A Formal Logic for Abductive Reasoning

JOKE MEHEUS and DIDERIK BATENS, *Centre for Logic and Philosophy of Science, Ghent University, Blandijnberg 2, B–9000 Gent, Belgium. E-mail: Joke.Meheus@UGent.be, Diderik.Batens@UGent.be*

Abstract

This paper presents and illustrates a formal logic for the abduction of singular hypotheses. The logic has a semantics and a dynamic proof theory that is sound and complete with respect to the semantics. The logic presupposes that, with respect to a specific application, the set of *explananda* and the set of possible *explanantia* are disjoint (but not necessarily exhaustive). Where an *explanandum* can be explained by different *explanantia*, the logic allows only for the abduction of their disjunction.

Keywords: abduction, adaptive logic, dynamic reasoning.

1 Introduction

Many logicians display disinterest or even suspicion with respect to abduction. The reason is twofold. The first is that abductive steps are of the form

$$B(\beta),\ (\forall\alpha)(A(\alpha) \supset B(\alpha))\ /\ A(\beta), \tag{1.1}$$

a fallacy known as Affirming the Consequent (combined with Universal Instantiation). The second is that many examples of purportedly sound abductions seem to rely on a hidden non-formal reasoning: the only sensible formal rule behind them seems to lead inevitably to a set of unsound and even inconsistent conclusions. For instance, from the explananda Qa and Ra and the generalizations $(\forall x)(Px \supset Qx)$ and $(\forall x)(\neg Px \supset Rx)$, (1.1) enables one to generate both Pa and $\neg Pa$.[1]

In this paper, we shall present a logic for the abduction of singular hypotheses, $\mathbf{LA}^r$. We were only able to forge this logic by introducing a restriction, and not a very original one. Where $\mathcal{W}$ is the set of closed formulas of the standard predicative language, we introduce two sets of truth functions of closed primitive formulas,[2] $\mathcal{W}^e$ and $\mathcal{W}^a$, requiring that no primitive formula occurs in a member of $\mathcal{W}^e$ as well as in a member of $\mathcal{W}^a$. The sets may but need not be combinatorially closed, in other words, they need not contain all subformulas of their members or all truth-functions of these subformulas.

[1] Outside the domain of formal logic, logic-based approaches to abduction are quite popular at the moment. In the domain of Artificial Intelligence, for instance, they recently led to an impressive number of systems for a wide variety of application contexts (such as diagnostic reasoning, text understanding, case-based reasoning, and planning)—see [19] for an overview. But also in the domain of cognitive science and philosophy of science they proved very fruitful. Examples are Hintikka's analysis of abduction in terms of the interrogative model of inquiry (see especially [15]), Aliseda's approach to abduction in terms of semantic tableaux (see [1]), Magnani's integration of results on diagnostic reasoning and scientific reasoning (see [16] and [17]), and Thagard's reconstruction of several important discoveries in the history of science and in the history of medicine (see, for instance, [20] and [21]). A reconstruction of these logic-based approaches in terms of ampliative adaptive logics is presented in [18].

[2] Primitive formulas are those that contain no logical symbols, except possibly for identity.

Intuitively, $\mathcal{W}^e$ is the set of *explananda*, formulas that are considered as requiring an explanation, whereas $\mathcal{W}^a$ is the set of *explanantia*, formulas that, if they can be abduced, form potential explanations for the *explananda*. The requirement that no primitive formula occurs in members of both sets can be easily justified with respect to applications. If one tries to abduce an explanation, one has in mind a phenomenon for which an explanation is sought, and the explanation should be logically independent of the explained phenomenon—everyone rejects (even partial) self-explanations. Similarly, one often looks for an explanation of a set of phenomena, for example the symptoms displayed by a patient. Here too the sought explanations will be in terms of diseases, or in terms of the past history of the patient, but not in terms of symptoms.

The sets $\mathcal{W}^e$ and $\mathcal{W}^a$ are seen as application dependent. After abducing $Ra \wedge Sa$ in order to explain $Pa \vee Qa$, nothing prevents one from seeking an explanation for Ra. So where $Ra \wedge Sa$ belonged to $\mathcal{W}^a$ for the first application, Ra belongs to $\mathcal{W}^e$ for the second application. Remark that we do not have to require that abducted knowledge has a lower degree of certainty than the original premises. The premises may be closed under $\mathbf{LA}^r$ with respect to a couple $\langle \mathcal{W}^e, \mathcal{W}^a \rangle$, and the resulting set may be closed under $\mathbf{LA}^r$ with respect to another such couple.

Some will claim that the reference to $\mathcal{W}^e$ and $\mathcal{W}^a$ turns $\mathbf{LA}^r$ into a non-formal logic, for example because uniformity (as standardly defined) fails. We consider such objections as mainly verbal. Consider the expression

$$\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle \vdash_{\mathbf{LA}^r} A\,.$$

If this expression is true, then so is the expression obtained by systematically replacing in the expression one schematic letter by a letter of the same sort (a sentential letter, a predicate of a certain rank, an individual constant or an individual variable) that does not occur in the original expression. The operation may be repeated on the result, etc. Put differently, if an expression as the above one is true, then so is every expression that shares all logical forms with it. This is as good a criterion for formality as any other.

Incidentally, for many application contexts, $\mathcal{W}^a$ may be taken to be a function of $\mathcal{W}^e$, viz. the set of all formulas that do not contain any primitive formula that occurs in $\mathcal{W}^e$. As the reader will see, the premise set may then be written as $\langle \Gamma, \mathcal{W}^e \rangle$, and the requirement that the abduced conclusions belong to the so defined $\mathcal{W}^a$ may be pushed into the logic itself (because it now became a purely formal matter). And $\langle \Gamma, \mathcal{W}^e \rangle$ may be interpreted as a set of declarative premises, Γ, together with a set of explanation questions, viz. why questions about the members of $\mathcal{W}^e$ that are $\mathbf{CL}$-derivable from Γ.

As one would expect, $\mathbf{LA}^r$ has some non-standard properties (it is non-monotonic, for instance). We shall also show that it adequately captures the main characteristics of abductive reasoning processes.

One such characteristic is that abductive steps are combined with deductive steps. Partly because of this combination, abductive reasoning processes are *dynamical.* For instance, a conclusion reached on the basis of an abductive step may be withdrawn when its negation is derived by deductive means.

An important property of $\mathbf{LA}^r$ is that it not only nicely integrates deductive and abductive steps, but that it moreover has a decent proof theory. This proof theory is dynamical, but warrants that the conclusions derived at a given stage are justified in view of the insight in the premises at that stage. Another advantage of the presented logic is that, as compared to other existing systems for abductive reasoning, it is very close to natural reasoning.

The logic presented in this paper will be based on Classical Logic—henceforth **CL**. So, all references to causality, laws of nature, and similar non-extensional concepts will be out of the picture. We do not doubt that more interesting results may be obtained from intensional logics. However, we want to keep the discussion as simple and transparant as possible. Moreover, that we are able to phrase an interesting abductive logic within an extensional context is rather fascinating in itself.

The results in the present paper are an outcome of the adaptive logic programme. Adaptive logics are a family of non-standard logics that are especially suited to study, in a formally exact way, reasoning processes that are non-monotonic and/or dynamical.[3] The first logic in this family was designed around 1980 (see [2]) and was meant to interpret (possibly) inconsistent theories *as consistently as possible*.[4] Later the notion of an adaptive logic was generalized in different ways (for instance, to capture ampliative forms of reasoning) and a whole variety of adaptive logics was designed—see [4] and [13] for a survey.

2 Preliminaries for a Logic of Abduction

As a first approximation, a logic may be called abductive if and only if it is obtained by extending **CL** with a suitably restricted version of rule (1.1). The restrictions will distinguish between sound and unsound applications of the rule. A first restriction is obviously that $B(\beta) \in \mathcal{W}^e$ and $A(\beta) \in \mathcal{W}^a$. We also have to require that $(\forall\alpha)(A(\alpha) \supset B(\alpha))$ is not a **CL**-theorem. In view of the first restriction, this rules out cases in which $A(\beta)$ is a contradiction or $B(\beta)$ is a tautology—nobody wants to seek an explanation for a tautology and nobody will accept an explanation by *ex falso quodlibet*.

There are some more restrictions. Given the formal character of **CL**, the claim that there are *formal* abductive logics commits one to the following statement:

> If some application of (1.1) is sound and some other application of it is not, then there should be a *formal difference* between the two.

Although this statement is correct, it is not free of ambiguity. Rules of logic are applied against the background of some set of premises, say Γ.[5] The formal character of a logic does not derive from the fact that there is a formal link between the premises of the application of some rule and its conclusion, but from the fact that there is a formal link between (a subset of) Γ and the last step in a proof (respectively, the semantic consequence). In the case of monotonic logics, the formal character of the logic warrants the formal character of the rules. In the case of non-monotonic logics, it does not. Here, the soundness of an application of a rule may depend on the set of premises. So the formal character of an abductive logic depends on whether there is a *formal* difference between sound and unsound applications of the following

[3] A reasoning pattern is called dynamical if the mere analysis of the premises may lead to the withdrawal of previously drawn conclusions. Not all dynamical reasoning patterns are non-monotonic. In [5], for instance, it is shown that the pure logic of relevant implication can be characterized by a dynamic proof theory.

[4] Logics that satisfy this property are referred to as *inconsistency-adaptive logics*.

[5] In order to keep the discussion as simple as possible, we comply with the usual supposition that all relevant knowledge about some domain may be considered as a set of premises Γ.

reformulation of (1.1):

$$\text{If } \Gamma \vdash_{\mathbf{CL}} B(\beta),\ \Gamma \vdash_{\mathbf{CL}} (\forall\alpha)(A(\alpha) \supset B(\alpha)),\ B(\beta) \in \mathcal{W}^e,\ A(\beta) \in \mathcal{W}^a, \\ \text{and } \nvdash_{\mathbf{CL}} (\forall\alpha)(A(\alpha) \supset B(\alpha)), \text{ then } \langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle \vdash_{\mathbf{LA}^r} A(\beta) \tag{2.1}$$

A reasonable further requirement on an abductive logic is that it consistently extends the **CL**-consequences of Γ. This requirement has some immediate consequences for our abductive logic.

Suppose that $\Gamma \vdash_{\mathbf{CL}} Pa$ and $\Gamma \vdash_{\mathbf{CL}} (\forall x)(Qx \supset Px)$ and that we want an explanatory hypothesis for Pa. Suppose further that $\mathcal{W}^e$ consists of all singular formulas that contain no other predicates than P and R, and that $\mathcal{W}^a$ comprises all remaining singular formulas. Applying (2.1) delivers Qa. But, quite obviously, we do not want to draw this conclusion from the above premises if $\neg Qa$ is **CL**-derivable from Γ.

What if $\neg Qa$ is *abductively* derivable from the Γ from the previous paragraph? The simplest case is where also $(\forall x)(\neg Qx \supset Px) \in \Gamma$. In this case, $\Gamma \vdash_{\mathbf{CL}} (\forall x)Px$, and hence $\Gamma \vdash_{\mathbf{CL}} (\forall x)(A(x) \supset Px)$ for any $A(x)$. So (given our extensional framework), it does not seem to make sense to abductively derive any explanation for Pa. In a slightly more complicated case, an attempt to explain Pa might lead to $\neg Qa$ by a series of applications of (2.1)—first to S_1a, from there to S_2a, ..., and from S_na to $\neg Qa$. However, this is only possible if $\Gamma \vdash_{\mathbf{CL}} (\forall x)(\neg Qa \supset Px)$, which brings us back to the previous situation.

A further complication is where, $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$ being as two paragraphs ago, $\Gamma \vdash_{\mathbf{CL}} Ra$ and $\Gamma \vdash_{\mathbf{CL}} (\forall x)(\neg Qx \supset Rx)$. If we look for an explanatory hypothesis for both Pa and Ra, (2.1) will enable us to arrive at both Qa and $\neg Qa$. Both Qa and $\neg Qa$ might be considered as sensible conclusions. Even if we do so, this does not force us to consider, say, Sa as a sensible conclusion. Nevertheless, at least one of the conclusions has to be rejected in view of the above requirement that abduction should lead to a consistent consequence set. As, in the present case, there is no formal difference between the applications of (2.1) that lead to Qa and $\neg Qa$ respectively, both applications have to be rejected in view of the requirement that our abductive logic be formal.

Things become more difficult once we consider more complex formulas. Where $\mathcal{W}^e$ and $\mathcal{W}^a$ are as before, consider a Γ such that:

$$\Gamma \vdash_{\mathbf{CL}} Pa \tag{2.2}$$

$$\Gamma \vdash_{\mathbf{CL}} (\forall x)((Qx \wedge Sx) \supset Px) \tag{2.3}$$

Applying (2.1) delivers $Qa \wedge Sa$, but is this a sound abduction? To see that it not always is, suppose that $\Gamma \vdash_{\mathbf{CL}} (\forall x)(Qx \supset Px)$. In this case, $\Gamma \vdash_{\mathbf{CL}} (\forall x)((Qx \wedge A(x)) \supset Px)$ for any $A(x)$. So, (2.2) and (2.3) cannot warrant that $Qa \wedge Sa$ is abductively derivable from Γ—if they did, the set of abductive consequences would be trivial. As we shall see below, also this case is adequately handled by the logic $\mathbf{LA}^r$.

Having introduced some general restrictions on (1.1), we turn to the purposes that an application of abduction may serve. There are at least two rather different ones. Consider the case of a patient a displaying some symptom P who consults a physician to get cured. Suppose that the physician's theoretical knowledge contains $(\forall x)(Qx \supset Px)$ and $(\forall x)(Rx \supset Px)$, and no other (sensible) candidates for an abductive step. It would be rather stupid of the physician to conclude to Qa and to act accordingly.

This would be stupid because, if Ra is the case, rather than Qa, the patient would not be cured. So, the appropriate behaviour for the physician would be to draw the conclusion $Qa \vee Ra$ and to test whether Qa, Ra or both are true, or to act in such a way that the patient gets cured in either case.

Compare this situation to one in which a ‘theoretician’ has the same knowledge (or a set of knowledge of the same logical form), but is merely interested in finding and testing explanatory hypotheses for Pa. In this case, there would be no harm if the theoretician derived, say, Qa and tested it. If it turns out true, an explanation is produced. If it turns out false, Ra might be the next hypothesis derived.

In this paper, we shall concentrate on the type of situation in which we have to act on the abducted conclusion and hence better take all possibilities into account. In line with this, our logic of abduction will lead to a set of explanatory hypotheses that are not only jointly compatible with the premises but also as weak as possible in view of them.

3 Illustration of the Logic

The general idea behind $\mathbf{LA}^r$ is extremely simple: it is allowed that (2.1) is applied “as much as possible”. For the moment, this ambiguous phrase may be interpreted as “unless and until $(\forall\alpha)(A(\alpha) \supset B(\alpha)) \wedge (B(\beta) \wedge \neg A(\beta))$ turns out to be **CL**-derivable from Γ”. So, whenever it is **CL**-derivable from Γ that, for some general rule $(\forall\alpha)(A(\alpha) \supset B(\alpha)$ and some explanandum $B(\beta)$, (2.1) cannot be applied consistently (because, $\neg A(\beta)$ is **CL**-derivable from Γ), the application of (2.1) is overruled. In view of what we have seen in the previous section, this is exactly what we want.

To save space, expressions of the form $(\forall\alpha)(A(\alpha) \supset B(\alpha)) \wedge (B(\beta) \wedge \neg A(\beta))$ will be abbreviated as $[\![B(\beta), \neg A(\beta)]\!]$ and, in line with what is common for adaptive logics, $[\![B(\beta), \neg A(\beta)]\!]$ will be called an “ abnormality”.[6] As we will see below, it is possible that a disjunction of abnormalities is **CL**-derivable from a set of premises Γ without any of its disjuncts being derivable from it. This fact will prove crucial to obtain an adequate logic for abduction.

We shall devote the sequel of this section to an illustration of the proof theory that we shall spell out in Section 4. We shall present a simple example, not worrying too much about technicalities, but concentrating on the way in which the requirements from the previous section are met. Suppose that our set of premises Γ consists of the following generalizations

$$(\forall x)(Px \supset Rx),\ (\forall x)(Px \supset Sx),\ (\forall x)(Qx \supset Sx),\ (\forall x)(Qx \supset Tx),$$
$$(\forall x)(\neg Px \supset Tx)$$

and the following data

$$Ra,\ Rb,\ \neg Sb,\ Sc,\ Sd,\ \neg Td,\ Re,\ Te.$$

Let $\mathcal{W}^e$ be the set of all singular formulas that are truth-functions of primitive formulas containing the predicates R, S and T, and $\mathcal{W}^a$ the set of all singular formulas that do not contain these predicates.

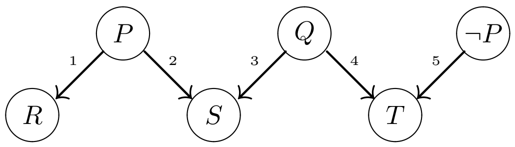
One way to start a $\mathbf{LA}^r$-proof from Γ is by entering all the premises:

[6] The term “abnormality” refers to formulas that overrule the application of some desired inference rule—in our case the abduction scheme (1.1).

1	$(\forall x)(Px \supset Rx)$	–	PREM	$\emptyset$
2	$(\forall x)(Px \supset Sx)$	–	PREM	$\emptyset$
3	$(\forall x)(Qx \supset Sx)$	–	PREM	$\emptyset$
4	$(\forall x)(Qx \supset Tx)$	–	PREM	$\emptyset$
5	$(\forall x)(\neg Px \supset Tx)$	–	PREM	$\emptyset$
6	Ra	–	PREM	$\emptyset$
7	Rb	–	PREM	$\emptyset$
8	$\neg Sb$	–	PREM	$\emptyset$
9	Sc	–	PREM	$\emptyset$
10	Sd	–	PREM	$\emptyset$
11	$\neg Td$	–	PREM	$\emptyset$
12	Re	–	PREM	$\emptyset$
13	Te	–	PREM	$\emptyset$

For each of these lines, the third and fourth element form the "justification" for the formula that constitutes the second element. The third element contains the line numbers of the formulas from which the formula is derived (obviously empty in the case of premises); the fourth element contains the name of the rule by means of which the formula is derived (in the above case the premise rule PREM). The empty sets at the end of each line can be ignored for the moment.

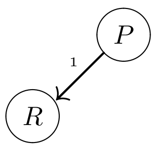
For reasons of transparency, we shall from now on represent the proof (as much as possible) in a diagrammatic way:



	P	Q	R	S	T
a			$+_6$		
b			$+_7$	$-_8$	
c				$+_9$	
d				$+_{10}$	$-_{11}$
e			$+_{12}$		$+_{13}$

The node-and-arrow-structure represents the generalizations (for instance, the first arrow stands for $(\forall x)(Px \supset Rx)$) and the array represents the data (the "+" in the first row stands for Ra, the "−" in the second row for $\neg Sb$). The numbers in the diagram refer to the stage at which the corresponding formula is entered in the proof.

We are now in a position to make inferences from the premises. Let us first concentrate on the explanandum Ra. As is easily observed, the generalization represented by the first arrow can be used to 'abduce' an explanatory hypothesis for Ra. In an $\mathbf{LA}^r$-proof from Γ, this is done by applying the rule RC:



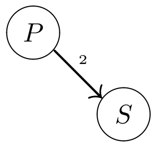
	P	Q	R	S	T
a			$+_6$		

14	Pa	$1, 6$	RC	$\{[\![Ra, \neg Pa]\!]\}$

RC is a conditional rule: it allows one to add abductive hypotheses to the proof, but only on a certain condition. This condition is represented by the fifth element of the line. Intuitively, line 14 can be read as: Pa is derivable from the formulas on lines 1 and 6, *unless and until* it can no longer be assumed (consistently) that $[\![Ra, \neg Pa]\!]$ is *false.*

Given our present insights in the premises (represented by the formulas that are explicitly written down in the proof), there is no reason to believe that $\neg Pa$ is true, and hence, it is consistent to assume that $[\![Ra, \neg Pa]\!]$ is false. This is why, at this stage of the proof,[7] Pa is considered to be derivable from the premises (in view of line 14). If, at a later stage of the proof, it would turn out that the condition of line 14 is no longer satisfied, then this line will be 'marked' and the formula that occurs on it will no longer be considered to be derived. (The marking of lines will be illustrated below.)

In view of the formula on line 14, the generalization represented by the second arrow allows one to infer the prediction Sa; this is done by means of the rule RU:

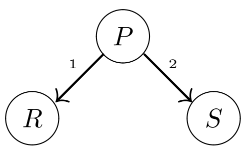


	P	Q	R	S	T
a	$+_{14}$				

15 Sa $2, 14$ RU $\{[\![Ra, \neg Pa]\!]\}$

RU is a generic rule that allows one to infer all **CL**-consequences: whenever some formula A is **CL**-derivable from a number of formulas $B_1, \ldots, B_n$ that are considered to be derived in the proof at some stage, then, at that stage, A can be added to the proof by means of RU. Note that RU is an *unconditional* rule: unlike RC, it does not lead to the introduction of new conditions. If, however, some of the B_i to which RU is applied are themselves derived on a non-empty condition, then these conditions are conjoined for the conclusion. Thus, as the formula of line 14 is used to derive the formula on line 15, the condition of the former is 'carried over' to the latter. This is obviously as it should be: if, at a later stage in the proof, the conclusion of line 14 is withdrawn because its condition is no longer satisfied, then all formulas that rely on it should also be withdrawn.

Let us now turn to the explanandum Rb. As in the previous case, the rule RC enables us to abduce an explanatory hypothesis for Rb (see line 16 below). However, this time, we are also able to infer, by means of RU, the *negation* of our explanatory hypothesis:

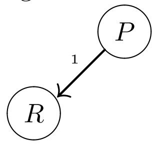


	P	Q	R	S	T
b			$+_{7}$	$-_{8}$	

16 Pb $1, 7$ RC $\{[\![Rb, \neg Pb]\!]\}$
17 $\neg Pb$ $2, 8$ RU $\emptyset$

Hence, we are able to infer the following abnormality:

[7]Remember that the proof theory of $\mathbf{LA}^r$ is dynamical: formulas that are considered to be derived at some stage in the proof, may no longer be considered as derived at a later stage of the proof. The dynamics will be illustrated below.



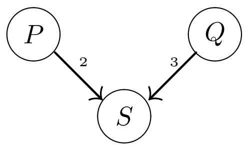
	P	Q	R	S	T
b	$+_{16} \| -_{17}$		$+_7$		

18 $[\![Rb, \neg Pb]\!]$ 1, 7, 17   RU $\emptyset$

At this stage in the proof, the condition of line 16 is no longer satisfied. As a consequence, the conclusion of line 16 is *withdrawn* from the proof. The withdrawal of a conclusion from the proof is recorded by *marking* the line on which the formula occurs. This is how the proof looks like at stage 18 (lines 1 to 15 are as before):

...				
16	Pb	1, 7	RC	$\{[\![Rb, \neg Pb]\!]\}$ $\checkmark^{18}$
17	$\neg Pb$	2, 8	RU	$\emptyset$
18	$[\![Rb, \neg Pb]\!]$	1, 7, 17	RU	$\emptyset$

We shall now show what happens when more than one explanatory hypothesis can be abduced for the same explanandum. Have a look at Sc:



	P	Q	R	S	T
c				$+_9$	

In view of the relevant generalizations, the proof can be extended as follows:

19	Pc	2, 9	RC	$\{[\![Sc, \neg Pc]\!]\}$
20	Qc	3, 9	RC	$\{[\![Sc, \neg Qc]\!]\}$

However, as the reader can verify, the following disjunctions of abnormalities are **CL**-derivable from the premises:

21	$[\![Sc, \neg Pc]\!] \vee [\![Sc, \neg(Qc \wedge \neg Pc)]\!]$	2, 3, 9	RU	$\emptyset$
22	$[\![Sc, \neg Qc]\!] \vee [\![Sc, \neg(Pc \wedge \neg Qc)]\!]$	2, 3, 9	RU	$\emptyset$

The formula on line 21 expresses that $[\![Sc, \neg Pc]\!]$ or $[\![Sc, \neg(Qc \wedge \neg Pc)]\!]$ is true. Hence, it cannot be assumed that both disjuncts are false.

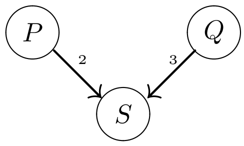
In view of such a disjunction of abnormalities, different strategies are possible. The one followed by $\mathbf{LA}^r$ is very cautious. As (at this stage of the proof) it is unclear which one of the two disjuncts is true, both disjuncts are (at this stage of the proof) considered as 'unreliable'. As a result, all formulas that are derived on the assumption that one of these disjuncts is false, are withdrawn. Thus, in our case, the formula on line 19 is withdrawn in view of the formula on line 21. By an analogous reasoning, the formula on line 20 is withdrawn in view of the formula on line 22:

...				
19	Pc	2, 9	RC	$\{[\![Sc, \neg Pc]\!]\}$ $\checkmark^{21}$
20	Qc	3, 9	RC	$\{[\![Sc, \neg Qc]\!]\}$ $\checkmark^{22}$
21	$[\![Sc, \neg Pc]\!] \vee [\![Sc, \neg(Qc \wedge \neg Pc)]\!]$	2, 3, 9	RU	$\emptyset$
22	$[\![Sc, \neg Qc]\!] \vee [\![Sc, \neg(Pc \wedge \neg Qc)]\!]$	2, 3, 9	RU	$\emptyset$

A mark may be removed at a later stage. Suppose, for example, that $[\![Sc, \neg(Qc \wedge \neg Pc)]\!]$ is **CL**-derivable from the premises, and is actually derived in the proof. So it would be clear which of the two disjuncts of the formula of line 21 is true, viz. the second one. As a result, line 19 would not be marked any more (unless $[\![Sc, \neg Pc]\!]$ is a disjunct of another disjunction of abnormalities).

We shall see that, apart from derivability at a stage, one can define a stable notion of derivability, viz. final derivability. Intuitively, a formula is finally derived on line i of a proof iff it is possible to extend the proof in such a way that line i is unmarked and remains unmarked in every further extension of the proof.

In view of the present premises, lines 19 and 20 will remain marked in any extension of the proof, whence neither Pc nor Qc is finally derivable from the premises. However, their disjunction $Pc \vee Qc$ is. This can be seen from the following extension of the proof:

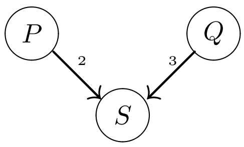


	P	Q	R	S	T
c				$+_9$	

23 $(\forall x)((Px \vee Qx) \supset Sx)$ 2, 3   RU $\emptyset$
24 $Pc \vee Qc$ 9, 23   RU $\{[\![Sc, \neg(Pc \vee Qc)]\!]\}$

As no minimal disjunction of abnormalities is derivable that has $[\![Sc, \neg(Pc \vee Qc)]\!]$ as one of its disjuncts, the formula on line 24 is finally derivable from the premises.[8]

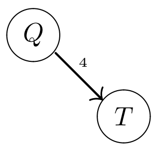
Also for the explanandum Sd the rule RC enables one to derive a disjunction of explanatory hypotheses:



	P	Q	R	S	T
d				$+_{10}$	

25 $Pd \vee Qd$ 2, 3, 10   RC $\{[\![Sd, \neg(Pd \vee Qd)]\!]\}$

This time, however, one of the disjuncts can be eliminated by pure deductive means:



	P	Q	R	S	T
d					$-_{11}$

26 $\neg Qd$ 4, 11   RU $\emptyset$
27 Pd 25, 26   RU $\{[\![Sd, \neg(Pd \vee Qd)]\!]\}$

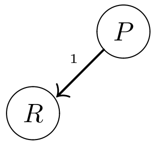
This nicely illustrates how $\mathbf{LA}^r$ allows for the integration of deductive and abductive steps.[9]

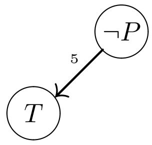
Let us finally turn to the situation where different explanatory hypotheses are mutually incompatible with the premises. As may be seen from the following extension

[8] Every disjunction of abnormalities can be seen as $\bigvee(\Delta)$ in which Δ is a finite set of abnormalities and $\bigvee(\Delta)$ is their disjunction. $\bigvee(\Delta)$ is a *minimal* disjunction of abnormalities that is derivable from the premises if and only if $\bigvee(\Delta)$ is derivable from the premises and there is no $\Delta' \subset \Delta$ such that $\bigvee(\Delta')$ is derivable from the premises.

[9] The attentive reader will have observed that Pd is *also* derivable on the condition $\{[\![Sd, \neg Pd]\!]\}$.

of the proof, this is the case for the explanatory hypotheses that are abducible for Re and Te:





	P	Q	R	S	T
e			$+_{12}$		$+_{13}$

28	Pe	$1, 12$	RC	$\{[\![Re, \neg Pe]\!]\}$
29	$\neg Pe$	$5, 13$	RC	$\{[\![Te, Pe]\!]\}$

Although both these hypotheses may be entered at some stage in the proof, neither of them is finally derivable from the premises. This is warranted by the following **CL**-derivable disjunction of abnormalities:

30	$[\![Re, \neg Pe]\!] \vee [\![Te, Pe]\!]$	$1, 5, 12, 13$	RU	$\emptyset$

As soon as the formula on line 30 is added to the proof, lines 28 and 29 are marked—they remain marked in any extension of the proof.

Note that this mechanism also comes into play when the antecedent of some generalization is arbitrarily strengthened. Suppose, for instance, that the proof is extended in the following way:

31	$(\forall x)((Px \wedge Ux) \supset Rx)$	1	RU	$\emptyset$
32	$(\forall x)((Px \wedge \neg Ux) \supset Rx)$	1	RU	$\emptyset$
33	$Pa \wedge Ua$	$6, 31$	RC	$\{[\![Ra, \neg(Pa \wedge Ua)]\!]\}$
34	$Pa \wedge \neg Ua$	$6, 32$	RC	$\{[\![Ra, \neg(Pa \wedge \neg Ua)]\!]\}$

If the formulas on lines 33 and 34 would be finally derivable from the premises, one would not only obtain explanatory hypotheses that are partly irrelevant, but even triviality:

35	Ua	33	RU	$\{[\![Ra, \neg(Pa \wedge Ua)]\!]\}$
36	$\neg Ua$	34	RU	$\{[\![Ra, \neg(Pa \wedge \neg Ua)]\!]\}$
37	p	$35, 36$	RU	$\{[\![Ra, \neg(Pa \wedge Ua)]\!], [\![Ra, \neg(Pa \wedge \neg Ua)]\!]\}$

However, the following disjunction of abnormalities is **CL**-derivable:[10]

38	$[\![Ra, \neg(Pa \wedge Ua)]\!] \vee [\![Ra, \neg(Pa \wedge \neg Ua)]\!]$	$1, 6$	RU	$\emptyset$

As soon as this line is added to the proof, lines 33–37 are marked.[11] It is moreover easily observed that these lines will remain marked in any further extension of the proof, and hence, that the formulas on these lines are not finally derivable.

The example discussed was rather simple. As we have no room to multiply examples, let us at least remark that there is no problem for $\mathbf{LA}^r$ to abduce from a theory an explanation for a complex statement, for example $Pa \wedge Qa$.

[10] Actually, the fact that the formulas on lines 33 and 34 are jointly incompatible with the premises warrants that the disjunction of the union of their conditions is **CL**-derivable from the premises.

[11] Even if 32, 34, 36 and 37 had not been derived, 38 would be derivable and would cause line 31 to be marked.

4 Precise Description of the Logic

We shall restrict the discussion to unary predicates (predicates expressing properties). It is very well possible that our result may be generalized, but we have no proof, at this moment, that it can.

$\mathbf{LA}^r$ can be formulated in the standard format from [6] and [9], which greatly simplifies the technical stuff. An adaptive logic **AL** is in standard format if it is characterized as a triple consisting of three elements: (i) **LLL**, a compact and monotonic lower limit logic, (ii) Ω, a set of abnormalities that all have the same logical form, and (iii) an adaptive strategy.

The lower limit logic **LLL** determines the part of the adaptive logic **AL** that is not subject to adaptation. From a proof theoretic point of view, the lower limit logic delineates the rules of inference that hold unexceptionally. From a semantic point of view, the adaptive models of a premise set Γ are a selection of the **LLL**-models of Γ. It follows that $Cn_{\mathbf{LLL}}(\Gamma) \subseteq Cn_{\mathbf{AL}}(\Gamma)$. The lower limit logic of $\mathbf{LA}^r$ is obviously **CL**, and remember that its premise set is $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$.

As we have seen, abnormalities are formulas that are presupposed to be false, unless and until proven otherwise. Ω comprises all formulas of a certain (possibly restricted) logical form. In the case of $\mathbf{LA}^r$ the restriction will refer to $\mathcal{W}^e$ and $\mathcal{W}^a$.[12]

For $\mathbf{LA}^r$ we define $\Omega = \{(\forall\alpha)(A(\alpha) \supset B(\alpha)) \wedge (B(\beta) \wedge (\neg A(\beta)) \mid A(\beta) \in \mathcal{W}^a;\ B(\beta) \in \mathcal{W}^e;\ \nvdash_{\mathbf{CL}} (\forall\alpha)(A(\alpha) \supset B(\alpha))\}$. In the present extensional framework, $(\forall\alpha)(A(\alpha) \supset B(\alpha))$ can be taken to express that A contains a (sufficient) cause for B—we write "'contains" because A may be itself a conjunction and some of its conjuncts may not be required for warranting B.[13] The second conjunct of an abnormality states that the specific sufficient cause $A(\beta)$ for $B(\beta)$ did not occur. The requirement that $(\forall\alpha)(A(\alpha) \supset B(\alpha))$ is not a **CL**-theorem has to be added in order to prevent that all models would display abnormalities, as we shall see when we come to the upper limit logic. However, as was explained in Section 2, this requirement is harmless. An adaptive logic presupposes that abnormalities are false unless and until proven otherwise. So, the presupposition of $\mathbf{LA}^r$ is that, if an effect did occur, then all its potential causes (in the weak, extensional, sense) did also occur.

The strategy will be Reliability. It is well-known that Minimal Abnormality would deliver a few more consequences in peculiar (and somewhat weird) cases. However, as its marking definition is rather tiresome, this would have unnecessarily complicated the discussion of the example in Section 3.

If one adds to the lower limit logic an axiom schema excluding that abnormalities occur, viz. an axiom schema that reduces abnormal premise sets to triviality, one obtains the so-called upper limit logic. The upper limit logic of $\mathbf{LA}^r$ is somewhat unusual as it refers to the sets $\mathcal{W}^e$ and $\mathcal{W}^a$. It is obtained by extending **CL** with the axiom schema $(\forall\alpha)(A(\alpha) \supset B(\alpha)) \supset (B(\beta) \supset A(\beta))$ *provided* $B(\beta) \in \mathcal{W}^e$ and $A(\beta) \in \mathcal{W}^a$. It is easily seen that this comes to the requirement that, if the proviso is met, $(\forall\alpha)(A(\alpha) \supset B(\alpha))$ is logically equivalent to $(\forall\alpha)(A(\alpha) \equiv B(\alpha))$. We shall not care to give this upper limit logic a name.

---

[12] The specific restriction that will be imposed causes $\mathbf{LA}^r$ not to be strictly in standard format. However, the metatheoretic claims that we want to derive from the standard format still can be proved. Moreover, the logic has a hardly different variant, in which $\mathcal{W}^a$ is a function of $\mathcal{W}^e$ as described in Section 1, that is in standard format.

[13] Every man who is too lazy to shave and wears spectacles has a beard, but the fact that the second author wears spectacles is not part of a potential cause for his having a beard.

In some cases, the upper limit logic for ampliative adaptive logics has no practical application context because none of its models corresponds to the actual world—for an example see uniform classical logic, **UCL**, from [10]. This does not apply to the upper limit logic of $\mathbf{LA}^r$. It is useful to discuss this briefly as it clarifies the circumstances under which abductions can be derived.

Our upper limit logic presupposes that every statement in $\mathcal{W}^e$ has a unique and maximally specific cause in $\mathcal{W}^a$. Let us consider an concrete example for a very simple language, in which only occur the predicates P, Q and R and, say, twenty individual constants $a_1, \ldots, a_{20}$. Suppose moreover that $\mathcal{W}^e$ contains the twenty formulas Pa_i $(1 \leq i \leq 20)$, and that $\mathcal{W}^a$ comprises all formulas in which P does not occur. Consider first a chaotic model, in which some elements of the domain have properties P, Q and R, some elements have properties P and Q but not R, and so on for all eight combinations. As this model verifies no contingent generalization of the form $(\forall x)(A(x) \supset Px)$, for $A(\beta) \in \mathcal{W}^a$, no abnormalities occur in it and hence it is an upper limit model (it is a **CL**-model that verifies the added axiom schema).

As this model is plainly uninteresting with respect to abduction, one wonders whether there are others, and indeed there are. Consider any model that verifies $(\forall x)((Qx \wedge Rx) \equiv Px)$, but falsifies $(\forall x)(\pm Qx \supset Px)$ as well as $(\forall x)(\pm Rx \supset Px)$ (in which each $\pm$ is either a negation or nothing)—hence it also falsifies $(\forall x)((Qx \wedge \neg Rx) \supset Px)$ and so on. It is easily seen that this model verifies not a single abnormality, and hence is an upper limit model. The same holds for any model that verifies $(\forall x)((\pm Qx \wedge \pm Rx) \equiv Px)$, but falsifies $(\forall x)(\pm Qx \supset Px)$ as well as $(\forall x)(\pm Rx \supset Px)$.

No upper limit model verifies $(\forall x)(Qx \supset Px)$ unless it also verifies $\neg(\exists x)Px$. Indeed, if a model would verify the former formula, it would also verify $(\forall x)((Qx \wedge Rx) \supset Px)$ as well as $(\forall x)((Qx \wedge \neg Rx) \supset Px)$, and hence, if some object, say a, had property P, the model would verify $[\![Pa, \neg(Qa \wedge Ra)]\!] \vee [\![Pa, \neg(Qa \wedge \neg Ra)]\!]$.

Of course, there are $\mathbf{LA}^r$-models that verify $(\forall x)(Qx \supset Px)$ as well as Pa. All of these will verify some abnormalities, but some do not verify a disjunction of abnormalities of which $[\![Pa, \neg Qa]\!]$ is a disjunct. This is precisely what makes adaptive logics interesting, viz. that they interpret abnormal premise sets as normally as possible.

Let us now turn to the proofs. If the deduction rules are formulated in generic format, they are identical for all adaptive logics in standard format. Where Γ contains the (declarative) premises as before,

$$A \quad \Delta$$

abbreviates that A occurs in the proof on the condition Δ, and $Dab(\Delta)$ is the disjunction of the members of a finite $\Delta \subset \Omega$, the rules may be phrased as follows:[14]

[14]The only rule that introduces non-empty conditions is RC. In other words, before RC is applied in a proof, the condition of every line will be $\emptyset$.

PREM If $A \in \Gamma$:

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

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

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

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

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

There is a striking correspondence between $\mathbf{LA}^r$-proofs and **CL**-proofs. Suppose that one transforms each line

$$A \quad \Delta$$

from the $\mathbf{LA}^r$-proof into

$$A \vee Dab(\Delta)\,,$$

where "$\vee Dab(\emptyset)$" is defined as the empty string. It is easy enough to establish, by an obvious induction on the length of the proof, that the resulting sequence of formulas is a **CL**-proof obtained by applications of PREM and RU only. This result is extremely useful from a metatheoretic point of view and clarifies what is going on in a dynamic proof.

We now turn to the *marking definition*. We shall say that $Dab(\Delta)$ is a *minimal Dab-formula* at stage s of a proof if, at that stage, $Dab(\Delta)$ occurs in the proof on the empty condition and, for any $\Delta' \subset \Delta$, $Dab(\Delta')$ does not occur in the proof on the empty condition. Where $Dab(\Delta_1)$, …, $Dab(\Delta_n)$ are the minimal Dab-formulas at stage s of the proof, $U_s(\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle) = \Delta_1 \cup \ldots \cup \Delta_n$ is the set of unreliable formulas at stage s. The marking definition for the Reliability strategy is as follows:

Definition 4.1
Line i is marked at stage s iff, where Δ is its condition, $\Delta \cap U_s(\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle) \neq \emptyset$.

If $Dab(\Delta)$ is a minimal Dab-formula at stage s of the proof, then, in as far as one knows in view of the proof at this stage, the premises require one of the abnormalities in Δ to be true but do not specify which one is true. The Reliability strategy considers all of them as unreliable. So the underlying idea is: if the understanding of the premises provided by the present stage of the proof is correct, the formulas occurring at unmarked lines are derivable from the premises, whereas the formulas occurring at marked lines are not.

Apart from the unstable derivability at a stage, one wants a stable kind of derivability, which is called final derivability.

Definition 4.2
A is *finally derived* from $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$ on line i of a 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) any extension of the proof in which line i is marked may be further extended in such a way that line i is unmarked.

Definition 4.3
$\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle \vdash_{\mathbf{LA}^r} A$ (A is *finally* $\mathbf{LA}^r$*-derivable* from Γ) iff A is finally derived on a line of a $\mathbf{LA}^r$-proof from $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$.

Remark that these are definitions, and that they are not intended to have a direct computational use.

The semantics of all adaptive logics is defined in the same way in terms of the lower limit logic, here **CL**, the set of abnormalities Ω and the strategy. $M \models A$ will denote that M assigns a designated value to A, in other words that M verifies A. $M \models \Gamma$ will denote that M verifies all members of Γ.

The abnormal part of a **CL**-model M will be defined as follows:

Definition 4.4
$Ab(M) = \{A \in \Omega \mid M \models A\}$

Where $Dab(\Delta_1)$, $Dab(\Delta_2)$, ... are the minimal Dab-consequences of $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$,[15] $U(\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle) = \Delta_1 \cup \Delta_2 \cup \ldots$ is the set of formulas that are unreliable with respect to $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$.[16]

Definition 4.5
A **CL**-model M of $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$ is *reliable* iff $Ab(M) \subseteq U(\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle)$.

Intuitively, $U(\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle)$ comprises the abnormalities that, in view of the Reliability strategy, cannot be avoided if all members of Γ are supposed to be true. A reliable model of $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$ is one that verifies at most the members of $U(\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle)$.

Definition 4.6
$\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle \models_{\mathbf{LA}^r} A$ iff A is verified by all reliable models of $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$.

It is provable by standard means that $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle \vdash_{\mathbf{LA}^r} A$ iff $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle \models_{\mathbf{LA}^r} A$ (that the syntax is sound and complete with respect to the semantics). Some further desirable properties are provable as well (see [7] and especially [9]). These properties include the essential property that $Cn_{\mathbf{LA^r}}(\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle)$ is consistent. Actually this is an easy consequence of Strict Reassurance (see [6]).

5 Concluding Remarks

The plot we had in mind is that, given a set of knowledge Γ, one abduces the (weakest) explanation of some fact or some set of facts (the members of $\mathcal{W}^e$ that are **CL**-derivable from Γ). This explanation should belong to a specific part of the language $\mathcal{W}^a$, possibly all sentences of the language that do not contain any primitive formula that occurs in $\mathcal{W}^e$. As we mentioned before, the abduced statements may be added to the premises, Γ thus being extended to Γ', and no harm results if $\mathbf{LA}^r$ is applied to Γ' with respect to a different set of explanatory questions.

We have looked into the properties of Γ that contain existentially quantified *explananda*, as in the case where Γ contains both $(\forall x)(Px \supset Qx)$ and $(\exists x)Qx$. Not much sensible can be derived from such premise sets, and we think this is quite all

[15] Obviously, the minimal Dab-consequences of $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$ may be semantically defined.

[16] In the following definition we write that "a **CL**-model M of $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$ is Reliable" because the Reliability of M depends on $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$ and not just on Γ. That M is a **CL**-model of $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$ obviously depends on Γ alone.

right. By all means, if one wants to explain the properties of some object that has no name, one can simply give it one—even if one did not observe the object, but merely knows that it exists.

In the sequel of this section, we shall mainly comment on some alternative logics that can easily be obtained from $\mathbf{LA}^r$ and look attractive. A first alternative was already suggested before, viz. that $\mathcal{W}^a$ is seen as a function of $\mathcal{W}^e$. The advantage of introducing $\mathcal{W}^a$ as an independent set is that it enables the user of the logic to restrict the explanations that may be abduced. This, however, is only a slight advantage. Nothing prevents one to be interested only in (and to seek to derive only) some of the statements that logically can be abduced.

There is a more different alternative for $\mathbf{LA}^r$ that one might prefer. Suppose that one seeks an explanation for Pa and that $Qa \vee Ra$ is finally $\mathbf{LA}^r$-derivable in view of the presence of $(\forall x)(Qx \supset Px)$ and $(\forall x)(Rx \supset Px)$ (and the absence of certain other generalizations). Suppose, however, that Γ contains also examples of cases where being Q or R is not a good explanation for being P, even if no generalization indicates which is the explanation for the P-hood of those examples. It is easy enough to produce such premise sets, for example premise sets that contain Pc, $\neg Qc$ and $\neg Rc$. The question is whether, in such cases, one is still prepared to abduce $Qa \vee Ra$.

We think there are some convincing arguments for answering the preceding question in the negative. In Section 2, we made a choice as to the purposes that an application of abduction may serve. In view of that choice, we wanted to derive the disjunction of all possible explanations for the explanandum, unless when some of these explanations are known (from Γ) to be false. Quite in line with this, one might reason that one should not abduce $Qa \vee Ra$ in the example of the previous section, because one knows that this is not the disjunction of all possible explanations for Pa. In other words, one will only abduce $Qa \vee Ra$ if one has no reason to believe that something else might be the explanation for P-hood (and not, as in the logic $\mathbf{LA}^r$ if one merely does not see another possible explanation for Pa).

It is not difficult to articulate a logic which agrees with this viewpoint. All one has to do is introduce a slight change in the definition of the set of abnormalities, viz. as follows: $\Omega = \{(\forall\alpha)(A(\alpha) \supset B(\alpha)) \wedge (\exists\alpha)(B(\alpha) \wedge \neg A(\alpha)) \mid \beta \in \mathcal{C};\ A(\beta) \in \mathcal{W}^a;\ B(\beta) \in \mathcal{W}^e;\ \nvdash_{\mathbf{CL}} (\forall\alpha)(A(\alpha) \supset B(\alpha))\}$.

Some readers will not be convinced by the properties of $\mathbf{LA}^r$. It may be shown that the set of finally derivable $\mathbf{LA}^r$-consequences of a premise set $\langle\Gamma, \mathcal{W}^e, \mathcal{W}^a\rangle$ contains exactly the formulas that one wants to abduce (together with the **CL**-consequences of the premises and the abduced statements). However, these readers will wonder, how can one be sure to have derived the finally derivable statements? Remark indeed that the formulas abduced at a stage of the proof may not include all formulas that are finally derivable and may moreover include some formulas that are not finally derivable (and will be marked at a later stage).

There are two kinds of answers to this objection. The first is that it may be justified to act on the basis of derivability at a stage. It has been shown in [3] that, as the dynamic proof proceeds, the insights in the premises may become better and never become worse (and that one can determine from the extension of the proof whether the insights in the premises did become better). This means that, at every stage of a dynamic proof, one confronts the choice between continuing the proof in order to improve the insights in the premises, or to act upon present insights. That there is such a choice is an intrinsic result of the kind of consequence relation we try to explicate—if a logic would avoid this choice, it would be a bad explication of the consequence relation.

The second answer is that, in the absence of a positive test,[17] there is no algorithm for establishing in general that A is finally derivable from Γ even if it is. This does not prevent the existence of *criteria* that enable one to establish, for a specific A, that it has been finally derived from the premises in a given proof. Some criteria were presented in [3], [11] and [12], and more criteria may be derived from results presented in those papers. Unfortunately, most of these criteria are awfully complex and only transparent for people that are well acquainted with the dynamic proofs. So we continued searching for something better. This was provided by recent work on goal-directed proofs. The idea is not to formulate a specific criterion, but rather to articulate a proof procedure that functions as a criterion. Whenever the proof procedure stops, it establishes that A is or is not finally derivable from the premises. Preparatory work on the propositional fragment of **CL** is presented in [14] and the proof procedure is applied to a (propositional) inconsistency-adaptive logic in [8]. Meanwhile the results for the predicative version are ready and it can easily be shown that these deliver criteria for final derivability with respect to any adaptive logic that has Reliability as its strategy.

The last paragraph does not make the considerations from the next-to-last one useless. Indeed, given the properties of the explicated consequence relation, no criterion can apply in all cases. So the reader might still feel unsatisfied. Suppose that one applies $\mathbf{LA}^r$ to $\langle \Gamma, \mathcal{W}^e, \mathcal{W}^a \rangle$—in other words that one tries to explain the members of $\mathcal{W}^e$ in as far as they are **CL**-derivable from Γ. Let Γ' be the union of Γ with the set of statements that have been abduced at some point in time from the previous application of $\mathbf{LA}^r$. Suppose next that one applies $\mathbf{LA}^r$ to $\langle \Gamma', \mathcal{W}^{e'}, \mathcal{W}^{a'} \rangle$. One of the troubles that might arise, is that one obtains insights that motivate one to revise conclusions from the first application of $\mathbf{LA}^r$.[18] This is not a problem for $\mathbf{LA}^r$, and we think this to be a very strong point. Indeed, several consecutive applications of $\mathbf{LA}^r$ may be combined with each other (in the same way as prioritized adaptive logics are obtained in [6]). The combined adaptive logic enables to to revise any abduction at any stage of the dynamic proof.

Several properties of $\mathbf{LA}^r$ have still insufficiently been studied. One of them concerns the effect of exchanging the order of consecutive applications of the logic. Still, we hope to have shown that $\mathbf{LA}^r$ is a sensible formal logic and that it leads to adequate results with respect to the kinds of abduction that we described in Section 2.

Acknowledgements

Research for this paper was supported by subventions from Ghent University and from the Fund for Scientific Research – Flanders.[19]

[17] We are discussing here the situation for the generalized version of $\mathbf{LA}^r$. If one restricts the language to unary predicates, one remains within a decidable fragment of the language.

[18] We think that it is this possibility which brought some scholars to the (in our view untenable) conclusion that the results of an abduction should forever have a lower degree of certainty than, for example, empirical data.

[19] Unpublished papers in the reference section (and many others) are available from the internet address `http://logica.UGent.be/centrum/writings/`.

References

[1] Atocha Aliseda. *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artifical Intelligence*. Phd thesis, Institute for Logic, Language and Computation (ILLC), University of Amsterdam, The Netherlands, 1997.

[2] Diderik Batens. Dynamic dialectical logics. In Graham Priest, Richard Routley, and Jean Norman, editors, *Paraconsistent Logic. Essays on the Inconsistent*, pages 187–217. Philosophia Verlag, München, 1989.

[3] Diderik Batens. Blocks. The clue to dynamic aspects of logic. *Logique et Analyse*, 150–152:285–328, 1995. Appeared 1997.

[4] Diderik Batens. A survey of inconsistency-adaptive logics. In Diderik Batens, Chris Mortensen, Graham Priest, and Jean Paul Van Bendegem, editors, *Frontiers of Paraconsistent Logic*, pages 49–73. Research Studies Press, Baldock, UK, 2000.

[5] Diderik Batens. A dynamic characterization of the pure logic of relevant implication. *Journal of Philosophical Logic*, 30:267–280, 2001.

[6] Diderik Batens. A general characterization of adaptive logics. *Logique et Analyse*, 173–175:45–68, 2001. Appeared 2003.

[7] Diderik Batens. The need for adaptive logics in epistemology. In Dov Gabbay, S. Rahman, J. Symons, and J. P. Van Bendegem, editors, *Logic, Epistemology and the Unity of Science*, pages 459–485. Dordrecht, Kluwer, 2004.

[8] Diderik Batens. A procedural criterion for final derivability in inconsistency-adaptive logics. *Journal of Applied Logic*, 3:221–250, 2005.

[9] Diderik Batens. A universal logic approach to adaptive logics. To appear.

[10] Diderik Batens and Lieven Haesaert. On classical adaptive logics of induction. *Logique et Analyse*, 173–175:255–290, 2001. Appeared 2003.

[11] Diderik Batens and Joke Meheus. A tableau method for inconsistency-adaptive logics. In Roy Dyckhoff, editor, *Automated Reasoning with Analytic Tableaux and Related Methods*, volume 1847 of *Lecture Notes in Artificial Intelligence*, pages 127–142. Springer, 2000.

[12] Diderik Batens and Joke Meheus. Shortcuts and dynamic marking in the tableau method for adaptive logics. *Studia Logica*, 69:221–248, 2001.

[13] Diderik Batens and Joke Meheus. Recent results by the inconsistency-adaptive labourers. In Jean-Yves Beziau and Walter A. Carnielli, editors, *Paraconsistent Logic with no Frontiers*, Studies in Logic and Practical Reasoning. North-Holland/Elsevier, in print.

[14] Diderik Batens and Dagmar Provijn. Pushing the search paths in the proofs. A study in proof heuristics. *Logique et Analyse*, 173–175:113–134, 2001. Appeared 2003.

[15] Jaakko Hintikka. What is abduction? The fundamental problem of contemporary epistemology. *Transactions of the Charles S. Peirce Society*, 34:503–533, 1998.

[16] Lorenzo Magnani. Model-based creative abduction. In Lorenzo Magnani, Nancy Nersessian, and Paul Thagard, editors, *Model-Based Reasoning in Scientific Discovery*, pages 219–238. Kluwer/Plenum, Dordrecht, 1999.

[17] Lorenzo Magnani. *Abduction, Reason, and Science. Processes of Discovery and Explanation.* Kluwer Academic / Plenum Publishers, New York, 2001.

[18] Joke Meheus, Liza Verhoeven, Maarten Van Dyck, and Dagmar Provijn. Ampliative adaptive logics and the foundation of logic-based approaches to abduction. In Lorenzo Magnani, Nancy J. Nersessian, and Claudio Pizzi, editors, *Logical and Computational Aspects of Model-Based Reasoning*, pages 39–71. Kluwer, Dordrecht, 2002.

[19] Gabriele Paul. AI approaches to abduction. In Dov M. Gabbay and Philippe Smets, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems. Volume 4: Abductive Reasoning and learning*, pages 35–98. Kluwer Academic Publishers, Dordrecht, 2000.

[20] Paul Thagard. *Computational Philosophy of Science.* MIT Press/Bradford Books, Cambridge, MA, 1988.

[21] Paul Thagard. *How Scientists Explain Disease.* Princeton University Press, 1999.

Received 14 March 2005