L&PS – Logic & Philosophy of Science
Vol. IV, No. 1, 2006, pp. 3-32

On a Logic of Induction*

Diderik Batens
Centre for Logic and Philosophy of Science,
University of Ghent (Belgium)
e-mail: diderik.batens@UGent.be

ABSTRACT. In this paper I present a simple and straightforward logic of induction: a consequence relation characterized by a proof theory and a semantics. This system will be called **LI**. The premises will be restricted to, on the one hand, a set of empirical data and, on the other hand, a set of background generalizations. Among the consequences will be generalizations as well as singular statements, some of which may serve as predictions and explanations.

1. Prelude

I published my first paper in English a long time ago. In the paper (Batens 1968) I compared Carnap's and Popper's approach to induction, and basically assigned each approach a context of application, except that a modification was proposed for Popper's corroboration function. I had sent the paper to Carnap, Popper, Hempel, Kemeny, and several other famous people. With one exception, all had returned a few encouraging lines. Not long thereafter, I received a letter, in Dutch, by someone I immediately recognized as Dutch be-

* This paper appeared in R. Festa, A. Aliseda and J. Peijnenburg (eds.), *Confirmation, Empirical Progress, and Truth Approximation* (*Poznań Studies in the Philosophy of the Sciences and the Humanities,* vol. 83), Amsterdam/New York, NY: Rodopi, 2005, pp. 221-247. As a result of an unhappy mistake, the book contains the uncorrected version, which is hardly understandable at some points. For this reason the corrected version is made available here.

Published by L&PS – Logic and Philosophy of Science
http://www.units.it/~episteme/ – ISSN: 1826-1043

cause he used an impressive number of middle initials – the Flemish use them in official documents only. The letter contained some questions and suggestions; a brief correspondence ensued.

I left the field later. However, for the sake of an old friendship, I dedicate this first logic of induction to Theo.

2. Aim of this paper

It is often said that there is no logic of induction. This view is mistaken: this paper contains one. It is not a contribution to the great tradition of Carnapian inductive logic – see Kuipers (2000, Ch. 4); it is a logic of induction in the most straightforward sense of the term, a logic that, from a set of empirical data and possibly a set of background generalizations, leads to a set of consequences that comprises generalizations and their consequences. Incidentally, the underlying ideas oppose the claims that were widespread in Carnap's tradition – see, for example, Bar-Hillel (1968).

LI is characterized by a proof theory and a semantics. Some people will take these properties to be insufficient for calling **LI** a logic. I shall not quarrel about this matter, which I take to be largely conventional. As far as I am concerned, any further occurrence of 'logic' may be read as 'giclo'. The essential point is that **LI** is characterized in a formally decent way, that its metatheory may be phrased in precise terms, and, most importantly, that **LI** may serve to explicate people's actual inductive reasoning.

LI takes as premises descriptions of empirical data as well as background generalizations that are formulated in the language of standard predicative logic. Its consequences follow either deductively or inductively from the premises. By deductive consequences I mean statements that follow from the premises by Classical Logic (**CL**). The main purpose of **LI** obviously concerns the inductive consequences. In this respect the proof of the pudding will be in the eating: the reader will have to read this paper to find out whether he or she considers **LI** as sensible with respect to the intended domain of application. For now, let me just mention that the inductive consequences of a set of empirical data and a set of background knowledge will, first and foremost, be empirical generalizations, and next, the deductive consequences of the empirical generalizations and the premises, including singular statements that may serve the purposes of prediction and explanation.

LI is only one member of a family of logics. It is severely restricted by the standard predicative language. This rules out statistical generalizations as well as quantitative predicates (lengths, weights, etc.). **LI** will not take account of degrees of confirmation or the number of confirming (and disconfirming) instances. **LI** will not deal with serious problems, usually connected to discovery and creativity, such as the genesis of new concepts and other forms of conceptual dynamics. Nor will **LI** deal with the historically frequent case of inconsistent background knowledge – see Brown (1990), Norton (1987; 1993), Smith (1988), Nersessian (2002), Meheus (1993; 2002), … **LI** is a bare backbone, a starting point.

More sophisticated inductive logics may be designed by modifying **LI**. Some of the required modifications are straightforward. But given the absence of any logic of induction of the kind, it seems advisable to present a simple system that applies in specific (although common) contexts. Incidentally, I shall also keep my remarks in defense and justification of **LI** as simple as possible. As most people reading the present book will be familiar with the literature on induction, they will easily see further arguments. It also seems wise, in defending a logic of induction, to refrain from siding with one of the many parties or schools in the research on induction. The logic **LI** is intended to please most of these parties. It should serve as a point of unification: this bit at least we all agree about, even if each explains it in his or her own way.

When working on this paper I wondered why a system as simple and clarifying as **LI** had not been presented a long time ago.[1] However, although **LI** is simple and straightforward to understand, its formulation presupposes familiarity with the adaptive logic programme. I shall not summarize this programme here because several easy introductions to its purpose and range are available, such as Batens (2000; 2004). Rather, I shall introduce the required adaptive elements as the paper proceeds. However, it is only fair to the reader to mention that the ideas underlying adaptive logics and dynamic proof theories have some pedigree and are by no means the outcome of the present research.

3. Tinkering with the dynamic proof theory

Children have a profound tendency to generalization. This tendency has a clear survival value. In a sense, our present scientific (and other) knowledge is the result of a sophistication of this tendency. Of course, we know today that all

[1] In the form of a formal logic, that is. Mill's canons come close. There are also clear connections with Reichenbach's straight rule, if restricted to general hypotheses, and with Popper's conjectures and refutations. Articulating the formal logic is worthwhile, as we shall see.

simple empirical generalizations are false – compare Popper (1973, p. 10). This insight is a result of experience, of systematization, of free inquiry, and of systematic research. Our present knowledge, however, is neither the result of an urge that is qualitatively different from children's tendency to systematization, nor the outcome of a form of reasoning that is qualitatively different from theirs.

Let us for a moment consider the case in which only a set of empirical data is available – I shall remove this utterly unrealistic supposition in the present section. Where these empirical data are our only premises, what shall we want to derive from them? Apart from the **CL**-consequences of the premises, we shall also want to introduce some general hypotheses. Only by doing so may we hope to get a grasp of the world – to understand the world and to act in it. And from our premises and hypotheses together we shall want to derive **CL**-consequences (to test the hypotheses, to predict facts, and to explain facts).

LI should define a consequence relation that connects the premises with their **CL**-consequences, with the generalizations, and with their common **CL**-consequences. Is there such a consequence relation? Of course there is. The consequence relation is obviously non-monotonic[2] – inductive reasoning is the oldest and most familiar form of non-monotonic reasoning.

Generalizations that are inductively derived from the set of premises, Γ, should be compatible with Γ. A further requirement on inductively derived statements is that they should be *jointly* compatible with Γ. The latter requirement is the harder one. The logic of compatibility – see Batens and Meheus (2000) – provides us with the set of all statements that are compatible with Γ. The problem of induction is, in its simplest guise, to narrow down this set in such a way that the second requirement is fulfilled. And yet, as I shall now explain, this problem is easy to solve.

Consider an (extremely simple) example of a **CL**-proof of the usual kind – for the time being, just disregard the ∅s at the end of the lines. As stated before, all premises will be singular statements.

1	$(Pa \wedge Pb) \wedge Pc$	PREM	∅
2	$Rb \vee \sim Qb$	PREM	∅
3	$Rb \supset \sim Pb$	PREM	∅

[2] A consequence relation '⊢' is non-monotonic iff a consequence of a set of premises need not be a consequence of an extension of this set. Formally: there is a Γ, a Δ, and an *A* such that $\Gamma \vdash A$ and $\Gamma \cup \Delta \nvdash A$.

4	$(Sa \wedge Sb) \wedge Qa$		PREM	$\emptyset$
5	Pa	1	RU	$\emptyset$
6	Pb	1	RU	$\emptyset$
7	Qa	4	RU	$\emptyset$
8	Sa	4	RU	$\emptyset$
9	Sb	4	RU	$\emptyset$

The rule applied in lines 5-9 is called **RU**. This name refers to the generic "unconditional rule". For the moment, just read it as: formula 5 is **CL**-derivable from formula 1, etc.

Suppose that our data comprise 1-4, and that we want to introduce an empirical generalization, for example $(\forall x)(Px \supset Sx)$. Obviously, this formula is not **CL**-derivable from 1-4. However, we may want to accept it *until and unless* is has been shown to be problematic – for example, because some P are not S. In other words, we may want to consider $(\forall x)(Px \supset Sx)$ as conditionally true in view of the premises. By a similar reasoning, we may want to consider $(\forall x)(Px \supset Qx)$ as conditionally true. This suggests that we add these universally quantified formulas to our proof, but attach a condition to them, indicating that the formulas will not be considered as derived if the condition shows false. So we extend the previous proof as follows:[3]

10	$(\forall x)(Px \supset Sx)$	RC	$\{(\forall x)(Px \supset Sx)\}$
$11^{L_{14}}$	$(\forall x)(Px \supset Qx)$	RC	$\{(\forall x)(Px \supset Qx)\}$

The set $\{(\forall x)(Px \supset Sx)\}$ will be called the *condition* of line 10. If some member of this set is contradicted by the data, the formula derived at line 10, which happens to be $(\forall x)(Px \supset Sx)$, should be withdrawn (considered as not derived). Conditionally derived formulas may obviously be combined by **RU**. As expected, the condition of the derived formula is the union of the conditions of the formulas from which it is derived. Here is an example:

$12^{L_{14}}$	$(\forall x)(Px \supset (Qx \wedge Sx))$	10, 11 RU	$\{(\forall x)(Px \supset Sx), (\forall x)(Px \supset Qx)\}$

The interpretation of the condition of line 12 is obviously that $(\forall x)(Px \supset (Qx \wedge Sx))$ should be considered as not derived if either $(\forall x)(Px \supset Sx)$ or $(\forall x)(Px \supset Qx)$ turns out to be problematic.

[3] The superscript L_{14} on line 11 is explained below.

Logicians not familiar with dynamic proofs will complain that the negation of 11 *is* derivable from 1-4. Let me first show them to be right:

13	$\sim Qb$	2, 3, 6	RU	$\emptyset$
14	$\sim(\forall x)(Px \supset Qx)$	6, 13	RU	$\emptyset$

As $(\forall x)(Px \supset Qx)$ is shown to be contradicted by the data, lines 11 and 12, which rely on the presupposition that $(\forall x)(Px \supset Qx)$ is not problematic, have to be *marked.* Formulas that occur in marked lines are considered as not being inductively derivable from the premises.[4]

Some logicians may still complain: 14 is **CL**-derivable from 1-4, and hence, they might reason, it was simply a mistake to add lines 11 and 12 to the proof. Here I strongly disagree. Moreover, the point touches an essential property of dynamic proofs; so let me explain the matter carefully.

Suppose that Γ is a finite set. In view of the restrictions on generalizations and on Γ, it is decidable whether a generalization (in the sense specified below) is or is not derivable, and hence it is decidable whether some singular statement is or is not derivable. So, indeed, one may avoid applications of RC that are later marked (if Γ is finite). However, nearly any variant of **LI** that overcomes some of the restrictions on **LI** – see earlier as well as subsequent sections – will be undecidable and, moreover, will lack a positive test for derivability.[5]

In view of this, and in preparation for those more fascinating variants, it seems rather pointless to try circumventing a dynamic proof theory for **LI**. There is a second argument and it should not be taken lightly. It is the purpose of the present paper to explicate actual inductive reasoning. Quite obviously, humans are unable to see all the relevant consequences of the available information. Given our finite brains it would be a bad policy to make inductive hy-

[4] When a line is marked I shall sometimes say that the formula that is its second element is marked. We shall see later that there are two kinds of marks, L and B. At stage 14, lines 11 and 12 have to be L-marked. Normally, one would just add an L to those lines. In order to avoid repeating the proof at each stage, I add L_{14} to indicate that the lines are L-marked at stage 14 of the proof.

[5] A logic is decidable iff there is an algorithm to find out, for any finite set of premises and for any formula, whether the formula is derivable from the premises. There is a positive test for derivability iff there is an algorithm that leads, for any recursive set of premises and for any formula, to the answer 'Yes' in case the formula is derivable from the premises. **CL**-derivability is decidable for the propositional fragment of **CL** and undecidable for full **CL**. Nevertheless, there is a positive test for **CL**-derivability. A standard reference for such matters is Boolos and Jeffrey (1989).

potheses contingent on complete deductive certainty. To do so would slow down our thinking and often paralyse it. This does not mean that we neglect deductive logic. It only means that we often base decisions on incomplete knowledge, including incomplete deductive knowledge – see Batens (1995) for a formal approach to the analysis of deductive information. The third (and last) argument is of a different nature. I shall show in this paper that the dynamic proof theory of **LI** is formally sound and leads, within the bounds described in Section 2, to the desired conclusions. All this seems to offer a good reason to continue our journey.

To find out whether the sketched proof procedure holds water, we should also have a look at its weirder applications. Let us consider a predicate that does not occur in our premises, and see what happens to generalizations in which it occurs.

$15^{L_{17}}$	$(\forall x)(Px \supset Tx)$		RC	$\{(\forall x)(Px \supset Tx)\}$

Obviously, 1-4 do not enable one to contradict $(\forall x)(Px \supset Tx)$. However, we may moreover add:

$16^{L_{17}}$	$(\forall x)(Px \supset {\sim}Tx)$		RC	$\{(\forall x)(Px \supset {\sim}Tx)\}$

And now we see that we are in trouble, as the proof may obviously be continued as follows:

17	${\sim}(\forall x)(Px \supset Tx) \vee {\sim}(\forall x)(Px \supset {\sim}Tx)$	5	RU	$\varnothing$

Although neither 15 nor 16 is contradicted by the empirical data, their conjunction is. The thing to do here is obvious (and well known from the Reliability strategy of adaptive logics). As 15 and 16 are on a par, both of them should be considered as *unreliable*, and hence lines 15 and 16 should both be marked in view of their conditions.[6]

Let me straighten this out and introduce some useful terminology. We suppose that generalizations are not problematic until and unless they are shown to be contradicted by the empirical data. So the normal case will be that a generalization is compatible with the data. In view of this, the (derivable) negation of a generalization will be called an *abnormality.* Sometimes abnormali-

[6] In comparison to the Reliability strategy, the Minimal Abnormality strategy leads to a slightly richer consequence set in some cases. I return to this point later.

ties are connected. Line 17 is a good example: the disjunction of two abnormalities is derivable, but neither of the abnormalities is. Derivable abnormalities and derivable disjunctions of abnormalities will be called *Dab*-formulas – an abnormality is itself a disjunction with one disjunct only. Where Δ is a finite set of generalizations, $Dab(\Delta)$ is a handy abbreviation for $\bigvee(\{{\sim}A \mid A \in \Delta\})$ (the disjunction of the negations of the members of Δ).

In view of the derivability of 17, both $(\forall x)(Px \supset Tx)$ and $(\forall x)(Px \supset {\sim}Tx)$ are unreliable. But the fact that $Dab(\Delta)$ is derivable does not indicate that all members of Δ are unreliable. Indeed,

$${\sim}(\forall x)(Px \supset Qx) \vee {\sim}(\forall x)(Px \supset Sx)$$

is derivable from 14, but adding this formula to the proof does not render $(\forall x)(Px \supset Sx)$ unreliable. The reason is that, even if the displayed formula were added to the proof, it would not be a minimal *Dab*-formula in view of 14 (in the sense that a formula obtained by removing one of its disjuncts has been derived). A is unreliable at some stage of a proof, iff there is a Δ such that $A \in \Delta$ and $Dab(\Delta)$ is a *minimal Dab*-formula that is unconditionally derived in the proof at that stage.[7] Here is a further illustration:

18	${\sim}(\forall x)(Px \supset Sx) \vee {\sim}(\forall x)(Px \supset {\sim}Sx)$	5	RU	$\varnothing$
19	${\sim}(\forall x)(Px \supset {\sim}Sx)$	5, 8	RU	$\varnothing$

At stage 18 of the proof, $(\forall x)(Px \supset Sx)$ is unreliable, and hence line 10 is marked. However, at stage 19, $(\forall x)(Px \supset Sx)$ is again reliable – 19 is a minimal *Dab*-formula at this stage, whereas 18 is not – and hence line 10 is unmarked.[8] This nicely illustrates both sides of the dynamics: formulas considered as derived at one stage may have to be considered as not derived at a later stage, and *vice versa.* All this may sound unfamiliar, or even weird. And yet, as we shall see in subsequent sections, everything is under control: ultimately the dynamics is bound to lead to stability, the stable situation is determined only by the premises (as the semantics illustrates), and there are heuristic means to speed up our journey towards the stable situation.

[7] The reason for considering only unconditionally derived formulas is straightforward. Indeed, from 17 one may derive ${\sim}(\forall x)(Px \supset {\sim}Tx)$ on the condition $\{(\forall x)(Px \supset Tx)\}$, but this obviously does not render $(\forall x)(Px \supset Tx)$ reliable. The disjunction 17 follows from the premises by **CL**, and neither of its disjuncts does.

[8] So line 10 is L-marked at stage 18 but not at stage 19.

Having made this precise – formal definitions follow in Section 4 – I leave it to the reader to check that the introduction of 'wild' hypotheses leads nowhere. As the predicates U and V do not occur in the premises 1-4, applying RC to add formulas such as $(\forall x)(Ux \supset Vx)$ to the proof, will lead to lines that are bound to be marked sooner or later – and very soon if some simple heuristic instructions are followed.

Before moving on to background knowledge, let me add some important comments. We have seen that $(\forall x)(Px \supset Qx)$ was not inductively derivable from 1-4. However, $(\forall x)((Px \wedge Rx) \supset Qx)$ is. Indeed, line 20 below is not marked in the present proof. In some artificial and clumsy extensions of the proof, line 20 may be marked. But it is easy enough to further extend the proof in such a way that line 20 is unmarked. This is an extremely important remark to which I return later.

20	$(\forall x)((Px \wedge Rx) \supset Qx)$	RC	$\{(\forall x)((Px \wedge Rx) \supset Qx)\}$

The next comment concerns the *form* of formulas derived by RC. All that was specified before is that these formulas should be universally quantified. However, a further restriction is required. Suppose that it is allowed to add

[21]	$(\forall x)((Qx \vee {\sim}Qx) \supset {\sim}Sc)$	RC	$\{(\forall x)((Qx \vee {\sim}Qx) \supset {\sim}Sc)\}$

to the proof. As

$${\sim}(\forall x)(Px \supset Sx) \vee {\sim}(\forall x)((Qx \vee {\sim}Qx) \supset {\sim}Sc)$$

is derivable from 1, not only line [21] but also line 10 would be marked in view of this formula. Similar troubles arise if it is allowed to introduce such hypotheses as $(\forall x)((Qx \vee {\sim}Qx) \supset (\exists x)(Px \wedge {\sim}Sx))$.

The way out of such troubles is simple enough. RC should not allow one to introduce singular statements or existentially quantified statements in disguise. Hence, we shall require that the generalizations introduced by RC consist of a sequence of universal quantifiers followed by a formula of the form $A \supset B$ in which no constants, propositional letters or quantifiers occur. From now on, '*generalization*' will refer to such formulas only.[9] Some people will raise a his-

[9] It is possible to devise a formal language in which the effect of this restriction is reduced to nil. This is immaterial if such languages are not used in the empirical sciences to which we want to apply **LI**. But indeed, the formal restriction hides one on content: all predicates should be well entrenched, and not abbreviate identity to an individual constant.

torical objection to this restriction. Kepler's laws explicitly refer to the sun, and Galileo's law of the free fall to the earth. This, however, is related to the fact that the earth, the sun, and the moon had a specific status in the Ptolemaic worldview, and were slowly losing that status in the days of Kepler and Galileo. In the Ptolemaic worldview, each of those three objects was taken, just like God, to be the only object of a specific kind. So those generalizations refer to kinds of objects, rather than to specific objects – by Newton's time, any possible doubt about this had been removed.[10]

Any *generalization* may be introduced by RC. This includes such formulas as 21 and 22, that are **CL**-equivalent to 23 and 24 respectively. So the implicative form of generalizations may be circumvented.

21	$(\forall x)((Qx \vee {\sim}Qx) \supset Px)$		RC	$\{(\forall x)((Qx \vee {\sim}Qx) \supset Px)\}$
22	$(\forall x)(Rx \supset (Qx \wedge {\sim}Qx))$		RC	$\{(\forall x)(Rx \supset (Qx \wedge {\sim}Qx))\}$
23	$(\forall x)Px$	21	RU	$\{(\forall x)((Qx \vee {\sim}Qx) \supset Px)\}$
24	$(\forall x){\sim}Rx$	22	RU	$\{(\forall x)(Rx \supset (Qx \wedge {\sim}Qx))\}$

Is the dynamics of the proofs bound to stop at some finite point? The answer to this question is not simple, but nevertheless satisfactory. However, I postpone the related discussion until we have gained a better grasp of **LI**.

Let us now move on to situations in which *background knowledge* is available. Clearly, background knowledge cannot be considered as unquestionable. For one thing, the empirical data might contradict it. If they do, we face an inconsistent set of premises, which leaves us nowhere on the present approach.[11] So we shall consider background knowledge as defeasible. It is taken for granted *unless and until* it is shown to be problematic.

This being settled, it is simple enough to integrate background knowledge in the dynamic proof format. Background knowledge is the result of inductive inferences made in the past, by ourselves our by our predecessors.[12] For this

[10] Even in the Ptolemaic era, those objects were identified in terms of well entrenched properties – properties relating to their kind, not to accidental qualities. The non-physical example is even more clear: God has *no* accidental properties.

[11] Scientists may justifiably stick to hypothetical knowledge that is falsified by the empirical data, for example because no non-falsified theory is available. Including such cases in the picture requires that we move to a paraconsistent logic. Although I have worked in this domain for some thirty years now, I fear that I would lose most of the readers if I were to open that Pandora's box. So let me stress that there is absolutely no problem in including the paraconsistent case in the logic of induction, but that I leave it out for reasons of space as well as for pedagogical reasons.

[12] Or rather, background knowledge is so interpreted for present purposes. This is a simpli-

reason, I shall restrict background knowledge to background generalizations – another simplification – and introduce them as *conditional premises.* Here is an example:

25	$(\forall x)(Qx \supset Rx)$		BK	$\{(\forall x)(Qx \supset Rx)\}$
26	Ra	7, 25	RU	$\{(\forall x)(Qx \supset Rx)\}$

The central difference between background generalizations and other generalizations – the latter will be called *local generalizations* from now on – is that the former are retained whenever possible. If $Dab(\Delta)$ is unconditionally derived, and each member of Δ is a background generalization, then, in the absence of further information, we have to consider all members of Δ as unreliable. So we shall mark all lines the condition of which overlaps with Δ. This includes the lines on which the background generalizations are introduced as conditional premises.[13]

If, however, we unconditionally derive $\sim A_1 \vee ... \vee \sim A_n \vee \sim B_1 \vee ... \vee \sim B_m$, and each A_i is a reliable background generalization (in the sense of the previous paragraph), then we should consider the local generalizations $B_1, ..., B_m$ as unreliable, and retain the background knowledge $A_1, ..., A_n$. Here is a simple example:

27^{L29}	$(\forall x)(Qx \supset \sim Rx)$		RC	$\{(\forall x)(Qx \supset \sim Rx)\}$
28^{L29}	$\sim Ra$	7, 27	RC	$\{(\forall x)(Qx \supset \sim Rx)\}$
29	$\sim(\forall x)(Qx \supset Rx) \vee \sim(\forall x)(Qx \supset \sim Rx)$	7	RU	$\varnothing$

$(\forall x)(Qx \supset Rx)$ is a background generalization and has not been shown to be an unreliable background generalization.[14] But the local generalization $(\forall x)(Qx \supset \sim Rx)$ is unreliable in view of 29. Hence, lines 27 and 28 are marked, but lines 25 and 26 are not, as desired.

In view of the asymmetry between background hypotheses and local hypotheses, **LI** is a *prioritized* adaptive logic. This means that the members of one set of defeasible formulas, the background hypotheses, are retained at the expense of the members of another set, the local generalizations.

---

fication. Humanity did not start from scratch at some point in time, not even with respect to scientific theories – see also Section 7.

[13] Here is a simple example. If $(\forall x)(Px \supset Qx)$ is a background generalization, one may introduce it by the rule BK. However, this background generalization (and anything derived from it) would be B-marked in view of line 14. See the next section for the precise definition.

[14] This agrees with the above criterion: there is no set of background generalizations Δ such that $(\forall x)(Qx \supset Rx) \in \Delta$ and $Dab(\Delta)$ is a minimal *Dab*-formula at stage 29 of the proof.

Before moving on to the precise formulation of the dynamic proof theory, let me intuitively explain some peculiarities of the proof format. Traditionally, a *proof* is seen as a list of formulas. This is not different for **LI**-proofs: the line numbers, the justification of the line (a set of line numbers and a rule), the conditions, and the marks are all introduced to make the proof more readable, but are not part of the proof itself. However, there is a central difference in this connection. In the dynamic case, one writes down a list of formulas, but the proof consists only of the *unmarked* formulas in the list. This does not make the marks part of the proof itself: which formulas are marked is determined by the empirical data, the background generalizations, and the list of formulas written down. Let us now continue to speak in terms of the annotated proof format.

What we are interested in are formulas that are *finally* derivable. On our way toward them, we have to go through the stages of a proof. Some formulas derived *at a stage* may not be finally derivable. As formulas that come with an empty condition (fifth element of the line) cannot possibly be marked at a later stage, they are sometimes called *unconditionally* derived. These formulas are deductively derived (by **CL**) from the empirical data. Formulas that have a non-empty condition are called *conditionally* derived. These formulas are inductively derived *only*. Of course, the interesting formulas are those that are inductively derived only, but nevertheless finally derived. In the present paper I offer a correct definition of final derivability, but cannot study the criteria that are useful from a computational point of view.

A last comment concerns the rules of inference. The *unconditional rules* of **LI** are those of Classical Logic, and they carry the conditions from their premises to their conclusion. The *conditional rules* BK and RC add a *new* element to the condition, and hence start off the dynamics of the proofs. As far as their structure is concerned, however, they are of the same type as the standard premise and axiom rules.

4. The dynamic proof theory

Our language will be that of predicative logic. Let $\forall A$ denote A preceded by a universal quantifier over any variable free in A. A *generalization* is a formula of the form $\forall(A \supset B)$ in which no individual constant, sentential letter or quantifier occurs in either A or B.

A dynamic proof theory consists of (i) a set of unconditional rules, (ii) a set of conditional rules, and (iii) a definition of *marked* lines. The rules allow one to add lines to a proof. Formulas derived on a line that is marked at a stage of the proof are considered as not inductively derived at that stage (from the premises and background generalizations).[15]

Lines in an annotated dynamic proof have five elements: (i) a line number, (ii) the formula derived, (iii) a set of line numbers (of the lines from which the formula is derived), (iv) a rule (by which the formula is derived), and (v) a set of conditions.

The logic **LI** operates on ordered sets of premises, $\Sigma = \langle \Gamma, \Gamma^* \rangle$, in which Γ is a set of *singular formulas* (the empirical data) and Γ^* is a set of *generalizations* (the background generalizations).

The rules of **LI** will be presented here in generic format. There are two unconditional rules, PREM and RU, and two conditional rules, BK and RC:

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) $\varnothing$.

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) $\{A\}$.

RU If $A_1, ..., A_n \vdash_{\mathbf{CL}} B$ and each of $A_1, ..., A_n$ occur in the proof on lines $i_1, ..., i_n$ that have conditions $\Delta_1, ..., \Delta_n$ respectively, one may add a line comprising the following elements: (i) an appropriate line number, (ii) B, (iii) $\mathrm{i}_1, ..., i_n$, (iv) RU, and (v) $\Delta_1 \cup ... \cup \Delta_n$.

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) $\{A\}$.

A proof constructed by these rules will be called an **LI**-proof from Σ. In such a proof, a formula is *unconditionally derived* iff it is derived at a line of which the fifth element is empty. It is conditionally derived otherwise.

An *abnormality* is the negation of a generalization. *Dab*-formulas are formulas of the form $Dab(\Delta) = \bigvee\{{\sim}A \mid A \in \Delta\}$, in which Δ is a finite set of generalizations.[16] $Dab(\Delta)$ is a *minimal Dab*-formula at stage s of a proof iff $Dab(\Delta)$

[15] That background generalizations may be B-marked themselves illustrates that they are defeasible premises.

[16] Note that $Dab(\Delta)$ refers to any formula that belongs to an equivalence class that is closed under permutation and association.

is unconditionally derived in the proof at stage s and there is no $\Delta' \subset \Delta$ such that $Dab(\Delta')$ is unconditionally derived in the proof at that stage.

DEFINITION 1

Where $Dab(\Delta_1)$, ..., $Dab(\Delta_n)$ *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. As lines that depend on unreliable background generalizations are B-marked, these generalizations are themselves removed from the proof. This is interpreted by not considering them as part of the background knowledge at that stage of the proof. What remains of the background knowledge at stage s will be denoted by $\Gamma^*_s = \Gamma^* - U^*_s(\Gamma)$.

Now we come to an important point. In order to determine which local generalizations are unreliable, we have to take the reliable background knowledge for granted. A *Dab*-formula *Dab* 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 $Dab(\Delta_1)$, ..., $Dab(\Delta_n)$ *are the minimal local Dab-formulas at stage s of a proof from* $\Sigma = \langle \Gamma, \Gamma^* \rangle$, $U^o_s(\Gamma) = \bigcup\{\Delta_{\mathrm{i}} - \Gamma^*_s \mid 1 \leq i \leq n\}$.

DEFINITION 4

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

$U^o_s(\Gamma)$ comprises the unreliable local generalizations at stage s. These generalizations may have been introduced by RC, they may be unreliable background generalizations, or they may be generalizations that do not occur in the proof (or occur as derived formulas only). Let me briefly clarify Definition 3.

Given the B-marks, we have to assess the hypotheses introduced by RC. Which of these are unreliable at stage s of the proof? The key to the answer to this question lies in the following theorem, the proof of which is obvious:

THEOREM 1

$Dab(\Delta \cup \Delta')$ is a minimal Dab-formula at stage s of a proof, iff a line may be added that has $Dab(\Delta)$ as its second, RC as its fourth, and Δ' as its fifth element.

Suppose that, in a proof at stage *s*, Δ' contains only reliable background generalizations, whereas no such background generalization is a member of Δ – that is, $\Delta' \subseteq \Gamma^*_s$ and $\Delta \cap \Gamma^*_s = \varnothing$. That $Dab(\Delta)$ is derivable on the condition Δ' indicates that some member of Δ is unreliable *if* the background generalizations in Δ' are reliable. Moreover, we consider the background generalizations to be more trustworthy than the local generalizations. So from the occurrence of the minimal local *Dab*-consequence $Dab(\Delta \cup \Delta')$ we should conclude that the members of Δ are unreliable.

Incidentally, an equivalent (and also very intuitive) proof theory is obtained by defining $U^o_s(\Gamma)$ in a different way. Let $Dab(\Delta_1)$, ..., $Dab(\Delta_n)$ be the minimal (in the usual, simple sense) *Dab*-formulas that have been derived at stage *s* on the conditions Θ_1, ..., Θ_n respectively, and for which $(\Delta_1 \cup ... \cup \Delta_n) \cap \Gamma^*_s = \varnothing$ and $\Theta_1 \cup ... \cup \Theta_n \subseteq \Gamma^*_s$. $U^o_s(\Gamma)$ may then be defined as $\Delta_1 \cup ... \cup \Delta_n$.[17] But let us stick to Definition 3 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

$\Sigma \vdash_{\mathbf{LI}} A$ *(A is finally* **LI***-derivable from Σ) iff A is derived at a stage s of a proof from Σ, say at line i, and, whenever line i is marked in an extension of the proof, it is unmarked in a further extension of the proof.*

This definition is the same as for other dynamic proof theories. The following theorem is helpful to get a grasp of **LI**-proofs. The formulation is somewhat clumsy because the line may be marked, in which case *A* cannot be said to be derivable.

[17] This alternative definition need not lead to the same results with respect to a specific proof at a stage, but it determines the same set of finally derivable consequences (see below in the text) in view of Theorem 1.

THEOREM 2

To an **LI**-*proof from* $\Sigma = \langle\Gamma, \Gamma^*\rangle$ *a (marked or unmarked) line may be added that has A as its second element and* Δ *as its fifth element, iff* $\Gamma \vdash_{\mathbf{CL}} A \vee Dab(\Delta)$.

The proof of the theorem is extremely simple. Let the **CL**-transform of an **LI**-proof from $\Sigma = \langle\Gamma, \Gamma^*\rangle$ be obtained by replacing any line that has B as its second and Θ as its fifth element, by an unconditional line that has $B \vee Dab(\Theta)$ as its second element. To see that this **CL**-transform is a **CL**-proof from Γ, it is sufficient to note the following: (i) the **CL**-transform of applications of PREM are justified by PREM, (ii) the **CL**-transform of applications of BK and RC are justified in that they contain a **CL**-theorem of the form $A \vee \sim A$, and (iii) the **CL**-transform of applications of RU are turned into applications of the **CL**-derivable (generic) rule "If $A_1, ..., A_n \vdash_{\mathbf{CL}} B$, then from $A_1 \vee C_1, ..., A_n \vee C_n$ to derive $B \vee C_1 \vee ... \vee C_n$". This establishes one direction of the theorem. The proof of the other direction is immediate in view of the **LI**-derivable rule: "Where all members of Δ are generalizations, to derive A on the condition Δ from $A \vee Dab(\Delta)$".

So, in a sense, **LI**-proofs are **CL**-proofs in disguise. We interpret them in a specific way in order to decide which generalizations should be selected.

In order to obtain a better grasp of final derivability, I first define the sets of unreliable formulas with respect to Γ, independently of the stage of a proof. First we need: $Dab(\Delta)$ is a *minimal Dab-consequence* of Γ iff $\Gamma \vdash_{\mathbf{CL}} Dab(\Delta)$ and, for no $\Delta' \subset \Delta$, $\Gamma \vdash_{\mathbf{CL}} Dab(\Delta')$.

DEFINITION 7

Where Ω^* *is the set of all minimal Dab-consequences of* Γ *in which occur only members of* Γ^*, $U^*(\Gamma) = \bigcup(\Omega^*)$.

This defines the set of background generalizations that are unreliable with respect to the empirical data Γ. The set of *retained* background generalizations is $\Gamma^*_\Gamma = \Gamma^* - U^*(\Gamma)$.

$Dab(\Delta)$ is a *minimal local Dab-consequence* of Γ iff $\Gamma \vdash_{\mathbf{CL}} Dab(\Delta)$ and, for no Δ', $\Gamma \vdash_{\mathbf{CL}} Dab(\Delta')$ and $(\Delta' - \Gamma^*_\Gamma) \subset (\Delta - \Gamma^*_\Gamma)$.[18]

DEFINITION 8

Where Ω *is the set of all minimal local Dab-consequences of* Γ, $U^\circ(\Gamma) = \bigcup(\Omega) - \Gamma^*_\Gamma$.

[18] As $\Gamma \vdash_{\mathbf{CL}} Dab(\Delta)$, $\Delta \neq \varnothing$ and, in view of Definition 7, $\Delta' - \Gamma^*_\Gamma \neq \varnothing$; similarly for Δ'.

This defines the set of local generalizations that are unreliable with respect to the empirical data Γ.

Given that **LI**-proofs are **CL**-proofs in disguise, the proofs of the following theorems can safely be left to the reader:

THEOREM 3

Where $\Sigma = \langle \Gamma, \Gamma^* \rangle$, $\Gamma^*_\Gamma = \{A \in \Gamma^* \mid \Sigma \vdash_{\mathbf{LI}} A\}$.

THEOREM 4

Where $\Sigma = \langle \Gamma, \Gamma^* \rangle$, $\Sigma \vdash_{\mathbf{LI}} A$, *A is finally* **LI**-*derivable from* Σ, *iff there is a* (*possibly empty*) Δ *such that* (*i*) $\Gamma \cup \Gamma^*_\Gamma \vdash_{\mathbf{CL}} A \vee Dab(\Delta)$, *and* (*ii*) $(\Delta - \Gamma^*_\Gamma) \cap U^\circ(\Gamma) = \varnothing$.

This sounds much simpler in words. A is an **LI**-consequence of Σ iff A is **CL**-derivable from Γ together with the reliable background generalizations, or, for some set Δ of reliable local generalizations,[19] $A \vee Dab(\Delta)$ is **CL**-derivable from Γ together with the reliable background generalizations.

The **LI**-consequence relation may be characterized in terms of compatibility – where is compatible with Δ' iff $\Delta \cup \Delta'$ is consistent (iff no inconsistency is **CL**-derivable from this set).[20] The characterization is remarkably simple, as appears from the following three theorems. The proof of the theorems is obvious in view of Definition 7 and Theorem 4.

THEOREM 5

$A \in \Gamma^*_\Gamma$ *iff* $A \in \Gamma^*$ *and* $\Delta \cup \{A\}$ *is compatible with* Γ *whenever* $\Delta \subseteq \Gamma^*$ *is compatible with* Γ.

A background generalization A is retained iff, whenever a set of background generalizations is compatible with the data, then A and Δ are jointly compatible with the data.

THEOREM 6

Where $\Sigma = \langle \Gamma, \Gamma^* \rangle$ *and A is a generalization*, $\Sigma \vdash_{\mathbf{LI}} A$ *iff* $\Delta \cup \{A\}$ *is compatible with* $\Gamma \cup \Gamma^*_\Gamma$, *whenever a set of generalizations* Δ *is compatible with* $\Gamma \cup \Gamma^*_\Gamma$.

[19] Remark that $\Gamma \cup \Gamma^*_\Gamma \vdash_{\mathbf{CL}} A \vee Dab(\Delta)$ iff $\Gamma \cup \Gamma^*_\Gamma \vdash_{\mathbf{CL}} A \vee Dab(\Delta - \Gamma^*_\Gamma)$.

[20] This definition presupposes that nothing is compatible with an inconsistent set – see also Batens and Meheus (2000), for an alternative.

A generalization A is inductively derivable iff, whenever a set Δ of generalizations is compatible with the data and retained background generalizations, then A and Δ are jointly compatible with the data and retained background generalizations. Let Σ^G be the set of generalizations that are inductively derivable from Σ.

THEOREM 7
Where $\Sigma = \langle \Gamma, \Gamma^* \rangle$, $\Sigma \vdash_{\mathbf{LI}} A$ *iff* $\Gamma \cup \Gamma^*_\Gamma \cup \Sigma^G \vdash_{\mathbf{CL}} A$.

A is inductively derivable from a set of data and background generalizations iff it is **CL**-derivable from the data, the reliable background generalizations, and the inductively derivable generalizations.

Let me finally mention, without proofs, some properties of the **LI**-consequence relation: Non-Monotonicity, Proof Invariance (any two proofs from Γ define the same set of final consequences), **CL**-Closure ($Cn_{\mathbf{CL}}(Cn_{\mathbf{I}}(\Sigma)) = Cn_{\mathbf{I}}(\Sigma)$),[21] Decidability of $\langle \Gamma, \Gamma^* \rangle \vdash_{\mathbf{LI}} A$ whenever Γ and Γ^* are finite and A is either a generalization or a singular formula. Cautious cut with respect to facts: where A is a singular statement, if $\langle \Gamma, \Gamma^* \rangle \vdash_{\mathbf{LI}} A$ and $\langle \Gamma \cup \{A\}, \Gamma^* \rangle \vdash_{\mathbf{LI}} B$, then $\langle \Gamma, \Gamma^* \rangle \vdash_{\mathbf{LI}} B$. Cautious monotonicity with respect to facts: where A is a singular statement, if $\langle \Gamma, \Gamma^* \rangle \vdash_{\mathbf{LI}} A$ and $\langle \Gamma, \Gamma^* \rangle \vdash_{\mathbf{LI}} B$, then $\langle \Gamma \cup \{A\}, \Gamma^* \rangle \vdash_{\mathbf{LI}} B$. By the last two: that inductively derivable predictions are verified, does not lead to new inductive consequences. Cautious cut with respect to generalizations: where A is a generalization, if $\langle \Gamma, \Gamma^* \rangle \vdash_{\mathbf{LI}} A$ and $\langle \Gamma, \Gamma^* \cup \{A\} \rangle \vdash_{\mathbf{LI}} B$, then $\langle \Gamma, \Gamma^* \rangle \vdash_{\mathbf{LI}} B$. Cautious monotonicity with respect to generalizations: where A is a generalization, if $\langle \Gamma, \Gamma^* \rangle \vdash_{\mathbf{LI}} A$ and $\langle \Gamma, \Gamma^* \rangle \vdash_{\mathbf{LI}} B$, then $\langle \Gamma, \Gamma^* \cup \{A\} \rangle \vdash_{\mathbf{LI}} B$. By the last two: if inductively derivable generalizations are accepted as background knowledge, no new inductive consequences follow.

5. The semantics

The previous sections merely considered the dynamic proof theory of **LI**. This proof theory is extremely important, as it enables us to explicate actual inductive reasoning – humans reach conclusions by finite sequences of steps. A logical semantics serves different purposes. Among other things, it provides insights into the conceptual machinery. Such insights increase our understanding of a logic, even if they are not directly relevant for the computational aspects.

Let $\mathcal{M}_\Gamma$ denote the set of **CL**-models of Γ. The **LI**-models of $\Sigma = \langle \Gamma, \Gamma^* \rangle$,

[21] $Cn_{\mathbf{L}}(\Gamma) = \{A \mid \Gamma \vdash_{\mathbf{L}} A\}$ as usual.

will be a subset of $\mathcal{M}_\Gamma$. This subset is defined in terms of the abnormal parts of models – see Batens (1986) for the first application of this idea (to a completely different kind of logic). The abnormal part of a model (the set of abnormalities verified by a model) is defined as follows. Let $\mathcal{G}$ denote the set of generalizations.

DEFINITION 9
$Ab(M) = \{\forall(A \supset B) \mid M \nvDash \forall(A \supset B); \forall(A \supset B) \in \mathcal{G}\}$.

In words: the abnormal part of a model is the set of generalizations it falsifies. Obviously, $Ab(M)$ is not empty for any model M. For example, either $(\forall x)((Px \vee {\sim}Px) \supset Qx) \in Ab(M)$ or $(\forall x)((Px \vee {\sim}Px) \supset {\sim}Qx) \in Ab(M)$. And if $M \vDash Pa$, then either $(\forall x)(Px \supset Qx) \in Ab(M)$ or $(\forall x)(Px \supset {\sim}Qx) \in Ab(M)$. However, in some models of Pa both $(\forall x)(Px \supset Qx)$ and $(\forall x)(Px \supset {\sim}Qx)$ belong to $Ab(M)$, whereas in others only one of them does.

Given that **CL** is sound and complete with respect to its semantics, $Dab(\Delta)$ is a *minimal Dab-consequence* of Γ iff all $M \in \mathcal{M}_\Gamma$ verify $Dab(\Delta)$ and no $\Delta' \subset \Delta$ is such that all $M \in \mathcal{M}_\Gamma$ verify $Dab(\Delta')$.

This semantic characterization of the minimal *Dab*-consequences of Γ immediately provides a semantic characterization of $U^*(\Gamma)$, of $U^\circ(\Gamma)$, and of Γ^*_Γ. This is sufficient to make the first required selection. The proof of Theorem 8 is obvious.

DEFINITION 10
$M \in \mathcal{M}_\Gamma$ *is* background-reliable *iff* $(Ab(M) \cap \Gamma^*) \subseteq U^*(\Gamma)$.

THEOREM 8
$M \in \mathcal{M}_\Gamma$ is background-reliable iff $M \vDash \Gamma^*_\Gamma$.

In words, the retained background knowledge consists of the members of Γ^* that are verified by all background-reliable models of Γ. So a model M of Γ is background-reliable iff it verifies all reliable background generalizations. For any consistent Γ and for any set of background generalizations Γ^*, there are background-reliable models of Γ.[22] This is warranted by the compactness of **CL**: Γ is compatible with Γ^*_Γ iff it is compatible with any finite subset of Γ^*_Γ.

I now proceed to the second selection of models of Γ.

[22] In the worst case, all background generalizations are unreliable, and hence all models of Γ are background-reliable.

DEFINITION 11

$M \in \mathcal{M}_\Gamma$ *is* reliable (*is an* **LI**-*model of* Σ)[23] *iff* $Ab(M) \subseteq U^{\circ}(\Gamma)$.

Since, in view of Definitions 7 and 8, $U^*(\Gamma) = U^{\circ}(\Gamma) \cap \Gamma^*$, it follows that:

THEOREM 9

All reliable models of Σ *are background reliable.*

One should not be misled by this. $Ab(M) \subseteq U^{\circ}(\Gamma)$ only warrants $(Ab(M) \cap \Gamma^*) \subseteq U^*(\Gamma)$ because the definition of $U^{\circ}(\Gamma)$ refers to the definition of $U^*(\Gamma)$.

DEFINITION 12

Where $\Sigma = \langle \Gamma, \Gamma^* \rangle$, $\Sigma \vDash_{\mathbf{LI}} A$ *iff all reliable models of* Γ *verify* A.

THEOREM 10

$\Sigma \vdash_{\mathbf{LI}} A$ *iff* $\Sigma \vDash_{\mathbf{LI}} A$. (*Soundness and Completeness*)

The proof is longwinded, especially its right-left direction, but follows exactly the reasoning of the proofs of Theorems 5.1 and 5.2 from Batens (1999). The present proof is simpler, however, as it concerns **CL**.

Some further provable properties: Strong Reassurance (if a **CL**-model M of Γ is not an **LI**-model of Σ, then some **LI**-model M' of Σ is such that $Ab(M') \subset Ab(M)$), and Determinism of final derivability (the co-extensive semantic consequence relation defines a unique consequence set for any Σ).

Although it is important to semantically characterize final **LI**-derivability in terms of a set of models of Σ, some might complain that the dynamics of the proofs does not appear in the semantics. However, there is a simple method to obtain a dynamic semantics for adaptive logics. This method, exemplified in Batens (1995), offers a dynamic semantics that is characteristic for derivability at a stage.

A slightly different (and richer) result would be obtained by applying the Minimal Abnormality strategy. I skip technicalities and merely mention the central difference from the Reliability strategy. In the presence of an instance[24] of Px and in the absence of instances of both $Px \wedge Qx$ and $Px \wedge {\sim}Qx$, the Reliability strategy leads to the rejection of both $(\forall x)(Px \supset Qx)$ and $(\forall x)(Px \supset {\sim}Qx)$ – if any

[23] As **LI** is an adaptive logic, it does not make sense to say that M is or is not an **LI**-model, but only that M is or is not an **LI**-model of some Σ.

[24] An instance of the open formula A, is any closed formula obtained by replacing each variable free in A by some constant.

of these formulas occurs in the fifth element of a line, the line is marked. It follows that even the disjunction of both generalizations will be marked. On the Minimal Abnormality strategy, both generalizations are marked but their disjunction is not. This supplementary consequence seems weak and pointless. Moreover, the Minimal Abnormality strategy, while leading to a very simple semantics, terribly complicates the proof theory. For this reason I shall not discuss it further here.

6. Heuristic matters and further comments

Some people think that all adaptive reasoning (including all non-monotonic reasoning) should be explicated in terms of heuristic moves rather than in terms of logic proper. For their instruction and confusion, I shall first spell out some basics of the heuristics of the adaptive logic **LI**. I leave it to the reader to compare both conceptions.

Suppose that one applies RC to introduce, on line i, the generalization $\forall(A \supset B)$ on the condition $\{\forall(A \supset B)\}$. As (1) is a **CL**-theorem, it may be derived in the proof and causes $\forall(A \supset B)$ to be L-marked.

$$\sim\forall(A \supset B) \vee \sim\forall(A \supset \sim B) \vee \sim\forall(\sim A \supset B) \vee \sim\forall(\sim A \supset \sim B) \quad (1)$$

So, in order to prevent $\forall(A \supset B)$ from being L-marked, one needs to unconditionally derive

$$\sim\forall(A \supset \sim B) \vee \sim\forall(\sim A \supset B) \vee \sim\forall(\sim A \supset \sim B)$$

or a "sub-disjunction" of it. How does one do so? An instance of A enables one to derive

$$\sim\forall(A \supset B) \vee \sim\forall(A \supset \sim B) \quad (2)$$

whereas an instance of $\sim A$ enables one to derive

$$\sim\forall(\sim A \supset B) \vee \sim\forall(\sim A \supset \sim B) \quad (3)$$

An instance of $A \wedge B$ enables one to derive

$$\sim\forall(A \supset \sim B) \quad (4)$$

and so on.

In view of this, it is obvious how one should proceed. Suppose that one is interested in the relation between A and B. It does not make sense to introduce by RC, for example, the generalization $\forall(A \supset B)$, if falsifying instances (instances of $A \wedge {\sim}B$) are derivable – if there are, the generalization is marked and will remain marked forever. Moreover, in order to prevent $\forall(A \supset B)$ from becoming marked in view of (1) or (2), one needs a confirming[25] instance (an instance of $A \wedge B$) and one needs to derive (4) from it. So two aims have to be pursued: (i) search for instances of $A \wedge {\sim}B$ in order to make sure that one did not introduce a falsified generalization, and (ii) search for instances of $A \wedge B$ in order to make sure that the generalization is not marked.

To see that the matter is not circular, note that it does not make sense, with respect to (ii) from the previous paragraph, to derive, say $B(a)$ from $A(a)$ together with the generalization $\forall(A \supset B)$ itself. Indeed, $B(a)$ will then be derived on the condition $\{\forall(A \supset B)\}$. (4) is derivable from $B(a)$, but again only on the condition $\{\forall(A \supset B)\}$. The only *Dab*-formula that can be unconditionally derived from (4) on the condition $\{\forall(A \supset B)\}$ is (2) – compare Theorem 2. In view of this, the line at which $\forall(A \supset B)$ was introduced by RC will still be marked.

But suppose that $A(a)$ and $C(a)$ occur unconditionally in the proof and that the generalization $\forall(C \supset B)$ was introduced by RC. If we derive $B(a)$ from these, it will be derived on the condition $\{\forall(C \supset B)\}$. So we are not able to unconditionally derive (4) from $A(a)$ and $B(a)$. All we can unconditionally derive along this road is

$$\sim\forall(A \supset {\sim}B) \vee {\sim}\forall(C \supset B) \tag{5}$$

and, in view of this, both $\forall(C \supset B)$ and $B(a)$ will be marked.

The reader might find this weird. There may be unconditional instances of $C \wedge B$ in the proof, and hence ${\sim}\forall(C \supset {\sim}B)$ may be unconditionally derived. This seems to warrant that $\forall(C \supset B)$ is finally derived, but obviously it does not. If such unexpected dependencies between abnormalities obtain, are we not losing control? Nothing very complicated is actually going on here. Control is provided by the following simple and intuitive fact:

(†) If the introduction of a local generalization G entails a falsifying instance of another generalization $\forall(A \supset B)$, and no falsifying instance of the latter is derivable from the empirical data together with the re-

[25] Obviously, 'confirming' is meant here in the qualitative sense – see Kuipers (2000, Ch. 2).

liable background knowledge, then $\sim G \vee \sim\forall(A \supset \sim B)$ is unconditionally derivable.

What does all this teach us? If we introduce a generalization, we want to find out whether it is finally derived in view of the present data. In order to do so, we should look for falsifying as well as for confirming instances, and we should look for falsifying instances of other generalizations, as specified in (†).[26] These instances may be derived from the union of the empirical data, the reliable background generalizations, and the reliable local generalizations. There is a clear bootstrapping effect here. At the level of the local generalizations the effect is weak, in that wild generalizations will not be finally derivable. At the level of the background generalizations, the effect is very strong – it is only annihilated by falsifying instances. However, at the level of the local generalizations, the bootstrapping effect does *not* reduce to a form of circularity.

So in order to speed up our journey towards the stable situation we need to look for the instances mentioned in the previous paragraph. As this statement may easily be misunderstood let me clarify it. Let the generalization introduced by RC be $\forall(A \supset B)$. (i) We need to find a confirming instance – if there are none, the generalization is bound to be marked.[27] (ii) We need to search for falsifying instances of the generalization and for falsifying instances of other generalizations that are novel with respect to the empirical data and reliable background generalizations – if there are falsifying instances of either kind, the generalization is bound to be marked. As a result of the search for falsifying instances (of either kind), we may find more confirming instances as well as a number of undetermined cases – individual constants for which there is an instance of A but not of either B or $\sim B$. When new empirical data become available, objects about which we had no information, or only partial information, may turn out to be falsifying, and so may objects about which we can only derive *conditionally* that they are confirming. So, (iii) we need to collect further data, by observation and experiment. At this point, confirmation theory enters the picture. Although **LI** does not take into account the number of confirming instances, only well-established hypotheses will convincingly eliminate potential falsifiers. Incidentally, I tend to side with Popper in this re-

[26] It is unlikely that effects like the one described by (†) will be discovered if one does not handle induction in terms of a logic. I have never seen such effects mentioned in the literature on induction, and they certainly are not mentioned in Kuipers (2000).

[27] This should be qualified. If there are instances of $\sim A$ and none of A, then $\forall(A \supset B)$ may be derivable and unmarked because $\forall\sim A$ is so.

spect: what is important is not the number of confirming instances, but rather the strength of the tests to which a generalization has been subjected. Whether this concept may be explicated within the present qualitative framework is dubious.

Although the heuristics of **LI** depends on confirmation theory in the sense described above, **LI** in itself enables us to spell out quite interesting heuristic maxims. Given a set of empirical data and a set of background generalizations, it is clear how we should proceed. Most of what was said above relates to that. If the given data and background knowledge do not allow one to finally derive any generalization concerning the relation between A and B because there is insufficient information, **LI** clearly instructs one about the kind of data that should be gathered to change the situation. In this sense, **LI** *does* guide empirical research. This guidance may be considered somewhat unsophisticated, but it is the basic guidance, the one that points out the most urgent empirical research.

I now turn to a different kind of heuristic maxims. In order to speed up our journey towards stability with respect to *given* empirical data and background generalizations, it is essential to derive as soon as possible the minimal *Dab*-consequences of Γ and to derive as soon as possible the minimal local *Dab-consequences* of Γ. Some **LI**-derivable rules are extremely helpful in this respect, and are related to deriving inconsistencies – the techniques to do so are well-known from the **CL**-heuristics. I mention only two examples. Suppose that, in an **LI**-proof from Σ, A is unconditionally derived, and that $\sim A$ is derived on the condition Δ. Then $Dab(\Delta)$ is unconditionally derivable in the proof. Similarly, if an inconsistency is derived on the condition Δ, $Dab(\Delta)$ is unconditionally derivable in the proof.

An equally helpful derivable rule was exemplified before (and is warranted by Theorem 2). If a *Dab*-formula $Dab(\Delta)$ is derived on the condition Δ', then $Dab(\Delta \cup \Delta')$ is unconditionally derivable. Similarly, if an instance of A is derived on the condition Δ and an instance of B is derived on the condition Δ', then $\sim\forall(A \supset \sim B) \vee Dab(\Delta \cup \Delta')$ is unconditionally derivable – either or both of Δ and Δ' may be empty.

A very rough summary reads as follows: derive all singular statements that lead to instances of formulas no instances of which have been derived, and derive *Dab*-formulas that change either the minimal *Dab*-formulas or the minimal local *Dab*-formulas. The first instruction requires the derivation of a few formulas only. The second may be guided by several considerations, (i) Whenever $Dab(\Delta)$ has been derived, one should try to unconditionally derive $Dab(\Delta')$ for all $\Delta' \subset \Delta$. This is a simple and decidable task. (ii) One should only try to derive $Dab(\Delta)$ when Δ consists of background generalizations, generalizations introduced by the rule RC, or "variants" of such generalizations – the variants of $\forall(A \supset B)$ being the four generalizations that occur in (1). This instruction may be further restricted. Given a background generalization or local generalization $\forall(A \supset B)$, one should first and foremost try to derive the $Dab(\Delta)$ for which Δ contains variants of $\forall(A \supset B)$. The only cases in which it pays to consider other *Dab*-formulas is the one described in (†).

Up to now I have considered the general heuristic maxims that apply to **LI**. However, **LI** has distinct application contexts, in which different aims are pursued and specific heuristic maxims apply. I shall consider only two very general application contexts.

If one tries to derive *Dab*-formulas that result in some lines being marked or unmarked, one basically checks whether the introduced generalizations are compatible with and confirmed by the available empirical data. However, one might also, after introducing a generalization, concentrate on its consequences by deriving singular statements from it. These singular statements will be derived conditionally. As said before, this may be taken to be a good reason to invoke observation and experiment in order to test them. This leaves room for a "Popperian" application of **LI**. Even if a generalization may be marked in view of derivable *Dab*-formulas, and even if it *is* marked in view of derived *Dab*-formulas, we may try to gather novel data that cause the generalization to be unmarked. Incidentally, the "stronger" generalizations in the sense of Popper (1935; 1963) are those from which a larger number of weaker generalizations are derivable, and hence have more potential falsifiers. Popper was quite right, too, to stress that it is advisable to infer the most general (the bolder) generalizations first. If they become marked, we may still retract to less general generalizations. As long as these are not marked, the less general generalizations are available for free because they are **CL**-consequences of the more general ones.

A distinction is useful in the present context. If an instance of Px is derivable from the empirical data together with the reliable background knowledge, but no instances of either $Px \wedge Qx$ or $Px \wedge {\sim}Qx$ are so derivable, then both $(\forall x)(Px \supset Qx)$ and $(\forall x)(Px \supset {\sim}Qx)$ may be marked *because we have no means to choose between them.* If instances of both $Px \wedge Qx$ and $Px \wedge {\sim}Qx$ are **CL**-derivable from the empirical data together with the reliable background knowledge, then both $(\forall x)(Px \supset Qx)$ and $(\forall x)(Px \supset {\sim}Qx)$ may be marked *because both are falsified.* The transition from the first situation to the second clearly indicates an increase in knowledge. Moreover, in the second situation it does not make sense to look for further confirming instances of either generalization. What does make sense in the second situation, and not in the first, is that one looks for less general hypotheses, for example $(\forall x)((Px \wedge Rx) \supset Qx)$ that may still be derivable.

This at once answers the objection that **LI** too severely restricts a scientist's freedom to launch hypotheses. **LI** does not in any way restrict the freedom to introduce generalizations. Rather, **LI** points out, if sensibly applied, which generalizations cannot be upheld, and which empirical research is desirable. A scientist's "freedom" to launch hypotheses is not a permission for dogmatism – to make a claim and stick to it. If it refers to anything, then it is to the freedom to break out of established conceptual schemes. Clearly, the introduction of new conceptual schemes goes far beyond the present simple logic of induction – I return to this in Section 7. Given the limits of **LI**, the set of **LI**-consequences of a given Σ should be determined by Σ and should be independent of any specific line of reasoning. In this respect the rule RC differs drastically from such rules as Hintikka's bracketing rule – see, for example, Hintikka (1999; forthcoming).

A very different application context concerns predictions derived in view of actions. It makes sense, in the Popperian context, to derive predictions from a generalization *A*, even before checking whether the proof can be extended in such a way that *A* is marked. In the present context, it does not. It would be foolish to act on the generalization $(\forall x)(Px \supset Qx)$ in the absence of confirming instances – such actions would be arbitrary. In action contexts, one should play the game in a safer way by introducing only well-confirmed generalizations, not bold ones. Thus $(\forall x)(Px \supset Qx)$ should be *derived* from safe generalizations, for example, $(\forall x)((Px \wedge Rx) \supset Qx)$ and $(\forall x)((Px \wedge {\sim}Rx) \supset Qx)$ if both of these happen to be safe.

In both contexts,[28] **LI** suggests a specific heuristic procedure. This procedure differs from one context to the other, and may be justified in view of the specific aims.

Some people may find it suspect that applications of the rule RC do not require the presence of any formulas in the proof. RC is a positing rule rather than a deduction rule. This is no reason to worry. **LI** has a dynamic proof theory. A proof at a stage should not be confused with a proof of a logic that has a (static) proof theory of the usual kind. The central question in an **LI**-proof is not whether a generalization can be introduced, but whether it can be retained – the aim is final derivability, not derivability at a stage. The preceding paragraphs make it sufficiently clear that final derivability is often difficult to reach, and that one needs

[28] This distinction between action contexts and contexts concerning theoretical inquiry was one of the points made in my (1968).

to follow a set of heuristic rules in order even to obtain a sensible estimate of final derivability – see also below. In this connection, it is instructive to see that $Cn_{\mathbf{I}}(\langle\emptyset, \emptyset\rangle) = Cn_{\mathbf{CL}}(\emptyset)$, that $Cn_{\mathbf{I}}(\langle\{Pa, \sim Pa\}, \emptyset\rangle) = Cn_{\mathbf{CL}}(\{Pa, \sim Pa\})$, and hence that neither of these comprises a non-tautological generalization.

A final comment concerns the nature of an adaptive logic. It would be foolish to build a logic that allows for some mistakes. Obviously, adaptive logics do not allow for mistakes: $Cn_{\mathbf{I}}(\Sigma)$ is a well-defined set that leaves no room for any choice or arbitrariness. The dynamic proof theory constitutes a way to search for $Cn_{\mathbf{I}}(\Sigma)$. A proof at a stage merely offers an estimate of $Cn_{\mathbf{I}}(\Sigma)$ – an estimate that is determined by the insights in the premises that are provided by the proof. We have seen that there are heuristic means to make these insights as rich and useful as possible. There also are criteria to decide, in some cases, whether a formula is finally derived in a proof – see Batens (2002). In the absence of a positive test, that is the best one can do in a computational respect.

For large fragments of the language, **LI**-derivability is decidable. This includes all generalizations, and hence all predictions and explanations. But even for undecidable fragments of the language, dynamic proofs at a stage offer a sensible estimate of $Cn_{\mathbf{I}}(\Sigma)$, the best estimate that is available from the proof – see Batens (1995). This means that an **LI**-proof at a stage is sufficient to take justified decisions: decisions that may be mistaken, but are justified in terms of our present best insights.

7. Further research

As announced, **LI** is very simple – only a starting point. In the present section I briefly point to some open problems. Some of these relate to alternatives for **LI**, others to desirable sophistication.

With respect to background generalizations, an interesting alternative approach is obtained by not introducing members of Γ^* but rather generalizations that belong to $Cn_{\mathbf{CL}}(\Gamma^*)$. Suppose that $(\forall x)(Px \supset Qx) \in \Gamma^*$, and that Pa, Ra and $\sim Qa$ are **CL**-consequences of Γ. According to **LI**, $(\forall x)(Px \supset Qx)$ is falsified, and hence not retained. According to the alternative, $(\forall x)((Px \wedge \sim Rx) \supset Qx)$ would, for all that has been said, be a retained background generalization. This certainly deserves further study, both from a technical point of view and with respect to application contexts.

LI is too empiricist, even too positivistic. Let me just mention some obvious sophistication that is clearly desirable. Sometimes our background knowledge is inconsistent and sometimes falsified generalizations are retained. As there is room for neither in **LI**, this logic is unfit to explicate certain episodes from the history of the sciences. It is not difficult to modify **LI** in such a way that both inconsistent background knowledge and the application of falsified generalizations are handled. Available (and published) results on inconsistency adaptive logics make this change a rather easy exercise.

Another weakness of **LI**, or rather of the way in which **LI** is presented in the present paper, is that there seems to be only room for theories in the simple sense of the term: sets of generalizations. This weakness concerns especially background theories – the design of new theories is not a simple inductive matter anyway. Several of the problems listed above are solved in Batens and Haesaert (2001); this paper contains also a variant of **LI** that follows the standard format for adaptive logics.

LI does not enable one to get a grasp of conceptual change or of similar phenomena that are often related to scientific creativity and discovery. This will be the hardest nut to crack. That it is not impossible to crack it will be obvious to readers of such papers as Meheus (1999a; 1999b; 2000).

Let me say no more about projected research. The basic result of the present paper is that there is now a logic of induction. It is simple, and even a bit old-fashioned, but it exists and may be applied in simple circumstances.[29]

Acknowledgments
Research for this paper was supported by subventions from Ghent University and from the Fund for Scientific Research – Flanders, and indirectly by the Flemish Minister responsible for Science and Technology (contract BIL98/37). I am indebted to Atocha Aliseda, Theo Kuipers, Dagmar Provijn, Ewout Vansteenkiste, and Liza Verhoeven for comments on previous drafts.

REFERENCES

BAR-HILLEL, Y. 1968: "The Acceptance Syndrome", in I. Lakatos (ed.), *The Problem of Inductive Logic*, Amsterdam: North-Holland, pp. 150-161.

BATENS, D. 1968: "Some Proposals for the Solution of the Carnap-Popper Discussion on Inductive Logic", *Studia Philosophica Gandensia*, 6, pp. 5-25.

——— 1986: "Dialectical Dynamics within Formal Logics", *Logique et Analyse*, 114, pp. 161-173.

[29] Unpublished papers by members of our research group are available from the Internet address http://logica.UGent.be/centrum/writings/.

——— 1995: “Blocks. The Clue to Dynamic Aspects of Logic”, *Logique et Analyse*, 150-152, pp. 285-328.

——— 1999: “Inconsistency-Adaptive Logics”, in E. Or³owska (ed.), *Logic at Work. Essays Dedicated to the Memory of Helena Rasiowa*, Heidelberg/New York: Physica Verlag Springer, pp. 445-472.

——— 2000: “A Survey of Inconsistency-Adaptive Logics”, in D. Batens, C. Mortensen, G. Priest, and J.P. Van Bendegem (eds.), *Frontiers of Paraconsistent Logic*, Baldock, UK: Research Studies Press, pp. 49-73.

——— 2002: “On a Partial Decision Method for Dynamic Proofs”, in H. Decker, J. Villadsen, and T. Waragai (eds.), *PCL 2002. Paraconsistent Computational Logic* (*Datalogiske Skrifter*, vol. 95), pp. 91-108. Also available as cs.LO/0207090 at http://arxiv.org/ archive/cs/intro.html.

——— 2004: “Extending the Realm of Logic. The Adaptive-Logic Programme”, in P. Weingartner (ed.), *Alternative Logics. Do Sciences Need Them?*, Berlin-Heidelberg: Springer Verlag, pp. 149-164.

BATENS, D., HAESAERT, L. 2001: “On Classical Adaptive Logics of Induction”, *Logique et Analyse*, 173-175, pp. 255-290.

BATENS, D., MEHEUS, J. 2000: “The Adaptive Logic of Compatibility”, *Studia Logica*, 66, pp. 327-348.

BOOLOS, G.S., JEFFREY, R.J. 1989: *Computability and Logic*, 3rd edition, Cambridge: Cambridge University Press.

BROWN, B. 1990: “How to Be Realistic about Inconsistency in Science”, *Studies in History and Philosophy of Science*, 21, pp. 281-294.

HINTIKKA, J. 1999: *Inquiry as Inquiry: A Logic of Scientific Discovery*, Dordrecht: Kluwer.

——— (forthcoming), “Argumentation in a Multicultural Context”.

KUIPERS, T. 2000: *From Instrumentalism to Constructive Realism*, Dordrecht: Kluwer.

MEHEUS, J. 1993: “Adaptive Logic in Scientific Discovery: The Case of Clausius”, *Logique et Analyse*, 143-144, pp. 359-389.

——— 1999a: “Deductive and Ampliative Adaptive Logics as Tools in the Study of Creativity”, *Foundations of Science*, 4, pp. 325-336.

——— 1999b: “Model-Based Reasoning in Creative Processes”, in L. Magnani, N. Nersessian, and P. Thagard (eds.), *Model-Based Reasoning in Scientific Discovery*, Dordrecht: Kluwer/Plenum, pp. 199-217.

——— 2000: “Analogical Reasoning in Creative Problem Solving Processes: Logico-Philosophical Perspectives”, in F. Hallyn (ed.), *Metaphor and Analogy in the Sciences*, Dordrecht: Kluwer, pp. 17-34.

——— 2002: “Inconsistencies in Scientific Discovery. Clausius’s Remarkable Derivation of Carnot’s Theorem”, in H. Krach, G. Vanpaemel, and P. Marage (eds.), *History of Modern Physics*, Brepols, Brepols, pp. 143-154.

NERSESSIAN, N. 2002: “Inconsistency, Generic Modeling, and Conceptual Change in Science”, in J. Meheus (ed.), *Inconsistency in Science*, Dordrecht: Kluwer, pp. 197-211.

NORTON, J. 1987: “The Logical Inconsistency of the Old Quantum Theory of Black Body Radiation”, *Philosophy of Science*, 54, pp. 327-350.

——— 1993: “A Paradox in Newtonian Gravitation Theory”, in *PSA 1992*, vol. II, pp. 421-420.

POPPER, K.R. 1935: *Logik der Forschung*, Wien: Verlag von Julius Springer.

——— 1959: *Logic of Scientific Discovery*, London: Hutchinson. English translation with new appendices of Popper 1935.

——— 1963: *Conjectures and Refutations*, London: Routledge & Keagan.

——— 1973: *Objective Knowledge*, Oxford: Clarendon.

SMITH, J. 1988: “Inconsistency and Scientific Reasoning”, *Studies in History and Philosophy of Science*, 19, pp. 429-445.