoning. These different styles can be recognized both in AI and in Philosophy of Science.

After giving a short history showing the different uses of the terms induction and abduction in Philosophy and in AI, we describe the general schema for explanatory reasoning and explain the different parameters. We end by connecting our approach with that of Peirce, Hempel, Gärdenfors, and Flach.

*Philosophy Department and Center for the Study of Language and Information. Stanford University. Stanford, CA 94305-4115. U.S.A.

## 2 SOME HISTORY

### 2.1 Artificial Intelligence

In the AI literature the term induction has generally being used for the process of learning from examples, and for creating a theory to explain facts [1], thus suggesting abduction as an instance of induction.

As for abduction, this term has been used in a variety of contexts. Although all approaches seem to agree in that $\Theta \not\Rightarrow \varphi$ (the observation is not explained by the theory) is a condition to search for an explanation, proposals differ in their attitude toward $\neg\varphi$, i.e. in their answer to the question: does $\Theta \Rightarrow \neg\varphi$?.

Most approaches in Logic Programming [2] focus on cases were neither the observation nor its negation is explained by the theory ($\Theta \not\Rightarrow \varphi$, $\Theta \not\Rightarrow \neg\varphi$). In contrast, in Diagnostic Reasoning [3] abduction is also used to explain cases in which the observation contradicts a prediction of the theory. That is, cases in which an anomaly between the observation and the theory is found, thus suggesting $\Theta \Rightarrow \neg\varphi$ as the extreme case.

Moreover, recent approaches [4] propose abductive reasoning to include visual aspects of hypothesis formation in the context of visual inference in archeology. Under this view, hypotheses may have complex structures, and not necessarily a sentential form as required by logical frameworks.

### 2.2 Philosophy

Since the time of Mill, the logical status given to all kinds of reasoning from data to explanation has been induction. This continues to be the case even in modern approaches, such as in computational philosophy of science in which induction is understood *"in the broad sense of any kind of inference that expands knowledge in the face of uncertainty; abduction, then, counts as a kind of induction."* [5]. Thagard' s induction is what Peirce renders *ampliative inference* – that which amplifies knowledge – identified with induction and abduction.

On the other hand, some approaches show induction as an instance of abduction. This is when abduction is understood as *inference to the best explanation*, considered by Harman [6] as the basic form of nondeductive inference, which encloses enumerative induction as one case of it. Still a more drastic view was held by Popper who neglected induction altogether, considering it a myth and not a scientific procedure: *"the actual procedure of science is to operate with conjectures: to jump to conclusions – often after one single observation"* [7].

## 3 A GENERAL SCHEMA

A general formulation for explanatory reasoning is the following:

Given a theory $\Theta$ (a set of formulae), and a formula $\varphi$ (which may be a literal or a conjunction of them), $\alpha$ is an explanation if:

1. $\Theta, \alpha \Rightarrow \varphi$
2. $\Theta, \alpha$ consistent.
3. $\Theta \not\Rightarrow \varphi$
   $\alpha \not\Rightarrow \varphi$

This schema is general in that $\Rightarrow$ may be interpreted in a variety of logical ways. However, only the first two conditions: inference and consistency are present in all and every logical account of explanation[1].

Conditions in 3 are interpreted as follows. The first one ($\Theta \not\Rightarrow \varphi$) attempts to capture the condition that $\Theta$ does not explain $\varphi$, thus making an explanation $\alpha$ needed to account for $\varphi$. As for $\alpha \not\Rightarrow \varphi$, it aims to avoid the trivial explanation and forces $\Theta$ to contribute to the explanation as well, so that $\alpha$ does not do all the explaining. This last condition is certainly not needed for enumerative induction, in which the inductive rule does all the explaining of the set of individual observations.

The core of this schema (the first two conditions) under the interpretation of $\Rightarrow$ as classical deduction, is generally identified with abduction in the AI literature. On the other hand, the schema with all three conditions, is very similar to Hempel's models of explanation for singular facts[8]. We will explain this last connection in detail later on.

[1] An additional condition of minimality is often required in formulations of explanation. We are however, only concern in characterizing explanation, not the *best explanation*.

# 4 PARAMETERS

## 4.1 Inference $\Rightarrow$

The notion of inference $\Rightarrow$ is not fixed with classical deduction or semantic entailment. It is indeed a parameter which may be alternatively identified with the notion of statistical inference ($\varphi$ follows from $\Theta$ and $\alpha$ with high probability). Under this interpretation, the schema has the flavour of an inductive argument (induction, in this case, understood as statistical inference).

There may also be other interpretations of $\Rightarrow$ such as dynamic inference, under which the notion of truth is replaced by that of information change potential, forcing $\alpha$ to introduce some informational change when added to $\Theta$.

## 4.2 Triggers ($\Theta \Rightarrow \neg\varphi$ ?)

Explanatory reasoning is triggered by a surprising phenomenon, by a fact in need of explanation. The condition of being *'surprising'* seems to have been already captured by one condition in our schema ($\Theta \not\Rightarrow \varphi$). However, the type of trigger for explanation concerns also the status of the negation of $\varphi$ (i.e. does $\Theta \Rightarrow \neg\varphi$?).

In science, new discoveries are not just the result of finding a novel phenomenon, but also the result of finding an anomaly, a phenomenon which puzzles the theory.

We distinguish two ways of being surprising. That is, two different triggers for explanatory reasoning. Trigger by *novelty* and trigger by *anomaly*. The first one consists of finding a novel phenomenon $\varphi$ which is consistent with the theory ($\Theta \not\Rightarrow \neg\varphi$). The second trigger consists of finding an anomaly, a phenomenon which is in conflict with some statement of the theory. An extreme case of the latter is when the theory actually explains the negation of the phenomenon observed ($\Theta \Rightarrow \neg\varphi$).

The case of finding a novelty has already been characterized in the computational literature of abduction as the condition for an *abductive problem* ($\Theta \not\Rightarrow \varphi$, $\Theta \not\Rightarrow \neg\varphi$). Therefore, my suggestion is to incorporate a second kind of trigger, that of trigger by anomaly ($\Theta \not\Rightarrow \varphi$, $\Theta \Rightarrow \neg\varphi$).

## 4.3 Outcomes

Explanations may have different forms and structures. Sometimes it might be enough to produce a simple fact to explain a surprising phenomenon. In other cases, a rule or a whole new theory has to be devised to account for a puzzling observation.

Let me illustrate this point with two sample programs in Prolog:

1. Abducing a Fact:
   - Program $\Theta$:
     lawn-wet :- rain.
     lawn-wet:- sprinklers-on.
   - Goal $\varphi$: lawn-wet.
   - Explanations $\alpha$'s:
     rain, sprinklers-on.
2. Abducing a Rule:
   - Program $\Theta$:
     rain : - nimbrostatus-clouds.
     nimbrostatus-clouds.
   - Goal $\varphi$: lawn-wet.
   - Explanation $\alpha$: lawn-wet :- rain.

In the first case, the facts 'rain' or 'sprinklers-on' are enough to account for the goal, while in the second a rule is needed to propose a connection between the lawn being wet and rain.

Triggers and outcomes are, respectively, the preconditions and the outputs of explanatory procedures, whatever these turn out to be. As Ai research has shown, these algorithms may be very diverse. Examples to generate hypotheses are linear resolution and the use of ATMS. Our aim however, is not to characterize different types of processes for computing explanations.

# 5 Related Work

## 5.1 Peirce

Although Peirce's abductive formulation has been the point of departure for studies on abductive reasoning in AI, no attention has been paid to what he said about logic and reasoning in his more mature work.

Peirce held a broad view of logic, considering on the one hand, abduction, induction, and deduction all logical in nature and as three stages of inquiry composing a method. In describing how inquiry proceeds, the order proposed is as follows: abduction, deduction, induction: *"from a suggestion deduction can draw a prediction which can be tested by induction"*[9, 5.171].

Peirce also distinguished different types and *degrees* of reasoning. He explicitly states three kinds of induction and two kinds of deduction[9, 7.204,7.208]. As for abduction, he uses the term in contexts as varied as perception and invention. For Peirce perceptual judgments are *"acritical abductions, an extreme case of abductive inference"*. [9, 5.181]. On the other hand, the abductive process of constructing a hypothesis to explain a scientific phenomenon is regarded to be much more elaborated [9, 5.513].

Peirce's writings present explanatory reasoning as a complex notion, covering a multitude of phenomena in which inquiry is called for. At first sight, Peirce's reflections do not seem to have much impact on AI, a field limited to issues to be described in computational terms. However, the ideas of having several degrees for each type of reasoning and of proposing an order in which they occur may give some insight to recent discussions in AI. These ideas have certainly inspired our own approach.

## 5.2 Hempel

The idea of a general schema for deductive and inductive explanation is given by Hempel through his Deductive-Nomological (D-N) and Inductive-Statistical (I-S) models[8]. His presentation is that of an argument form in which the structure is essentially the same for both types of explanation, the only change being for I-S explanation that the theory $\Theta$ has statistical rules rather than universal, thus forcing the inference to be only highly certain.

In addition to the general argument form, Hempel sets the conditions for explanation (our conditions 3). Therefore, our schema captures both of Hempel's models: the Deductive-Nomological when $\Rightarrow$ is interpreted as classical entailment, the Inductive-Statistical one when $\Rightarrow$ is interpreted as statistical inference.

In the case of an inductive-statistical explanation there is also the requirement of maximal specificity, which Hempel's treats outside of his schema. We believe, however that it may be incorporated as part of a notion of consistency for statistical inference (actually Hempel does not require a condition of consistency explicitly, though he assumes it).

## 5.3 Gärdenfors

The framework exposed in this paper has strong connections with belief change, as originally defined by Gärdenfors [10] through the operations of expansion and revision.

If an explanation $\alpha$ is to be incorporated into the theory $\Theta$, the case of the novel phenomenon $\varphi$ induces an expansion, i.e. $\alpha$ is incorporated by a simple addition. On the other hand, the case of finding an anomaly $\varphi$ induces a revision. For $\alpha$ to be incorporated into $\Theta$, a revision of $\Theta$ is needed first in order to prevent $\Theta \Rightarrow \neg\varphi$ (i.e. find a $\beta$ such that $(\Theta - \beta) \not\Rightarrow \neg\varphi$). Both triggers activate a search of an explanation, but their different nature point to quite different operations for incorporating an explanation.

### 5.4 Flach

Flach's approach to explanatory reasoning is also very broad, and he indeed suggests the inference relationship to be a logical parameter: *'a more comprehensive view of inductive hypothesis formation can be obtained by viewing the logical relation between evidence and hypothesis as a parameter, that can be instantiatiated to explanation by deductive entailment as proposed by Peirce, but also to, for instance, explanation by plausible entailment, or to quite different relations, such as the relation of confirmation studied by Hempel.*[11].

Our approach as well as his seems to be very much inspired by Hempel's work. It is interesting to note however, that while Flach uses Hempel's notion of confirmation, ours concentrates in the notion of explanation.

Moreover, the different styles of explanatory reasoning that can be produced with the different parameters described so far, may be studied by logical means. That is, by analyzing their logical structural properties.

This line of analysis is proposed and developed by Flach[11] building on the approach proposed in [12] for non-monotonic consequence relations. Ours is built as proposed for dynamic styles of inference[13][2]. However, these two approaches have the same focus. Indeed, it would be interesting to compare the different notions of explanatory consequence that we have independently set forward.

## 6 Conclusions

We proposed abductive and inductive reasoning as instances of explanatory reasoning described by a general schema in which the inferential relationship between explananda and explanandum is a parameter which can be either deductive or statistical[3].

Moreover, types of explanatory reasoning can be characterized by their triggers and outcomes. We distinguished two types of triggers (novelty and anomaly) and three degrees of outcomes (facts, rules, theories). Combination of all these items gives us six types altogether (for each interpretation of $\Rightarrow$): for each kind of trigger, three degrees are distinguished. An explanatory process may go from facts to facts, from facts to rules or from facts to theories.

This characterization is important because if we want to be able to combine abduction and induction, knowing the triggers for each kind of reasoning as well as what they produce, may help design procedures that puts them in sequence.

Still we have not said explicitly where does abduction and induction, – as understood in AI – fit in our analysis?. Enumerative induction is the case when the reasoning is triggered by repeatedly finding a novelty in a set of observations. And the outcome is a rule that generalizes the novel property found in the observations. The inferential relationship is statistical. Abduction, on the other hand, covers the cases when the reasoning is triggered by a single fact (either by novelty or anomaly) and the outcome is also a fact.

Several extensions to existing proposals in the AI literature may be accommodated

[2] Though in this paper we have not presented such results.

[3] The details of how to define the latter type of inference were not explained here, however.

within our approach, such as allowing to abduce rules in Logic Programming or deal with the problem of anomalies by incorporating the role of revision in explanation construction.

This brief note also attempts to show that Philosophy of Science and AI share some of their concerns and that they may benefit from each other. The philosophical tradition reminds AI of the complexity of explanatory phenomena, research in AI proposes specific algorithms for explanatory models.

# ACKNOWLEDGEMENTS

I would like to thank the organizing committee of this Workshop for their comments, and Johan van Benthem for his guidance and encouragement while writing this paper. In addition, I am very greatful to the Archimedes project at CSLI, Stanford University for letting me use their speech recognition system, which allowed me to 'talk' this paper since the first draft.

# References

[1] Shapiro, E. Inductive Inference of Theories from Facts. Contained in *Computational Logic: Essays in Honor of Alan Robinson.* Ed. by J. L. Lassez and G. Plotkin. Cambridge, Mass. MIT. 1991.

[2] Kakas, A.C., Kowalski, R.A., Toni, F. Abductive Logic Programming. *Journal of Logic and Computation* 2(6) (1993) 719-770 (1992).

[3] Peng, Y. and Reggia, J. *Abductive Inference Models for Diagnostic Problem Solving.* Springer Verlag, New York. 1990.

[4] Thagard, P. *Abductive Reasoning: Logic, Visual Thinking, and Coherence.* University of Waterloo. Ontario, Canada. 1995.

[5] Thagard, P. *Conceptual Revolutions.* Princeton, NJ, PUP. 1992.

[6] Harman, G. The Inference to the Best Explanation. *Philosophical Review.* 74 88-95 (1965).

[7] Popper, K. *Conjectures and Refutations. The Growth of Scientific Knowledge.* 5th ed. London and New York. Routledge. 1963.

[8] Carl Hempel. *Aspects of Scientific Explanation.* Free Press, New York. 1965.

[9] Peirce, C. S. *Collected Papers of Charles Sanders Peirce.* Charles Hartshorne, Paul Weiss, and Arthur W. Burks (eds.) 8 vols. Cambridge, MA: Belknap Press, Harvard University Press, 1931.

[10] P.Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States.* MIT Press, 1988.

[11] P.A. Flach. *Conjectures: an inquiry concerning the logic of induction.* PhD Thesis, Tilburg University, 1995.

[12] S. Kraus, D. Lehmann, M. Magidor. Nonmonotonic Reasoning Preferential Models and Cumulative Logics. *Artificial Intelligence,* 44 (1990) 167–207.

[13] J. van Benthem. General Dynamics. *ILLC Prepublication Series* LP-90-11. Institute for Logic, Language, and Information. University of Amsterdam. 1990.