Logique & Analyse **173–174–175** (2001), 255–290

ON CLASSICAL ADAPTIVE LOGICS OF INDUCTION*

DIDERIK BATENS AND LIEVEN HAESAERT

Abstract
This paper concerns the inference of inductive generalizations and of predictions derived from them. It improves on the adaptive logic of induction from [6] by presenting logics that are formulated strictly according to the usual adaptive standards. It moreover extends that paper with respect to background knowledge.

We present logics that handle inductive generalizations as well as logics that handle prioritized background knowledge of three kinds: background generalizations, pragmatic background generalizations (the instances of which may be invoked even after the generalizations are falsified), and background theories. All logics may be combined into a single system.

1. *Aim of this Paper*

This paper provides further evidence that inductive reasoning is a combination of forms of reasoning, each of which is characterized by a simple and intuitive adaptive logic. The adaptive logics can be combined in a straightforward way and the combined system characterizes inductive reasoning; its dynamic proof theory explicates actual inductive reasoning. It is essential for the dynamic proof theory that the combination of logics forms a single system in that all tasks are handled at the same time. In other words, the data are *not* closed by one consequence relation — for example the one settling which background theories are retained — before another consequence relation is applied — for example the one by which local generalizations are derived.

An adaptive logic of induction LI was presented in [6]. In its simplest guise, this logic leads from data to the right inductive generalizations and to

*Research for this paper was supported by subventions from Ghent University and from the Fund for Scientific Research – Flanders, and indirectly by the the Flemish Minister responsible for Science and Technology (contract BIL01/80). We are indebted to Dagmar Provijn and Liza Verhoeven for comments on a previous draft.

the consequences that are derivable by CL (Classical Logic) from the data together with the derived generalizations — these are called *local generalizations*. LI also handles *background generalizations*. The underlying idea is that a background generalization is retained iff it is compatible with the data. Background generalizations that are falsified by the data are discarded. When background generalizations are present, the 'local' generalizations are derived from the data together with the retained background generalizations.

We shall briefly present LI in Section 2. In the subsequent sections, we modify and extend LI. A central modification is that the adaptive logic that handles inductive generalization is clearly separated from adaptive logics that take care of different kinds of background knowledge. A central reason for a further modification is technical in nature. As [6] was written for the philosophy of science community, the presentation avoids some technical complications. As a result, LI does not strictly fit the usual adaptive logic format. We shall repair this in Section 3 for the logic of inductive generalization. This is not only important for systematic reasons, but also because it is fairly easy to handle the metatheory of adaptive logics that are formulated in the standard format — see [5] for some evidence for this claim.

Thus formulated, adaptive logics are also more transparent from a logical point of view. It will turn out that the system obtained by this reformulation is slightly richer than LI. Moreover, both the Reliability strategy and the Minimal Abnormality strategy turn out to be sensible from the new perspective.[1] The new systems for inductive generalization will be called IL^r and IL^m respectively and will be presented in Section 3 — the first uses the Reliability strategy, the second the Minimal Abnormality strategy. The relation with LI will be discussed.

Further reasons to modify LI relate to the way in which background generalizations are handled. First, not all background generalizations need to be assigned the same *priority*. If two of them are jointly incompatible with the data, the priorities may determine that the one is retained while the other is discarded. Next, background generalizations are often retained in a weak sense, at least until better replacements are found, even after being falsified — see for example [12]. To be more precise, instances of such background generalizations are supposed to apply, even if the generalization is falsified, unless and until the instances themselves are shown incompatible with the data. Let us call them *pragmatic* background *generalizations*.[2] Remark that

[1] LI proceeds in terms of Reliability and the variant obtained by the minimal Abnormality strategy does not seem very sensible. Adaptive strategies and their effects will be discussed in Section 3.

[2] Just to avoid confusion: a pragmatic background generalization is stronger than a non-pragmatic one in that, even if it is falsified, its instances are retained wherever possible. However, we do not suppose that all pragmatic background generalizations are falsified.

different pragmatic background generalizations may also be assigned a different priority. Thus, if both $(\forall x)(Px \supset Qx)$ and $(\forall x)(Rx \supset \sim Qx)$ are pragmatic background generalizations, and the data contain both Pa and Ra (and neither Qa nor $\sim Qa$), the priorities may determine that Qa applies whereas $\sim Qa$ does not. The logics that handle background generalizations as well as pragmatic background generalizations will be spelled out in Sections 5 and 6. They will be called IL^{gr} and IL^{gm} — the "g" in the superscripts refers to the fact that background generalizations are handled by these logics.

Incidentally, we shall not allow for *local* (that is: inductively derived) pragmatic background generalizations. Pragmatic background generalizations usually find their origin in either a theory, a worldview or a conceptual frame. If this theory, worldview or conceptual frame is known to be mistaken, in view of falsifications, one may hope that it will be partially retained in a non-falsified replacement (which usually will require some restructuring or even a conceptual shift). This is why a pragmatic background generalization $(\forall x)(Px \supset Qx)$ should not be interpreted as "most P are Q". The idea is rather that there is an as yet unknown P', which is closely related to P, and an as yet unknown Q', which is closely related to Q, and that *all* P' are Q'.

A realistic approach to background knowledge requires that one also handles background *theories*, rather than isolated generalizations. We discuss the matter in Section 7. This results in two logics which we shall call LI^{tr} and LI^{tm} — the "t" in the superscripts obviously refers to "theories".

In Section 8, all these logics are combined into a single logic. All background knowledge that is jointly compatible with the data has effects on predictions from the data. The joint compatibility is judged in terms of the priorities of the background knowledge. Moreover, the data and all retained background knowledge have priority over the local generalizations. Put differently, local generalizations are inductively derived from the data supplemented with the retained background knowledge.

Some open problems are listed in Section 7 of [6]. It seems useful to mention from the outset that inconsistent background knowledge will be disregarded throughout the whole paper because it requires an approach that is by no means difficult — inconsistency-adaptive logics are well mastered — but that departs drastically from the present logics — that all rely on CL and modal extensions of it.

2. *The Inductive Logic* LI

The first adaptive logic of induction was LI from [6]. In this section we briefly present it. Although LI does not fully agree with the standard adaptive logic format, it is characterized by a dynamic proof theory. Dynamic proof theories proceed in terms of conditional rules, unconditional rules and a Marking definition. The unconditional rules are those of CL, which is the so-called lower limit logic. The conditional rules are responsible for the ampliative character of the logic: they lead to generalizations. To distinguish such 'derived' generalizations from the background generalizations, they are called *local* generalizations.

The lines of a dynamic proof have five elements: (i) a line number, (ii) a derived formula A, (iii) the line numbers of the formulas from which A is derived, (iv) the rule by which A is derived, and (v) the set of formulas that should behave normally in order for A to be so derivable. All this will become clear below.

LI is formulated for the standard predicative language (with identity). It takes ordered sets as premises, such as $\Sigma = \langle\Gamma, \Gamma^*\rangle$, in which Γ is a set of empirical data and Γ^* is a set of background generalizations (conditional premises). The data are singular statements (of any form). The background generalizations are considered as the result of inductive inferences made in the past. For this reason, they are *generalizations* in the technical sense of the term: closed formulas of the form $\forall(A \supset B)$, in which $\forall$ abbreviates a sequence of universal quantifiers and $A \supset B$ is *purely functional* — no individual constant, sentential letter or quantifier occurs in it. We refer the reader to [6] for the justification of this decision.

We shall illustrate the LI-proof theory by means of a (very) simple example: let $\Gamma = \{(Pa \wedge {\sim}Qa) \wedge {\sim}Ra, {\sim}Pb \wedge (Qb \wedge Rb), Pc \wedge Rc, Qd \wedge {\sim}Pe\}$ and $\Gamma^* = \{(\forall x)(Px \supset {\sim}Qx), (\forall x)(Qx \supset Rx), (\forall x)(Px \supset Rx)\}$.

Γ contains the empirical data, which function just as usual premises. The background generalizations in Γ^* are also a sort of premises, but they are handled in a special way because they are *defeasible*. We shall disregard them for a while, and only introduce members of Γ by the premise rule:

PREM If $A \in \Gamma$, one may add a line comprising the following elements: (i) an appropriate line number, (ii) A, (iii) $-$, (iv) PREM, and (v) $\emptyset$.

We shall present the rules of inference in generic form,[3] and call them RU (unconditional rule) and RC (conditional rule) respectively. RU simply comes to applying CL, with one small difference.

[3] This simplifies the task of listing the rules, but there obviously is nothing essential to it.

RU If $A_1, \ldots, A_n \vdash_{\mathsf{CL}} B$ and each of A_1, …, A_n occur in the proof on lines i_1, …, i_n that have conditions Δ_1, …, Δ_n respectively, one may add a line comprising the following elements: (i) an appropriate line number, (ii) B, (iii) $i_1, \ldots, i_n$, (iv) RU, and (v) $\Delta_1 \cup \ldots \cup \Delta_n$.

Remark that RU simply adds the union of the conditions of its premises to its conclusion. Conditions are introduced by the rules RC and BK — the latter serves to introduce background generalizations.

RC Where A is a *generalization*, one may add a line comprising the following elements: (i) an appropriate line number, (ii) A, (iii) $-$, (iv) RC, and (v) $\{A\}$.

Unlike what is the case for the conditional rule of most other adaptive logics, the present RC enables one to write down just any generalization at any point in the proof. We shall soon see that this is harmless in view of the Marking definition. At present, let us look at our example.

1	$(Pa \wedge \sim Qa) \wedge \sim Ra$	$-$	PREM	$\emptyset$
2	$\sim Pb \wedge (Qb \wedge Rb)$	$-$	PREM	$\emptyset$
3	$Pc \wedge Rc$	$-$	PREM	$\emptyset$
4	$Qd \wedge \sim Pe$	$-$	PREM	$\emptyset$
5	$(\forall x)(Qx \supset Rx)$	$-$	RC	$\{(\forall x)(Qx \supset Rx)\}$
6	Rd	4, 5	RU	$\{(\forall x)(Qx \supset Rx)\}$
7	$(\forall x)(\sim Px \supset Qx)$	$-$	RC	$\{(\forall x)(\sim Px \supset Qx)\}$
8	Qe	4, 7	RU	$\{(\forall x)(\sim Px \supset Qx)\}$

So, we have introduced two generalizations, and we have derived conclusions from the data together with the generalizations. In order to obtain Rd, we relied on the supposition that $(\forall x)(Qx \supset Rx)$ *behaves normally*. For the time being, read this as: its negation has not been derived on the empty condition (there is no line that has $\sim(\forall x)(Qx \supset Rx)$ as its second element and $\emptyset$ as its fifth element). In general, LI presupposes that generalizations behave normally, unless and until they are shown to behave abnormally.

9^{L10}	$(\forall x)(Px \supset \sim Rx)$	$-$	RC	$\{(\forall x)(Px \supset \sim Rx)\}$
10	$\sim(\forall x)(Px \supset \sim Rx)$	3	RU	$\emptyset$

Here we see the Marking definition at work.[4] At line 9, we introduced another generalization. At line 10, however, we obtained the negation of

[4] We postpone the exact formulation of the Marking definition until we explained the complication introduced by background generalizations.

the generalization on the empty condition. As a result, all lines that have $(\forall x)(Px \supset {\sim}Rx)$ in their fifth element are marked — the superscripted L10 indicates that line 9 is marked because it relies on a local (whence the "L") generalization that was shown to behave abnormally at stage 10 of the proof.[5] Let us at once consider some more complex applications of the Marking definition.

11^{L15}	$(\forall x)(Px \supset {\sim}Qx)$	–	RC	$\{(\forall x)(Px \supset {\sim}Qx)\}$
12^{L15}	${\sim}Qc$	3, 11	RU	$\{(\forall x)(Px \supset {\sim}Qx)\}$
13^{L15}	$(\forall x)(Rx \supset Qx)$	–	RC	$\{(\forall x)(Rx \supset Qx)\}$
14^{L15}	Qc	3, 13	RU	$\{(\forall x)(Rx \supset Qx)\}$
15	${\sim}(\forall x)(Px \supset {\sim}Qx) \vee {\sim}(\forall x)(Rx \supset Qx)$	3	RU	$\emptyset$

What is going on here? On line 15, it was shown that either $(\forall x)(Px \supset {\sim}Qx)$ or $(\forall x)(Rx \supset Qx)$ behaves abnormally. As we have no clue as to which of both behaves abnormally, we consider both as unreliable. This enables us to make the Marking definition more precise. Remember that an abnormality is the negation of a generalization. A disjunction of abnormalities will be called a Dab-formula. It is handy to use the form $Dab(\Delta)$ to abbreviate the disjunction of the finite set of abnormalities Δ. $Dab(\Delta)$ is a *minimal* Dab-formula at a stage of a proof iff $Dab(\Delta)$ occurs as the second element of some line that has $\emptyset$ as its fifth element, and there is no $\Delta' \subset \Delta$ such that $Dab(\Delta')$ occurs as the second element of some line that has $\emptyset$ as its fifth element. A generalization A is *unreliable* at a stage of a proof iff ${\sim}A$ is a disjunct of a minimal Dab-formula at that stage. The Marking definition will make sure that a line is marked if an unreliable (local) generalization is a member of the fifth element of the line. This is why lines 11–14 are marked. Incidentally, the marks of lines 11 and 12 would be removed at a later stage, if the proof then would contain a line that has ${\sim}(\forall x)(Rx \supset Qx)$ as its second element and $\emptyset$ as its fifth element.[6]

It is instructive to see what happens to generalizations that contain non-instantiated predicates.

16^{L17}	$(\forall x)(Px \supset Sx)$	–	RC	$\{(\forall x)(Px \supset Sx)\}$
17	${\sim}(\forall x)(Px \supset Sx) \vee {\sim}(\forall x)(Px \supset {\sim}Sx)$	1	RU	$\emptyset$

[5] Stage i of a proof is obtained at the moment that line i is added to the proof.

[6] Clearly, ${\sim}(\forall x)(Rx \supset Qx)$ is not a CL-consequence of the premises. We shall see, however, that ${\sim}(\forall x)(Rx \supset Qx)$ is LI-derivable from the premises and retained background generalizations.

Let us now have a look at background generalizations. As announced, these are like premises, except for being defeasible. To be more precise, they may be falsified by the data. This is why they are introduced by a special rule, BK, which attaches a condition to them.

BK If $A \in \Gamma^*$, one may add a line comprising the following elements: (i) an appropriate line number, (ii) A, (iii) $-$, (iv) BK, and (v) $\{A\}$.

As expected, the condition should be interpreted as: provided the generalization does not behave abnormally. Remark that, in the continuation of the proof, one background generalization is at once marked.

18	$(\forall x)(Px \supset {\sim}Qx)$	$-$	BK	$\{(\forall x)(Px \supset {\sim}Qx)\}$
19	$(\forall x)(Qx \supset Rx)$	$-$	BK	$\{(\forall x)(Qx \supset Rx)\}$
20^{B21}	$(\forall x)(Px \supset Rx)$	$-$	BK	$\{(\forall x)(Px \supset Rx)\}$
21	${\sim}(\forall x)(Px \supset Rx)$	1	RU	$\emptyset$

The mark of line 20 reads "B21" because a background generalization is shown unreliable on line 21.

Now we come to an extremely important point which touches upon the marking of background generalizations and upon their effect on the marking of local generalizations. First, a background generalization is only unreliable at a stage of a proof if, at that stage, there is a minimal Dab-formula, $Dab(\Delta)$ such that $A \in \Delta$ *and* $\Delta \subseteq \Gamma^*$. This is why line 20 is marked in view of line 21, but line 18 is *not* marked in view of line 15. Indeed, $(\forall x)(Rx \supset Qx)$ is not a background generalization.

The upshot is that line 15 is interpreted as follows: as $(\forall x)(Px \supset {\sim}Qx) \in \Gamma^*$ and $(\forall x)(Rx \supset Qx) \notin \Gamma^*$, $(\forall x)(Px \supset {\sim}Qx)$ is considered as a reliable background generalization and (hence) $(\forall x)(Rx \supset Qx)$ is considered as an unreliable local generalization. As a result, the marks have to be modified at stage 18 of the proof. Lines 13 and 14 remain marked because $(\forall x)(Rx \supset Qx)$ is considered as an unreliable local generalization. However, lines 11 and 12 are unmarked (from stage 18 on) because $(\forall x)(Px \supset {\sim}Qx)$ is a reliable background generalization. Hence, ${\sim}Qc$ is considered as derived at stage 18 of the proof — it will remain considered derived at all later stages as line 18 cannot be marked on the present premises).[7]

We now present the precise Marking definitions, starting with the marks deriving from unreliable background generalizations.

[7] That 11 is unmarked should not cause puzzlement. As $(\forall x)(Px \supset {\sim}Qx)$ is a retained background generalization (from stage 18 on), one obviously can also introduce it as a local generalization.

Definition 1: *Where $Dab(\Delta_1), \ldots, Dab(\Delta_n)$ are the minimal Dab-formulas at stage s of a proof from $\Sigma = \langle \Gamma, \Gamma^* \rangle$, $U^*_s(\Gamma) = \bigcup\{\Delta_i \subseteq \Gamma^* \mid 1 \leq i \leq n\}$.*

Definition 2: *Where Δ is the fifth element of line i, line i is* B*-marked iff $\Delta \cap U^*_s(\Gamma) \neq \emptyset$.*

$U^*_s(\Gamma)$ comprises the background generalizations that are unreliable at stage s of the proof. What remains of the background knowledge at stage s will be denoted by $\Gamma^*_s = \Gamma^* - U^*_s(\Gamma)$ (the reliable background knowledge at stage s). Intuitively, the background knowledge is restricted to Γ^*_s at stage s of the proof.

Which local generalizations are unreliable is determined by the data and the reliable background knowledge. A Dab-formula $Dab(\Delta)$ will be called a *minimal local Dab-formula* iff no formula $Dab(\Delta')$ occurs in the proof such that $(\Delta' - \Gamma^*_s) \subset (\Delta - \Gamma^*_s)$.

Definition 3: *Where $Dab(\Delta_1), \ldots, Dab(\Delta_n)$ are the minimal local Dab-formulas at stage s of a proof from $\Sigma = \langle \Gamma, \Gamma^* \rangle$, $U^\circ_s(\Gamma) = \bigcup\{\Delta_i - \Gamma^*_s \mid 1 \leq i \leq n\}$.*

Definition 4: *Where Δ is the fifth element of a line i that is not* B*-marked, line i is* L*-marked iff $\Delta \cap U^\circ_s(\Gamma) \neq \emptyset$.*

$U^\circ_s(\Gamma)$ comprises the unreliable local generalizations at stage s. These generalizations may have been introduced by RC, they may be unreliable background generalizations, and they may be generalizations that have not even been introduced by either BK or RC in the proof.

To complete the dynamic proof theory of LI, we still have to present the definition of final derivability. The three following definitions are the same as for other adaptive logics. This is why we present them without referring to the specific logic (here LI) and shall not repeat them in the sequel.

Definition 5: *A formula A is derived at stage s of a proof from Σ iff A is the second element of a non-marked line at stage s.*

Definition 6: *A is* finally derived *from Γ 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 7: $\Gamma \vdash A$ *(A is* finally derivable *from Γ) iff A is finally derived on a line of a proof from Γ.*

3. *The Basic Inductive Logic*

Let us for a moment disregard background knowledge, which requires a prioritized adaptive logic. Thus restricted, we are discussing a flat adaptive logic ("flat" meaning "non-prioritized").

The standard way to characterize a flat adaptive logic is by specifying a lower limit logic, a set of abnormalities, which is characterized by some logical form,[8] and an adaptive strategy. The lower limit logic LLL and the set of abnormalities Ω should specify an upper limit logic ULL in that a characterization of ULL is obtained by adding to the characterization of LLL an axiom, or rule, or semantic clause, etc. (depending on the type of the characterization) that rules out abnormalities. Moreover, the lower limit logic and the upper limit logic should be connected by the *Derivability Adjustment Theorem*, viz.

$$\Gamma \vdash_{\mathsf{ULL}} A \text{ iff there is a } Dab(\Delta) \text{ such that } \Gamma \vdash_{\mathsf{LLL}} A \vee Dab(\Delta)$$

in which $Dab(\Delta)$ is a disjunction of abnormalities as before and "$\vee$" is the standard CL-disjunction.[9]

The effect of this construction is for example that the set of lower limit models that do not verify any abnormality forms a semantics that is characteristic for the upper limit logic. Also, if no abnormalities are derivable by the lower limit logic from some Γ, then the adaptive consequence set of Γ is identical to the upper limit consequence set of Γ.[10]

Seen in this respect, LI clearly has some odd properties. The lower limit logic is obviously CL. But if the generalizations (as defined in Section 2) form the set of abnormalities, then every lower limit model verifies some abnormalities — remember that $\sim\forall(A \supset B) \vee \sim\forall(A \supset \sim B) \vee \sim\forall(\sim A \supset B) \vee \sim\forall(\sim A \supset \sim B)$ is a CL-theorem. So, if LI is forced into the above format, its upper limit logic is the trivial logic (which has any formula as a theorem) and no set of premises is normal (not even the empty set).

[8] This logical form may be restricted. For example, the metavariables that occur in it may be restricted to primitive formulas or to formulas that are purely functional.

[9] If the lower limit logic requires that disjunction behaves abnormally, a standard CL-disjunction should be added — see [4] for examples.

[10] By some historical accident, the first adaptive logics — see [3], the oldest paper, and [2] — were such that the lower limit logic and the upper limit logic determine a unique set of abnormalities. In this case, the set of abnormalities is a function of the lower limit logic and the upper limit logic. This caused some confusion which was only cleared up when the Ghent logic group started studying ampliative logics — see for example [7]. Then it was realized that a lower limit logic may be combined with many different sets of abnormalities to obtain the same upper limit logic.

And yet, it is quite obvious, especially from a semantic point of view, that there is a nice upper limit logic that removes the above oddities. Remembering the connection between induction and the uniformity of models as described for example in [10], the upper limit logic should be characterized by the CL-models that are completely uniform. Let us call these the UCL-models (uniform classical logic models). These are the models in which the interpretation of any predicate of adicity n is either the empty set or the set of all n-tuples of members of the domain. Axiomatically, this logic may be characterized by extending CL with, for example, the axiom

$$\exists A \supset \forall A \tag{1}$$

in which $\exists A$ abbreviates the existential closure of A. Let us call this logic UCL.[11] Where derivability and semantic consequence are defined as usual, we leave it to the reader to prove that UCL is sound and strongly complete with respect to the UCL-semantics.

Of course, we have put cart before horse. We now have to find a set of abnormalities that, together with the lower limit logic CL, delivers UCL. This problem, however, is easily solved. We shall simply take as abnormalities the formulas of the form

$$\exists A \wedge \exists{\sim}A \tag{2}$$

in which A is purely functional — see [6] for the justification of this requirement. So, where $\mathcal{F}^{\circ}$ is the set of purely functional formulas, the set of abnormalities will be

$$\Omega = \{\exists A \wedge \exists{\sim}A \mid A \in \mathcal{F}^{\circ}\}$$

As before, a Dab-formula will be a disjunction of members of Ω. An expression of the form $Dab(\Delta)$ will always denote a formula with $\Delta \subseteq \Omega$ and the conditions (fifth elements) of lines in the dynamic proof will always be finite subsets of Ω.

We now prove a theorem that is essential for the dynamic proof theory of adaptive logics. We first need a simple Lemma, the proof of which is obvious.

Lemma 1: UCL *is equivalent to the system obtained by restricting axioms of the form of axioma schema (1) in such a way that* $A \in \mathcal{F}^{\circ}$.

[11] UCL is one of the many extensions of CL that fulfil all the traditional requirements. It is monotonic, structural, transitive, etc. The only trouble with UCL is that it is no good for being applied to the real world.

Theorem 1: $\Gamma \vdash_{\mathsf{UCL}} A$ *iff there is a* $\Delta \subseteq \Omega$ *such that* $\Gamma \vdash_{\mathsf{CL}} A \vee Dab(\Delta)$. *(Derivability Adjustment Theorem.)*

Proof. For the left–right direction, suppose that $\Gamma \vdash_{\mathsf{UCL}} A$ and consider an UCL-axiomatic proof of A from Γ in which all axioms of the form (1) are restricted as specified in Lemma 1. Let Θ be the set of all instances of the axiom (1) that are invoked in the proof. Obviously, Θ is finite and $\Gamma \cup \Theta \vdash_{\mathsf{CL}} A$. It follows that, where $\Theta^{\sim} = \{\exists A \wedge \exists{\sim}A \mid \exists A \supset \forall A \in \Theta\}$, $\Theta^{\sim} \subseteq \Omega$ and $\Gamma \vdash_{\mathsf{CL}} A \vee Dab(\Theta^{\sim})$.

For the right–left direction, suppose that $\Gamma \vdash_{\mathsf{CL}} A \vee Dab(\Delta)$, and hence that $\Gamma \vDash_{\mathsf{CL}} A \vee Dab(\Delta)$. It follows that all CL-models of Γ, and hence all UCL-models of Γ, verify $A \vee Dab(\Delta)$. But all UCL-models of Γ falsify $Dab(\Delta)$. Consequently, $\Gamma \vdash_{\mathsf{UCL}} A$. ■

From the lower limit logic CL and the set of abnormalities Ω, we now define the adaptive logics IL^r and IL^m, by the Reliability strategy and the Minimal Abnormality strategy respectively. Given that their premises comprise data only (and no background knowledge), Theorem 1 enables one to define these logics. As is usual, the difference between both logics appears only in the Marking definition. Here are the common rules of inference:

PREM If $A \in \Gamma$, one may add a line comprising the following elements: (i) an appropriate line number, (ii) A, (iii) $-$, (iv) PREM, and (v) $\emptyset$.

RU If $A_1, \ldots, A_n \vdash_{\mathsf{CL}} B$ and each of $A_1, \ldots, A_n$ occur in the proof on lines $i_1, \ldots, i_n$ that have conditions $\Delta_1, \ldots, \Delta_n$ respectively, one may add a line comprising the following elements: (i) an appropriate line number, (ii) B, (iii) $i_1, \ldots, i_n$, (iv) RU, and (v) $\Delta_1 \cup \ldots \cup \Delta_n$.

RC If $A_1, \ldots, A_n \vdash_{\mathsf{CL}} B \vee Dab(\Theta)$ and each of $A_1, \ldots, A_n$ occur in the proof on lines $i_1, \ldots, i_n$ that have conditions $\Delta_1, \ldots, \Delta_n$ respectively, one may add a line comprising the following elements: (i) an appropriate line number, (ii) B, (iii) $i_1, \ldots, i_n$, (iv) RC, and (v) $\Delta_1 \cup \ldots \cup \Delta_n \cup \Theta$.

At any stage of the proof, zero or more Dab-formulas will be derived. Some of them are minimal (at that stage). Let $U_s(\Gamma)$ be the union of all Δ for which $Dab(\Delta)$ is a minimal Dab-formula at stage s. Let $\Phi^{\circ}_s(\Gamma)$ be the set of all sets that contain one disjunct out of each minimal Dab-formula at stage s, and let $\Phi_s(\Gamma)$ contain those members of $\Phi^{\circ}_s(\Gamma)$ that are not proper supersets of other members of $\Phi^{\circ}_s(\Gamma)$.

Definition 8: *Marking for* IL^r*: Line i is marked at stage s iff, where Δ is its fifth element,* $\Delta \cap U_s(\Gamma) \neq \emptyset$.

Definition 9: *Marking for* IL^m*: Line* i *is marked at stage* s *iff, where* A *is the second element and* Δ *the fifth element of line* i*, (i) there is no* $\varphi \in \Phi_s(\Gamma)$ *such that* $\varphi \cap \Delta = \emptyset$*, or (ii) for some* $\varphi \in \Phi_s(\Gamma)$*, there is no line* k *that has* A *as its second element and has as its fifth element some* Θ *such that* $\varphi \cap \Theta = \emptyset$*.*

In view of these Marking definitions, IL^r-derivability and IL^m-derivability are defined as usual, viz. by Definitions 5, 6 and 7.

Let us at once present an example to illustrate the logics. In the example, the Reliability strategy and the Minimal Abnormality strategy lead to the same results. The distinction between both has been illustrated in many papers.

To keep the proofs within margins, we shall abbreviate abnormalities by the formulas they derive from. In the fifth element of lines, A will abbreviate $\exists A \wedge \exists{\sim}A$. Thus a fifth element $\{Px \supset Qx\}$ abbreviates

$$\{(\exists x)(Px \supset Qx) \wedge (\exists x){\sim}(Px \supset Qx)\}.$$[12]

Similarly, $Dab\{Px \supset Qx, Rx \supset Qx\}$ abbreviates

$$((\exists x)(Px \supset Qx) \wedge ((\exists x){\sim}(Px \supset Qx)) \\ \vee((\exists x)(Rx \supset Qx) \wedge ((\exists x){\sim}(Rx \supset Qx))$$

Here is the example:

1	$(Pa \wedge {\sim}Qa) \wedge {\sim}Ra$	–	PREM	$\emptyset$
2	${\sim}Pb \wedge (Qb \wedge Rb)$	–	PREM	$\emptyset$
3	$Pc \wedge Rc$	–	PREM	$\emptyset$
4	$Qd \wedge {\sim}Pe$	–	PREM	$\emptyset$
5	$(\forall x)(Qx \supset Rx)$	2	RC	$\{Qx \supset Rx\}$
6	Rd	4, 5	RU	$\{Qx \supset Rx\}$
7	$(\forall x)({\sim}Px \supset Qx)$	2	RC	$\{{\sim}Px \supset Qx\}$
8	Qe	4, 7	RU	$\{{\sim}Px \supset Qx\}$
9^{L10}	$(\forall x)(Px \supset {\sim}Rx)$	1	RC	$\{Px \supset {\sim}Rx\}$
10	$Dab(Px \supset {\sim}Rx)$	1, 3	RU	$\emptyset$
11^{L17}	$(\forall x)(Px \supset {\sim}Qx)$	1	RC	$\{Px \supset {\sim}Qx\}$
12^{L17}	${\sim}Qc$	3, 11	RU	$\{Px \supset {\sim}Qx\}$
13^{L17}	$(\forall x)(Rx \supset Qx)$	2	RC	$\{Rx \supset Qx\}$
14^{L17}	Qc	3, 13	RU	$\{Rx \supset Qx\}$

[12] This is obviously CL-equivalent to $\{(\exists x)({\sim}Px \vee Qx) \wedge (\exists x)(Px \wedge {\sim}Qx)\}$.

15	$(\exists x){\sim}(Px \supset {\sim}Qx) \vee (\exists x){\sim}(Rx \supset Qx)$	3	RU	$\emptyset$
16	$(\exists x)(Px \supset {\sim}Qx) \wedge (\exists x)(Rx \supset Qx)$	1, 2	RU	$\emptyset$
17	$Dab\{Px \supset {\sim}Qx, Rx \supset Qx\}$	15, 16	RU	$\emptyset$
18^{L22}	$(\forall x)(Px \supset Sx)$	4	RC	$\{Px \supset Sx\}$
19^{L22}	Sa	1, 18	RU	$\{Px \supset Sx\}$
20	$(\exists x){\sim}(Px \supset Sx) \vee (\exists x){\sim}(Px \supset {\sim}Sx)$	3	RU	$\emptyset$
21	$(\exists x)(Px \supset Sx) \wedge (\exists x)(Px \supset {\sim}Sx)$	4	RU	$\emptyset$
22	$Dab\{Px \supset Sx, Px \supset {\sim}Sx\}$	20, 21	RU	$\emptyset$

It is instructive to have a look at the marked generalizations. 9 is not finally derivable because it is falsified by 3 — it is incompatible with the data. However, several generalizations that are themselves compatible with the data may be jointly incompatible with them. This is the case for 11, 13 and 18. Among these, 18 is a somewhat special case. As S does not occur in the data, $(\forall x)(Px \supset Sx)$ and $(\forall x)(Px \supset {\sim}Sx)$ are compatible with them, but are jointly incompatible with them in view of the presence of Pa. In general, generalizations that contain a 'new' predicate will be marked, except when they are **CL**-consequences of finally derivable generalizations.[13] Obviously, line 13 is not needed to derive 17 (and to mark line 11). However, line 13 enables us to illustrate a useful point: that two generalizations are jointly incompatible with the data may be found out by deriving inconsistent predictions from them — in this case 12 and 14.

The semantics of IL^r and IL^m follows the usual lines of adaptive logics and soundness and completeness are provable along the standard road. For each **CL**-model M, we define $Ab(M) = \{A \mid M \models A;\ A \in \Omega\}$. Where $\Delta_1, \Delta_2, \ldots$ are the subsets of Ω for which $Dab(\Delta_i)$ is a minimal Dab-consequence of Γ,[14] $U(\Gamma) = \Delta_1 \cup \Delta_2 \cup \ldots$.

A **CL**-model M of Γ is an IL^r-model of Γ iff $Ab(M) \subseteq U(\Gamma)$ and $\Gamma \models_{\mathsf{IL}^r} A$ iff A is verified by all IL^r-models of Γ. A **CL**-model M of Γ is an IL^m-model of Γ iff there is no **CL**-model M' of Γ such that $Ab(M') \subset Ab(M)$ and $\Gamma \models_{\mathsf{IL}^m} A$ iff A is verified by all IL^m-models of Γ.

This seems the right place to briefly discuss the relation between IL^r and **LI**. In a sense, the former explicates the latter. Consider a case where, in a **LI**-proof from Γ, the hypothesis $(\forall x)(Px \supset Qx)$ is introduced at line i. If $(\exists x)(Px \supset Qx)$ is not derivable from Γ, ${\sim}(\forall x)(Px \supset Qx)$ is a disjunct of a minimal Dab-consequence (of **LI**) of Γ and hence line i is marked in the **LI**-proof. Suppose then that $(\exists x)(Px \supset Qx)$ is derivable from Γ. In

[13] For example, as $(\forall x)({\sim}Px \supset Qx)$ happens to be finally derivable, so is $(\forall x)(({\sim}Px \wedge Sx) \supset Qx)$.

[14] This simply means that all **CL**-models of Γ verify $Dab(\Delta_i)$ and that there is no $\Theta \subset \Delta_i$ such that all **CL**-models of Γ verify $Dab(\Theta)$.

this case, $(\forall x)(Px \supset Qx)$ can be derived on the condition $\{(\exists x)(Px \supset Qx) \wedge (\exists x){\sim}(Px \supset Qx)\}$ in an IL^r-proof from Γ. This line will be marked just in case $(\exists x)(Px \supset Qx) \wedge (\exists x){\sim}(Px \supset Qx)$ is a disjunct of a minimal Dab-consequence of Γ. As $(\exists x)(Px \supset Qx)$ is derivable from Γ, it seems that this will be the case iff, in the context of LI, ${\sim}(\forall x)(Px \supset Qx) \in U_s^{\circ}(\Gamma)$. If this is correct $(\forall x)(Px \supset Qx)$ is finally LI-derivable from Γ just in case it is IL^r-derivable from Γ.

We did not try to prove the above hypothesis, because the matter is not very important. More important is that IL^r is clearly distinct from LI in that, from some sets of premises, it enables one to derive *disjunctions* of formulas of the form $\forall A$ that are not LI-derivable from them. Here is a simple example:

1 Pa	–	PREM	$\emptyset$
2 $(Pa \supset Qa) \vee (Pa \supset {\sim}Qa)$	1	RU	$\emptyset$
3 $(\forall x)(Px \supset Qx) \vee (Pa \supset {\sim}Qa)$	2	RC	$\{Px \supset Qx\}$
4 $(\forall x)(Px \supset Qx) \vee (\forall x)(Px \supset {\sim}Qx)$	3	RC	$\{Px \supset Qx, Px \supset {\sim}Qx\}$

The premise has UCL-models and hence no Dab-formula is derivable from it. It follows that line 4 will not be marked in any extension of the proof. However, $(\forall x)(Px \supset Qx) \vee (\forall x)(Px \supset {\sim}Qx)$ is not a LI-final consequence of Pa because $(\exists x)(Px \wedge Qx) \vee (\exists x)(Px \wedge {\sim}Qx)$ is CL-derivable from Pa.[15]

It is easily seen that $(\forall x)Qx \vee (\forall x){\sim}Qx$ is also finally IL^r-derivable from Pa (as well as from, for example, the empty premise set) whereas it is not finally LI-derivable from Pa (or from the empty premise set). It is simple enough to see reasons for this — we consider the Reliability strategy only. In LI, $(\forall x)Qx \vee (\forall x){\sim}Qx$ can only be derived on the condition $\{(\forall x)Qx\}$ or on the condition $\{(\forall x){\sim}Qx\}$. However, if neither condition is derivable from the data, then the CL-theorem ${\sim}(\forall x)Qx \vee {\sim}(\forall x){\sim}Qx$ functions as a disjunction of abnormalities. Nothing similar prevents the derivation of $(\forall x)Qx \vee (\forall x){\sim}Qx$ by IL^r. For example,

$$((\exists x)Qx \wedge (\exists x){\sim}Qx) \vee ((\exists x){\sim}Qx \wedge (\exists x)Qx),$$

which is CL-equivalent to $(\exists x)Qx \wedge (\exists x){\sim}Qx$, is not a CL-theorem.

[15] The only way to obtain $(\forall x)(Px \supset Qx) \vee (\forall x)(Px \supset {\sim}Qx)$ in LI, is by deriving it from one of its disjuncts (or from formulas that are equivalent or stronger than these disjuncts). So, in LI, the disjunction is only derivable on lines that have $\{(\forall x)(Px \wedge {\sim}Qx)\}$ or $\{(\forall x)(Px \wedge Qx)\}$ as their fifth element. But both kinds of lines are marked. A similar reasoning applies if the disjunction is derived from formulas that are equivalent or stronger than these disjuncts.

This is exactly as it should be. Even according to LI, $(\exists x)Qx$ is sufficient to inductively derive $(\forall x)Qx$ and $(\exists x){\sim}Qx$ is sufficient to inductively derive $(\forall x){\sim}Qx$, unless $(\exists x)Qx \wedge (\exists x){\sim}Qx$ is CL-derivable from the premises. The same inductive presupposition enables one to inductively derive $(\forall x)Qx \vee (\forall x){\sim}Qx$ if neither $(\exists x)Qx$ nor $(\exists x){\sim}Qx$ is a CL-consequence of the premises. In words, if one did not observe any object to have either property Q or property ${\sim}Q$, one nevertheless may presuppose that all objects are identical with respect to Q-hood.

Apart from agreeing with the usual adaptive format, IL^r and IL^m seem to improve upon LI in two respects. First, they enable one to derive generalizations *from* instances rather than enabling one to posit generalizations for no reason at all. Next, for some sets of premises, they enable one to derive certain disjunctions of generalizations that are not derivable according to LI, but should be derivable from an intuitive viewpoint.

4. *Two Modal Logics*

LI is a prioritized adaptive logic in that its set of premises is $\Sigma = \langle \Gamma, \Gamma^* \rangle$ and Γ, the set of data, is given a higher priority than Γ^*, the set of background generalizations. However, LI handles background generalizations in too poor a way because it puts them all on a par. If two background generalizations are compatible with the data, but jointly incompatible, LI rejects both. In real life, background generalizations may receive different degrees of priority. In the following three sections, we shall introduce ways to handle three kinds of background knowledge in a prioritized way. We shall introduce separate adaptive logics, and discuss their combination in Section 8.

The premises of a prioritized logic are often represented by a tuple, say:

$$\Sigma = \langle \Gamma_0, \Gamma_1, \ldots, \Gamma_n \rangle \,, \tag{3}$$

in which Γ_0 represents the data, which receive the maximal priority, and the other sets contain background generalizations that receive a lower priority, viz. a priority that is lower as i is larger.

In Section 7, we shall need a somewhat more complex case, in which the premises are represented by[16]

$$\Sigma = \{\Gamma_0^0, \Gamma_1^{i_1}, \ldots, \Gamma_n^{i_n}\} \quad (i_1, \ldots, i_n \in \mathbb{N} - \{0\})\,. \tag{4}$$

[16] No background knowledge should be considered as absolutely certain, whence the indices are members of $\mathbb{N} - \{0\}$.

Here the priority of a Γ_i is determined by its superscript — we shall take the priority to be lower as the superscript is higher. The set of data Γ_0^0 obviously receives maximal priority. As we need to combine the different adaptive logics, and as the priorities of the different kinds of background knowledge have to be commensurable, we shall throughout concentrate on (4). In other words, premises given in the form (3) will be first rephrased in the form (4) and then ‘translated’.

Some like to characterize (flat as well as prioritized) adaptive logics as formula preferential systems — see [13] and [1]. For reasons which we cannot discuss here, we prefer to stick to the original intuition behind adaptive logics, and to characterize them by a set of abnormalities Ω (or by a tuple of sets of abnormalities) which is (or are) defined by a *logical form*. Moreover, we want the official *formulation* of our premises to express the degree of priority. The easiest way to realize this is by ‘translating’ the Σ from (4) to the standard modal language, for example as follows: $\Sigma^{\Diamond} = \{\Diamond^i A \mid A \in \Gamma_j^i;\ 0 \leq j \leq n\}$ in which $\Diamond^i A$ denotes A preceded by i occurrences of $\Diamond$.

This approach requires that some modal logic is chosen. We shall briefly (semantically) characterize two modal logics that will prove useful in subsequent sections. It is not difficult to see that we need modal logics in which the accessibility relation is not transitive or not reflexive (for otherwise $\Diamond^i A$ collapses into $\Diamond A$ for all i). We shall employ two logics in which the accessibility relation is not transitive. One of them is a predicative version of the modal logic T of Feys (which is von Wright’s M). We take this system from [9]. The other is a rather non-standard modal logic. As we shall need to combine these logics in Section 8, we shall employ two different symbols for “possibly”. Incidentally, in the present context $\Diamond^i A$ and $\diamond^i A$ are better read as “A is accepted unless and until falsified” — or perhaps “A is likely to be true” where the “likely” becomes weaker as i is larger.

Let $\mathcal{L}$ be the standard predicative language with $\mathcal{S}$, $\mathcal{P}^r$, $\mathcal{C}$, $\mathcal{F}$, and $\mathcal{W}$ the sets of sentential letters, predicative letters of rank r, constants, formulas, and wffs (closed formulas). Let $\mathcal{L}^M$ be the standard predicative modal language and $\mathcal{W}^M$ the set of its closed formulas. Negation, disjunction, the existential quantifier and possibility (and identity) will be considered as primitive; other logical symbols are defined in the usual way.

To simplify the semantic metalanguage, we introduce a set of pseudo-constants $\mathcal{O}$, *requiring that any element of the domain D is named by at least one member of $\mathcal{C} \cup \mathcal{O}$*.[17] Let $\mathcal{W}^{M+}$ denote the set of wffs of the pseudo-language $\mathcal{L}^{M+}$ (defined by letting $\mathcal{C} \cup \mathcal{O}$ play the role played by $\mathcal{C}$ in $\mathcal{L}^M$). The function of $\mathcal{O}$ is to simplify the clauses for the quantifiers.

[17] $\mathcal{O}$ should have at least the cardinality of the largest model considered — if there is no such model, one selects a suitable $\mathcal{O}$ for each model.

A T-model M is a quintuple $\langle W, w_0, R, D, v\rangle$ in which W is a set of worlds, $w_0 \in W$ the real world, R a reflexive binary relation on W, D a non-empty set and v an assignment function. The assignment function v is as follows:

C1.1 $v : \mathcal{S} \times W \mapsto \{0, 1\}$
C1.2 $v : \mathcal{C} \cup \mathcal{O} \times W \mapsto D$
C1.3 $v : \mathcal{P}^r \times W \mapsto \wp(D^r)$ (the power set of the r-th Cartesian product of D)

The valuation function, $v_M : \mathcal{W}^{M+} \times W \mapsto \{0, 1\}$, determined by the model M is defined by:

C2.1 where $A \in \mathcal{S}$, $v_M(A, w) = v(A, w)$
C2.2 $v_M(\pi^r\alpha_1 \ldots \alpha_r, w) = 1$ iff $\langle v(\alpha_1, w), \ldots, v(\alpha_r, w)\rangle \in v(\pi^r, w)$
C2.3 $v_M(\alpha = \beta, w) = 1$ iff $v(\alpha, w) = v(\beta, w)$
C2.4 $v_M(\sim A, w) = 1$ iff $v_M(A, w) = 0$
C2.5 $v_M(A \vee B, w) = 1$ iff $v_M(A, w) = 1$ or $v_M(B, w) = 1$
C2.6 $v_M((\exists\alpha)A(\alpha), w) = 1$ iff $v_M(A(\beta), w) = 1$ for at least one $\beta \in \mathcal{C} \cup \mathcal{O}$
C2.7 $v_M(\Diamond A, w) = 1$ iff $v_M(A, w') = 1$ for at least one w' such that Rww'.

A model M verifies $A \in \mathcal{W}^M$ iff $v_M(A, w_0) = 1$. $\Gamma \vDash A$ (A is a semantic consequence of Γ) iff all models of Γ verify A. $\vDash A$ (A is valid) iff A is verified by all models.

A bit of explanation seems useful. One may define a function d that assigns to each $w \in W$ its domain $d(w) = \{v(\alpha, w) \mid \alpha \in \mathcal{C} \cup \mathcal{O}\}$. If an element of an r-tuple of $v(\pi^r, w)$ does not belong to $d(w)$, then the r-tuple does not have any effect on the valuation.[18] Remark also that, if Rww', the question whether $v(\alpha, w)$ is or is not a member of $d(w')$ is immaterial for any $v(A, w)$. For example, the value of $v_M(\Diamond Pa, w)$ is determined by the values of $v(a, w')$ and $v(P, w')$ for those w' for which Rww'. Obviously, the semantics may be rephrased as a counterpart semantics: $\mathsf{a} \in d(w)$ is a counterpart of $\mathsf{b} \in d(w')$ just in case there is an $\alpha \in \mathcal{C} \cup \mathcal{O}$ such that $v(\alpha, w) = \mathsf{a}$ and $v(\alpha, w') = \mathsf{b}$. An $\alpha \in \mathcal{C} \cup \mathcal{O}$ may be seen as picking a specific counterpart 'path' on W.[19] T is axiomatically characterized by axioms for CL with Replacement of Identicals invalid within the scope of a modality, plus the axioms of the propositional logic T, plus the Barcan Formula.

[18] This means that C1.3 may be replaced by "$v : \mathcal{P}^r \times W \mapsto \wp((d(w))^r)$".

[19] The technique to handle quantifiers in terms of $\mathcal{C} \cup \mathcal{O}$ is itself not related to modal logic — see, for example, the semantics for P in [8].

In some cases, a very non-standard modal logic is more suitable. We shall call it IM (because it isolates modal formulas). We first present its semantics and then offer some explanation.

An IM-model M is a quintuple $\langle W, w_0, R, D, v\rangle$ in which W is a set of worlds, $w_0 \in W$ the real world, R a reflexive binary relation on W, D a non-empty set and v an assignment function. The assignment function v is such that:

C1.1 $v : \mathcal{S} \times W \mapsto \{0, 1\}$
C1.2 $v : \mathcal{C} \cup \mathcal{O} \times W \mapsto D$
C1.3 $v : \mathcal{P}^r \times W \mapsto \wp(D^r)$
C1.4 $v : \mathcal{W} \times W \mapsto \{0, 1\}$

The valuation function, $v_M : \mathcal{W}^{M+} \times W \mapsto \{0, 1\}$, determined by the model M is defined by:

C2.1 where $A \in \mathcal{S}$, $v_M(A, w_0) = v(A, w_0)$
C2.2 $v_M(\pi^r\alpha_1 \ldots \alpha_r, w_0) = 1$ iff $\langle v(\alpha_1, w_0), \ldots, v(\alpha_r, w_0)\rangle \in v(\pi^r, w_0)$
C2.3 $v_M(\alpha = \beta, w_0) = 1$ iff $v(\alpha, w_0) = v(\beta, w_0)$
C2.4 $v_M({\sim}A, w_0) = 1$ iff $v_M(A, w_0) = 0$
C2.5 $v_M(A \vee B, w_0) = 1$ iff $v_M(A, w_0) = 1$ or $v_M(B, w_0) = 1$
C2.6 $v_M((\exists\alpha)A(\alpha), w_0) = 1$ iff $v_M(A(\beta), w_0) = 1$ for at least one $\beta \in \mathcal{C} \cup \mathcal{O}$
C2.7 if $w \in W - \{w_0\}$ and $A \in \mathcal{W}$, then $v_M(A, w) = v(A, w)$
C2.8 $v_M(\diamond A, w) = 1$ iff $v_M(A, w') = 1$ for at least one w' such that Rww'.

A model M verifies $A \in \mathcal{W}^M$ iff $v_M(A, w_0) = 1$. Semantic consequence and validity are defined as before. IM is axiomatically characterized by axioms for CL with Replacement of Identicals invalid within the scope of a modality, plus the axiom $A \supset \diamond A$ plus the Barcan Formula.

Remark that C1.1–3 have only effect in w_0 whereas C1.4 has no effect in w_0. In other words, non-modal formulas (members of $\mathcal{W}$) are assigned arbitrary values in worlds that are different from w_0. Formulas of the form $\diamond A$ are governed by C2.8 in all worlds. The main effect of the construction is that 'modal formulas have no modal consequences'. Thus $({\sim}\diamond{\sim}(A \supset B) \wedge \diamond A) \supset \diamond B$ and $\diamond(A \wedge B) \supset \diamond A$ are not theorems of IM. It also is not excluded that IM verifies A even if ${\sim}A$ is a CL-theorem — we shall neither need nor use this feature below.[20]

In view of subsequent sections, it is useful to see the following. Where Σ is as in (4) and $\Gamma_0^0 \cup \Gamma_1^{i_1} \cup \ldots \cup \Gamma_n^{i_n}$ is consistent, $\Sigma^\diamond = \{\diamond^i A \mid A \in \Gamma_j^i;\ 0 \leq$

[20] Where $A \supset B =_{df} {\sim}A \vee B$, a model that verifies $\diamond({\sim}A \vee B)$ always verifies $\diamond(A \supset B)$. This feature will cause no trouble for our adaptive logics.

$j \leq n\}$ has IM-models in which $W = \{w_0\}$; similarly (replacing $\diamond$ by $\Diamond$) for T-models.

5. *Prioritized Background Generalizations*

In line with Section 3, we shall consider background generalizations to be formulas of the form $\forall A$ with $A \in \mathcal{F}^\circ$. As announced in Section 4, we shall take every background generalization to have a certain priority. Σ will be as in (4), with Γ_0^0 the set of data and the other Γ_j^i sets of background generalizations. The members of Γ_j^i $(1 \leq j \leq n)$ receive a priority which is lower according as i is higher.

The idea underlying our approach is that a background generalization in $\Gamma_1^{i_1}$ will be retained unless it is, separately or jointly, incompatible with Γ_0^0; a background generalization in $\Gamma_2^{i_2}$ will be retained unless it is, separately or jointly, incompatible with the union of Γ_0^0 and the retained background generalizations from $\Gamma_1^{i_1}$; etc.

As announced, the official formulation of the premises will be a set of formulas of the form $\diamond^i \forall B$ in which $B \in \mathcal{F}^\circ$ is a non-modal formula and $i \in \mathbb{N} - \{0\}$. Thus, the usual formulation $\Sigma = \langle \Gamma_0^0, \Gamma_1^{i_1}, \ldots, \Gamma_n^{i_n} \rangle$ will be given a modal 'translation', viz. $\Sigma^\diamond = \{\diamond^i A \mid A \in \Gamma_j^i;\ 0 \leq j \leq n\}$ in which $\diamond^i A$ denotes A preceded by i occurrences of $\diamond$.

The adaptive logics will proceed in terms of the modal logic IM. The abnormalities will be of the form $\diamond^i A \wedge \sim A$. As we are after a prioritized adaptive logic, a set of abnormalities should be associated with each priority level. Thus, for any $i \in \mathbb{N} - \{0\}$, $\Omega^i = \{\diamond^i A \wedge \sim A \mid A \in \mathcal{W}\}$, and $\Omega = \Omega^1 \cup \Omega^2 \cup \ldots$. As before, a Dab-formula will be a disjunction of members of Ω. An expression of the form $Dab(\Delta)$ will always denote a formula for which $\Delta \subseteq \Omega$ and the conditions (fifth elements) of lines in the dynamic proof will always be finite subsets of Ω. In practice, the fifth elements of lines of a sensible proof from $\Sigma^\diamond$ will be formulas of the form $\diamond^i A \wedge \sim A$ in which A is of the form $\forall B$ and $B \in \mathcal{F}^\circ$.[21]

The upper limit logic defined by IM is the modal logic Triv, which is obtained, for example, by adding to IM the axiom schema $\diamond A \supset A$. Where Σ

[21] Proceeding in a non-sensible way, one might introduce a premise $\diamond\diamond(\forall x)(Px \supset Qx)$ (on the condition $\emptyset$), derive from this $\diamond\diamond(\forall x)(Px \supset Qx) \vee \diamond(Pa)$ (still on the condition $\emptyset$), and from this derive $\diamond\diamond(\forall x)(Px \supset Qx) \vee (Pa)$ on the condition $\{\diamond(Pa) \wedge \sim Pa\}$. This leads nowhere but is not forbidden.

is as in (4), we obviously have

$$\Sigma^{\diamond} \vdash_{\mathsf{Triv}} A \quad \text{iff} \quad \Gamma_0^0 \cup \Gamma_1^{i_1} \cup \ldots \cup \Gamma_n^{i_n} \vdash_{\mathsf{CL}} A\,.$$

If $\forall A \in \Gamma_j^i$, then $\diamond^i \forall A \in \Sigma^{\diamond}$. In agreement with the Derivability Adjustment Theorem, we then have:

$$\diamond^i \forall A \vdash_{\mathsf{IM}} \forall A \vee (\diamond^i \forall A \wedge \sim\!\forall A)\,.$$

So, if the premises have a (systematic) maximally normal interpretation[22] according to which $\diamond^i \forall A \wedge \sim\!\forall A$ is false, $\forall A$ will be adaptively derivable from the premises if $\diamond^i \forall A$ is IM-derivable from them.

It is easily proved that

$$\Sigma^{\diamond} \vdash_{\mathsf{Triv}} A \text{ iff there is a } Dab(\Delta) \text{ such that } \Sigma^{\diamond} \vdash_{\mathsf{IM}} A \vee Dab(\Delta)$$

which is the Derivability Adjustment Theorem. This provides the motor for the dynamic proof theory, which we now present.

The adaptive logics will be called IL^{gr} and IL^{gm} — the "g" refers to background generalization, the "r" and "m" refer (as before) to the Reliability strategy and the Minimal Abnormality strategy respectively. We first list their common rules of inference.

PREM If $A \in \Sigma^{\diamond}$, one may add a line comprising the following elements: (i) an appropriate line number, (ii) A, (iii) $-$, (iv) PREM, and (v) $\emptyset$.

RU If $A_1, \ldots, A_n \vdash_{\mathsf{IM}} B$ and each of $A_1, \ldots, A_n$ occur in the proof on lines $i_1, \ldots, i_n$ that have conditions $\Delta_1, \ldots, \Delta_n$ respectively, one may add a line comprising the following elements: (i) an appropriate line number, (ii) B, (iii) $i_1, \ldots, i_n$, (iv) RU, and (v) $\Delta_1 \cup \ldots \cup \Delta_n$.

RC If $A_1, \ldots, A_n \vdash_{\mathsf{IM}} B \vee Dab(\Theta)$ and each of $A_1, \ldots, A_n$ occur in the proof on lines $i_1, \ldots, i_n$ that have conditions $\Delta_1, \ldots, \Delta_n$ respectively, one may add a line comprising the following elements: (i) an appropriate line number, (ii) B, (iii) $i_1, \ldots, i_n$, (iv) RC, and (v) $\Delta_1 \cup \ldots \cup \Delta_n \cup \Theta$.

The marking definitions require a bit of care, but are very transparent and intuitive.

Remember that $\Omega = \Omega^1 \cup \Omega^2 \cup \ldots$. Consider some $Dab(\Delta)$. If $\Delta \subseteq \Omega^i$, we shall write the formula as $Dab^i(\Delta)$. This notation enables us to write any Dab-formula as a disjunction of Dab^i-formulas. Thus, if $\Delta = \Delta_1 \cup$

[22] The meaning of this expression will depend on the chosen strategy.

$\Delta_2 \cup \Delta_3$ and $\Delta_1 \subseteq \Omega^4$, $\Delta_2 \subseteq \Omega^7$ and $\Delta_3 \subseteq \Omega^8$, $Dab(\Delta)$ may be written as $Dab^4(\Delta_1) \vee Dab^7(\Delta_2) \vee Dab^8(\Delta_3)$.

Consider a proof from $\Sigma^\diamond$. Suppose that $Dab^4(\Delta_1)\vee Dab^7(\Delta_2)\vee Dab^8(\Delta_3)$ has been derived on the empty condition in the proof. In other words, there is a line of the form

i	$Dab^4(\Delta_1) \vee Dab^7(\Delta_2) \vee Dab^8(\Delta_3)$	...	...	$\emptyset$

The proof obviously can be extended with the following line:

j	$Dab^8(\Delta_3)$	i	RC	$\Delta_1 \cup \Delta_2$

Remark that the sequences of symbols $\diamond$ in Θ count *less than* 8 members. If $\Delta_1 \cup \Delta_2$ is such that line j is unmarked, in other words, if the premises are such that the members of $\Delta_1 \cup \Delta_2$ may be taken to be false, then one should conclude that at least one member of Δ_3 is true on the premises.

Let us say that line j is clean. In general, a *clean line* is one that has as its second element a formula $Dab^i(\Delta)$ and as its fifth element a set $\Theta \subseteq \Omega^1 \cup \ldots \cup \Omega^{i-1}$. Put differently, the second element is a disjunction of abnormalities of the same level, and the fifth element contains only abnormalities of a higher level (that is, with shorter sequences of diamonds).[23]

We shall say that $Dab^i(\Delta)$ is a *minimal* Dab^i-formula at stage s of the proof if it has been derived at a clean and unmarked line whereas, at stage s, $Dab^i(\Delta')$ is not derived at a clean and unmarked line for any $\Delta' \subset \Delta$.

At any stage of the proof, zero or more Dab^i-formulas will be derived. Some of them are minimal (at that stage). Let, for each $i \in \mathbb{N} - \{0\}$, $U^i_s(\Sigma^\diamond)$ be the union of all Δ for which $Dab^i(\Delta)$ is a minimal Dab^i-formula at stage s. Let $\Phi^{i\circ}_s(\Gamma)$ be the set of all sets that contain one disjunct out of each minimal Dab^i-formula at stage s, and let $\Phi^i_s(\Gamma)$ contain those members of $\Phi^{i\circ}_s(\Gamma)$ that are not proper supersets of other members of $\Phi^{i\circ}_s(\Gamma)$.

Definition 10: *Marking for* IL^{gr}*: Where* Δ *is the fifth element of line* j*, line* j *is marked at stage* s *iff,* $\Delta \cap U^i_s(\Gamma) \neq \emptyset$ *for some* $i \in \mathbb{N} - \{0\}$.

Definition 11: *Marking for* IL^{gm}*: Where* A *is the second element and* Δ *the fifth element of line* j*, line* j *is marked at stage* s *iff, for some* $i \in \mathbb{N} - \{0\}$*, (i) there is no* $\varphi \in \Phi^i_s(\Gamma)$ *such that* $\varphi \cap \Delta = \emptyset$*, or (ii) for some* $\varphi \in \Phi^i_s(\Gamma)$*, there is no line* k *that has* A *as its second element and has as its fifth element some* Θ *such that* $\varphi \cap \Theta = \emptyset$.

[23] If the second element of a clean line is a Dab^1-formula, then its fifth element is obviously empty.

Next, IL^{gr}-derivability and IL^{gm}-derivability are defined as usual, viz. by Definitions 5, 6 and 7.

Here is a simple example of a proof. The Reliability strategy and the Minimal Abnormality strategy lead to the same marks. Abbreviations are as before.

1	$(Pa \wedge Ra) \wedge Sa$	–	PREM	$\emptyset$
2	$Qb \wedge {\sim}Sb$	–	PREM	$\emptyset$
3	$Rd \wedge Pc$	–	PREM	$\emptyset$
4	$\diamond(\forall x)(Rx \supset Sx)$	–	PREM	$\emptyset$
5	$\diamond(\forall x)(Qx \supset Rx)$	–	PREM	$\emptyset$
6	$\diamond^2(\forall x)(Px \supset {\sim}Rx)$	–	PREM	$\emptyset$
7	$\diamond^3(\forall x)(Rx \supset {\sim}Qx)$	–	PREM	$\emptyset$
8	$\diamond^4(\forall x)(Px \supset Qx)$	–	PREM	$\emptyset$
9^{B15}	$(\forall x)(Rx \supset Sx)$	4	RC	$\{\diamond(R \supset S)\}$
10^{B15}	Sd	3, 9	RU	$\{\diamond(R \supset S)\}$
11^{B15}	${\sim}Rb$	2, 9	RU	$\{\diamond(R \supset S)\}$
12^{B15}	$(\forall x)(Qx \supset Rx)$	5	RC	$\{\diamond(Q \supset R)\}$
13^{B15}	Rb	2, 12	RU	$\{\diamond(Q \supset R)\}$
14	${\sim}(\forall x)(Rx \supset Sx) \vee$ ${\sim}(\forall x)(Qx \supset Rx)$	2	RU	$\emptyset$
15	$Dab\{\diamond(R \supset S), \diamond(Q \supset R)\}$	4, 5, 14	RU	$\emptyset$
16^{B19}	$(\forall x)(Px \supset {\sim}Rx)$	6	RC	$\{\diamond^2(P \supset {\sim}R)\}$
17^{B19}	${\sim}Rc$	3, 16	RU	$\{\diamond^2(P \supset {\sim}R)\}$
18	${\sim}(\forall x)(Px \supset {\sim}Rx)$	1	RU	$\emptyset$
19	$Dab\{\diamond^2(P \supset {\sim}R)\}$	6, 18	RU	$\emptyset$
20	$(\forall x)(Rx \supset {\sim}Qx)$	7	RC	$\{\diamond^3(R \supset {\sim}Q)\}$
21	${\sim}Qa$	1, 20	RU	$\{\diamond^3(R \supset {\sim}Q)\}$
22^{B26}	$(\forall x)(Px \supset Qx)$	8	RC	$\{\diamond^4(P \supset Q)\}$
23^{B26}	Qa	1, 22	RU	$\{\diamond^4(P \supset Q)\}$
24^{B26}	Qc	3, 22	RU	$\{\diamond^4(P \supset Q)\}$
25	${\sim}(\forall x)(Px \supset Qx) \vee$ ${\sim}(\forall x)(Rx \supset {\sim}Qx)$	1	RU	$\emptyset$
26	$Dab\{\diamond^4(P \supset Q)\}$	7, 8, 25	RU	$\{\diamond^3(R \supset {\sim}Q)\}$

Finally, we briefly spell out the semantics. For any IM-model M of $\Sigma^\diamond$ and for every $i \in \mathbb{N} - \{0\}$, $Ab^i(M) =_{df} \{\diamond^i A \wedge {\sim}A \mid M \models \diamond^i A \wedge {\sim}A\}$. $Dab^i(\Delta)$ is a *minimal* Dab^i*-consequence* of $\Sigma^\diamond$ iff $Dab^i(\Delta) \vee Dab(\Theta)$ is a minimal Dab-consequence of $\Sigma^\diamond$ for some (possibly empty) $\Theta \subseteq \Omega^1 \cup \ldots \cup \Omega^{i-1}$. Where $Dab^i(\Delta_1), Dab^i(\Delta_2), \ldots$ are the minimal Dab^i-consequences of $\Sigma^\diamond$, $U^i(\Sigma^\Delta) =_{df} \Delta_1 \cup \Delta_2 \cup \ldots$.

The IL^{gr}-models of $\Sigma^\diamond$ are defined by a sequence of selections of models as follows:

$$\mathcal{M}_0^r =_{df} \{M \mid M \models \Sigma^\diamond\}$$

and

$$\mathcal{M}_{i+1}^r =_{df} \{M \in \mathcal{M}_i^r \mid Ab^{i+1}(M) \subseteq U^{i+1}(\Sigma^\diamond)\}\,.$$

Finally, $\Sigma^\diamond \vDash_{\mathsf{IL}^{gr}} A$ iff A is verified by every IL^{gr}-model of $\Sigma^\diamond$.

The IL^{gm}-models of $\Sigma^\diamond$ are defined by a similar sequence of selections of models:

$$\mathcal{M}_0^m =_{df} \{M \mid M \models \Sigma^\diamond\}$$

and

$$\mathcal{M}_{i+1}^m =_{df} \{M \in \mathcal{M}_i^m \mid \text{ for no } M' \in \mathcal{M}_i^m,\ Ab^{i+1}(M') \subset Ab^{i+1}(M)\}$$

$\Sigma^\diamond \vDash_{\mathsf{IL}^{gm}} A$ iff A is verified by every IL^{gm}-model of $\Sigma^\diamond$.

6. *Prioritized Pragmatic Background Generalizations*

The reason to introduce pragmatic background generalizations was discussed in Section 1: some generalizations are retained even after they have been falsified. What this comes to is that, even if the generalization itself has been falsified, the instances of the generalizations are retained, unless and until they are themselves falsified. The previous sentence contains the full plot. Pragmatic background generalizations have to be handled as sets of instances, rather than as generalizations. A different matter is that, just like normal background generalizations, pragmatic background generalizations should be prioritized. Let A and B be instances of two different pragmatic background generalizations. If the data contradict $A \wedge B$, either A or B may nevertheless be retained depending on the priority of the pragmatic generalization from which it is an instance.

With this settled, handling pragmatic background generalizations is simple enough. All we have to change is the modal 'translation' of Σ. Let us proceed in the crudest possible way. All pragmatic background generalizations are of the form $\forall A$ in which A is purely functional. Suppose that $x_1, \ldots, x_n$ are the individual variables that occur in A. By an instance of A we obviously mean a formula in which each x_i is systematically replaced by an individual constant a_i. Let $A^{x_1,\ldots,x_n}_{a_1,\ldots,a_n}$ denote an instance of A. Where Σ is as in (4), we define (quite differently from the previous section)[24]

$$\Sigma^{\diamond} = \{A \mid A \in \Gamma_0^0\} \cup \{\diamond^i A^{x_1,\ldots,x_n}_{\alpha_1,\ldots,\alpha_n} \mid \forall A \in \Gamma^i_j; 0 < j \leq n; \alpha_1, \ldots, \alpha_n \in \mathcal{C}\}.$$

The logics IL^{gr} and IL^{gm} adequately handle pragmatic background generalizations. We have the same lower limit logic, the same set of abnormalities (and hence the same upper limit logic) and the same semantics and proof theory. Under the modal translation from the previous section, we faced a take-it-or-leave-it situation with respect to a generalization $\forall A$. To be more precise, any background generalization $\forall A$ was separately and jointly (with other generalizations of the same level) compatible with the data and the retained generalizations of lower levels, or it had to go. Under the present translation, instances of pragmatic background generalizations are retained, provided they are separately and jointly (with instances of other generalizations of the same level) compatible with the data and the retained instances of generalizations of lower levels. Once the premises are correctly phrased, the same logics handle them adequately. To illustrate the matter, we present an example of a simple proof. The instances of pragmatic background generalizations are all instances of $\diamond(\forall x)(Rx \supset Sx)$, $\diamond(\forall x)(Qx \supset Rx)$, $\diamond\diamond(\forall x)(Px \supset {\sim}Rx)$, $\diamond\diamond\diamond(\forall x)(Rx \supset {\sim}Qx)$, and $\diamond\diamond\diamond\diamond(\forall x)(Px \supset Qx)$.

1	$(Pa \wedge Ra) \wedge Sa$	–	PREM	$\emptyset$
2	$Qb \wedge {\sim}Sb$	–	PREM	$\emptyset$
3	$Rd \wedge Pc$	–	PREM	$\emptyset$
4	$\diamond(Rb \supset Sb)$	–	PREM	$\emptyset$
5	$\diamond(Rd \supset Sd)$	–	PREM	$\emptyset$
6^{B14}	$Rb \supset Sb$	4	RC	$\{\Diamond(Rb \supset Sb)\}$
7^{B14}	${\sim}Rb$	2, 6	RU	$\{\Diamond(Rb \supset Sb)\}$
8	$Rd \supset Sd$	5	RC	$\{\Diamond(Rd \supset Sd)\}$
9	Sd	3, 8	RU	$\{\Diamond(Rd \supset Sd)\}$
10	$\diamond(Qb \supset Rb)$	–	PREM	$\emptyset$
11^{B14}	$Qb \supset Rb$	10	RC	$\{\Diamond(Qb \supset Rb)\}$
12^{B14}	Rb	2, 11	RU	$\{\Diamond(Qb \supset Rb)\}$
13	${\sim}(Qb \supset Rb) \vee {\sim}(Rb \supset Sb)$	2	RU	$\emptyset$
14	$Dab\{\diamond(Qb \supset Rb), \diamond(Rb \supset Sb)\}$	4, 10, 13	RU	$\emptyset$
15	$\diamond^2(Pa \supset {\sim}Ra)$	–	PREM	$\emptyset$

[24] Even if Σ is finite, $\Sigma^{\diamond}$ is infinite but characterized by finitely many forms. For those who dislike this, it is simple enough to apply a variant of IM in which all $w \in W - \{w_0\}$ have the property $v_M((\forall\alpha)A(\alpha), w) = 1$ iff $v((\forall\alpha)A(\alpha), w) = 1$ and $v_M(A(\beta), w) = 1$ for all $\beta \in \mathcal{C} \cup \mathcal{O}$.

16^{B18}	$Pa \supset \sim Ra$	15	RC	$\{\Diamond^2(Pa \supset \sim Ra)\}$
17	$\sim(Pa \supset \sim Ra)$	1	RU	$\emptyset$
18	$Dab\{\Diamond^2(Pa \supset \sim Ra)\}$	15, 17	RU	$\emptyset$
19	$\Diamond^2(Pc \supset \sim Rc)$	–	PREM	$\emptyset$
20	$Pc \supset \sim Rc$	19	RC	$\{\Diamond^2(Pc \supset \sim Rc)\}$
21	$\sim Rc$	3, 20	RU	$\{\Diamond^2(Pc \supset \sim Rc)\}$
22	$\Diamond^3(Ra \supset \sim Qa)$	–	PREM	$\emptyset$
23	$Ra \supset \sim Qa$	22	RC	$\{\Diamond^3(Ra \supset \sim Qa)\}$
24	$\sim Qa$	1, 23	RU	$\{\Diamond^3(Ra \supset \sim Qa)\}$
25	$\Diamond^4(Pa \supset Qa)$	–	PREM	$\emptyset$
26^{B29}	$Pa \supset Qa$	25	RC	$\{\Diamond^4(Pa \supset Qa)\}$
27^{B29}	Qa	1, 26	RU	$\{\Diamond^4(Pa \supset Qa)\}$
28	$\sim(Pa \supset Qa) \vee \sim(Ra \supset \sim Qa)$	1	RU	$\emptyset$
29	$Dab\{\Diamond^4(Pa \supset Qa)\}$	22, 25, 28	RU	$\{\Diamond^3(Ra \supset \sim Qa)\}$
30	$\Diamond^4(Pc \supset Qc)$	–	PREM	$\emptyset$
31	$Pc \supset Qc$	30	RC	$\{\Diamond^4(Pc \supset Qc)\}$
32	Qc	3, 31	RU	$\{\Diamond^4(Pc \supset Qc)\}$

7. *Background Theories*

We shall consider theories as sets of statements that form connected wholes in at least two respects. First, the consequences of a theory are not just the consequences of the statements that make up the theory, but the consequences of the set. Next, if some consequence of a theory is falsified, the theory is rejected as a whole, and not just those members of the theory from which the consequence is derivable.

Of course, this approach is a simplification. First of all, consequences of a rejected theory will usually be retained as background generalizations or as pragmatic background generalizations. However, the choice of those consequences is not a function of the theory itself or of its degree of priority with respect to other theories. Hence, this problem falls beyond the scope of the logics of induction discussed in the present paper. Moreover, falsified theories are often not rejected at all, but are retained for all applications except for those at which the falsification applies. The underlying reasoning is clearly captured by an adaptive logic. However, this logic is necessarily *inconsistency-adaptive* and cannot have CL as its lower limit logic. For this reason, we do not discuss this logic in the present paper but postpone its discussion to a paper about the many inconsistency-adaptive forms of reasoning that are connected to inductive reasoning.

Given the present confines, how should a logic handle background theories? It is important to realize that two different theories may be assigned the same degree of priority. This is why the usual notation should be as the Σ in (4). The official formulation of the premises will again be a 'modal translation' of Σ. As the consequences of a theory are the consequences of the set of statements (for example $\Gamma_j^{i_j}$), and not of the individual statements, the 'translation' will read[25]

$$\Sigma^{\Diamond} = \{\Diamond^i(A_1 \wedge \ldots \wedge A_m) \mid A_1, \ldots, A_m \in \Gamma_j^i \text{ for some } \Gamma_j^i \in \Sigma\}. \qquad (5)$$

in which the modal symbols are settled by the logic T from Section 4. This warrants at once that $\Diamond^i A$ is T-derivable from the premises iff there is a Γ_j^i such that $\Gamma_j^i \vdash_{\mathsf{CL}} A$. The idea for the adaptive logic is obviously that A is derivable from $\Diamond^i A$, unless an abnormality prevents this.

Let us now turn to the adaptive logic that should handle the relation between the data Γ_0^0 and a *single* theory, say Γ_1^1 (the priority index is of course arbitrary provided it is larger than 0). On the official formulation, the set of premises is $\Sigma^{\Diamond} = \Gamma_0^0 \cup \{\Diamond(A_1 \wedge \ldots \wedge A_m) \mid A_1, \ldots, A_m \in \Gamma_1^1\}$. As the whole theory has to be left or kept according as it is falsified or not, we typically need a *flip-flop* logic. Flip-flop logics are a border case of adaptive logics with the following property: the adaptive consequences are identical to the lower limit consequences if an abnormality is derivable from the premises (by the lower limit logic), and are identical to the upper limit consequences otherwise.[26] This is precisely what a flip-flop logic will deliver: either all CL-consequences of the theory Γ_1^1 are retained, or none of them is (except when it follows from the data and retained theories).

The technical implementation of this idea is easily obtained. Let us define two adaptive logics T^{ftr} and T^{ftm}. These are *flat* adaptive logics to handle a single background *theory* by, respectively, the *Reliability* strategy and the *Minimal Abnormality* strategy — later we shall upgrade these logics to respectively T^{tr} and T^{tm} to handle a prioritized set of theories. The set of

[25] Sometimes a theory is represented by a single formula (for example in [11]), viz. the conjunction of the finite set of statements that make up the theory. However, most interesting theories will at least include Arithmetic, and the first order induction 'axiom' is actually a schema, representing infinitely many axioms.

[26] Some people think that all adaptive logics are flip-flop logics, but this is obviously mistaken. Most adaptive logics assign to an abnormal premise set more consequences than the lower limit logic.

abnormalities Ω comprises all formulas of the form $\Diamond A \wedge \sim A$ and a Dab-formula is a disjunction of members of Ω as before. The upper limit logic is obviously Triv.

It is easily seen that these logics provide exactly the desired result. We shall show this for T^{ftr}. If $\Sigma^{\Diamond}$ is normal, then obviously $Cn_{\mathsf{T}^{ftr}}(\Sigma^{\Diamond}) = Cn_{\mathsf{Triv}}(\Sigma^{\Diamond})$; in other words, all of the theory is retained. If $\Sigma^{\Diamond}$ is abnormal, then, as we now shall show, $Cn_{\mathsf{T}^{ftr}}(\Sigma^{\Diamond}) = Cn_{\mathsf{T}}(\Sigma^{\Diamond})$; in other words, the whole theory is rejected.

Suppose then that $\Sigma^{\Diamond}$ is abnormal. This means that there are $B_1, \ldots, B_n$ $(n > 0)$ such that $(\Diamond B_1 \wedge \sim B_1) \vee \ldots \vee (\Diamond B_n \wedge \sim B_n)$ is T-derivable from $\Sigma^{\Diamond}$. Consider an arbitrary A such that $\Sigma^{\Diamond} \vdash_{\mathsf{T}} \Diamond A$. If $\Sigma^{\Diamond} \vdash_{\mathsf{T}} A$, then $\Sigma^{\Diamond} \vdash_{\mathsf{T}} \Diamond A$ is inconsequential. If $\Sigma^{\Diamond} \vdash_{\mathsf{T}} \sim A$, then $\Sigma^{\Diamond} \vdash_{\mathsf{T}} \Diamond A \wedge \sim A$, and hence $\Diamond A \wedge \sim A \in U(\Sigma^{\Diamond})$. If $\Sigma^{\Diamond} \nvdash_{\mathsf{T}} A$ and $\Sigma^{\Diamond} \nvdash_{\mathsf{T}} \sim A$, then

$$(\Diamond A \wedge \sim A) \vee (\Diamond(A \supset B_1) \wedge \sim(A \supset B_1)) \vee \ldots \vee (\Diamond(A \supset B_n) \wedge \sim(A \supset B_n))$$

is a minimal Dab-consequence of $\Sigma^{\Diamond}$ and hence $\Diamond A \wedge \sim A \in U(\Sigma^{\Diamond})$.[27] Summarizing, if $\Sigma^{\Diamond}$ is abnormal and A is a CL-consequence of the theory (hence $\Diamond A$ is a T-consequence of $\Sigma^{\Diamond}$), then A is a T^{ftr}-consequence of $\Sigma^{\Diamond}$ iff it is a CL-consequence of $\Sigma^{\Diamond}$ (and hence is a CL-consequence of the data). A similar result may be established for T^{ftm}.

Given these results, we shall present a direct dynamic proof format (one in which the modalities are suppressed). In a standard proof, one would have a conditional rule allowing for the following transition, where $A \in \Gamma^1_1$:

i	$\Diamond A$	$\ldots$	$\ldots$	$\emptyset$
$i+1$	A	i	RC	$\{\Diamond A \wedge \sim A\}$

If Γ^1_1 is contradicted by the data, then line $i+1$ will be marked because the adaptive logic is a flip-flop. This means that we may just as well write:

i	$\Diamond A$	$\ldots$	$\ldots$	$\emptyset$
$i+1$	A	i	RC	$\{\Gamma^1_1\}$

and mark all lines that have Γ^1_1 in their condition as soon as some $B \wedge \sim B$ has been derived on the condition $\{\Gamma^1_1\}$.[28] Remark that we do not change the logic, but only the notation, and that we mark in view of a derivable rule.

[27] As $\Sigma^{\Diamond} \nvdash_{\mathsf{T}} A$ and $\Sigma^{\Diamond} \nvdash_{\mathsf{T}} \sim A$, neither $\Diamond A \wedge \sim A$ nor $(\Diamond(A \supset B_1) \wedge \sim(A \supset B_1)) \vee \ldots \vee (\Diamond(A \supset B_n) \wedge \sim(A \supset B_n))$ are T-derivable from $\Sigma^{\Diamond}$.

[28] If both Γ^0_0 and Γ^1_1 are consistent, B will be derived from one of these and $\sim B$ from the other.

As the next step, we upgrade the logics to respectively T^{tr} and T^{tm} in order to handle a prioritized set of theories. The lower logic is T, the set of abnormalities $\Omega = \{\Diamond^i A \wedge {\sim} A \mid A \in \mathcal{W};\ i \in \mathbb{N} - \{0\}\}$, and the upper limit logic Triv. The set of premises has the form $\Sigma^{\Diamond}$ as in (5).

Let us start with the semantics. For any T-model M of $\Sigma^{\Diamond}$ and for every $i \in \mathbb{N} - \{0\}$, we define $Ab^i(M) =_{df} \{\Diamond^i A \wedge {\sim}A \mid M \models \Diamond^i A \wedge {\sim}A\}$. For every $i \in \mathbb{N} - \{0\}$, we define $U^i(\Sigma^{\Delta}) =_{df} \Delta_1 \cup \Delta_2 \cup \ldots$, where $Dab^i(\Delta_1), Dab^i(\Delta_2), \ldots$ are the minimal Dab^i-consequences of $\Sigma^{\Diamond}$.

The IL^{tr}-models of $\Sigma^{\Diamond}$ are defined by a sequence of selections of models as follows:

$$\mathcal{M}_0^r =_{df} \{M \mid M \models \Sigma^{\Diamond}\}$$

and

$$\mathcal{M}_{i+1}^r =_{df} \{M \in \mathcal{M}_i^r \mid Ab^{i+1}(M) \subseteq U^{i+1}(\Sigma^{\Diamond})\}\,.$$

Finally, $\Sigma^{\Diamond} \vDash_{\mathsf{IL}^{tr}} A$ iff A is verified by every IL^{tr}-model of $\Sigma^{\Diamond}$.

The IL^{tm}-models of $\Sigma^{\Diamond}$ are defined in a similar way:

$$\mathcal{M}_0^m =_{df} \{M \mid M \models \Sigma^{\Diamond}\}$$

and

$$\mathcal{M}_{i+1}^m =_{df} \{M \in \mathcal{M}_i^m \mid \text{ for no } M' \in \mathcal{M}_i^m,\ Ab^{i+1}(M') \subset Ab^{i+1}(M)\}$$

$\Sigma^{\Diamond} \vDash_{\mathsf{IL}^{tm}} A$ iff A is verified by every IL^{tm}-model of $\Sigma^{\Diamond}$.

In the direct dynamic proofs, the conditions (fifth elements) of the lines will be sets of theories rather than sets of abnormalities. That A is derived on a condition like $\{\Gamma_5^2, \Gamma_7^1\}$ is correctly interpreted as: A is derivable from the premises (the data and the theories) unless an abnormality is CL-derivable from $\Gamma_0^0 \cup \Gamma_5^2 \cup \Gamma_7^1$ — remark that the members of Γ_0^0 are always available. As before, that an inconsistency is derived on some condition Δ indicates that a disjunction of abnormalities is derivable from $\Gamma_0^0 \cup \bigcup(\Delta)$. Here are the rules for the dynamic proof theories:

PREM If $A \in \Sigma^{\Diamond} \cap \mathcal{W}$ (in other words, $A \in \Gamma_0^0$), one may add a line comprising the following elements: (i) an appropriate line number, (ii) A, (iii) $-$, (iv) PREM, and (v) $\emptyset$.

CPREM If $B_1, \ldots, B_n \in \Gamma_j^i$ and hence $\Diamond^i(B_1 \wedge \ldots \wedge B_n) \in \Sigma^{\Diamond}$, one may add a line comprising the following elements: (i) an appropriate line number, (ii) $B_1 \wedge \ldots \wedge B_n$, (iii) $-$, (iv) CPREM, and (v) $\{\Gamma_j^i\}$.[29]

RU If $A_1, \ldots, A_n \vdash_{\mathsf{T}} B$ and each of $A_1, \ldots, A_n$ occur in the proof on lines $i_1, \ldots, i_n$ that have conditions $\Delta_1, \ldots, \Delta_n$ respectively, one may add a line comprising the following elements: (i) an appropriate line number, (ii) B, (iii) $i_1, \ldots, i_n$, (iv) RU, and (v) $\Delta_1 \cup \ldots \cup \Delta_n$.

Remark that the direct dynamic proof theory does not require a conditional rule of the usual type. CPREM is the only conditional rule. The marking rules are very simple but require some explanation. Suppose that, on line k, an inconsistency is derived on the condition $\{\Gamma^{i_1}_{j_1}, \ldots \Gamma^{i_n}_{j_n}\}$. This indicates that one of the theories is unreliable. Remember, however, that a lower superscript expresses that a theory has a higher priority. So, line k indicates that one should reject one of the theories in the condition that have the highest superscript, unless of course another theory is rejected for a different reason. Such a reason would be that the theory occurs in the condition of an inconsistency and that its superscript is maximal in that condition — we illustrate this by an example below.[30] A simple example is where an inconsistency is derived on condition $\{\Gamma^2_4, \Gamma^4_3\}$ but an inconsistency is also derived on the condition $\{\Gamma^2_4, \Gamma^1_2\}$. The latter line indicates that Γ^2_4 should be rejected, whence the former line (is already marked and) does not provide one with a reason to reject Γ^4_3.

We now specify what corresponds to a disjunction of abnormalities on the normal construction. Given a proof at a stage, let L be the set of lines at which an inconsistency is derived. If the condition of some line in L is the empty set, the data are inconsistent and even the lower limit logic T-consequence set of $\Sigma^\Diamond$ is trivial, whence nay continuation of the proof is useless. Next, one locates the lines in L for which 1 is the highest superscript in the condition. For each such line, its condition Δ is an *abnormal set of theories*. Next one locates the the lines in L for which 2 is the highest superscript in the condition. For each such line, one defines from its condition Δ the set $\Delta^2 = \{\Gamma^i_j \in \Delta \mid i = 2\}$; these Δ^2 are *abnormal sets of theories*. And so one continues up to the highest priority number (denoting the lowest priority). Remark that all members of an abnormal set of theories have the same superscript. A *minimal abnormal set of theories* at stage s is

[29] We let CPREM introduce conjunctions to remind the reader of the official formulation of the proofs.

[30] It is easily checked that the official formulation of the proof theory, in which the conditions are sets of abnormalities rather than sets of theories, leads to exactly the same result. Actually, the rule in the text is obtained by 'translating' the marking rules for those proof theories to the shortcut format.

an abnormal set of theories at stage s that is not a proper superset of another abnormal set of theories at stage s.

For the Reliability strategy, one defines $U^i_s(\Sigma^\Diamond)$ as the *union* of the minimal abnormal sets of theories (at stage s) of which the members have superscript i.

Definition 12: *Marking for* LI^{tr}*, for all* $i \in \mathbb{N} - \{0\}$*, starting with the lowest one: Where* Δ *is the fifth element of line* j*, line* j *is marked at stage* s *iff,* $\Delta \cap U^i_s(\Gamma) \neq \emptyset$.

The idea behind the Minimal Abnormality strategy is equally simple. For each priority level i, one defines $\Phi^{i\circ}_s(\Gamma)$ as the set of all sets that contain one member out of each minimal set of abnormal set of theories, and one defines $\Phi^i_s(\Gamma)$ as those members of $\Phi^{i\circ}_s(\Gamma)$ that are not proper supersets of other members of $\Phi^{i\circ}_s(\Gamma)$.

Definition 13: *Marking for* IL^{tm}*, for all* $i \in \mathbb{N} - \{0\}$*, starting with the lowest one: Where* A *is the second element and* Δ *the fifth element of line* j*, line* j *is marked at stage* s *iff, (i) there is no* $\varphi \in \Phi^i_s(\Gamma)$ *such that* $\varphi \cap \Delta = \emptyset$*, or (ii) for some* $\varphi \in \Phi^i_s(\Gamma)$*, there is no line* k *that has* A *as its second element and has as its fifth element some* Θ *such that* $\varphi \cap \Theta = \emptyset$.

As an example, we list a fragment of a proof to illustrate the marking rules (both again leading to the same marks in the example). Let $(Pa \wedge Qa) \wedge {\sim}Ra, Rb \wedge {\sim}Qb \in \Gamma^0_0$ and let the background theories be such that $\Gamma^1_1 \vdash_{\mathsf{CL}} (\forall x)(Qx \supset {\sim}Sx)$, $\Gamma^2_4 \vdash_{\mathsf{CL}} (\forall x)(Px \supset Sx)$, $\Gamma^2_4 \vdash_{\mathsf{CL}} (\forall x)(Rx \supset {\sim}Tx)$, $\Gamma^3_5 \vdash_{\mathsf{CL}} (\forall x)({\sim}Qx \supset Tx)$. The following lines may then occur in a proof from those premises:

1	$(Pa \wedge Qa) \wedge {\sim}Ra$	$-$	Prem	$\emptyset$
2	$Rb \wedge {\sim}Qb$	$-$	Prem	$\emptyset$
17	$(\forall x)(Qx \supset {\sim}Sx)$	$\ldots$	RC	$\{\Gamma^1_1\}$
18	$(\forall x)(Px \supset Sx)$	$\ldots$	RC	$\{\Gamma^2_4\}$
19	$(\forall x)(Rx \supset {\sim}Tx)$	$\ldots$	RC	$\{\Gamma^2_4\}$
20	$(\forall x)({\sim}Qx \supset Tx)$	$\ldots$	RC	$\{\Gamma^3_5\}$

Suppose that the proof is continued as follows:

21	Tb	2, 20	RU	$\{\Gamma^3_5\}$
22	Sa	1, 18	RU	$\{\Gamma^2_4\}$
23	${\sim}Tb$	2, 19	RU	$\{\Gamma^2_4\}$
24	$Tb \wedge {\sim}Tb$	21, 23	RU	$\{\Gamma^3_5, \Gamma^2_4\}$

At this point (at stage 24 of the proof), Γ_5^3 is considered as falsified and hence lines 20, 21 and 24 are marked. Next, let the proof continue as follows:

25	$\sim Sa$	1, 17	RU	$\{\Gamma_1^1\}$
26	$Sa \wedge \sim Sa$	22, 25	RU	$\{\Gamma_4^2, \Gamma_1^1\}$

At this stage, Γ_4^2 is considered as falsified, whence lines 18, 19, 22–24 and 26 are marked. It follows that Γ_5^3 is not any more considered as falsified, and hence that no line is marked except for those mentioned in the previous sentence. How the proof may continue further will of course depend on the other premises.

8. *Putting the logics Together*

It is very easy to combine the previously mentioned logics, but the matter is somewhat longwinded. This is why we shall discuss it in a rather informal way.

We need a modal logic that combines T and IM. Let us call it TIM. A TIM-model is a sextuple $\langle W, W', w_0, R, D, v\rangle$ in which W and W' are (disjoint) sets of worlds, $w_0 \in W$ the real world, R a reflexive binary relation on $W \cup W'$, D a non-empty set and v an assignment function. The assignment function v is as for IM-models. For all $A \in \mathcal{W}$, one defines $v_M(A, w)$ as in the IM-semantics if $w \in W'$, and one defines $v_M(A, w)$ as in the T-semantics if $w \in W$ (in which case C1.4 has no effect). $v_M(\Diamond A, w) = 1$ iff $v_M(A, w') = 1$ for some $w' \in W$ such that Rww'. $v_M(\diamond A, w) = 1$ iff $v_M(A, w') = 1$ for some $w' \in W'$ such that Rww'.

The set of premises $\Sigma^{\Diamond}$ will contain: (i) modal free formulas, which are the data, (ii) formulas of the form $\diamond^i \forall A$, which are background generalizations of priority i, (ii) sets of formulas $\{\diamond^i A^{x_1,\ldots,x_n}_{\alpha_1,\ldots,\alpha_n} \mid \alpha_1, \ldots, \alpha_n \in \mathcal{C}\}$, each of which contains all the instances of some pragmatic background generalization, and (iv) formulas of the form $\Diamond^i A$ in which A is a conjunction of members of the same background theory of priority i.

The adaptive logics IL^{+r} and IL^{+m} are defined as follows. The lower logic is obviously TIM and the set of abnormalities Ω is the union of all $\Omega^i = \{\diamond^i A \mid i > 0;\ A \in \mathcal{W}\} \cup \{\Diamond^i A \mid i > 0;\ A \in \mathcal{W}\} \cup \{\exists A \wedge \exists{\sim}A \mid A \in \mathcal{F}^{\circ}\}$. The upper limit logic is UTriv, obtained by adding to TIM the axioms $\diamond A \supset A$, $\Diamond A \supset A$ and $\exists A \supset \forall A$.

Let us start with the semantics. It is most handy to proceed in two steps, the first taking care of background knowledge, the second of the local generalizations.

For every TIM-model M of $\Sigma^{\Diamond}$ and for every $i \in \mathbb{N} - \{0\}$, we define $Ab^i(M) =_{df} \{\diamond^i A \wedge \sim A \mid M \models \diamond^i A \wedge \sim A\} \cup \{\Diamond^i A \wedge \sim A \mid M \models \Diamond^i A \wedge$

$\sim A\}$. For every $i \in \mathbb{N} - \{0\}$, we define $U^i(\Sigma^{\Delta}) =_{df} \Delta_1 \cup \Delta_2 \cup \ldots$, where $Dab^i(\Delta_1), Dab^i(\Delta_2), \ldots$ are the minimal Dab^i-consequences of $\Sigma^{\Diamond}$.[31]

For every TIM-model M of $\Sigma^{\Diamond}$ we also define $Ab^0(M) =_{df} \{\exists A \wedge \exists{\sim}A \mid M \models \exists A \wedge \exists{\sim}A;\ A \in \mathcal{F}^{\circ}\}$.

That a model is reliable with respect to the background knowledge is defined by a sequence of selections of models:

$$\mathcal{M}_0^r =_{df} \{M \mid M \models \Sigma^{\Diamond}\}$$

and

$$\mathcal{M}_{i+1}^r =_{df} \{M \in \mathcal{M}_i^r \mid Ab^{i+1}(M) \subseteq U^{i+1}(\Sigma^{\Diamond})\}\,.$$

Let $\mathcal{M}_b^r$ be the models obtained by this selection. This set of models defines a set of minimal Dab^0-consequences of $\Sigma^{\Diamond}$ (disjunctions of formulas of the form $\exists A \wedge \exists{\sim}A$ with $A \in \mathcal{F}^{\circ}$). Let $U^0(\Sigma^{\Diamond})$ be their union.

The IL^{+r}-models of $\Sigma^{\Diamond}$ are defined by

$$\mathcal{M}^{+r} =_{df} \{M \in \mathcal{M}_b^r \mid Ab^0(M) \subseteq U^0(\Sigma^{\Diamond})\}\,.$$

Finally, $\Sigma^{\Diamond} \vDash_{\mathsf{IL}^{tr}} A$ iff A is verified by every IL^{+r}-model of $\Sigma^{\Diamond}$.

We proceed in a similar way for IL^{+m}. First we define the models of $\Sigma^{\Diamond}$ that are minimally abnormal with respect to the background knowledge:

$$\mathcal{M}_0^m =_{df} \{M \mid M \models \Sigma^{\Diamond}\}$$

and

$$\mathcal{M}_{i+1}^m =_{df} \{M \in \mathcal{M}_i^m \mid \text{ for no } M' \in \mathcal{M}_i^m,\ Ab^{i+1}(M') \subset Ab^{i+1}(M)\}$$

Let $\mathcal{M}_b^m$ be the models obtained by this selection.

The IL^{+m}-models of $\Sigma^{\Diamond}$ are defined by

$$\mathcal{M}^{+r} =_{df} \{M \in \mathcal{M}_b^r \mid \text{ for no } M' \in \mathcal{M}_b^m,\ Ab^0(M') \subset Ab^0(M)\}\,.$$

Finally, $\Sigma^{\Diamond} \vDash_{\mathsf{IL}^{+m}} A$ iff A is verified by every IL^{+m}-model of $\Sigma^{\Diamond}$.

[31] Remark that $Dab^i(\Delta)$ is a minimal Dab^i-consequences of $\Sigma^{\Diamond}$ iff it is verified by any TIM-model of $\Sigma^{\Diamond}$ and, for all $\Delta' \subset \Delta$, $Dab^i(\Delta')$ is falsified by at least one TIM-model of $\Sigma^{\Diamond}$.

We now turn to the proof theory. The rules PREM, RU and RC are as for IL^{gr} and IL^{gm} — see Section 5 — except that IM is replaced by TIM.[32] For the marking definitions we shall follow an approach that is as simple as possible in view of the logics that have been combined.

Remember the order in which abnormalities are avoided. The logics first avoid abnormalities that derive from background items with priority 1; let us call these abnormalities of level 1. Next, the logics avoid abnormalities that derive from background items with priority 2 (abnormalities of level 2). And so on. Only with all background knowledge thus taken into account, the logics avoid abnormalities that pertain to local generalizations — remember that these have a different form than the others, viz. $\exists A \wedge \exists{\sim}A$.

Some warning may be useful here. First, the abnormalities of priority level i contain abnormalities of the form $\Diamond^i A \wedge {\sim}A$ as well as abnormalities of the form $\diamond^i A \wedge {\sim}A$. Next, disjunctions of abnormalities of the form $\exists A \wedge \exists{\sim}A$ (called Dab^0-formulas because no possibility operator occurs in them) may be thought of as abnormalities of level 0, *provided* one remembers that while, for all other numbers, the logics avoid first abnormalities of the lower levels, they avoid abnormalities of level 0 only after abnormalities of all other levels. To avoid confusion in this respect we shall say that priority level 1 is *better* than level 2, etc., but that all levels are better than level 0.

A *clean line* is defined as in Section 5, except that we now require that the abnormalities in the fifth element are *better* than the disjuncts of the second element. Here are two examples of clean lines:

i	$\Diamond\Diamond A \wedge {\sim}A$	$\dots$	$\dots$	$\{\Diamond B \wedge {\sim}B, \diamond C \wedge {\sim}C\}$
j	$\exists A \wedge \exists{\sim}A$	$\dots$	$\dots$	$\{\Diamond\Diamond B \wedge {\sim}B, \diamond C \wedge {\sim}C\}$

The definition of a *minimal* Dab^i*-formula* (at a stage) is extended accordingly to Dab^0-formulas. Given this, $U^i_s(\Sigma^\Diamond)$ and $\Phi^i_s(\Sigma^\Diamond)$ are defined as before (see for example Section 5) in terms of the minimal Dab^i-formulas at stage s. The marking rules are exactly as in Section 5, except that, at each stage of a proof, one starts marking in view of the best priority level, which is 1, next marks in view of priority level 2, etc., and finally one marks in view of priority level 0.

The logics IL^{+r} and IL^{+m} handle data, background generalizations, pragmatic background generalizations, and background theories. One way to look at these logics is by saying that their consequence sets extend the data

[32] Remark that RC handles any Dab-formula, even a mixed one, in which abnormalities of all sorts and all levels occur. Needless to say, our present construction will handle background theories in the official way. Combining the direct proof theory from Section 7 with the proof theories of the other sections would complicate the exposition (although not the proofs).

in the following way. First, all CL-consequences of the data are in the adaptive consequence set. Next, for each priority level and starting with the best one (represented by the lowest number), certain background generalizations, instances of pragmatic background generalizations, and theories are added to the set, and the result is closed under CL. Remark that there is no intrinsic difference in priority between all those elements (with the exception of the data). In other words, if one judges that, for a given set of knowledge, theories should have precedence on pragmatic generalizations, one has to express this by the way in which the premises are represented, viz. by a difference in modalities in the premises. The thus extended data determine which *local* generalizations will be added to the consequence set.

An essential feature of the proofs is that (unlike what the previous paragraph might suggest) applications of the rules may occur *in any order*. This is essential because the consequence sets mentioned in the previous paragraph may be undecidable. Suppose that one first derives some local generalizations by RC, and only then applies RC to bring some background items to level 0. In view of the background items that are retained, lines at which local generalizations are derived may be marked. Certain heuristic strategies may obviously improve the efficiency of a proof, but this a different matter which we shall not discuss here.

9. *Open Problems*

It goes without saying that the properties of the discussed logics should be studied better, and that technical simplifications should be introduced wherever possible. One interesting property that deserves a high priority is the relation between the basic induction logics and compatibility. Roughly stated, a generalization is inductively derivable from a set of data iff it is compatible with the union of the data and any set of generalizations that is itself compatible with the data.

An important problem that is rather easy to answer but could not be discussed in the present paper concerns predictions on the basis of data that are not uniform. Suppose that, amongst the data, most P are Q. One then clearly wants to predict that the next P will be Q. The logic leading to the desired consequence is simple, but quite different from the ones presented in previous sections. For this reason, it was not discussed here. A related problem concerns the derivation of statistical hypotheses, say of the form $P(A/B) = r$ in which A and B are purely functional and r is a real number between 0 and 1 included. For some related results we refer to [14].

Very different open problems concern inconsistent data and inconsistent background knowledge. The logics discussed in this paper handle such cases, but they do not do so in an adequate way. Inconsistent data result in triviality, inconsistent background knowledge is simply discarded. It is not difficult to devise more adequate logics for such situations, and the logics presented in this paper are special cases of them. However, the inconsistent situations require that one combines induction logics with inconsistency-adaptive logics, and anything relating to inconsistency-adaptive logics was avoided in the present paper.

Finally, quite some problems that belong to the philosophy of science are as yet unsolved. For example, there are situations in which a scientist will introduce local generalizations that conflict with background knowledge in the hope to reach a new theory. We by no means underestimate such problems. The main reason for not discussing them here is that we consider the present results as a necessary condition for that discussion.[33]

Centre for Logic and Philosophy of Science
Universiteit Gent, Belgium
E-mail: {Diderik.Batens, Lieven.Haesaert}@rug.ac.be

REFERENCES

[1] Arnon Avron and Iddo Lev. Formula-preferential systems for paraconsistent non-monotonic reasoning (an extended abstract). To appear.

[2] Diderik Batens. Dialectical dynamics within formal logics. *Logique et Analyse*, 114:161–173, 1986.

[3] 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.

[4] Diderik Batens. Zero logic adding up to classical logic. *Logical Studies*, 2:15, 1999. (Electronic Journal: http://www.logic.ru/LogStud/02/LS2.html).

[5] Diderik Batens. The need for adaptive logics in epistemology. To appear.

[6] Diderik Batens. On a logic of induction. To appear.

[7] Diderik Batens and Joke Meheus. The adaptive logic of compatibility. *Studia Logica*, 66:327–348, 2000.

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

[9] Diderik Batens, Joke Meheus, Dagmar Provijn, and Liza Verhoeven. Some adaptive logics for diagnosis. *Logique et Analyse*, To appear.

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

[10] Rudolf Carnap. *The Continuum of Inductive Methods*. University of Chicago Press, Chicago, 1952.
[11] Jaakko Hintikka. *Inquiry as Inquiry: A Logic of Scientific Discovery*. Kluwer, Dordrecht, 1999.
[12] Theo A. F. Kuipers. *From Instrumentalism to Constructive Realism*, volume 287 of *Synthese Library*. Kluwer, Dordrecht, 2000.
[13] Iddo Lev. Preferential systems for plausible non-classical reasoning. Master’s thesis, Department of Computer Science, Tel-Aviv University, 2000.
[14] Maarten Van Dyck and Erik Weber. Adaptive logic and covering law explanation. *Logique et Analyse*, to appear.