Abduction of Generalizations

November 15, 2011

Abstract

Abduction of generalizations is the process in which explanatory hypotheses are formed for an observed, yet puzzling generalization such as "pineapples taste sweet" or "rainbows appear when the sun breaks through the rain". This phenomenon has received little attention in formal logic and philosophy of science. The current paper remedies this lacuna by first giving an overview of some general characteristics of this process, elaborating on its ubiquity in scientific and daily life reasoning. Second, the adaptive logic $\mathbf{LA}_{\forall}$ is presented to explicate this process formally.

keywords: hypothesis formation – abduction patterns – adaptive logics

1 Introduction

Abduction is generally defined as "the process of forming an explanatory hypothesis" [22, p. 216]. In this paper we will focus on a specific "pattern of abduction" (to use a phrase introduced by Schurz [23]). Consider the following example [23, p. 212]:

$$\begin{array}{l} \text{Pineapples taste sweet.} \\ \text{Everything that contains sugar, tastes sweet.} \\ \hline \text{Pineapples contain sugar.} \end{array}$$

Inferences of this kind have been called "law abduction" [23], or "rule abduction" [24]. As "law" and "rule" are heavily debated concepts in philosophy (of science), we will use the more neutral term *abduction of generalizations* (henceforth AG) for this specific pattern. More examples and a general characterization of AG will be presented in Section 2. It will be argued that this pattern is ubiquitous in both everyday and scientific reasoning, and is commonly recognized as a useful – be it fallible – way to extend one's knowledge.

Notwithstanding the importance of AG, little effort has been made so far to study the characteristics of this inference pattern, and to explicate it by means of a formal logic. As will be explained in Section 2.2, most scholars in AI and formal logic have focused on singular fact abduction, whereas philosophers of science have taken a more general, but informal point of view on abduction. It is our aim to treat AG as a distinct subject matter, and to see how one may understand and formalize it.

Outline A first analysis of AG is provided in Section 2. We describe this pattern informally, showing that it is a widespread inference pattern; secondly, we explain why it has been neglected in formal logic and philosophy of science; finally, we argue for the specific importance of AG in scientific contexts.

In Section 3, we turn our focus to problems that emerge when representing AG formally. We argue that a distinction in the object language is needed between what we call *mere generalizations* and the *explanatory framework* for any logic that models AG; and, moreover, that this distinction is useful in any logic for abduction. As AG is a non-monotonic inference form, we also discuss how the dynamic features can be represented.

In the last section before the conclusion, Section 4, the logic $\mathbf{LA}_\forall$ is presented. This is a logic for AG, formulated in the standard format of adaptive logics. After we argue why this framework is well-suited for the current application, we will illustrate the proof theory of $\mathbf{LA}_\forall$, which allows us to model the dynamic interaction of AG and classical inferences.

Preliminaries Let $\mathcal{L}$ be the standard language of classical first-order predicate logic, obtained from a set of constants $\mathcal{C} = \{a, b, c, \ldots\}$, a set of variables $\mathcal{V} = \{x, y, z, \ldots\}$, a set of predicates $\mathcal{P} = \{P, Q, R, \ldots\}$, the connectives $\neg, \vee, \wedge, \supset, \equiv$ and quantifiers $\forall, \exists$. $\mathcal{W}$ is the set of formulas in $\mathcal{L}$. Depending on the context, A, B, C are used either as metavariables for members of $\mathcal{W}$, or for (conglomerates of) predicates, e.g. $(P \wedge Q) \vee (\neg R)$. The metavariables $\alpha, \beta, \ldots$ refer to constants and variables.

2 Abduction of General Explanations

2.1 The phenomenon

We define *abduction of generalizations* (AG) as every inference that fits the following pattern:

> It is observed that all A are B.[1]
> Also, being C is regarded as an explanation for being B.
> Therefore, the hypothesis that all A are C is raised.

Hence, by AG we generate hypotheses that explain why all observed objects of a certain class have a specific property. In Section 3, we will explain how this definition can be operationalized in a first-order modal language. But first, let us point out some general characteristics of AG.

First of all, AG is a specific form of classical abduction as defined by Peirce [21, 5.189]:

> The surprising fact, X, is observed;
> But if Y were true, X would be a matter of course,
> Hence, there is reason to suspect that Y is true.[2]

[1]Strictly speaking, this is shorthand for "All observed A are B, and therefore it is believed that all A are B."

[2]To avoid confusion, the schematic letters A and C originally used by Peirce are replaced by X and Y.

In an AG both the surprising fact X and the hypothesis Y are generalizations, respectively "all A are B", and "all A are C".

This definition leads to an important consideration about the *Peircean* or *classical* notion of abduction: it is defined in a *deterministic* way, i.e. the truth of Y implies X. Although we do not suggest that this notion of abduction cannot be meaningfully extended to other accounts in which the motivation to adopt the abductive hypothesis is, for instance, probabilistic ($P(X|Y)$ is high) or comparative ($P(X|Y) > P(X|\neg Y)$), we restrict ourselves in this paper, as most of the literature on abduction does, to the classical case. As it is also assumed that Y explains X,[3] this restriction will have consequences for the formalization of AG in Section 3.1.

Secondly, AG is distinct from what is called *singular fact abduction*, in which both the surprising fact and the hypothesis are singular facts. In a first-order language, the surprising fact of a singular fact abduction is modeled as an object having a certain property (such as Pa). In contrast, in AG it will be modeled by a generalization (such as $(\forall x)(Px \supset Qx)$). Existing models for abduction usually limit themselves to singular fact abduction, as we will see in the next section.

Thirdly, AG is not a novel reasoning pattern. It has been known at least since Aristotle who treats something similar in his *Posterior Analytica* when he considers the "middle term" of a definition. This pattern is, according to his view, the essence of a good definition: it should not only say what the *definiendum* (A) is, it should also be an explanation (C) for its observed properties (B). As an example, he explains why horned animals (A) lack upper incisors (B) by defining horned animals as a subclass of animals that have inflected hard material from their mouth to their heads (C). According to Aristotle, this is a good definition of a class because it explains the properties of that class.[4] However, the reasoning pattern we are considering is much broader than what Aristotle had in mind. A, B and C can be any properties, and neither should A be a definiendum, nor C a definiens.

Fourthly, AG is frequently applied in human reasoning, often in combination with or following an instance of singular fact abduction. For instance, people do not only wonder why their heads hurt (they drank too much last night) or why there is a thunderstorm (it was very hot during the day). Not much of a reflective mind is needed to also start asking questions such as why it is that every time one drinks a bit too much, one suffers from headaches, or why thunderstorms often follow hot days. In other words, people do not only wonder why certain facts are the case, they also wonder why certain regularities occur.

[3] Applying the above schema as such is only justified in case of abduction, i.e. the formation of explanatory hypotheses. If Y does not explain X, flagrant examples of the logical fallacy *affirming the consequent* that have little value qua hypothesis will be obtained.

[4] See [2, II.10] for Aristotle's distinction between two types of definitions and [2, II.12-14] for his view on the role of the middle term in a definition. A good treatment of the analogy between Aristotelian definitions and Peircean abduction can be found in [10]. In our opinion, Schurz refers in [23] to the wrong concept when he links AG (in his words: law abduction) to Aristotle. The concept "hitting upon the middle term" is only employed in the definition of quick wit [2, I.34], in which it is illustrated with an example of a singular fact abduction. In our view, a predecessor of AG can only be found in Aristotle's treatment of the role of the middle term in definitions.

2.2 The Lack of Models for AG

The lack of models for AG will be explained by pointing out how the application of the concept of abduction in a variety of fields has caused a growing divergence in definitions and interpretations. This will also clarify the relation between our current project and the literature on abduction.

Broadly speaking, two main currents in research on abduction can be discerned. On the one hand, research in AI and formal logic mostly focuses on a *syllogistic* interpretation of Peirce's work, in which abduction is introduced as part of a tripod that is clarified with the following famous beans-example of Peirce [21, 2.623]:

All the beans from this bag are white. (Rule)
These beans are from this bag. (Case)
These beans are white. (Result)

All reasoning deriving a result from a case and a rule is called *deductive*, all reasoning deriving a rule from a case and a result *inductive*, and all reasoning deriving a case from a rule and a result *abductive*. Having this schema in mind, researchers in AI or formal logic generally focus on instances of singular fact abduction, which are variations on the following pattern:

$$B\alpha, \forall\beta(A\beta \to B\beta)/A\alpha$$

This pattern is usually combined with the condition that the hypothesis should be explanatory. Aliseda even adds a further condition suggested by Peirce, i.e. that the observed fact should be surprising (in the sense that $B\alpha$ cannot be derived from the background theory alone) [1]. One noteable exception to the exclusive focus on singular fact abduction is Thagard [24]. He obtains a reconstruction of AG, which he calls "rule abduction", by adding to his logic program PI the ability to generalize the results of a singular fact abduction.

On the other hand, research in philosophy of science usually departs from a *methodological* interpretation of Peirce. In his later writings Peirce distinguishes abduction, induction and deduction as different steps in a methodology of science [22, p. 212–218]. Abduction is the process of forming an explanatory hypothesis, from which deduction can draw predictions, which then can be tested by induction.[5] Research in this tradition, see e.g. [17, 23], considers abduction as a very broad concept including analogical reasoning, visual abduction, common cause reasoning, etc. Here, Peirce's definition of abduction (see Section 2.1) is seen as an expression in metalanguage, in which a "fact" could be any proposition. Some, see e.g. [12, 16, 9], still try to capture the concept of abduction under the single schema of *inference to the best explanation* (IBE).[6] However, these attempts to reduce the broadness of the considered concept prevent the discovery of interesting features of more specific patterns of abduction. Schurz explains this as follows [23, p. 205]:

[5] It is generally acknowledged (see e.g. [11, p. 5–8]) that both interpretations can be found in Peirce's work, although they are not fully compatible. They represent an evolution in his thinking, as he hinted himself when he remarked that he "was too much taken up in considering syllogistic forms" [21, 2.102].

[6] These scholars consider Peirce's remark that abduction should be as economical as possible [21, 7.220], as an essential and crucial condition.

> The majority of the recent literature on abduction has aimed at *one most general* schema of abduction (for example IBE) which matches every particular case. I do not think that good heuristic rules for generating explanatory hypotheses can be found along this route, because these rules are dependent of the specific type of abduction scenario.

In this article, Schurz subsequently presents a taxonomy of distinct patterns of abduction. Having this in mind, we think that it is best to remain pluralistic on the logical form of abduction. We should maintain the rich concept of abduction as it is understood in the philosophy of science, but, in order to provide the formal rigor which is characteristic of the logic and AI community, we have to focus on each of the different specific forms of abduction separately.

2.3 The Ubiquity of AG in Scientific Practice

At the end of Section 2.1, we mentioned several examples in which abduction of generalizations is triggered by a question concerning the result of a singular fact abduction. This question is brought up by a need for a deeper understanding of the observed relations. We can recognize this curious spirit in the endeavors of many scientists. For instance, Descartes was not satisfied with the folk explanation of the rainbow, i.e. that a rainbow appears because the sun breaks through shortly after a rain shower. He wanted to understand why rainbows appear whenever the sun shines while it rains. We will argue that AG is at least as important in scientific practice as singular fact abduction by considering two general characteristics of this practice.[7]

Firstly, in scientific practice one attempts to formulate theories, which have both a *universal* and *falsifiable* nature.[8] One does not want an explanation why, for instance, this particular person suffers from this disease. One wants to understand why and how this disease is transmitted in general. Formulating theories about particularities is seldom considered as good scientific practice; such theories are often labeled as *ad hoc*. Theories are thus mainly formulated for a whole class of objects and, by consequence, formulated in terms of generalizations. These generalizations allow us to derive singular fact predictions by means of which theories can be tested. Therefore, in the formation process of such theories, reasoning methods resulting in generalizations, such as inductive generalization or AG, are essential.

Secondly, augmented *unification* (as characterized, for instance, by [15]) is generally seen as an indicator of scientific progress.[9] Each application of AG is in essence a unification step, because it explains an observational generalization, e.g. "All A are B", by characterizing its antecedent (A) as a subclass of a more general class (C) for which the observed properties (B) hold. Therefore, AG is a key method to enhance unification in scientific practice. The most

[7]This claim is about *scientific practice* and not about *scientific explanation*. In scientific explanation, a scientific theory is employed to explain a certain fact (which can be either a singular fact or a generalization). Scientific practice is the activity of forming such scientific theories and expanding current scientific knowledge.

[8]Universality should not be taken as an absolute notion, but as an achievable level of generality that is relative to the methods and scope of the specific field.

[9]Both the instrumentalist and realist view concerning the nature of scientific progress seem to agree on this point [20].

interesting examples in the history of science can be found when a new theory is proposed as a solution for some anomalies of an existing theory. In that case, the proponents of the new theory also need to show that most of the already known and well-tested observational laws, which are explained by the old theory, can be explained by the new theory. For instance, Newton could explain Huygens' pendulum law using his general laws of motion by pointing out how the different parameters of the pendulum law could be translated into his general mathematical framework. In the same way, Bohr could explain by means of his atomic model why the wave lengths of the visible emission spectrum of hydrogen can be calculated by the Balmer formula.

3 Introducing the Formal Framework

3.1 The Explanatory Framework

According to the definition of AG from Section 2, this pattern could be formally explicated as follows:

(P1) $\forall x(Ax \supset Bx)$
(P2) $\forall x(Cx \supset Bx)$
(H) $\forall x(Ax \supset Cx)$

However, we must be careful here: the definition stipulates that C-hood explains B-hood, not just that everything that has the property C also has the property B. In other words, where (P1) and (H) can be of any kind, the set of possible candidates for (P2) is restricted.[10] We call this set the *explanatory framework*. It consists of all generalizations of the form $\forall x(Fx \supset Gx)$ where being F provides an explanation for being G. Whether or not a generalization belongs to the explanatory framework, may depend on the phenomenon we are trying to explain. In other words, it is contextually defined. All we assume is that it is clear for each generalization, given the abductive problem at hand, whether it is a member of the explanatory framework or not. In the latter case we call it a *mere generalization*.

With this new terminology, we are now able to characterize all the lines of the above schema: (P1) is the *explanandum*, i.e. the mere generalization that is to be explained; (P2) is a generalization that is part of the explanatory framework for the current abductive context; (H) is the *explanatory hypothesis*. An *explanation* or *explanans* for (P1) consists of an explanatory hypothesis together with one or more elements of the explanatory framework that connect the hypothesis to the explanandum.

Now what does it actually mean that F-hood explains G-hood? Needless to say, the philosophical literature abounds in theories of explanation. However, as we chose to restrict ourselves to classical abduction, certain preconditions apply. First, F-hood does not only explain G-hood, F-hood should also imply G-hood. Second, as abduction is an inference, only argumentative accounts of explanation are relevant. Hence, the choices to explicate the notion of "explanation" in the definition of the *explanatory framework of a (classical) abductive problem*

[10] In our opinion, Schurz [23] puts too little emphasis on this point in his discussion of AG, or "law abduction" as he calls it. In his schema, (P2) is called a "background law", but as far as we see, no explicit definition or circumscription is provided.

are limited to accounts of explanation that have the structure of a deductive argument such as a DN-argument (e.g. Hempel [14]), a causal argument (e.g. Hausman [13]) or an augmented unification argument (e.g. Kitcher [15]).[11]

In any of these accounts, (P2) has a specific status – it must be either lawlike, refer to an underlying causal mechanism, or be a more general argumentation scheme. We use the more abstract term *explanatory framework* to express this status of (P2). This specific status turns AG into a fundamentally *asymmetric* inference. It is not possible to derive $\forall x(Cx \supset Ax)$ from the same premises, since A-hood does not explain B-hood. Hence, if a logic explicates AG, it should be able to represent this asymmetry between (P1) and (P2) in its object language.

Before we explain how this can be done, let us briefly give an extra reason to motivate the distinction between the explanatory framework and mere generalizations as a valuable asset for any logic that models abductive processes in general. Mere generalizations are often used in abductions that involve knowledge about methods or procedures. Consider the following premises:

(P1) The Geiger counter produces audible clicks close to the object a.
(P2) If the Geiger counter produces audible clicks, β-radiation is present.
(P3) If an object contains C-14, β-radiation is emitted.

Without the distinction between the explanatory framework and mere generalizations, a logic for singular fact abduction treats (P2) and (P3) as having the same formal structure. But a physicist interested in explaining the presence of β-radiation is only interested in the hypothesis suggested by (P3), as the behaviour of the Geiger counter provides no explanation. On the other hand, (P2) is needed to derive the fact that there is β-radiation in the first place (as it is not directly observable). Hence, (P2) cannot be omitted from this abductive reasoning context. Only a logic that is able to represent explanatory frameworks can handle this case properly.

3.2 A Modal Approach

In Section 4, we will present the logic $\mathbf{LA}_{\forall}$. This system is a non-monotonic extension of the well-known logic $\mathbf{T}$, and allows us to model instances of AG in the modal language $\mathcal{L}^{\Box}$. Therefore, let us first define $\mathcal{L}^{\Box}$ and $\mathbf{T}$ formally, after which we add some comments on our choice for them.

Let $\mathcal{L}^{\Box}$ denote the extension of $\mathcal{L}$ with the modal necessity operator $\Box$. The set of formulas $\mathcal{W}^{\Box}$ is the smallest set for which the following holds:

For all $A \in \mathcal{W}$: $A, \Box A \in \mathcal{W}^{\Box}$
For all $A, B \in \mathcal{W}^{\Box}$: $\neg A, A \vee B, A \wedge B, A \supset B, A \equiv B \in \mathcal{W}^{\Box}$

Hence, $\Box\forall x(Px \supset Qx)$ and $Pa \vee \Box\exists x(\neg Rx)$ are, for instance, members of $\mathcal{W}^{\Box}$, whereas $\Box\Box\forall xPx$ or $\forall x\Box Px$ are not.

An axiomatization for the predicative version of $\mathbf{T}$ over the language $\mathcal{W}^{\Box}$ is obtained by taking the axioms of classical predicate logic (henceforth $\mathbf{CL}$), and adding the following axioms (closed under *modus ponens*):

[11] It is not implied that there are no other valuable accounts of explanation. We only claim that (classical) abductive hypotheses (the only ones that are our concern here) are part of a deductive argument that forms an explanation for the explanandum.

K $\quad \Box(A \supset B) \supset (\Box A \supset \Box B)$
RN $\quad$ if $\vdash A$ then $\vdash \Box A$
T $\quad \Box A \supset A$

A semantics of **T** that is sound and complete with this axiomatization can be found in [8, pp. 46-47].

The language $\mathcal{L}^{\Box}$ allows us to represent the premises involved in abductive reasoning processes with the expressive power of classical first-order logic, but gives us the extra operator $\Box$, which allows us to indicate at the object level that a certain generalization is in the explanatory framework. Let $\mathcal{F}^{\circ}$ denote the set of *purely functional formulas*, i.e. formulas that do not contain individual constants, quantifiers, or sentential letters. For example, $Px \wedge (Qxy \vee Rx)$ is a purely functional formula, whereas $Pa \vee Qxy$ and $Px \wedge \exists y Qxy$ are not. Where $A \in \mathcal{F}^{\circ}$, let $\forall A$ be the universal quantification over every variable that is free in A. The logic $\mathbf{LA}_{\forall}$ treats any formula of the form $\Box\forall(A \supset B)$ with $A, B \in \mathcal{F}^{\circ}$ as an element of the explanatory framework.

The choice for **T** as the monotonic core of $\mathbf{LA}_{\forall}$ has two important consequences. First of all, since **T** is an extension of **CL**, our logic for AG assumes both our factual knowledge and our explanatory framework to be internally consistent. Moreover, since the axioms of **CL** are valid in the scope of $\Box$, classical logic consequences of the explanatory framework may themselves be used to generate explanatory hypotheses. For instance, if $\Box\forall x(Px \supset Qx)$ and $\Box\forall x(Qx \supset Rx)$ are premises of a particular abductive problem, not only these formulas but also $\Box\forall x(Px \supset Rx)$ will be part of the explanatory framework. Second, in view of the axiom T ($\Box A \supset A$), a generalization that is part of the explanatory framework is also assumed to be true as such. This is the formal expression of our restriction to the classical account of abduction, where "A explains B" implies "A implies B".

Our logic for AG is in a sense minimal: iterations of boxes are excluded, and explanation is expressed by rather simple formal tools. It is a topic for further research whether our model can be meaningfully extended to include specific, more fine-grained accounts of explanation (e.g. adding asymmetric axioms to specify causal arguments in the sense of Hausman [13]).

3.3 The Dynamics of AG

Apart from the distinction between the explanatory framework and mere generalizations, several other difficulties arise when we try to model abduction in general, and AG in particular. First of all, abduction is a non-monotonic reasoning method: new information may contradict the hypotheses we have raised. Moreover, it may not always be clear whether the currently available information contradicts some of these hypotheses - this requires classical inferences, which might not yet have been drawn. As a result, we can discern a double dynamics in abductive reasoning: previously drawn inferences can become retracted in view of additional premises, but also in view of further inferences from the same body of evidence. A formal logic for AG should hence be able to frame this double dynamics, yet still define a sensible and stable output for any given premise set.

Second, every realistic model of AG should allow us to combine deductive (or classical) inferences with ampliative (or supraclassical) steps. That is, it should allow the user to draw new inferences on the basis of previously inferred hypotheses, and it should allow the classical consequences of the evidence to falsify such hypotheses (and whatever we derived from them). This relates to a third important desideratum, i.e. that the hypotheses yielded by a formal logic for AG should be mutually consistent with the evidence and the explanatory framework. Ampliative reasoning should not only allow us to go beyond the mere deductive consequences of our knowledge, but it should also remain within the boundaries of consistency.

The fourth problem is specific to the context of abduction: explanatory hypotheses should be as parsimonious as possible. That is, if Y suffices to explain the puzzling fact X, then we should not raise the explanatory hypothesis "Y and Z".

Finally, any logic for abduction should be able to handle cases of multiple explanatory hypotheses in a consistent and uniform way – see Section 4.3 where we discuss two distinct ways in which this can be done.

We chose to use the framework of adaptive logics (henceforth ALs) to formulate the logic $\mathbf{LA}_\forall$ for AG. ALs are powerful formal systems that explicate various forms of defeasible reasoning such as reasoning on the basis of inconsistent premises [5], inductive generalization [7], reasoning on the basis of conflicting norms [3], etc. Several adaptive logics have also already been developed for singular fact abduction [18, 4], and all of them were shown to meet the above desiderata.

One of the most important developments within the AL program is the definition of a canonical format, the so-called *standard format* for ALs. This format encompasses a generic dynamic proof theory and a selection semantics. A rich and attractive metatheory has been shown to hold generically for all ALs in standard format (see [6]): they are sound and complete, have the reassurance property, their consequence relation is idempotent, cautiously monotonic, etc. Most ALs have been successfully expressed within this format, so it provides a good basis for a unifying study of defeasible reasoning forms in general, and patterns of abduction in particular.

The main motivation to choose this non-monotonic framework is its dynamic proof theory, which enables us to construct proofs that are very similar to actual human reasoning processes, as will become clear from the examples in Section 4. There we will also argue that each of the other desiderata from the current section are met by $\mathbf{LA}_\forall$.

4 The Logic $\mathbf{LA}_\forall$

4.1 The Definition of $\mathbf{LA}_\forall$

Let us briefly explain the general characteristics of adaptive logics in standard format. These are characterized by a triple $\langle \mathbf{LLL}, \Omega, \mathbf{x} \rangle$. The so-called *lower limit logic* **LLL** is a monotonic, reflexive and transitive logic, the rules of which are unconditionally valid in the AL. Ω is called the *set of abnormalities*; this set is specified in terms of a logical form. Every AL strengthens its **LLL** by allowing for a specific kind of defeasible inferences, which are determined by Ω and the *strategy* **x**. This will be clarified below.

The adaptive logic $\mathbf{LA}_\forall$ employs **T** as its lower limit logic. The set of abnormalities of $\mathbf{LA}_\forall$ requires a bit more explanation. Consider once more the inference schema of AG, using the formal representations introduced in Section 3:

(P1) $\forall(A \supset B)$
(P2) $\Box\forall(C \supset B)$
(H) $\forall(A \supset C)$

The $\mathbf{LA}_\forall$-abnormalities are all formulas which imply that the premises in the above schema are true, whereas its conclusion is false, for a particular A, B and C. To simplify notation, we introduce the following abbreviation:

$$A \not\rightarrow_C B =_{\mathsf{def}} \forall(A \supset B) \wedge \Box\forall(C \supset B) \wedge \neg\forall(A \supset C)$$

According to this definition, $A \not\rightarrow_C B$ can be read as: "although all A are B, and although C-hood explains B-hood, it is not the case that all A are C". Using this abbreviation, we can now define the set of abnormalities:

$$\Omega = \{A \not\rightarrow_C B \mid A, B, C \in \mathcal{F}^\circ \text{ and } A, B, C \text{ share no predicates } \}$$

The restriction that A, B and C share no predicates is added to avoid self-explanations.[12]

The strategy of $\mathbf{LA}_\forall$ is *reliability* – we will explain its role in Section 4.2.[13] There we will focus on the proof theory of $\mathbf{LA}_\forall$, which allows us to explicate the intereaction of AG and classical inferences. As for all ALs in standard format, the $\mathbf{LA}_\forall$-semantics is obtained from the same triple $\langle \mathbf{T}, \Omega, \textit{reliability} \rangle$ – we refer to [6] for a generic definition of the AL-semantics. In Section 4.3, we will present some particular features of $\mathbf{LA}_\forall$ that show how it meets the desiderata from Section 3.3.

4.2 The Proof Theory of $\mathbf{LA}_\forall$

The $\mathbf{LA}_\forall$-proof theory is a mere instantiation of the generic proof theory for ALs in standard format – see [6]. As spelled out before, inferences such as AG are by definition defeasible. Hence if we want to formalize them, we should be able to model the retraction of previously drawn conclusions in view of later insights. For this purpose, a *line* in an $\mathbf{LA}_\forall$-proof has – apart from a line number, a formula and a justification – a fourth element, the *condition*. This condition consists of n ($\in \mathbb{N}$) members of Ω, and it specifies the assumptions on which the formula of that line is derived. More precisely, the formulas in this set are assumed to be false until and unless proven otherwise. A line becomes *defeated*, if one of these assumptions becomes untenable in light of further derivations in the proof.

The inference rules of $\mathbf{LA}_\forall$-proofs reduce to three generic rules. This requires some notational conventions. For any finite $\Theta \subset \Omega$, let $Dab(\Theta)$ denote the classical disjunction of the members of Θ. In general, we use the term *Dab-formula* to refer to finite disjunctions of abnormalities. Where Γ is a premise set, and where

[12] See e.g. [19, p. 221-222] where a similar restriction is motivated for the logic $\mathbf{LA}^{\mathbf{r}}$, an adaptive logic for singular fact abduction.

[13] ALs in standard format can also use the minimal abnormality strategy. However, as the difference between both strategies is not relevant here, we restrict ourselves to the (slightly less complicated) reliability-variant of $\mathbf{LA}_\forall$.

$$A \quad \Delta$$

abbreviates that A occurs in the proof on a line with the condition Δ, the inference rules are given by the following generic rules:

PREM If $A \in \Gamma$:

$$\begin{array}{ll} \vdots & \vdots \\ \hline A & \emptyset \end{array}$$

RU If $A_1, ..., A_n \vdash_{\mathbf{T}} B$:

$$\begin{array}{ll} A_1 & \Delta_1 \\ \vdots & \vdots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \ldots \cup \Delta_n \end{array}$$

RC If $A_1, ..., A_n \vdash_{\mathbf{T}} B \vee Dab(\Theta)$

$$\begin{array}{ll} A_1 & \Delta_1 \\ \vdots & \vdots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \ldots \cup \Delta_n \cup \Theta \end{array}$$

The premise rule PREM states that a premise may be introduced at any line of a proof on the empty condition. The unconditional inference rule RU states that, if $A_1, \ldots, A_n \vdash_{\mathbf{T}} B$ and $A_1, \ldots, A_n$ occur in the proof on the conditions $\Delta_1, \ldots, \Delta_n$, we may add B on the condition $\Delta_1 \cup \ldots \cup \Delta_n$. The strength of an adaptive logic comes with the third rule, the conditional inference rule RC, which works analogously to RU, but allows us to push abnormalities from the formula to the condition. Hence RC allows us to take defeasible steps based on the assumption that the abnormalities are false.

To get an idea of how these generic rules allow us to model AG, consider the formalization of the pineapple-example from the introduction:

$$\Gamma_1 = \{\forall x(Px \supset Qx), \Box\forall x(Rx \supset Qx), \exists x Px\}$$

The last premise is added to avoid certain unwelcome results – see Section 4.3. Note that in view of the interpretation of the premises, this is a harmless addition: if we want to explain the fact that all pineapples taste sweet, then it seems evident that we also know that pineapples exist.

We start an $\mathbf{LA}_{\forall}$-proof from Γ_1 by writing down the premises:

1	$\forall x(Px \supset Qx)$	PREM	$\emptyset$
2	$\Box\forall x(Rx \supset Qx)$	PREM	$\emptyset$

Note that $\{\forall x(Px \supset Qx), \Box\forall x(Rx \supset Qx)\} \vdash_{\mathbf{T}} \forall x(Px \supset Rx) \vee (P \not\rightarrow_R Q)$. Hence we may apply the rule RU to derive $\forall x(Px \supset Rx) \vee (P \not\rightarrow_R Q)$, and from the latter, derive that all P are R by RC:

3	$\forall x(Px \supset Rx) \vee (P \not\rightarrow_R Q)$	1,2;RU	$\emptyset$
4	$\forall x(Px \supset Rx)$	3;RC	$\{P \not\rightarrow_R Q\}$

The interesting aspect of adaptive logics is of course their dynamic flavor, which can only be illustrated if we add more premises to Γ_1. Suppose that we learn about a genetically modified pineapple a, which contains no sugar, but nevertheless tastes sweet because it has been injected with a synthetic sweetener. So we have to add the premise $Pa \wedge \neg Ra$, which contradicts the derived explanation. Let us call the extended premise set Γ_2. A nice advantage of ALs is that, since the proofs are dynamic, we need not start a proof all over again whenever premises are added; we can just pick up where we ended our line of thought. Hence we may continue our proof as follows:

$\vdots$	$\vdots$	$\vdots$	$\vdots$
4	$\forall x(Px \supset Rx)$	3;RC	$\{P \not\rightarrow_R Q\}$ ✓[7]
5	$Pa \wedge \neg Ra$	PREM	$\emptyset$
6	$\neg\forall x(Px \supset Rx)$	5;RU	$\emptyset$
7	$P \not\rightarrow_R Q$	1,2,6;RU	$\emptyset$

For the time being, ignore the ✓[7]-sign on line 4 – this will be clarified below. At line 7, we have reached the insight that $P \not\rightarrow_R Q$ follows from our premises by **T**. Hence, we need a way to indicate that there is something wrong with the condition of line 4. This is done by a marking criterion, which depends on what we have derived so far. Let a *stage of a proof* be a list of lines, obtained by application of the three generic rules. A *proof* is then a list of subsequent stages. At every stage s, a *marking definition* stipulates which lines are marked and which are not, and, hence, which have become marked or unmarked with respect to the previous stage. If A is derived on an unmarked line at stage s, we say that A is derived at stage s; otherwise, A is not derived at stage s.

For $\mathbf{LA}_\forall$, a line is marked at stage s, whenever at this stage, a member of its condition is derived on the empty condition, either by itself, or as part of a *minimal* disjunction of abnormalities. In that case, we say that the condition is unreliable at stage s. Putting everything together, we obtain the following standard definitions:

Definition 1 *A Dab-formula $Dab(\Delta)$ is a* minimal Dab-formula at stage *s iff $Dab(\Delta)$ is derived on the empty condition at stage s, and there is no $\Delta' \subset \Delta$ for which $Dab(\Delta')$ is derived on the empty condition at stage s.*

Definition 2 *The* set of unreliable formulas *$U_s(\Gamma)$ at stage s is the union of all Δ for which $Dab(\Delta)$ is a minimal Dab-formula at stage s.*

Definition 3 (Marking for Reliability) *The line i with condition Δ is marked at stage s iff $\Delta \cap U_s(\Gamma) \neq \emptyset$.*

In view of these definitions, line 4 is marked at stage 7 of the above proof, as indicated by the ✓[7]-sign.

It is important to remark that, despite the dynamic character of the proofs, adaptive logics are proper proof-invariant logics. Given a certain premise set, an adaptive logic determines for every formula unambiguously whether it is part of the consequence set. To avoid confusion with formulas that are derivable at a certain stage of a proof (but can be defeated at a later stage), formulas in the consequence set are called *finally derivable*. The final derivability relation of an adaptive logic is defined as follows:

Definition 4 *A formula A is* finally derived from Γ *at stage s of a proof if and only if A is derived at line i, line i is not marked at stage s and every extension of the proof in which i is marked may be further extended in such a way that line i is unmarked.*

Definition 5 (Final Derivability) *$\Gamma \vdash_{\mathbf{LA}_\forall} A$ ($A \in Cn_{\mathbf{LA}_\forall}(\Gamma)$) if and only if A is finally derived in a $\mathbf{LA}_\forall$-proof from Γ.*

To illustrate the above definitions, consider again our proof from Γ_2. Since the Dab-formula at line 7 is minimal, line 4 will remain marked in every extension of the proof. More generally, $\forall x(Px \supset Rx)$ is not finally derivable from Γ_2, i.e. there is no proof in which we can finally derive $\forall x(Px \supset Rx)$ from this premise set.

4.3 Some Salient Features of the Logic

We end this section with a brief survey of the ways in which $\mathbf{LA}_\forall$ solves some typical problems for any formal model of abduction. First of all, as any AL in standard format, $\mathbf{LA}_\forall$ has the Reassurance property [6, Corollary 1]:

Theorem 1 *If Γ is not* **T***-trivial, then neither is $Cn_{\mathbf{LA}_\forall}(\Gamma)$. (Reassurance)*

Theorem 1 implies that if our explanatory framework and our factual knowledge are mutually consistent, then $\mathbf{LA}_\forall$ will always yield a consistent set of explanatory hypotheses. This is an immediate consequence of the fact that **CL** is included in **T** and of the axiom T.

Second, as explained in Section 3.3, a logic for abduction should only yield the most parsimonious hypotheses. Consider the following proof from Γ_1:

1	$\forall x(Px \supset Qx)$	PREM	$\emptyset$	
2	$\Box\forall x(Rx \supset Qx)$	PREM	$\emptyset$	
3	$\Box\forall x((Rx \wedge Sx) \supset Qx)$	2;RU	$\emptyset$	
4	$\forall x(Px \supset (Rx \wedge Sx))$	1,3;RC	$\{P \not\rightarrow_{R\wedge S} Q\}$	$\checkmark^{9}$
5	$\forall x(Px \supset Sx)$	4;RU	$\{P \not\rightarrow_{R\wedge S} Q\}$	$\checkmark^{9}$
6	$\exists x Px$	PREM	$\emptyset$	
7	$\exists x(Px \wedge \neg Sx) \vee \exists x(Px \wedge Sx)$	6;RU	$\emptyset$	
8	$\neg\forall x(Px \supset (Rx \wedge Sx)) \vee \neg\forall x(Px \supset (Rx \wedge \neg Sx))$	7; RU	$\emptyset$	
9	$(P \not\rightarrow_{R\wedge S} Q) \vee (P \not\rightarrow_{R\wedge\neg S} Q)$	1,2,8;RU	$\emptyset$	

As the material implication has the property $A \supset B \vdash (A \wedge C) \supset B$ (*strengthening the antecedent*), the hypothesis on line 5, which states that anything that is P also has the random property S, could be derived. However, using the premise $\exists x Px$, we can derive the *Dab*-formula on line 9 which defeats lines 4 and 5.

Third, the dynamic proof theory and the form of the abnormalities also ensure that no hypotheses can be finally derived from tautologies, and that no contradictions can be finally derived as a hypothesis. The following proof from Γ_1 illustrates how the logic enables us to defeat both self-contradictory hypotheses and hypotheses derived from tautotologies – see line 3, resp. lines 7 and 8:

1	$\forall x(Px \supset Qx)$	PREM	$\emptyset$	
2	$\Box\forall x((Sx \wedge \neg Sx) \supset Qx)$	-;RU	$\emptyset$	
3	$\forall x(Px \supset (Sx \wedge \neg Sx))$	1,2;RC	$\{P \not\rightarrow_{S\wedge\neg S} Q\}$	$\checkmark^{5}$
4	$\exists x Px$	PREM	$\emptyset$	
5	$\exists x(Px \wedge \neg(Sx \wedge \neg Sx))$	4;RU	$\emptyset$	
6	$P \not\rightarrow_{S\wedge\neg S} Q$	1,2,5;RU	$\emptyset$	
7	$\forall x(Px \supset (Sx \vee \neg Sx))$	-;RU	$\emptyset$	

8	$\Box(\forall x)(Tx \supset (Sx \vee \neg Sx)))$	-;RU	$\emptyset$	
9	$\forall x(Px \supset Tx)$	1,2;RC	$\{P \not\twoheadrightarrow_T (S \vee \neg S)\}$	$\checkmark^{10}$
10	$(P \not\twoheadrightarrow_T S \vee \neg S) \vee (P \not\twoheadrightarrow_{\neg T} S \vee \neg S)$	4;RU	$\emptyset$	

The final feature that will be illustrated is how this logic handles multiple explanatory hypotheses. Suppose that we learn about a property S, which explains Q-hood just as well as the property R does. Hence we have to add the premise $\Box\forall x(Sx \supset Qx)$ to Γ_1, which results in the following set:

$$\Gamma_3 = \{\forall x(Px \supset Qx), \Box\forall x(Rx \supset Qx, \Box\forall x(Sx \supset Qx), \exists xPx\}$$

At first sight, both the hypotheses $\forall x(Px \supset Rx)$ and $\forall x(Px \supset Sx)$ can be derived from Γ_3. But, as shown in the proof below, these two formulas are not finally derivable.The composed hypothesis $\forall x(Px \supset (Rx \vee Sx))$ is, however, finally derivable from Γ_3.

1	$\forall x(Px \supset Qx)$	PREM	$\emptyset$	
2	$\Box\forall x(Rx \supset Qx)$	PREM	$\emptyset$	
3	$\Box\forall x(Sx \supset Qx)$	PREM	$\emptyset$	
4	$\forall x(Px \supset Rx)$	1,2;RC	$\{P \not\twoheadrightarrow_R Q\}$	$\checkmark^{7}$
5	$\forall x(Px \supset Sx)$	1,3;RC	$\{P \not\twoheadrightarrow_S Q\}$	$\checkmark^{8}$
6	$\exists xPx$	PREM	$\emptyset$	
7	$(P \not\twoheadrightarrow_R Q) \vee (P \not\twoheadrightarrow_{S\wedge\neg R} Q)$	1,2,3,6;RU	$\emptyset$	
8	$(P \not\twoheadrightarrow_S Q) \vee (P \not\twoheadrightarrow_{R\wedge\neg S} Q)$	1,2,3,6;RU	$\emptyset$	
9	$\Box\forall x((Rx \vee Sx) \supset Qx)$	2,3;RU	$\emptyset$	
10	$\forall x(Px \supset (Rx \vee Sx))$	1,9;RC	$\{P \not\twoheadrightarrow_{R\vee S} Q\}$	

In view of this last feature, $\mathbf{LA}_\forall$ models a kind of *practical abduction*: whenever multiple explanatory hypotheses are available, $\mathbf{LA}_\forall$ only allows for the (undefeated) derivation of a disjunctive combination of these hypotheses. This is opposed to *theoretical abduction*, in which each of the individual hypotheses can be seperately derived. For a thorough discussion of this distinction, see [4].

5 Conclusion

As argued in this paper, abduction of generalizations (AG) is ubiquitous in everyday and scientific reasoning. We provided a first general analysis of this pattern, and argued that the notion of an explanatory framework should be embodied in any formal model for AG. This suggestion was implemented in $\mathbf{LA}_\forall$, which is a very intuitive and well-behaved formal logic that allows us to apply AG, and to withdraw its applications in those cases where a conflict with the other premises occurs.

Several enrichments of our formal model can be studied in future research, in order to deal with e.g. probabilistic information (see Section 2.1), causal arguments (see Section 3.1), and abductive anomalies.[14] Also, it seems worthwhile to develop ways in which singular fact abduction and AG can be integrated in the framework of adaptive logics. Finally, case studies of some of the examples mentioned in Section 2 may shed new light on the relation between AG, unification and other patterns of abduction.

[14] In Aliseda's terminology, an anomaly is a fact, the negation of which follows from our background theory.

References

[1] Atocha Aliseda. *Abductive Reasoning. Logical Investigations into Discovery and Explanation.* Springer, Dordrecht, 2006.

[2] Aristotle. *Posterior Analytics.* Online Publication, `http://www.logoslibrary.org/aristotle/posterior/index.html`, n.d.

[3] Author, 2011.

[4] Author, 2011.

[5] Diderik Batens. Inconsistency-adaptive logics. In *Logic at Work. Essays dedicated to the memory of Helena Rasiowa*, pages 445–472. Springer, 1999.

[6] Diderik Batens. A universal logic approach to adaptive logics. *Logica Universalis*, 1:221–242, 2007.

[7] Diderik Batens. Logics for qualitative inductive generalization. *Studia Logica*, 97:61–80, 2011.

[8] Diderik Batens, Joke Meheus, Dagmar Provijn, and Liza Verhoeven. Some adaptive logics for diagnosis. *Logic and Logical Philosophy*, 11/12:39–65, 2003.

[9] Igor Douven. Abduction. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Spring 2011 edition, 2011.

[10] Umberto Eco. Horns, Hooves, Insteps: Some Hypotheses on three types of Abduction. In U. Eco and A. Sebeok, editors, *The Sign of Three: Dupin, Holmes, Peirce*, pages 198–220. Indiana University Press, Bloomington and Indianapolis, 1983.

[11] Peter A. Flach and Antonis C. Kakas. Abductive and Inductive Reasoning: Background and Issues. In Peter A. Flach and Antonis C. Kakas, editors, *Abduction and Induction. Essays on their Relation and their Integration*, pages 1–27. Kluwer Academic Publishers, Dordrecht, 2000.

[12] Gilbert H. Harman. The inference to the best explanation. *Philosophical Review*, 74(1):88–95, 1965.

[13] Daniel M. Hausman. *Causal Asymmetries.* Cambridge University Press, 1998.

[14] Carl G. Hempel. *Aspects of Scientific Explanation and other essays in the Philosophy of Science.* Free Press, New York, 1965.

[15] P. Kitcher. *The Advancement of Science: Science without Legend, Objectivity without Illusions.* Oxford University Press, 1993.

[16] Peter Lipton. *Inference to the best explanation (2nd edition).* Routledge/Taylor and Francis Group, London, 2004.

[17] Lorenzo Magnani. *Abductive Cognition: The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning.* Springer, Berlin Heidelberg, 2009.

[18] Joke Meheus. A Formal Logic for The Abduction of Singluar Hypothesis. unpublished paper, available at `http://logica.ugent.be/centrum/preprints/fma3.pdf`.

[19] Joke Meheus and Diderik Batens. A formal logic for abductive reasoning. *Logic Journal of the IGPL*, 14:221–236, 2006.

[20] Ilkka Niiniluoto. Scientific progress. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Summer 2011 edition, 2011.

[21] Charles S. Peirce. *Collected Papers*. Belknap Press of Harvard University Press, Cambridge, Massachusetts, 1958.

[22] Charles S. Peirce. *The Essential Peirce*, volume 2. Indiana University Press, Bloomington and Indianapolis, 1998.

[23] Gerhard Schurz. Patterns of Abduction. *Synthese*, 164:201–234, 2008.

[24] Paul Thagard. *Computational Philosophy of Science*. MIT Press, Cambridge, Massachussetts, 1988.