A formal explication of the search for explanations: the adaptive logics approach to abductive reasoning

HANS LYCKE, *Centre for Logic and Philosophy of Science, Ghent University, Blandijnberg 2, 9000 Gent, Belgium*
e-mail: hans.lycke@ugent.be

Abstract

Most logic–based approaches characterize abduction as a kind of *backwards deduction plus additional conditions*, which means that a number of conditions is specified that enable one to decide whether or not a particular abductive inference is sound (one of those conditions may e.g. be that abductive consequences have to be compatible with the background theory). Despite the fact that these approaches succeed in specifying which formulas count as valid consequences of abductive inference steps, they do not explicate the way people actually reason by means of abductive inferences. This is most clearly shown by the absence of a decent proof theory. Instead, search procedures are provided that enable one to determine the right abductive consequences. However, these do not by far resemble human reasoning. In order to explicate abductive reasoning more realistically, an alternative approach will be provided in this article, namely, one that is based on the *adaptive logics programme*. Proof theoretically, this approach interprets the argumentation schema *affirming the consequent* (**AC**: $A \supset B, B \vdash A$) as a defeasible rule of inference. This comes down to the fact that the abductive consequences obtained by means of **AC** are accepted only for as long as certain conditions are satisfied—e.g. as long as their negation has not been derived from the background theory. In the end, only the *unproblematic* applications of **AC** are retained, while the problematic ones are rejected. In this way, the adaptive logics approach to abduction succeeds to provide a more realistic explication of the way people reason by means of abductive inferences. Moreover, as multiple abduction processes will be characterized, this article may be considered as the first step in the direction of a general formal approach to abduction based on the adaptive logics programme.

Keywords: Abduction, abductive reasoning, explanation, defeasible reasoning, adaptive logics, adaptive logics programme.

1 Introduction

When searching an explanation for a (puzzling) phenomenon, people often reason backwards, from the phenomenon to be explained to possible explanations. When they do so, they perform abductive inferences, inferences based on the argumentation schema known as *affirming the consequent*:

(**AC**) $\quad A \supset B, B \vdash A$

Clearly, **AC** is not deductively valid. In the context of *classical logic* (**CL**), its acceptance would even lead to triviality. Nevertheless, people make use of **AC**. Though, to avoid the derivation of unsound consequences, they do so in a defeasible way. As a consequence, abductive explanations remain provisional, and, in the end, some are rejected. The reasons for doing so may be external or internal to the background theory from which these explanations were derived. In the external case, new information is obtained that forces the rejection of some abductive explanations—e.g. in case the results of further research are incompatible with these explanations. In formal terms: abductive reasoning is non–monotonic. In the internal case, new (deductive) consequences are derived from the background theory that necessitate the rejection of some abductive explanations—e.g. in case it

Vol. 20 No. 2,
doi:10.1093/jigpal/jzq053 Advance Access published 8 January 2011

TABLE 1. Consistent and explanatory abductive explanation

Given Θ (a set of formulae) and ϕ (a sentence), α is a consistent and explanatory abductive explanation of ϕ iff

(i)	$\Theta \cup \{\alpha\} \vdash \phi$
(ii)	$\Theta \nvdash \neg\alpha$
(iii)	$\Theta \nvdash \phi$
(iv)	$\alpha \nvdash \phi$

turns out the background theory already provides a perfectly good explanation for the phenomenon at hand. This kind of (internal) defeasibility results from the fact that people lack logical omniscience (people do not have complete insight in the theories they reason from). As such, when a better insight is gained in those theories, some of the earlier drawn consequences might have to be withdrawn.

Most logic–based approaches characterize abduction as a kind of *backwards deduction plus additional conditions*—see e.g. Aliseda–Llera [1, 2], Mayer and Pirri [3, 4], McIlraith [5] and Gabriele [6]. In these approaches, a number of conditions is specified that enable one to decide whether or not a particular abductive inference is sound. Moreover, different kinds of abduction are characterized by different sets of such conditions. For example, in Table 1, the conditions are stated that were given in Aliseda–Llera [1, pp. 48–49] to characterize abductive reasoning that is (in terms of Aliseda–Llera) both consistent and explanatory.

Although the traditional logic-based approaches to abduction succeed in specifying which formulas may count as valid consequences of abductive inference steps, they do not explicate the way in which people actually reason by means of abductive inferences—hence, the focus is on abductive consequence, not on abductive reasoning. This is most clearly shown by the absence of a decent proof theory, i.e. a proof theory that explicates abductive reasoning steps as described above, namely as defeasible applications of the inference rule **AC**. Instead, search procedures are provided that enable one to determine the right abductive consequences—e.g. the tableaux methods presented in Aliseda–Llera [1, 2] and Mayer and Pirri [3, 4]. However, these explicate **AC** only at the semantic or the metatheoretic level. As a consequence, these search procedures do not resemble human reasoning at all.[1]

In order to explicate abductive reasoning more realistically, an alternative approach will be provided in this article, namely, one that is based on the *adaptive logics programme*.[2] In the accompanying proof theory, the argumentation schema **AC** is really interpreted as a defeasible rule of inference. More specifically, the consequences obtained by applying **AC** are accepted only for as long as certain conditions are satisfied—e.g. as long as their negation has not been derived from the background theory. In short, adaptive logics for abduction only retain the unproblematic *applications* of **AC**, while they reject the problematic ones. Hence, in comparison to the traditional logic-based approaches, the adaptive logics approach more realistically captures the way people make abductive inferences. Nonetheless, I will show that all conditions stated by the traditional logic-based approaches are still satisfied by the adaptive logics approach. Finally, as the adaptive logics approach is not restricted to a particular kind of abduction process (multiple kinds of abduction will be explicated), this article should be considered as the first step in the direction of a general approach towards the explication of abductive reasoning.

[1] In some cases, one might even doubt whether these search procedures even obtain the right abductive consequences—see Meheus and Provijn [7].

[2] A thorough introduction into adaptive logics can be found in Batens [8, 9], and an overview of the adaptive logics programme can be found on the Adaptive Logics Homepage (`http://logica.ugent.be/adlog`).

TABLE 2. The Languages $\mathcal{L}$ and $\mathcal{L}^{\mathcal{M}}$

Language	Letters	Logical Symbols	Set of Formulas
$\mathcal{L}$	$\mathcal{S}$	$\neg, \wedge, \vee, \supset, \equiv$	$\mathcal{W}$
$\mathcal{L}^{\mathcal{M}}$	$\mathcal{S}$	$\neg, \wedge, \vee, \supset, \equiv, , \Box_n, \Diamond_n, \Box_e, \Diamond_e$	$\mathcal{W}^{\mathcal{M}}$

2 The deductive frame

As spelled out in the previous section, abduction validates some arguments that are not deductively valid—in casu, applications of **AC**. Hence, abductive reasoning goes beyond deductive reasoning. Nevertheless, abduction is constrained by deductive reasoning, for some abductive consequences of a premise set might have to be withdrawn in view of its deductive consequences—e.g. in case these abductive consequences are incompatible with the deductive ones. Hence, abduction and deduction go hand in hand, the latter serving as the *deductive frame* of the former.

2.1 A modal frame

In most logic-based approaches to abduction, deductive reasoning is explicated by means of *classical logic*—see e.g. Aliseda–Llera [1, 2], Meheus and Batens [10] and Meheus [11]. In this article though, the deductive frame is captured by the modal logic **RBK**.[3] Contrary to the classical frame, a modal frame makes it possible to capture some intensional elements of abductive reasoning contexts. Moreover, the modal frame will turn out quite useful to capture some of the metatheoretic conditions on abductive explanation put forward by the traditional logic-based approaches (see Table 1). Both these claims will be discussed more thoroughly later on. First though, the logic **RBK** will be characterized, both semantically as well as proof theoretically.

2.1.1 Language schema

The logic **RBK** is a standard bimodal logic extending (propositional) classical logic with the modal operators $\Box_n$ and $\Box_e$. As a consequence, the modal language $\mathcal{L}^{\mathcal{M}}$ of **RBK** is obtained by adding both these necessity operators, together with the corresponding possibility operators, to the standard propositional language $\mathcal{L}$ (see Table 2 for an overview). The set of well-formed formulas $\mathcal{W}^{\mathcal{M}}$ of the language $\mathcal{L}^{\mathcal{M}}$ is defined in the usual way.

In the remaining of this article, only negation, disjunction and both necessity operators are taken as primitive. The other logical symbols are defined in the standard way.

2.1.2 Semantic characterization

An **RBK**-model M is a 5 tuple $\langle W, w_0, R^n, R^e, v\rangle$. The set W is a (non-empty) set of worlds, with $w_0 \in W$ the actual world. R^n and R^e are both accessibility relations on W, the former of which is both reflexive and transitive, while the latter is merely reflexive. Moreover, the following relation holds between both accessibility relations:

CEI For all $w, w' \in W$: if $R^e ww'$ then also $R^n ww'$.

Finally, v is an assignment function, for which the following condition holds:

C1.1 v: $\mathcal{S} \times W \mapsto \{0,1\}$.

[3] Actually, in Meheus *et al*. [12], the deductive frame is captured by the logic $\mathbf{S5^2}$ that is quite similar to the logic **RBK**.

Table 3. Additional axiom schemas, rules and definitions of **RBK**

AM1n	$\Box_n(A \supset B) \supset (\Box_n A \supset \Box_n B)$	AM1e	$\Box_e(A \supset B) \supset (\Box_e A \supset \Box_e B)$
AM2n	$\Box_n A \supset A$	AM2e	$\Box_e A \supset A$
AM3n	$\Box_n A \supset \Box_n \Box_n A$		
AEI	$\Box_n A \supset \Box_e A$		
NECn	From $\vdash A$ conclude to $\vdash \Box_n A$	NECe	From $\vdash A$ conclude to $\vdash \Box_e A$
Dfn	$\Diamond_n A =_{df} \neg \Box_n \neg A$	Dfe	$\Diamond_e A =_{df} \neg \Box_e \neg A$

The valuation function v_M, determined by the model M, is now defined as follows:

C2.0 $v_M : \mathcal{W}^{\mathcal{M}} \times W \mapsto \{0,1\}$.
C2.1 Where $A \in \mathcal{S}$, $v_M(A,w) = 1$ iff $v(A,w) = 1$.
C2.3 $v_M(\neg A,w) = 1$ iff $v_M(A,w) = 0$.
C2.4 $v_M(A \vee B,w) = 1$ iff $v_M(A,w) = 1$ or $v_M(B,w) = 1$.
C2.5 $v_M(\Box_n A,w) = 1$ iff for all $w' \in W$: if $R^n ww'$ then $v_M(A,w') = 1$.
C2.6 $v_M(\Box_e A,w) = 1$ iff for all $w' \in W$: if $R^e ww'$ then $v_M(A,w') = 1$.

A model M verifies a formula $A \in \mathcal{W}^{\mathcal{M}}$ iff $v_M(A,w_0) = 1$. Moreover, a model M is a model of a premise set Γ iff, for all $B \in \Gamma$, $v_M(B,w_0) = 1$.

Definition 1
$\vDash_{\mathbf{RBK}} A$ (A is valid) iff A is verified by all **RBK**–models.

Definition 2
$\Gamma \vDash_{\mathbf{RBK}} A$ (A is a semantic consequence of Γ) iff all **RBK**–models of Γ verify A.

Some remarks seem to be necessary. First of all, as the accessibility relation R^n is both reflexive and transitive, the modal operator $\Box_n$ corresponds to the necessity operator of the (normal) modal logic **S4**. Secondly, the accessibility relation R^e is reflexive but not transitive, meaning that the modal operator $\Box_e$ corresponds to the necessity operator of the (normal) modal logic **KT**. Finally, because of the specific relation between R^n and R^e, as expressed by the condition **CEI**, the truth of a formula $\Box_n A$ in a world w yields the truth of the formula $\Box_e A$ in that world. In the proof theoretic characterization below, this is expressed by the axiom schema **AEI** (see Table 3).

2.1.3 Proof theoretic characterization

The **RBK**-proof theory is obtained by adding the axiom schemas, inference rules and definitions stated in Table 3 to the axiom system of (propositional) classical logic.

2.1.4 Soundness and completeness

As both soundness and completeness for **RBK** are proven by standard means, the proofs are left to the reader.

Theorem 1
$\Gamma \vdash_{\mathbf{RBK}} A$ iff $\Gamma \vDash_{\mathbf{RBK}} A$.

2.2 Representing abductive contexts

Because of the higher expressive power of the **RBK**-language $\mathcal{L}^{\mathcal{M}}$ (as compared to the language of classical logic), the logic **RBK** not only enables one to capture deductive reasoning as such but also enables one to capture some intensional elements of reasoning contexts.

2.2.1 Background knowledge

The modal operators $\Box_n$ and $\Box_e$ will be used to express both the *nomological* and *empirical background knowledge* held by a reasoner in a reasoning context. First, the nomological background knowledge is represented by elements of the set $\mathcal{W}^{\mathcal{N}} \subset \mathcal{W}^{\mathcal{M}}$—see Definition 3. More specifically, a formula $\Box_n A \in \mathcal{W}^{\mathcal{N}}$ states that A is considered a nomological fact by the reasoner in the reasoning context at hand. Second, the empirical background knowledge is represented by elements of the set $\mathcal{W}^{\mathcal{E}} \subset \mathcal{W}^{\mathcal{M}}$—see definition 4. A formula $\Box_e A \in \mathcal{W}^{\mathcal{E}}$ states that A is taken to be an empirical fact by the reasoner in the reasoning context.

Definition 3
$\mathcal{W}^{\mathcal{N}} = \{\Box_n A \mid A \in \mathcal{W}\}$.

Definition 4
$\mathcal{W}^{\mathcal{E}} = \{\Box_e A \mid A \in \mathcal{S} \cup \mathcal{S}^{\neg}\}$.[4]

Two remarks are needed at this point. Firstly, in view of axiom schema **AEI** (see Table 3), nomological background knowledge can be combined with empirical background knowledge in order to derive further empirical information—e.g. in the process of making predictions. Hence, nomological information may be said to have *empirical impact*.

More importantly, one might wonder what the modal operators are taken to express. In accordance with the epistemological framework presented in Batens [15, 16], the elements of the background knowledge are classified as (part of the) *contextual certainties* and *relevant premises* of a given context (which is defined as a problem-solving situation). Without going into the details, both the contextual certainties and the relevant premises of a context are considered as true in that context, and helpful in order to solve the problem at hand (for more details, the reader is referred to the cited literature). Hence, the necessity operators occurring in the elements of $\mathcal{W}^{\mathcal{N}}$ and $\mathcal{W}^{\mathcal{E}}$ capture the fact that the elements of the background knowledge are considered as unproblematic in the given context (thus, the necessities have to be interpreted epistemologically, not ontologically).[5]

2.2.2 Abductive contexts

The contexts considered in this article are *abductive contexts*, namely, problem-solving situations in which possible explanations are sought for puzzling (empirical) phenomena. Given the above

[4] Firstly, the elements of the set $\mathcal{S}^{\neg}$ are the negations of the elements of the set $\mathcal{S}$ — $\mathcal{S}^{\neg} =_{df} \{\neg A \mid A \in \mathcal{S}\}$. Secondly, one of the anonymous referees rightly remarked that also conjunctions of elements of the set $\mathcal{S} \cup \mathcal{S}^{\neg}$ may be considered as empirical facts, so that the empirical background knowledge is better explicated by the set $\mathcal{W}^{\mathcal{E}'} = \{\Box_e(A_1 \wedge \ldots \wedge A_n) \mid A_1, \ldots, A_n \in \mathcal{S} \cup \mathcal{S}^{\neg}\}$. However, the elements of $\mathcal{W}^{\mathcal{E}'}$ are all derivable from the elements of $\mathcal{W}^{\mathcal{E}}$ by means of the logics presented in this article. For the purposes of this article, this implies that replacing $\mathcal{W}^{\mathcal{E}}$ by $\mathcal{W}^{\mathcal{E}'}$ would not make a difference. Hence, to keep things as simple as possible, I will stick to $\mathcal{W}^{\mathcal{E}}$.

[5] From this clearly follows that $\Box_n$ and $\Box_e$ are interpreted as a kind of *epistemic* operators. Because of the disagreement on the characteristics of epistemic necessity in the standard literature—see e.g. Garson [13] and Hendricks [14], the characterization of $\Box_n$ and $\Box_e$ as an **S4**– and a **KT**–necessity, respectively, may be considered as consistent with tradition in epistemic logic.

elaboration of the meaning of the modal operators, the elements of the background knowledge are considered unproblematic, which in abductive contexts means that they are not in need of any explanation.

Besides background knowledge, abductive contexts obviously also contain some elements that are in need of an explanation. The latter will be represented by elements of the set $\mathcal{W}^{\mathcal{O}}$, i.e. the set of observed (empirical) phenomena—see Definition 5.

Definition 5
$\mathcal{W}^{\mathcal{O}} = \mathcal{S} \cup \mathcal{S}^{\neg}$.

Surely, not all elements of $\mathcal{W}^{\mathcal{O}}$ express (empirical) phenomena in need of an explanation (henceforth, puzzling phenomena). Nonetheless, a formula $A \in \mathcal{W}^{\mathcal{O}}$ will be taken to express a puzzling phenomenon in an abductive context in case A is not considered as unproblematic by the reasoner in that abductive context—in other words, in case $\Box_e A$ is not derivable in that abductive context (for otherwise, A would be considered as unproblematic by the reasoner).

2.2.3 Final remark

In the remaining of this article, premise sets will be taken to express abductive contexts. As such, premise sets will be restricted to formulas that express the background knowledge of a reasoner (i.e. elements of $\mathcal{W}^{\mathcal{N}} \cup \mathcal{W}^{\mathcal{E}}$) and formulas that express observed (empirical) phenomena (i.e. elements of $\mathcal{W}^{\mathcal{O}}$). Obviously, in case the latter express puzzling phenomena, they will trigger abductive inferences.

3 On defeasible inference

As stated in Section 1, to capture abduction in a realistic way, abductive inference steps have to be captured proof theoretically by the (defeasible) inference rule **AC**. However, because the deductive frame is captured in modal terms (as set out in the previous section), abductive inference has to be captured in modal terms as well. As a consequence, the inference rule **AC** will be restricted to the following schema (A and B are formulas, and Δ is a set of formulas):

(**AC^m**) $\Box_n(A \supset B), B, \Delta \vdash A$

Some clarification is called for. First of all, **AC^m** expresses that a formula A can only be considered as a possible explanation for a phenomenon B in view of the formula $\Box_n(A \supset B)$. The latter expresses that the reasoner in the abductive context takes B to be nomologically dependent upon A. As a consequence, to capture abductive inference by means of the inference rule **AC^m** in a sense resembles Hempel's account of explanation—see Hempel and Oppenheim [17].

Secondly, the explanandum B may not express empirical background knowledge of the reasoner, i.e. B is not allowed to be a modal formula of the form $\Box_e C$, for otherwise it cannot be considered to trigger abductive inferences (remember that the background knowledge is here taken to be accepted beyond doubt, and hence, in no need of explanation)!

Finally, dependent on the particular abduction process one intends to capture, certain additional conditions have to be satisfied before the defeasible inference rule **AC^m** may be applied. In the representation of **AC^m** above, these conditions are represented by the elements of Δ. Some important remarks have to be made with respect to Δ. Firstly, the elements of Δ correspond to some of the conditions stated by the traditional logic-based approaches to abduction (see Section 1). This is not surprising as both approaches tend to capture the same reasoning patterns, albeit by distinct means.

Secondly, the elements of Δ can only be presumed in a defeasible way themselves, namely, for as long as there is no information that forces us to reject them. This might be more surprising to some. Hence, some explanation is required. For example, consider the standard condition stating that an explanandum may only trigger abductive inferences in case it is not derivable from the background theory alone (see Table 1, condition (iii)). Clearly, that the explanandum is derivable from the background theory alone is only known by a reasoner in case all deductive consequences of the background theory are known by that reasoner (i.e. in case logical omniscience is presupposed). As this is nearly ever the case, a reasoner will normally base the decision to perform abductive inferences on the insight she has in the background theory at a particular moment (given that I am trying to capture abductive reasoning in a realistic way, logical omniscience is not presupposed in this article). In other words, a reasoner will presuppose the background theory does not provide an explanation for the explanandum unless or until one is actually found. If later it turns out this presupposition was wrong headed, the abductive explanations obtained by relying on that presupposition will be withdrawn. In the modal framework I presented in the previous section, the explananda are puzzling (empirical) phenomena. More specifically, a formula $B \in \mathcal{W}^{\mathcal{O}}$ is an explanandum in an abductive context only if $\Box_e B$ is not derivable in that abductive context. Obviously, this is the case as long as the formula $\neg\Box_e B$ may be presupposed to be true in that abductive reasoning context, for this expresses that the reasoner has no reason to suppose that his background theory already provides an explanation for the phenomenon represented by the formula B. However, as stated before, this may change, meaning that the derivation of $\neg\Box_e B$ is obtained in a defeasible way, by applying a defeasible inference rule, namely, the inference rule **NEN**:

(**NEN**) $\vdash \neg\Box_e A$

Actually, this is only one of the defeasible inference rules of which the consequences may enter the set Δ in **AC$^{\mathbf{m}}$**. The other defeasible inference rule is the inference rule **NNN**:

(**NNN**) $\vdash \neg\Box_n(A \supset B)$

This inference rule will be used to express that an abductive explanation is minimal, which will be explicated later on (in Section 4). Nonetheless, a final remark about these inference rules is asked for. Notice that the defeasible inferences based on **NNN** and **NEN** are actually prior to abductive inferences, for the consequences of these additional inference steps are necessary to be able to apply **AC$^{\mathbf{m}}$**. This clearly shows that abduction processes are layered processes, formally explicated by specific combinations of multiple defeasible inference steps.

4 Enter adaptive logics

Adaptive logics (**AL**) are formal logics that have primarily been developed to capture consequence relations that lack a positive test, i.e. consequence relations for which there are no finite means to determine whether a formula belongs to the consequence set of a particular premise set—see e.g. Batens [8, 9]. Well now, consequence relations that capture abduction processes clearly lack a positive test, which is why abductive inferences are characterized proof theoretically by means of defeasible inference rules. As a consequence, **AL** seem particularly well suited to capture these consequence relations.

As abduction processes are formally explicated by specific combinations of defeasible inference rules (as set out in the previous section), the **AL** that will be used to capture these processes will be so-called *prioritized adaptive logics*, i.e. adaptive logics that combine defeasible inference rules in a certain way—see e.g. Batens [8], and Batens *et al*. [18]. In this article, the prioritized adaptive logics $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$ will be characterized. These capture different kinds of abductive reasoning, such as *practical* and *theoretical* abduction. The difference between both in fact comes down to the following: in case there are multiple possible explanations for a phenomenon, $\mathbf{AbL^p}$ will merely enable one to derive the disjunction of these possible explanations (practical abduction), while $\mathbf{AbL^t}$ will enable one to derive all possible explanations (theoretical abduction).[6]

4.1 Previous attempts

This is not the first attempt to explicate abductive reasoning by means of the adaptive logics programme. Despite the fact that some nice results were obtained, the earlier attempts remained unsatisfactory. In Meheus *et al*. [12], a proof theory was provided for the traditional logic-based approaches to abduction. It is based on the characteristics of the adaptive logics-proof theory, and also incorporates some extra logical features. As such, only a proof theory for abduction was provided, not a formal logic. On the other hand, in Meheus and Batens [10] and Meheus [11], two (actual) adaptive logics were provided to explicate abduction, namely, the logics $\mathbf{LA^r}$ and $\mathbf{LA^r_s}$. Nonetheless, $\mathbf{LA^r}$ and $\mathbf{LA^r_s}$ only capture abductive reasoning in a limited way. First of all, these logics do not allow abductive inferences at the purely propositional level. Secondly, only practical abduction could be characterized by the approach presented in [10] and [11], which is most likely due to the fact that the deductive frame of $\mathbf{LA^r}$ and $\mathbf{LA^r_s}$ is based on classical logic. Thirdly, the logics $\mathbf{LA^r}$ and $\mathbf{LA^r_s}$ lack some properties that seem to be necessary to capture abductive explanation in a decent way. Most importantly, in case an explanandum is explained by the background theory alone, $\mathbf{LA^r}$ and $\mathbf{LA^r_s}$ go on to provide possible abductive explanations, despite the fact that none are needed.[7] Neither of these shortcomings also applies to the approach presented in this article.

4.2 The standard format

All adaptive logics (**AL**) have a uniform characterization. This characterization is called the *standard format* of adaptive logics and was presented most thoroughly in Batens [8, 9]. The main advantage of the standard format consists in the fact that all **AL** characterized accordingly have a common semantic and proof theoretic characterization. Moreover, a lot of metatheoretic properties have been proven for **AL** in standard format (most importantly, soundness and completeness).[8] Below, I will first give a general characterization of the standard format. Afterwards, I will present the proof theory of **AL** in standard format. The semantics of **AL** in standard format will not be spelled out. Nothing fundamental is lost though, for the focus of this article is on the proof theory (see Section 1). Moreover, the interested reader can find the semantic characterization of **AL** in standard format in Batens [8, 9] and Batens *et al*. [19].

[6] For an intuitive justification of these abductive processes, see Meheus and Batens [10, pp. 224–225].

[7] To be fair, in a lecture at the University of Utrecht (20 Octobre 2009), Joke Meheus showed how to overcome the second shortcoming of $\mathbf{LA^r}$ and $\mathbf{LA^r_s}$. Nevertheless, the other shortcomings still remain and are not likely to be overcome anytime.

[8] Proofs for these metatheoretic properties are provided in Batens [9].

4.2.1 General characterization

All *flat* adaptive logics in standard format are characterized by means of the following three elements:

- A *lower limit logic* (**LLL**): a reflexive, transitive, monotonic and compact logic that has a characteristic semantics (with no trivial models) and contains classical logic.
- A *set of abnormalities* Ω: a set of formulas characterized by a (possibly restricted) logical form $\mathtt{F}$ that is **LLL**-contingent and contains at least one logical symbol.
- An *adaptive strategy*.

Remark that the **AL** that will be presented later on are not flat **AL**, but prioritized adaptive logics (**PAL**). The latter can also be characterized by means of the standard format, though some slight modifications are necessary.[9] More specifically, the set of abnormalities Ω is replaced by a structurally ordered sequence $\Omega_>$ of sets of abnormalities:

Definition 6
$\Omega_> = \Omega_1 > \Omega_2 > \ldots$[10]

The order imposed on the sequence $\Omega_>$ expresses a priority relation: in case $\Omega_i > \Omega_j$, the priority of the elements of Ω_i is higher than the priority of the elements of Ω_j. For reasons of convenience, I will use Ω to refer to the union of the sets $\Omega_1, \Omega_2, \ldots$, while $\Omega_>$ will be used to refer to the sequence $\Omega_1 > \Omega_2 > \ldots$.

4.2.2 The adaptive consequence relation

The adaptive consequences of a premise set are obtained by the interplay between the three constituting elements of a (prioritized) adaptive logic. This will be explicated by characterizing the **PAL**-consequence relation in general. Where the expression $Dab(\Delta)$ is used to represent a finite disjunction of abnormalities (elements of Ω), the **PAL**-consequence relation is defined as follows.

Definition 7
$\Gamma \vdash_{\mathbf{PAL}} A$ iff there is a finite $\Delta \subset \Omega$ such that $\Gamma \vdash_{\mathbf{LLL}} A \vee Dab(\Delta)$ and $\mathrm{FALSE}_{\mathrm{AS}}(\Delta)$.

The above definition tells us that a formula A is **PAL**-derivable from a premise set Γ iff $A \vee Dab(\Delta)$ is **LLL** derivable from Γ and Δ satisfies the additional condition $\mathrm{FALSE}_{\mathrm{AS}}(\Delta)$. Intuitively, the latter means that one is allowed to derive A from $A \vee Dab(\Delta)$ in case all elements of Δ can safely be interpreted as *false*—metaphorically, one might consider this as a metatheoretic application of disjunctive syllogism. As a consequence, abnormalities are falsified as much as possible. In other words, premise sets are interpreted as normally as possible with respect to some standard of normality.

Definition 7 has some interesting consequences. In case $\Delta = \emptyset$, no abnormalities have to be falsified in order to derive the formula A from the premise set Γ. Hence, in case the formula A is a **LLL** consequence of Γ, it is an adaptive consequence of Γ as well. In general, this implies that a (prioritized) adaptive logic derives more consequences from a premise set than the lower limit logic it is based on (more specifically, the adaptive consequence set of a premise set is a superset of the **LLL** consequence set of that premise set). On the other hand, in case $\Delta \neq \emptyset$, the formula A is only an adaptive consequence of the premise set, in case all elements of Δ may be interpreted as false (if not,

[9] Prioritized adaptive logics are well studied in the literature, see e.g. Batens [8], Batens *et al*. [18], and Verhoeven [20].
[10] No upper bound is necessary, as is most clearly explained in [8, pp. 52–54]. However, all **PAL** that will be considered in this article do have an upper bound.

the formula A cannot safely be interpreted as true). For as long as it has not been determined whether or not all elements of Δ may indeed be interpreted as false, the formula A is called a *conditional consequence* of the premise set Γ. Obviously, this intermediate phase of conditional acceptance of consequences corresponds to the proof theoretic derivation of consequences by means of defeasible inference rules.

Whether a conditional consequence of a premise set is a final consequence as well, depends on the condition $\text{FALSE}_{\text{AS}}(\Delta)$. Whether this condition is satisfied for a particular Δ is determined by the adaptive strategy of an adaptive logic. For the adaptive logics I will present below, this is either the *reliability* strategy or the *normal selections* strategy.[11] Both strategies base the decision to reject (or to retain) a conditional consequence of a premise set on the minimal *Dab*-consequences of that premise set—see Definition 8. In advance though, it is important to notice that the minimal *Dab*-consequences of a premise set are defined in a *stepwise* manner: where the expression $Dab^i(\Delta)$ is used to represent finite disjunctions of abnormalities of priority i (elements of Ω_i), the minimal *Dab*-consequences of the form $Dab^1(\Delta)$ are determined first, then the minimal *Dab*-consequences of the form $Dab^2(\Delta)$,...

Definition 8
$Dab^i(\Delta)$ is a minimal *Dab*-consequence of a premise set Γ iff (1) there is a finite $\Theta \subset \Omega_1 \cup \ldots \cup \Omega_{i-1}$ such that $\Gamma \vdash_{\mathbf{LLL}} Dab^i(\Delta) \vee Dab(\Theta)$, (2) there is no $\Sigma \subset \Omega_j$ such that $\Omega_j > \Omega_i$, $Dab^j(\Sigma)$ is a minimal *Dab*-consequence of Γ and $\Sigma \cap \Theta \neq \emptyset$ and (3) there is no $\Delta' \subset \Delta$ such that (1) and (2) apply to $Dab^i(\Delta')$ as well.

Not all abnormalities occurring in a minimal *Dab*-consequence of a premise set, may be interpreted as false. Hence, some of the conditional consequences derived by interpreting certain of these abnormalities as false have to be rejected. First of all, the reliability strategy will reject all conditional consequences that were derived by interpreting some of the abnormalities occurring in a minimal *Dab*-consequence as false. As a consequence, the condition $\text{FALSE}_{\text{AS}}(\Delta)$ for the reliability strategy (henceforth, $\text{FALSE}_{\text{R}}(\Delta)$) is defined as follows.

Definition 9
For $\Delta \subset \Omega$, $\text{FALSE}_{\text{R}}(\Delta)$ iff, for all Ω_i in $\Omega_>$, there is no finite $\Theta \subset \Omega_i$ such that $Dab^i(\Theta)$ is a minimal *Dab*-consequence of Γ and $\Theta \cap \Delta \neq \emptyset$.

On the other hand, the normal selections strategy will only reject those conditional consequences that were derived by interpreting as false all abnormalities of a minimal *Dab*-consequence of the premise set. In other words, the condition $\text{FALSE}_{\text{AS}}(\Delta)$ for the normal selections strategy (henceforth, $\text{FALSE}_{\text{NS}}(\Delta)$) is defined as follows.

Definition 10
For $\Delta \subset \Omega$, $\text{FALSE}_{\text{NS}}(\Delta)$ iff, for all Ω_i in $\Omega_>$, there is no finite $\Theta \subset \Omega_i$ such that $Dab^i(\Theta)$ is a minimal *Dab*-consequence of Γ and $\Theta \subset \Delta$.

4.2.3 Dynamic behavior

Because of the conditional status of some of the **PAL** consequences, **PAL** display an external as well as an internal dynamics. Firstly, the external dynamics comes down to non–monotonicity: if the premise set is extended, some conditionally derived **PAL** consequences of the premise set

[11] For more information on alternative adaptive strategies, see e.g. Batens [9] (for the *minimal abnormality* strategy) and Meheus and Primiero [21] (for the *counting* strategy).

may not be derivable anymore. Secondly, the internal dynamics is a strictly proof theoretic feature: growing insights in the premises, obtained by deriving new consequences from the premises (in casu *Dab*-consequences), may lead to the withdrawal of earlier reached conclusions or to the rehabilitation of earlier withdrawn conclusions.

The dynamic behaviour of **PAL** resembles the dynamics present in abductive reasoning (see Sections 1 and 3). Consequently, **PAL** seem particularly well suited to explicate abductive reasoning.

4.2.4 Proof theory

As **PAL** are standard adaptive logics, the **PAL**-proof theory has some characteristic features shared by all adaptive logics. First of all, a **PAL**-proof is a succession of stages, each consisting of a sequence of lines. Adding a line to a proof means to move on to the next stage of the proof. Secondly, the lines of a **PAL**-proof consist of four elements (instead of the usual three): a line number, a formula, a justification and an adaptive condition. The latter is a finite subset of Ω (the union of the sets of abnormalities of a prioritized adaptive logic). Finally, the **PAL**-proof theory consists of both *deduction rules* and a *marking criterion*. Both of these will be discussed below.

4.2.5 Deduction rules

The deduction rules determine how new lines may be added to a proof. Below, the deduction rules are listed in shorthand notation, with

$$A \qquad \Delta$$

expressing that the formula A occurs in the proof on a line with condition Δ.

PREM If $A \in \Gamma$:

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

RU If $A_1, \ldots, A_n \vdash_{\mathbf{LLL}} B$:

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

RC If $A_1, \ldots, A_n \vdash_{\mathbf{LLL}} B \vee Dab(\Theta)$:

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

The adaptive condition of a line i expresses that as long as all its elements can be considered as false, the formula on that line may be considered as derivable from the premise set. Secondly, in order to indicate that not all elements of the adaptive condition of a line i can be considered as false, line i is marked—formally, this is done by placing the symbol ✓ next to the adaptive condition. Obviously, when a line is marked, the formula on that line may not be considered as derivable anymore. Finally, the marking in a **PAL**-proof is dynamic: at some stage of the proof, a line might be marked (resp. unmarked), while at a later stage, it might become unmarked (resp. marked) again.

4.2.6 Marking criterion

At every stage of a **PAL**-proof, the marking criterion determines which lines have to be marked. To determine whether a line has to be marked at a stage s of a **PAL**-proof, both the reliability strategy as well as the normal selections strategy first determine the *minimal Dab-consequences* of the premise set at stage s.

Definition 11
$Dab^i(\Delta)$ is a minimal *Dab*-consequence of a premise set Γ at stage s of a proof, iff (i) $Dab^i(\Delta)$ occurs on an unmarked line k at stage s, (ii) all members of the adaptive condition of line k belong to some Ω_j in the sequence $\Omega_>$ such that $\Omega_j > \Omega_i$ and (iii) there is no $\Delta' \subset \Delta$ such that (i) and (ii) apply to $Dab(\Delta')$ as well.

It is important to notice that the minimal *Dab*-consequences of a premise set at a stage s are determined in a *stepwise* manner: first for priority level 1, then for priority level 2,...

Well now, the marking definitions for **PAL** based on the reliability strategy and the normal selections strategy are the following.

Definition 12 (Reliability)
Line i with adaptive condition Δ is marked at stage s iff $Dab^i(\Theta)$ is a minimal *Dab*-consequence of Γ at stage s and $\Theta \cap \Delta \neq \emptyset$.

Definition 13 (Normal selections)
Line i with adaptive condition Δ is marked at stage s iff $Dab^i(\Theta)$ is a minimal *Dab*-consequence of Γ at stage s, and $\Theta \subset \Delta$.

4.2.7 Defining derivability

A formula is considered as derivable from a premise set Γ, in case it occurs as the second element of an unmarked line in a proof from Γ.

Definition 14 (Derivability)
The formula A is *derived* from Γ at stage s of a **PAL**-proof iff A is the second element of an unmarked line at stage s.

Because of the dynamic nature of adaptive proofs, markings may change at every stage. Hence, at every stage of a proof, it has to be reconsidered whether or not a formula is derivable. In other words, derivability is stage dependent. Although this may seem problematic at first, it nevertheless reflects the way people treat abductive consequences. For, given that abductive consequences are provisional consequences, conclusions drawn by relying on abductive inference steps are hardly ever conclusive. Hence, at any moment, two options are available to people, namely, to keep on reasoning until conclusiveness has been reached or to base one's actions on the provisional conclusions. As the first option may take more time than available, the latter option will be the only viable one in a lot of cases.[12]

[12] For a more extensive justification of this claim, see Batens *et al*. [22]. In [22], the (un)decidability of **AL** is discussed as well. In short, **AL** are undecidable. This is not surprising, for **AL** are intended to capture consequence relations that are undecidable such as for example the abductive consequence relations in this article. Of course, one may consider only

Besides a stage-dependent notion of derivability, a stable notion of derivability can be defined as well. It is called *final derivability*, which refers to the fact that, for some formulas, derivability is only decided at the final stage of a proof.

DEFINITION 15 (Final Derivability)
The formula A is *finally derived* from Γ on line i of a **PAL**-proof at stage s iff (i) A is the second element of line i, (ii) line i is not marked at stage s and (iii) every extension of the proof in which line i is marked may be further extended in such a way that line i is unmarked again.

Because of its stability, the notion of final derivability is used to define **PAL**-derivability.

DEFINITION 16
$\Gamma \vdash_{\mathbf{PAL}} A$ (A is **PAL**–derivable from Γ) iff A is finally derived on a line of a **PAL**-proof from Γ.

4.3 The prioritized adaptive logics $\mathbf{AbL^p}$ *and* $\mathbf{AbL^t}$

In this final section, I will characterize the prioritized adaptive logics $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$. First, I will show how these logics fit the standard format. Secondly, I will show that both logics characterize abductive reasoning as a combination of multiple defeasible inference rules. Thirdly, I will argue that the abductive consequences of both logics still satisfy the conditions put forward by the traditional logic-based approaches to abduction. To conclude, I will also point out the main difference between both logics.

4.3.1 Preliminary remark

I will limit myself to the propositional fragment of both $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$. However, extending these logics to their full predicate versions is completely straightforward, and hence can safely be left to the reader.

4.3.2 Characterizing $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$

As both $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$ are prioritized adaptive logics in standard format, they are characterized by means of a lower limit logic, an ordered sequence of sets of abnormalities and an adaptive strategy. First, consider those characterizing the logic $\mathbf{AbL^p}$:

- The **LLL** of $\mathbf{AbL^p}$ is the modal logic **RBK** (see Section 2).
- The abnormalities of $\mathbf{AbL^p}$ are characterized by the ordered sequence $\Omega_{>} = \Omega_{bk} > \Omega_p$, with

$$\Omega_{bk} \;=\; \{\Box_x A \mid x \in \{n, e\} \text{ and } A \in \mathcal{W}\}.$$
$$\Omega_p \;=\; \{\Box_n(A \supset B) \wedge B \wedge \neg\Box_e B \wedge \neg A \mid B \in \mathcal{S} \cup \mathcal{S}^{\neg},\ A \text{ in Conjunctive Normal Form, and } B \text{ is not a subformula of } A\}.$$

- The adaptive strategy of $\mathbf{AbL^p}$ is the reliability strategy.

Next, consider the elements characterizing the logic $\mathbf{AbL^t}$. These differ from those of $\mathbf{AbL^p}$ with respect to the sequence $\Omega_{>}$ as well as with respect to the adaptive strategy.

decidable fragments of these consequence relations, a popular trick in **AI** approaches to abduction—see Gabriele [6]. However, as a lot of interesting theories are undecidable, this is not always the most interesting option—see Meheus [23] for a more extensive discussion.

– The **LLL** of $\mathbf{AbL^t}$ is the modal logic **RBK** (again, see Section 2).
– The abnormalities of $\mathbf{AbL^t}$ are characterized by the ordered sequence $\Omega_{>} = \Omega_{bk} > \Omega_t$, with

$$\Omega_{bk} = \{ \Box_x A \mid x \in \{n,e\} \text{ and } A \in \mathcal{W} \}.$$
$$\Omega_t = \{ \Box_n(A \supset B) \wedge \neg\Box_n B \wedge B \wedge \neg\Box_e B \wedge \neg A \mid A, B \in \mathcal{S} \cup \mathcal{S}^{\neg}, \text{ and } B \text{ is not a subformula of } A \} \cup \{ \Box_n((A_1 \wedge \ldots \wedge A_n) \supset B) \wedge \neg\Box_n((A_2 \wedge \ldots \wedge A_n) \supset B) \wedge \neg\Box_n((A_1 \wedge A_3 \wedge \ldots \wedge A_n) \supset B) \wedge \ldots \wedge \neg\Box_n((A_1 \wedge \ldots \wedge A_{n-1}) \supset B) \wedge B \wedge \neg\Box_e B \wedge \neg(A_1 \wedge \ldots \wedge A_n) \mid A_1, \ldots, A_n, B \in \mathcal{S} \cup \mathcal{S}^{\neg}, \text{ and } B \text{ is not a subformula of } A_1 \wedge \ldots \wedge A_n \}.$$

– The adaptive strategy of $\mathbf{AbL^t}$ is the normal selections strategy.

In view of the standard format of (prioritized) adaptive logics outlined above, a semantic or proof theoretic characterization for these logics need not be provided anymore.

4.3.3 A formal explication of abductive explanation

Earlier on, I stated that to capture the layered character of abduction adaptive logics characterize abduction processes proof theoretically as specific combinations of multiple defeasible inference rules. For the kinds of abductive explanation explicated by the logics $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$, these defeasible inference rules are **NEN**, **NNN** and $\mathbf{AC^m}$ (see Section 3).

4.3.4 Preliminary remarks

Because of space limitations, all proofs presented below are $\mathbf{AbL^p}$-proof as well as $\mathbf{AbL^t}$-proofs. In order to make a clear distinction between both kinds of proofs, lines in a proof are given two adaptive conditions, one for each logic. Markings related to the logics $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$ are placed next to the corresponding adaptive condition. Some abbreviations are introduced as well. First of all, $\Box_n(A \supset B) \wedge \ldots \in \Omega_p$ is abbreviated as $\langle A, B \rangle^p$. Analogously, $\Box_n(A \supset B) \wedge \ldots \in \Omega_t$ is abbreviated as $\langle A, B \rangle^t$. When the ambiguous $\langle A, B \rangle^{p/t}$ is used, $\langle A, B \rangle^p$ is meant in the $\mathbf{AbL^p}$-proof, while $\langle A, B \rangle^t$ is meant in the $\mathbf{AbL^t}$-proof. Finally, $\Omega_{i_1,\ldots,i_n}$ is used to refer to the union of the adaptive conditions of lines $i_1, \ldots, i_n$.

4.3.5 Combining defeasible inference rules

In order to show how $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$ combine multiple defeasible inference rules to explicate abductive explanation, consider the adaptive proof below. It is based on the premise set $\Gamma = \{\Box_n(p \supset q),\ q\}$.

1	$\Box_n(p \supset q)$	–;PREM	$\emptyset$	$\emptyset$
2	q	–;PREM	$\emptyset$	$\emptyset$

At this stage of the proof, the premises have been introduced. These clearly show that p is a possible explanation for q. In order to derive p as an abductive consequence of Γ, the formulas $\neg\Box_e q$ and $\neg\Box_n q$ have to be derived first. This is done as follows:

3	$\Box_e q \vee \neg\Box_e q$	–;RU	$\emptyset$	$\emptyset$
4	$\Box_n q \vee \neg\Box_n q$	–;RU	$\emptyset$	$\emptyset$
5	$\neg\Box_e q$	3;RC	$\{\Box_e q\}$	$\{\Box_e q\}$
6	$\neg\Box_n q$	4;RC	$\{\Box_n q\}$	$\{\Box_n q\}$

Both $\neg\Box_e q$ and $\neg\Box_n q$ are conditional consequences of the premise set Γ. As is shown below, their derivation is necessary in order to derive p as an abductive consequence of the premise set Γ.[13]

7	$p\vee\neg p$	–;RU	$\emptyset$	$\emptyset$
8	$p\vee\langle p,q\rangle^{p/t}$	1,2,5,(6,)7;RU	Ω_5	$\Omega_{5,6}$
9	p	8;RC	$\Omega_5\cup\{\langle p,q\rangle^p\}$	$\Omega_{5,6}\cup\{\langle p,q\rangle^t\}$

At this stage of the proof, the formula p has been conditionally derived on line 9. Hence, as long as all elements of the adaptive condition of line 9 may be interpreted as false, p may be considered a (conditional) abductive consequence of the premise set Γ.

The first two applications of **RC** in the proof above, namely, those resulting in the conditional derivation of the formulas $\neg\Box_e q$ and $\neg\Box_n q$, clearly correspond to applications of the defeasible inference rules **NEN** and **NNN**. On the other hand, the third application of **RC** in the proof, namely, the one resulting in the conditional derivation of the formula p, corresponds to an application of the defeasible inference rule $\mathbf{AC^m}$. Remark that the formulas obtained by the first two applications of **RC** are required for the third application of **RC**. In other words, the formulas obtained by applying the inference rules **NEN** and **NNN** correspond to some of the conditions that have to be satisfied before the inference rule $\mathbf{AC^m}$ may be applied (see Section 3). However, notice that $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$ require slightly different conditions to be satisfied. For $\mathbf{AbL^p}$ only formulas of the form $\neg\Box_e A$ are required. For $\mathbf{AbL^t}$ on the other hand, also formulas of the form $\neg\Box_n A$ are required. In the proof above, this is clear from the fact that in the $\mathbf{AbL^t}$-version the formula on line 6 is necessary for the derivation of p on line 9, while it is not in the $\mathbf{AbL^p}$-version.

4.3.6 Meaning of the additional conditions

Let us have a closer look at the additional conditions captured by the consequences of the inference rules **NEN** and **NNN**. Firstly, consider the formulas of the form $\neg\Box_e A$ obtained by means of the defeasible inference rule **NEN**. Both $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$ only allow applications of $\mathbf{AC^m}$ in case $\neg\Box_e A$ is derivable for the explanandum A. The reason for this is quite simple: the formula $\neg\Box_e A$ guarantees that the explanandum A cannot be explained by means of the background theory alone. Hence, A is in need of an explanation and is allowed to trigger abductive inferences. It is easily verified that in case the explanandum A is derivable from the background theory alone, $\Box_e A$ will be derivable from the premises. As a consequence, $\neg\Box_e A$ will be withdrawn, as will all abductive consequences triggered by A. To illustrate this, consider again the proof above, but suppose that the formula $\Box_e p$ is added to the premise set Γ. As a consequence, the proof can be extended in the following way:

...	...	...	...		...	
5	$\neg\Box_e q$	3;RC	$\{\Box_e q\}$	✓	$\{\Box_e q\}$	✓
...	...	...	...		...	
8	$p\vee\langle p,q\rangle^{p/t}$	1,2,5,(6,)7;RU	Ω_5	✓	$\Omega_{5,6}$	✓
9	p	8;RC	$\Omega_5\cup\{\langle p,q\rangle^p\}$	✓	$\Omega_{5,6}\cup\{\langle p,q\rangle^t\}$	✓
10	$\Box_e p$	–;PREM	$\emptyset$		$\emptyset$	
11	$\Box_e q$	1,10;RU	$\emptyset$		$\emptyset$	

[13] In the justification of line 8, the reference to line 6 is placed between brackets in order to express that it is only necessary in the $\mathbf{AbL^t}$ version of the proof.

At this stage of the proof, the formula $\Box_e q$ has been derived on line 11. As this is a minimal, *Dab*-consequence of Γ at stage 11, lines 5, 8 and 9 are marked. Consequently, the formula $\neg\Box_e q$ cannot be considered as derivable anymore, and neither can the formula p.

Secondly, consider the formulas of the form $\neg\Box_n A$ obtained by means of the defeasible inference rule **NNN**. These are used in the logic $\mathbf{AbL^t}$ to guarantee that the inference step known as *strengthening the antecedent* does not enable one to derive abductive explanations containing irrelevant parts. For, the logic $\mathbf{AbL^t}$ only validates an abductive inference based on a nomological formula $\Box_n((A_1 \wedge \ldots \wedge A_n) \supset q)$ in case the formulas $\neg\Box_n((A_2 \wedge \ldots \wedge A_n) \supset q)$, $\neg\Box_n((A_1 \wedge A_3 \wedge \ldots \wedge A_n) \supset q)$,...,$\neg\Box_n((A_1 \wedge \ldots \wedge A_{n-1}) \supset q)$ are derivable. It is easily verified that this will never be the case when the formula $\Box_n((A_1 \wedge \ldots \wedge A_n) \supset q)$ has been obtained by means of strengthening the antecedent. To illustrate this, suppose the proof above (the original proof, i.e. lines 1–9) is extended in the following way:

10 $\Box_n((p \wedge r) \supset q)$	1;RU	$\emptyset$		$\emptyset$	
11 $\neg\Box_n(p \supset q)$	–;RC	$\{\Box_n(p \supset q)\}$	✓	$\{\Box_n(p \supset q)\}$	✓
12 $\neg\Box_n(r \supset q)$	–;RC	$\{\Box_n(r \supset q)\}$		$\{\Box_n(r \supset q)\}$	
13 $p \wedge r$	2,5,10(,11,12);RC	$\Omega_5 \cup \{\langle p \wedge r, q\rangle^p\}$		$\Omega_{5,11,12} \cup \{\langle p \wedge r, q\rangle^t\}$	✓

Clearly, the antecedent of $\Box_n((p \wedge r) \supset q)$ contains an irrelevant part, namely, r. The latter has been added to the antecedent of the nomological formula on line 1 by an application of strengthening the antecedent. However, as $\Box_n(p \supset q)$ is a minimal *Dab*-consequence of the premise set Γ, lines 11 and 13 are marked. Hence, neither $\neg\Box_n(p \supset q)$ nor $p \wedge r$ are considered as derivable from Γ at stage 13 of the proof (and at all later stages of the proof). A small digression is necessary at this point. Line 13 is only marked for the logic $\mathbf{AbL^t}$. Hence, the formula $p \wedge r$ seems to be derivable from Γ by the logic $\mathbf{AbL^p}$. This is not the case though, for $\mathbf{AbL^p}$ also blocks abductive inferences based on nomological statements obtained by strengthening the antecedent. Only, $\mathbf{AbL^p}$ does not need formulas of the form $\neg\Box_n A$ to do so, as the following extension of the proof above shows.

13 $p \wedge r$	2,5,10(,11,12);RC	$\Omega_5 \cup \{\langle p \wedge r, q\rangle^p\}$	✓	$\Omega_{5,11,12} \cup \{\langle p \wedge r, q\rangle^t\}$	✓
14 $\langle p \wedge r, q\rangle^p \vee \langle p \wedge \neg r, q\rangle^p$	1,5;RC	Ω_3		Ω_3	

To conclude, suppose that in the example above the formulas $\Box_n(p \supset q)$ and $\Box_n((p \wedge r) \supset q)$ are both elements of the premise set Γ. Hence, the latter is not obtained from the former by application of the inference rule strengthening the antecedent. Nonetheless, this does not change the status of the formula $p \wedge r$, for it is still not finally derivable from Γ by $\mathbf{AbL^p}$ or $\mathbf{AbL^t}$. Actually, this means that these logics not only block the abductive inferences of explanations based on nomological background knowledge implications that *are* in fact obtained by means of strengthening of the antecedent, but also those that are based on nomological background knowledge implications that *could have been* obtained by means of strengthening the antecedent. In other words, these logics only enable one to derive the *minimal explanations* for a given phenomenon.

4.3.7 Comparison with the 'backwards deduction'-approaches

In order to show that both $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$ capture abductive explanation in an adequate way, I will show that both logics satisfy the conditions for abductive explanation put forward by the traditional 'backwards deduction'-approaches to abduction (see Section 1, Table 1).

Condition (i) states that the background knowledge extended by an abductive explanation has to yield the explanandum. Given the dependency of the defeasible inference rule $\mathbf{AC^m}$ on nomological statements derivable from the background knowledge, this condition is satisfied a priori.

Condition (ii) states that an abductive explanation has to be compatible with the background knowledge. It is easily verified that this will be the case for both $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$. For, in case an abductive consequence derived on a line *i* in a proof is incompatible with the background knowledge, line *i* will irrevocably be marked at some stage of the proof. For example, consider the premise set $\Gamma = \{\Box_n(p \supset q), q, \Box_e \neg p\}$. As in the proof above (lines 1–9), the formula p is conditionally derivable from the premise set Γ. However, the proof can be extended in such a way that line 9 is marked.

...	...	...	...		...	
9	p	8;RC	$\Omega_5 \cup \{\langle p,q\rangle^p\}$	✓	$\Omega_{5,6} \cup \{\langle p,q\rangle^t\}$	✓
10	$\Box_e \neg p$	–;PREM	$\emptyset$		$\emptyset$	
11	$\langle p,q\rangle^{p/t}$	1,2,5,(6,)10;RU	Ω_5		$\Omega_{5,6}$	

At stage 11 of the proof, the formula $\langle p,q\rangle^{p/t}$ is a minimal *Dab*-consequence of the premise set Γ. As a consequence, line 9 is marked.

Condition (iii) states that the explanandum may not be derivable from the background knowledge alone. As I have shown above, this condition is satisfied for both $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$.

Condition (iv) states that an abductive explanation may not yield the explanandum by itself. Actually, this is satisfied by the fact that applications of $\mathbf{AC^m}$ are only validated conditionally in $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$ in case the nomological statements involved are of a specific syntactic form. This is a consequence of the way Ω_p and Ω_t were defined. For example, the elements of Ω_t are of the form $\Box_n((A_1 \wedge \ldots \wedge A_n) \supset B) \wedge \ldots \wedge \neg(A_1 \wedge \ldots \wedge A_n)$. However, B is not allowed to occur in $A_1 \wedge \ldots \wedge A_n$ (check the definition of Ω_t above)! As a consequence, it is impossible for $A_1 \wedge \ldots \wedge A_n$ to yield B by itself. The same reasoning also applies to the elements of Ω_p.

Finally, as I have shown above, the abductive explanations obtained by $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$ are minimal explanations (and hence, never contain irrelevant parts). Although Aliseda–Llera did not state this as a necessary condition for consistent and explanatory abduction in [1], it is easily verified that this should be a necessary condition (and in a lot of traditional logic based approaches, it also is).

4.3.8 Practical versus theoretical abductive explanation

The logics $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$ explicate different kinds of abductive explanation, namely, practical abductive explanation and theoretical abductive explanation, respectively. As stated at the beginning of this section, in case a puzzling phenomenon has multiple possible explanations, practical abduction only yields the disjunction of these explanations, while theoretical abduction yields all explanations separately.[14] As a consequence, the former is more cautious than the latter, for practical abduction would not enable one to act on a single possible explanation in case there are multiple. This is appropriate for contexts in which it is important that no possible explanations are overlooked, e.g. when trying to diagnose the disease causing a patient's symptoms—in case there is more than one possibility, acting on a single one would be foolish, for this could leave the patient uncured. On the other hand, in some contexts one might want to derive all possible explanations, e.g. in case one wishes to compare the predictions yielded by various scientific explanations. On the basis of this comparison, one may then decide which explanation should be favoured.

[14] The distinction between both kinds of abduction was introduced by Meheus and Batens [10, p. 224].

Example

To illustrate the different kinds of abductive explanation explicated by the logics $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$, respectively, consider the example below, based on the premise set $\Gamma = \{\Box_n(p \supset q), \Box_n(r \supset q), \Box_n \neg(p \wedge r), q\}$. For the premise set Γ, the logic $\mathbf{AbL^p}$ should enable one to derive the disjunction $p \vee r$, while the logic $\mathbf{AbL^t}$ should enable one to derive both p and r separately. As a matter of fact, this is exactly what happens.

1	$\Box_n(p \supset q)$	–;PREM	$\emptyset$	$\emptyset$
2	$\Box_n(r \supset q)$	–;PREM	$\emptyset$	$\emptyset$
3	q	–;PREM	$\emptyset$	$\emptyset$
4	p	1,3;RC	$\{\Box_e q, \langle p, q\rangle^p\}$	$\{\Box_e q, \Box_n q, \langle p, q\rangle^t\}$
5	r	2,3;RC	$\{\Box_e q, \langle r, q\rangle^p\}$	$\{\Box_e q, \Box_n q, \langle r, q\rangle^t\}$

At stage 5 of the proof, p and r have been derived on an unmarked line in both the $\mathbf{AbL^p}$-proof and the $\mathbf{AbL^t}$-proof. Both proofs now proceed differently. Hence, I will consider them separately, starting with the $\mathbf{AbL^p}$-proof.

...	...	...	...		–
4	p	1,3;RC	$\{\Box_e q, \langle p, q\rangle^p\}$	✓	–
5	r	2,3;RC	$\{\Box_e q, \langle r, q\rangle^p\}$	✓	–
6	$\langle p, q\rangle^p \vee \langle r \wedge \neg p, q\rangle^p$	1–3;RC	$\{\Box_e q\}$		–
7	$\langle r, q\rangle^p \vee \langle p \wedge \neg r, q\rangle^p$	1–3;RC	$\{\Box_e q\}$		–
8	$p \vee r$	1–3;RC	$\{\Box_e q, \langle p \vee r, q\rangle^p\}$		–

At stage 10 of the $\mathbf{AbL^p}$-proof, two minimal *Dab*-consequences of the premise set Γ have been derived, namely, $\langle p, q\rangle^p \vee \langle r \wedge \neg p, q\rangle^p$ on line 8 and $\langle r, q\rangle^p \vee \langle p \wedge \neg r, q\rangle^p$ on line 9. As a consequence, lines 6 and 7 are marked, which implies that neither p nor r is considered as derivable anymore. However, the disjunction of p and r is considered as derivable, for the formula $p \vee r$ occurs on an unmarked line of the proof (line 10 to be precise). Moreover, it is easily verified that line 10 will remain unmarked in any extension of the proof. Hence, the formula $p \vee r$ is a final abductive $\mathbf{AbL^p}$-consequence of Γ.

Now, consider the $\mathbf{AbL^t}$-proof below. At first, this proof seems to proceed as the $\mathbf{AbL^p}$-proof above. However, because the logic $\mathbf{AbL^t}$ is based on the normal selections strategy instead of the reliability strategy, neither line 6 nor line 7 is marked (nor will these lines be marked in any extension of the proof). As a consequence, both p and r are final abductive $\mathbf{AbL^t}$-consequences of Γ.

...	...	...	–	...	
6	p	1,3;RC	–	$\{\Box_e q, \Box_n q, \langle p, q\rangle^t\}$	
7	r	2,3;RC	–	$\{\Box_e q, \Box_n q, \langle r, q\rangle^t\}$	
8	$\langle p, q\rangle^t \vee \langle r \wedge \neg p, q\rangle^t$	1–3;RC	–	$\{\Box_e q, \Box_n q, \Box_n(r \supset q), \Box_n(\neg p \supset q)\}$	✓
9	$\langle r, q\rangle^t \vee \langle p \wedge \neg r, q\rangle^t$	1–3;RC	–	$\{\Box_e q, \Box_n q, \Box_n(p \supset q), \Box_n(\neg r \supset q)\}$	✓
10	$p \wedge r$	6,7;RU	–	$\{\Box_e q, \Box_n q, \langle p, q\rangle^t, \langle r, q\rangle^t\}$	

To conclude, consider the formula on line 10 of the $\mathbf{AbL^t}$-proof above. This is the formula $p \wedge r$, namely, the conjunction of both possible explanations for q. As line 10 is unmarked at this stage of the proof, the formula $p \wedge r$ is a conditional consequence of the premise set Γ. As distinct possible explanations are usually considered as mutually exclusive, this clearly is absurd. However, different possible explanations for the same phenomenon do not have to be mutually exclusive, for one of these may yield the other(s)—in the example above, this would be the case if $\Box_n(r \supset p)$ would have been an element of the premise set Γ. In this case, the derivation of the conjunction of multiple possible explanations makes perfect sense. Nonetheless, in case the possible explanations are mutually exclusive, their conjunction should not be derivable. As is shown below, this is exactly what happens in $\mathbf{AbL^t}$-proofs. Given that $\Box_n \neg(p \wedge r) \in \Gamma$, p and r are mutually exclusive. Hence, the line on which their conjunction occurs will get marked eventually. For example, in case the proof is extended as follows.

...	...	...	–	...	
10	$p \wedge r$	6,7;RU	–	$\{\Box_e q, \Box_n q, \langle p,q\rangle^t, \langle r,q\rangle^t\}$	✓
11	$\Box_n \neg(p \wedge r)$	–;PREM	–	$\emptyset$	
12	$\langle p,q\rangle^t \vee \langle r,q\rangle^t$	1–3,11;RC	–	$\{\Box_e q, \Box_n q\}$	

5 Conclusion

The (prioritized) adaptive logics $\mathbf{AbL^p}$ and $\mathbf{AbL^t}$ provide a formal explication of practical and theoretical abductive explanation, respectively. In contradistinction to the traditional logic-based approaches to abduction, these logics not only capture abductive explanation metatheoretically and/or semantically, but also proof theoretically, namely, as a combination of multiple defeasible inference rules. In general, this shows that logics for abduction based on the adaptive logics programme provide a more realistic explication of abductive explanation than most traditional logic-based approaches.

5.1 Further research

In this article, I only provided a formal explication of practical and theoretical abductive explanation. These are not the only kinds of abductive reasoning though, for a lot of other abduction processes have been characterized in the literature—e.g. preferential abductive explanation, abductive explanation triggered by an anomaly (a formula contradicting the background theory),... —for an overview, see e.g. Aliseda–Llera [2]. The formal explication of these abduction processes is left for further research. As a consequence, the logics presented in this article should be considered as the first step in the direction of a general formal approach to abductive reasoning based on the adaptive logics programme.

Acknowledgements

The author is a Postdoctoral Fellow of the Special Research Fund of Ghent University. I am indebted to Diderik Batens and Joke Meheus for their helpful suggestions that significantly improved the logics presented in this article. Moreover, I also wish to thank the anonymous referees for pointing out some unclarities in a previous version of this article, as well as Bert Leuridan for his suggestions on how to state things more clearly. Of course, all remaining unclarities are mine.

References

[1] A. Aliseda-Llera. *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence*. PhD Thesis, University of Amsterdam, 1997.

[2] A. Aliseda-Llera. *Abductive Reasoning. Logical Investigations into Discovery and Explanation*. Vol. 330 of *Synthese Library*. Kluwer, 2006.

[3] M. C. Mayer and F. Pirri. First order abduction via tableau and sequent calculi. *Logic Journal of the IGPL*, **1**, 99–117, 1993.

[4] M. C. Mayer and F. Pirri. Propositional abduction in modal logic. *Logic Journal of the IGPL*, **3**, 907–919, 1995.

[5] S. McIlraith. Logic–based abductive inference. *Technical Report Number KSL-98-19*. Knowledge Systems Laboratory, 1998.

[6] P. Gabriele. AI approaches to abduction. In *Abductive Reasoning and Learning*. D. Gabbay and R. Kruse, eds, Vol. 4 of *Handbook of Defeasable Reasoning and Uncertainty Management Systems*. pp. 35–98. Kluwer, 2000.

[7] J. Meheus and D. Provijn. Abduction through semantic tableaux versus abduction through goal-directed proofs. *Theoria*, **60**, 295–304, 2007.

[8] D. Batens. A general characterization of adaptive logics. *Logique et Analyse*, **173–175**, 45–68, 2001.

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

[10] J. Meheus and D. Batens. A formal logic for abductive reasoning. *Logic Journal of the IGPL*, **14**, 221–236, 2006.

[11] J. Meheus. A formal logic for the abduction of singular hypotheses. (Forthcoming).

[12] J. Meheus, L. Verhoeven, M. Van Dyck, D. Provijn. Ampliative adaptive logics and the foundation of logic-based approaches to abduction. In *Logical and Computational Aspects of Model-Based Reasoning*. L. Magnani, N. J. Nersessian, and C. Pizzi, eds, pp. 39–71. Kluwer, 2002.

[13] J. Garson. *Modal Logic for Philosophers*. Cambridge University Press, 2006.

[14] V. Hendricks. Epistemic Logic. In E. N. Zalta, ed. *The Stanford Encyclopedia of Philosophy*, 2009. Availabe at http://plato.stanford.edu/archives/spr2009/entries/logic-epistemic/

[15] D. Batens. *Menselijke Kennis. Pleidooi voor een Bruikbare Rationaliteit*. Garant, Antwerpen/Apeldoorn, 2004.

[16] D. Batens. Do we need a hierarchical model of science? In *Inference, Explanation, and Other Frustrations: Essays in the Philosophy of Science*, J. Earman, ed. pp. 199–215. University of California Press, 1992.

[17] C. Hempel and P. Oppenheim. Studies in the logic of explanation. *Philosophy of Science*, **15**, 135–175, 1948.

[18] D. Batens, J. Meheus, D. Provijn, and L. Verhoeven. Some adaptive logics for diagnosis. *Logic and Logical Philosophy*, **11–12**, 39–65, 2003.

[19] D. Batens, J. Meheus, and D. Provijn. An adaptive characterization of signed systems for paraconsistent reasoning, to appear.

[20] L. Verhoeven. All premisses are equal, but some are more equal than others. *Logique et Analyse*, **173–174–175**, 165–188, 2001.

[21] G. Primiero, and J. Meheus. Majority merging by adaptive counting. *Synthese*, **165**, 203–223, 2008.

[22] D. Batens, K. De Clercq, P. Verdée, and J. Meheus. Yes fellows, most human reasoning is complex. *Synthese*, **166**, 113–131, 2009.

[23] J. Meheus. Adaptive logics for abduction and the explication of explanation–seeking processes. In O. Pombo, and A. Gerner, eds, *Abduction and the Process of Scientific Discovery*, pp. 97–119. Centro de Filosofia das Ciências da U. de Lisboa, 2007.