On the Logic and Pragmatics of the Process of Explanation*

Diderik Batens Joke Meheus†
Centre for Logic and Philosophy of Science
Universiteit Gent, Belgium
Diderik.Batens@rug.ac.be Joke.Meheus@rug.ac.be

Abstract

In this paper, we present mainly two logical systems that clarify pragmatic aspects of the process of explanation. The first concerns a proof theory that leads to the derivation of possible initial conditions from an *explanandum* and a given theory. The second logic concerns the derivation of questions in view of the verification of some possible initial condition, or of one out of several possible initial conditions. It is essential that the latter derivation proceeds in terms of all available knowledge, and not in terms of the explaining theory. It is shown that the second logic provides useful information for explicating further pragmatic aspects of the process of explanation. Several extensions of the logics are argued to be both useful and rather easy to obtain.

1 Pragmatic Aspects of Explanation

Two central themes in Matti Sintonen's work on explanation, are (i) the urge to pay more attention to the *process* of explanation, rather than to its result, and (ii) the importance of pragmatic aspects of explanation. In order to understand scientific explanation, the important question is not to delineate the conditions under which a set of empirical data together with a set of theories constitute an explanation of some phenomenon, but rather to describe the way in which scientists proceed in search for an explanation of some phenomenon. Needless to say, this search is heavily laden with pragmatic aspects. We are on Sintonen's side in this respect. We are even on his side in a further respect, which is that, nevertheless, logic is essential in the study of explanation.

*The research for this paper was financed by the Fund for Scientific Research – Flanders, and indirectly by the Flemish Minister responsible for Science and Technology (contract BIL98/37).

†Joke Meheus is a Postdoctoral Fellow of the Fund for Scientific Research – Flanders (Belgium).

A brief warning seems useful in this respect. The role of logic for explanation is often identified with the definition of an explanation (as a product). This is presumably a consequence of the fact that Hempel's popular and (mainly) product oriented theory of explanation from (especially) [16], describes an explanation as a *deduction* of the *explanandum* from a theory together with empirical data, and that Hempel qualifies this theory as non-pragmatic. Writes Hempel: "This ideal intent suggests the problem of constructing a nonpragmatic concept of scientific explanation—a concept which is abstracted, as it were, from the pragmatic one, and which does not require relativization with respect to questioning individuals [...]" ([16, p. 426]), and he compares the concept of explanation with the concept of mathematical proof.

This paper concerns the process of explanation, including its pragmatic aspects. We shall present some simple logical tools that we consider not only useful but even required in this respect. We are not sure whether Sintonen will be fully in sympathy with our approach. In several respects, our approach diverges from the one that is apparently advocated by Sintonen. Where Hintikka and his associates, e.g., in [20], (and also Atocha Aliseda in [1]) proceed in terms of tableaux-like structures, we proceed in terms of proofs.[1] Moreover, we often push into the definition of the consequence relation (and hence into the 'definitory' rules) part (or all) of the reasoning that Hintikka and his associates locate in the strategy (see, for example [18]). As a result of this, our logics have often a consequence relation that is not Tarskian,[2] and most of our logics have *dynamic proof theories*, which enable one to explicate a person's dynamic reasoning. We shall not defend our approach here—we hope to do so in a different paper soon—but rather illustrate it.

In the subsequent sections we shall mainly concentrate on two important aspects of the process of explanation. First, in search of an explanation of some phenomenon in view of some theory, one needs to identify certain factual statements as possible explanations, or rather possible *initial conditions*. We shall show that this identification, which involves many pragmatic aspects, *is* a matter of logic (Section 3—some brief technical preparation is contained in Section 2). We shall present a specific proof format and a simple strategic rule that, when applied to a set of theories and a why-question, deliver the possible initial conditions. We shall even outline (briefly in order to avoid technicalities) a logic that does not require the aforementioned heuristic rules—put differently: for which the strategic rule is pushed down into the definitory rules.

Given one or more possible initial conditions, one has to find out whether it, or one of them, is true and hence leads to an explanation. This second important aspect of the process of explanation is often neglected. We shall handle it by

[1] So, we disagree with Sintonen where he writes, in [27], that reasoning about models has certain advantages over proofs from premises. Moreover, we have studied tableau methods for inconsistency-adaptive logics, in [11] and [12], and there the dynamic proofs are heuristically superior to all tableau methods we were able to devise. As far as we can see, this result may be generalized.

[2] A Tarskian consequence relation is reflexive ($A \vdash A$), transitive (if $\Gamma \vdash B$ and $\Delta \cup \{B\} \vdash A$, then $\Gamma \cup \Delta \vdash A$), and monotonic (if $\Gamma \vdash A$, then $\Gamma \cup \Delta \vdash A$).

phrasing possible initial conditions as 'starting questions' of the form $?A$ (is A the case?). One may obtain the answer to $?A$ by deriving A or $\sim A$ from one's knowledge—and, with one small proviso, *all* of one's knowledge may be relied upon here—or by observation and experiment. More often than not, however, $?A$ cannot be answered in a direct way, and in this case one has to derive other questions from $?A$ and from one's knowledge—again, *all* of one's knowledge—in order to obtain an answer to $?A$. We shall first present a simple logic for deriving questions in Section 4, and next upgrade it, in Section 5, to an adaptive logic[3] **Q** for deriving the right questions (in view of one or more explanatory goals). The logic **Q** enables one to "identify good research questions" in the sense of [28, p. 128], and implements a central idea from [27], namely that why-questions are tackled by deriving and answering a series of factual questions.

In the last three sections, we consider the combination of the explanation seeking deduction with the logic **Q**, discuss some further pragmatic aspects, and briefly consider some obvious extensions of the logics presented in this paper.

2 Informative and Analysing Moves

In the subsequent sections we need a distinction that is most easily made clear in terms of the block-approach from [4] and [5]. For present purposes, we shall take **CL** to be the standard of deduction. Any annotated **CL**-proof determines a *block-analysis*—if the proof is not annotated, only finitely many annotations are compatible with it. The block analysis of the proof is determined by the discriminations and identifications the author of the proof has *minimally* made in the formulas of the proof in order to construct the proof (according to the annotation). An example will clarify the matter. Applying the Premise rule comes to introducing a block: there is no need for any insight in the structure of the formula in order to write it down in the proof.

1	$[[(p \supset \sim q) \supset (p\&(\sim r \vee \sim p))]]^1$	Premise
2	$[[p \supset \sim q]]^2$	Premise

In order to apply Modus Ponens to these premises, one needs to *analyse* the first block as an implication between two blocks, and one needs to *identify* the implicans block with block 2; the derived block is of course the implicatum block of 1:

1	$[[(p \supset \sim q)]]^2 \supset [[(p\&(\sim r \vee \sim p))]]^3$	Premise
2	$[[p \supset \sim q]]^2$	Premise
3	$[[p\&(\sim r \vee \sim p)]]^3$	1, 2; Modus Ponens

[3]This kind of logics was developed in the Ghent Centre for Logic and Philosophy of Science. The name refers to a special feature of these logics: they adapt to the specific set of premises. Consequences of the premises are supposed to behave in a certain way with respect to the premises, unless and until proven otherwise. As a result, there are some rules of inference that are not validated or invalidated in general, but of which some applications are validated while others are not, depending on the premises. See, for example [6] for a survey of adaptive logics. A home page bibliography on the topic is available at `http://logica.rug.ac.be/adlog/`.

If the proof is continued by deriving p by Simplification from 3, then the block analysis will read:

1	$[\![(p \supset \sim q)]\!]^2 \supset ([\![p]\!]^4 \& [\![\sim r \vee \sim p]\!]^5)$	Premise
2	$[\![p \supset \sim q]\!]^2$	Premise
3	$[\![p]\!]^4 \& [\![\sim r \vee \sim p]\!]^5$	1, 2; Modus Ponens
4	$[\![p]\!]^4$	3; Simplification

Remark that block 3 is split up everywhere in the proof—if it were not split up at line 1, the justification of line 3 would be mistaken.

If the proof is continued by first deriving $\sim r \vee \sim p$ by Simplification, and next deriving $p \vee \sim r$ by Addition from 4, we obtain:

1	$[\![(p \supset \sim q)]\!]^2 \supset ([\![p]\!]^4 \& [\![\sim r \vee \sim p]\!]^5)$	Premise
2	$[\![p \supset \sim q]\!]^2$	Premise
3	$[\![p]\!]^4 \& [\![\sim r \vee \sim p]\!]^5$	1, 2; Modus Ponens
4	$[\![p]\!]^4$	3; Simplification
5	$[\![\sim r \vee \sim p]\!]^5$	3; Simplification
6	$[\![p]\!]^4 \vee [\![\sim r]\!]^6$	4; Addition

By now, the idea should be clear. Writing down a proof requires that one analyses some formulas (discriminates its subformulas) and that one identifies some formulas or subformulas. The effect of the discriminations is that blocks are replaced by block formulas (formulas formed from blocks by logical symbols and possibly parentheses). The effect of the identifications may be that two blocks that previously had a different number, receive the same number. We repeat that the block analysis of a proof reveals the discriminations and identifications that the author of the proof has minimally made in the formulas of the proof in order to construct the proof (according to the annotation). In [4] and [5], the block analysis is spelled out decently for predicative logics—these involve some complications that are not essential here. (Especially the first paper contains several interesting applications of the method.)

The block analysis enables one to distinguish between informative and non-informative moves in a proof. A step is *informative* iff it requires that the premises are further analysed. Steps 3 and 4 are obvious examples. For reasons that will become clear when we consider the semantic criterion (see below), applications of the premise rule are also seen as informative moves. So, steps 1–4 in the proof are informative, while steps 5 and 6 are not.

There is an interesting difference between step 5 and step 6. If $[\![\sim r \vee \sim p]\!]^5$ had been derived at line 4, its derivation would have been an informative move. Line 5 is not the result of an informative move simply because $[\![p]\!]^4$ was already derived from 3. So, in some proofs from premises 1 and 2, $[\![\sim r \vee \sim p]\!]^5$ is derived by an informative move. There is, however, no way to derive 6 from 4 by an informative move.

Remark that, in order to construct the proof up to line 6, there is no need to see that the 'content' of block 6 is a subformula of block 5. Even if the content of block 6 were identical to the content of block 5, it would not be required that the author of the proof has made the identification between blocks 5 and 6 (whence they should still have different numbers).

The distinction between informative and non-informative moves may also be characterized in semantical terms. Suppose that we have a semantics for block formulas—the matter is spelled out in [4] and [5]. Let a model for a block proof at a stage be a model that verifies the premises, as analysed in the present *and* previous stages of the proof—a stage of a proof is simply the number of its last line. In the transition from stage 2 to stage 3 in the proof displayed above, the formula $[[(p \supset \sim q) \supset (p\&(\sim r \vee \sim p))]]^1$ was analysed as $[[(p \supset \sim q)]]^2 \supset [[(p\&(\sim r \vee \sim p))]]^3$. At stage 1 and 2 of the proof, block 3 has value 1 (true) in some models of the proof and value 0 (false) in others. At stage 3 of the proof, only models that assign value 1 to block 3 are models of the premises.[4] In other words, with each *informative* move, the models of the premises are narrowed down (to a subset of the models of the previous stage).[5] This clarifies at once why we consider applications of the premise rule as informative moves. With each premise one adds to the proof, the set of models is narrowed down (to the models that verify the newly introduced block).

This semantic criterion is important because it does not depend on the set of formulas that actually constitute the proof, but only on the block formulas that occur in the proof. In the proof displayed above, $[[\sim r \vee \sim p]]^5$ is not yet derived at stage 4, but nevertheless is a semantic consequence of the premises (as analysed at stage 4 of the proof).

Some informative steps contain non-informative substeps. Thus, the step from $p\&r$ to $p \vee q$ is informative (one has to analyse $p\&r$ as a conjunction), but is composed of an (informative) application of Simplification and a (non-informative) application of Addition. For reasons that will appear later, we are interested in moves that are not only informative, but moreover *analysing*. Intuitively, they should not contain any non-informative substeps. Obviously, this can be made precise with respect to a specific set of rules of inference, but we want a more general characterization that classifies all derivable moves as analysing or non-analysing. The block approach provides us with a criterion.

By the *complexity* of a meta-language formula we shall mean the number of logical symbols that occur in the formula. Thus the complexity of A is 0, the complexity of $\sim A$ and of $A \supset B$ is 1, etc.

A sound rule of inference

$$X_1, \ldots, X_n \;/\; Y \tag{1}$$

will be called *non-redundant* iff, for any block analysis that makes it a sound rule, (i) the content of each block is a meta-variable (and nothing else) and (ii) no premise block formula is superfluous.[6] Rule (1) is *analysing* iff (i) it

[4] The reason is obviously that block 2 is also a premise, and that an implication is true only if its antecedent is false or its consequent is true.

[5] This enables one to 'measure' the information (about the premises) that is provided by a proof at a stage.

[6] Thus both $A,\ A \supset (B\&C) \;/\; B\&C$ and $A,\ A \supset B,\ C \;/\; B$ are redundant.

is informative and non-redundant, and (ii) there is no sound non-redundant rule $X_1, \ldots, X_n$ / Z in which the complexity of Z is smaller than the complexity of Y. Adjunction and A / $\sim\sim A$ are not analysing because they are not informative.[7] Addition and Irrelevance (A / $B \supset A$) are non-redundant but are not analysing because they are not informative and do not fulfil (ii). The rule $A\&B$ / $A \vee C$ is non-redundant and is informative, but does not fulfil (ii). Some examples of analysing rules: Modus Ponens, Modus Tollens,[8] Disjunctive Syllogism, Dilemma, $\sim\sim A$ / A, and $((A \vee B)\&C) \supset D$ / $(A\&C) \supset D$.

We now introduce a specific consequence relation derived from **CL**. We shall say that

$$A_1, \ldots, A_n \vdash_{\mathbf{CL}^a} B$$

iff there is a **CL**-proof of B from $A_1, \ldots, A_n$ in which only the Premise rule and analysing rules have been applied.[9]

3 Explanation Seeking Deduction

In this section, we illustrate an explanation seeking process that proceeds in terms of a proof. The results of this section depend on (i) the 'natural' proof search procedure for **CL** (and for a few other systems) presented in [3] and (ii) the results on conditional derivations that will be written up in a paper by Diderik Batens and Dagmar Provijn.[10] While there is no need to repeat any results from (i), we need to explain the basic idea behind (ii).

Formulas that occur in a proof may be analysed by deriving subformulas (and instances) from them. Thus, a formula of the form $A\&B$ is analysed by applying Simplification. Some formulas (basically those of the forms $A \supset B$ and $A \vee B$) can only be analysed with the help of a 'minor premise' (as in applications of Modus Ponens and Modus Tollens for formulas of the form $A \supset B$, and in applications of Disjunctive Syllogism for formulas of the form $A \vee B$). Precisely these formulas cause trouble in complex cases.[11] Interestingly, the formulas that cause trouble are typically those that are essential in an explanation seeking

[7]Rules obtained from an analysing rule by the metalinguistic equivalent of Uniform Substitution (in the way A, $A \supset (B\&C)$ / $B\&C$ is obtained from Modus Ponens) are not analysing because they are redundant. Their applications are obviously also applications of the analysing rule.

[8]All variants of Modus Tollens are analysing. For example $A \supset \sim B$, B / $\sim A$ fulfils (ii) because $A \supset \sim B$, B / B is redundant.

[9]An analysing (and hence informative) rule may have non-informative applications—see line 5 of the block-proof example. It follows that a $\mathbf{CL}^a$-proof may contain non-informative moves, provided each of them is the result of an application of an analysing rule.

[10]The proof search procedure meant in (i) is intended for students of elementary logic classes, and presupposes that the goal (the formula to be derived from the premises) is given. For the propositional part of **CL**, the instructions, which are algorithmic for this fragment, are discussed and justified in [2]. The instructions are also implemented in a logic training programme, earlier written in Turbo Pascal, now extended and rewritten in Delphi.

[11]Where a proof is directed at deriving a conclusion, the complex cases are those in which the antecedent (or the negation of the consequent) of an implicative formula is not obtained in the proof by analysing steps, but is derivable from available formulas. (Do not forget to include theorems.)

deduction. This is why an instrument, which was available for different reasons, proves extremely helpful in the present paper.

If a proof contains troublemakers of the aforementioned kind, we have to apply heuristic reasoning. This may be very complex, and many search paths may not lead to success. Moreover, if one applies a standard proof format, the heuristic reasoning cannot be written down in the proof. So, in order to make explicit the heuristic reasoning that leads to adding new steps to a proof, a specific format was devised (relying on results from adaptive logics) to push the heuristic reasoning down into the proof. Suppose that a proof contains a line of the following form:

i	$A \supset B$	[some justification]	

If we are interested in deriving B, we should try to derive A (in order to apply Modus Ponens). Our new proof format allows us to write this in the proof as follows:

$i+1$	B	i, ?; Modus Ponens	$\{A\}$

The question mark in the 'justification' indicates that a line at which A occurs is missing *at this stage of the proof.* The A in the condition indicates that, if we are able to derive A, then we can remove the condition and replace the question mark (with the number of the line at which A is derived), and hence we will have derived B from the premises.[12]

From line i, we may also derive:

$i+2$	$\sim A$	i, ?; Modus Tollens	$\{\sim B\}$

and similar conditional moves may be made if the analysed formula is a disjunction (this time applying Disjunctive Syllogism).

Nothing prevents one to operate on a member of the condition. Thus, if the same proof contains

j	C	[some justification]	$\{D, B\}$

then one may add

$j+1$	C	$i+1$, j, C-Trans	$\{D, A\}$

Rule C-Trans applies transitivity to the condition: as $\{D, B\}$ is sufficient to obtain C and $\{A\}$ is sufficient to obtain B, $\{D, A\}$ is sufficient to obtain C.

All this should be easy to understand. The criterion for E being derivable from the premises on the condition Δ is simply that E is derivable from the premises together with the elements of Δ. In view of this, it is equally obvious that complex elements of a condition may themselves be analysed. If the condition contains a conjunction, this may be replaced by its conjuncts. The matter is equally simple if the condition contains a disjunction: from

[12] According to the standard use in adaptive logics, all lines that have a (non-empty) condition in the present kind of construction should be marked. We do not write these marks because they would be redundant anyway.

k	F	[some justification]	$\{G \vee H\}$

one may derive both

$k+1$	F	k, C-$\vee$-E	$\{G\}$
$k+2$	F	k, C-$\vee$-E	$\{H\}$

in which C-$\vee$-E indicates that we eliminated a disjunction in the condition of line k.

The reader who thinks to recognize all this is quite right. Our proof format is clearly related to tableaux as well as to Gentzen-style proofs.[13] However, our proof format has the advantage to avoid the (sometimes) complex trees required by tableau methods or Gentzen-style proofs.[14]

In order to verify the correctness of some conditional derivation (of any complexity), it is sufficient to realize that A is derivable on the condition Δ just in case $\bigwedge(\Delta) \supset A$ is derivable unconditionally from the premises.

With the promise that details and metatheoretic proofs on this proof method will be disclosed in the announced paper, we take it that the general idea is sufficiently clear and proceed to an example of an explanation seeking proof. In order to keep things as simple as possible, we present the matter in terms of a **CL**-proof and a specific strategy. At the end of this section, we shall argue that the strategy may be pushed down into the proof format itself.

Suppose that we are looking for an explanation of Qa, that we consider some theory as a potential candidate for delivering the required explanation (for one thing, because the predicate Q occurs in it), and set out to search for the initial conditions that, together with the theory, form an explanation of Qa. We start our proof by writing the axioms of the theory as premises—we chose an example that is not terribly complex.

1	$(\forall x)(Sx \supset (Tx \supset Px))$	Premise
2	$(\forall x)((Px \& {\sim}Qx) \supset Rx)$	Premise
3	$(\forall x)(Px \supset {\sim}Rx)$	Premise
4	$(\forall x)(Tx \supset Rx)$	Premise

The question is "Why Qa?". Given our format, we may note the goal within the proof, viz. by adding:

5	Qa	Goal	$\{Qa\}$

We might first have derived $Qa \supset Qa$, and next have derived 5 from that. To do so, however, would be more confusing than helpful. Indeed, we might have derived $A \supset A$ for any A, whereas only Qa is our goal. The only relation with

[13] It is also related to Fitch-style proofs. Indeed, as any condition is finite, E is derivable from the premises on the condition Δ just in case E is derivable from the premises together with the hypothesis $\bigwedge(\Delta)$.

[14] And it does not lead to the oddity that would result from a multiplicity of non-closed subproofs of the same level in a Fitch-style proof.

$Qa \supset Qa$ is that this theorem justifies line 5.[15] Line 5 states that one may derive Qa on the condition Qa. But clearly, this condition does not provide an explanation of Qa—we shall take this to be defined by Hintikka's criteria as stated in [19] or [15] and elsewhere.[16] So, our aim is to obtain from 1–5 a line on which Qa is derived on some condition such that its members, together with the theory 1–4, form an explanation of Qa.

Now we come to the *strategic* moves. With all due respect, and ready to blame any misunderstanding on ourselves, we were unable to find out which strategy either Hintikka and his associates or Aliseda would consider advisable in the present situation. So, we go by our own lights (but by no means claim originality here).

Let us call a formula *interesting* iff it brings us 'closer to' (that is: contains) some element of a condition.[17] Our *Golden Rule* will be: (i) simplify the members of the conditions and (ii) only derive (conditionally or unconditionally) *interesting* formulas.

Given the present stage of the proof, the only interesting formulas are those that contain Qa (as a positive part). This has an important effect: all formulas that we shall derive will be useful for finding an explanation of Qa (if there is one). Put differently, each conditional line in the proof will be a step towards deriving Qa on some condition that, together with the theory 1–4, forms an explanation of Qa.

Even an elementary insight in **CL** reveals that the only way to obtain Qa is by using premise 2: it is the only premise that contains the predicate Q. So, given that we are looking for Qa (as the only condition reveals), we first derive the a-instance of 2 and next conditionally derive the subformula that contains Qa (as a positive part):

6	$(Pa \& {\sim} Qa) \supset Ra$	2; $\forall$-Instantiation	
7	${\sim}(Pa \& {\sim} Qa)$	6, ?; Modus Tollens	$\{{\sim} Ra\}$

At this stage, two formulas occur in some condition, Qa and ${\sim}Ra$. By the Golden Rule, we either continue to analyse 7, or operate on 3. Actually, it does not matter what we do first; let us first analyse 7.

[15] Incidentally, our proof format enables one to introduce, without any harm, a multiplicity of goals (or *explananda*) within the same proof. To do so in a tableau method would cause the tableau to split in as many branches as there are goals—for Beth-tableaux: would cause the construction to split in as many tableaux as there are goals. Next, each of these branches (or tableaux) have to be split separately in order to handle formulas that, in our proofs, are analysed within the same proof.

[16] There is (in general) no positive test (see, for example, [13]) for two of these criteria: the initial condition should be compatible with the theory, and the negation of the *explanandum* should be compatible with the theory. Hence, our logic will only apply to theories formulated in decidable fragments of the language of **CL**. This is easily repaired by moving to an adaptive logic—an easy exercise in view of the adaptive logic of compatibility presented in [10].

[17] This must be specified: A brings us 'closer to' a B iff B is a positive part in A and at least some subformula of A of which B is a positive part is not derived on the same or a weaker condition. There are two technical bits here. A *positive part* of a formula is an unnegated disjunct of the prenex conjunctive normal form of the formula—this may be defined in a way that is simpler but more longwinded. Condition Δ is *weaker* than condition Δ' iff $\Delta \subset \Delta'$.

8	$\sim Pa \vee Qa$	7; $\sim$-&-Elimination	$\{\sim Ra\}$

in which we retain the condition of 7 for reasons that should be clear from what precedes.

From 8 one *may* take two conditional steps by Disjunctive Syllogism, but only one is interesting (at the present stage of the proof) in view of the Golden Rule:

9	Qa	8, ?; Disjunctive Syllogism	$\{Pa, \sim Ra\}$

It is easy to check that theory 1–4 together with initial condition $Pa\&\sim Ra$ forms an explanation of Qa by the criteria of [19] or [15].

While we might stop the explanation seeking deduction at this point, we still have something on our agenda: to operate on premise 3 in view of $\sim Ra$. This leads to the following steps:

10	$Pa \supset \sim Ra$	3; $\forall$-Instantiation	
11	$\sim Ra$	10, ?; Modus Ponens	$\{Pa\}$
12	Qa	9, 11; C-Trans	$\{Pa\}$

The last step invokes two (very different) comments. First, there is absolutely no need to keep in mind the question marks (referring to lines) that occur in previous steps in the proof. By deriving Qa on the (sole) condition Pa, we have established that $Pa \supset Qa$ is derivable from the theory 1–4. Next, and more importantly with respect to the process of explanation, we found out that Pa in itself is sufficient to explain Qa (the theory 1–4 together with the initial condition Pa explains Qa according to the criteria from [19] and [15]). However difficult, expensive, time consuming, or whatever—see Section 7—the test for Ra, this test is neither necessary nor useful to establish Qa. We stress this because it is important from a *pragmatic* point of view.

We have first seen that Qa may be explained by $Pa\&\sim Ra$. Next, we have 'simplified' this to: Qa may be explained by Pa. At this point we might decide to stop looking for further explanations of Qa from the theory 1–4. However, there may be other such explanations and some of them may be pragmatically preferable to Pa—see Section 7. So, let us continue.

If we continue, we should know which formulas are still interesting. Qa and Pa are certainly among them, but what about $\sim Ra$? Avoiding a technical explanation, we just state that $\sim Ra$ is not interesting for the continuation of our present deductive search. We shall see, however, that $\sim Ra$ may very well be interesting as soon as we enter the second part of our endeavour, which will concern the derivation of (factual) questions.[18]

To see what will be the next move, it is important to realize that premises 2 and 3 have been completely analysed *with respect to interesting formulas.* So, we shall have to operate on premise 1, which indeed contains P. This leads to:

[18] As we shall remark later, the separation of both reasoning processes is artificial, and introduced here only for reasons of exposition.

13	$Sa \supset (Ta \supset Pa)$	1; ∀-Instantiation	
14	$Ta \supset Pa$	13, ?; Modus Ponens	$\{Sa\}$
15	Pa	14, ?; Modus Ponens	$\{Sa, Ta\}$

Given that Qa is our (only) main goal (as is revealed by line 5), we continue:

16	Qa	12, 15; C-Trans	$\{Sa, Ta\}$

At this point, the theory 1–4 is exhausted with respect to the explanation seeking deduction. No further statement is derivable in view of the Golden Rule.

The explanation seeking deduction has led us to three 'initial conditions' that, together with 1–4, form an explanation of Qa: Pa, $Pa\&{\sim}Ra$, and $Sa\&Ta$. There is absolutely nothing wrong with the second initial condition. From a pragmatic viewpoint, however, we have to keep in mind that any empirical information needed to establish the first initial condition, is necessary but insufficient to establish the second initial condition. The situation for the third initial condition is rather different. Our theory warrants that establishing $Sa\&Ta$ is one of the ways to establish Pa, but it may be very sensible to set out testing $Sa\&Ta$. Obviously, other theories may lead to initial conditions that are independent, both logically and with respect to the theory.

By the reasoning exemplified before, the possible initial conditions are located. A quite different task is now before us: to empirically establish one of the initial conditions. This bit we shall explicate in terms of a logic of questions. The *starting questions* will concern the initial conditions, in the example: $?Pa$, $?(Pa\&{\sim}Ra)$, and $?(Sa\&Ta)$.

Before moving on to this task, let us briefly comment on the previous deduction. Up to now, we have shown that the possible initial conditions may be located in a proof that is guided by heuristic rules. We did not reason in terms of models (or tableaux), and by introducing the conditional lines, we were able to push the bookkeeping of the heuristic reasoning into the proof itself. Also, we hope that the reader who ever tried to reach the same effect by means of tableaux will appreciate the elegance and goal-directedness of the result.

However, we may do better. There is a consequence relation that defines $\{Pa, Pa\&{\sim}Ra, Sa\&Ta\}$ as the consequence set of 1–5. In order not to interrupt our main line of reasoning, we characterize it in stenographic form.

Using the same format as before (with conditional and unconditional lines), the only rules one is permitted to apply are (i) the Premise and Goal rules, and (ii) other rules provided they lead to splitting up members of a condition or deriving interesting formulas. In other words, we push the Golden Rule into the definitory rules themselves. The thus obtained proofs are a subset of the $\mathbf{CL}^a$-proofs.[19] The *consequences* of 1–5 (in general: of a theory and a main goal) are the $\bigwedge(\Delta)$ such that (i) the main goal is derived on the condition Δ, and (ii) the triple consisting of the theory, $\bigwedge(\Delta)$, and the main goal (*explanandum*)

[19]For example, the move from A on the condition $\Delta \cup \{B \vee C\}$ to A on the condition $\Delta \cup \{B\}$ corresponds to an application of the analysing rule $(D\&(B \vee C)) \supset A$ / $(D\&B) \supset A$.

fulfils the criteria as stated in [19] or [15].[20] Even if one would refuse the name "consequence relation" to this relation, one will admit that we are explicating *reasoning*. And actually, we are explicating *sound reasoning*. For example, to derive Ra on the condition Ta from 1–5 is not only heuristically mistaken but is bad reasoning plain and simple.[21]

4 A Logic of Questions

Let us return to the starting questions that we obtained in the previous section. While answering the starting questions is a pragmatic process, there are some logical aspects to it. If a question cannot be answered in a direct way, we shall have to obtain an answer from derived questions, and this derivation should be sound.

We shall present a simple and novel logic of questions, restricted to the derivation of questions from other questions. For our insights in questions, we are indebted to many, but especially to Jaakko Hintikka, who has been pressing the subject for more than twenty years now, and to Andrzej Wiśniewski who has cracked many hard nuts, for example in [32]. Our proposal will be different from theirs, but we shall not discuss the differences here.

Our logic of questions enables one to derive many questions from a given question. A first idea is that $?A$ is derivable from $\Gamma \cup \{?B\}$ iff one of the possible answers to $?B$ is derivable from one of the possible answers to $?A$ together with a (possibly empty) set of (accepted) descriptive premises Γ. Thus, one might consider $?q$ to be derivable from $?p$ and $p \supset q$ because $\sim q, p \supset q \vdash \sim p$. Taken literally, this approach leads to the derivation of irrelevant questions—for example, from $?p$ would follow $?(p \vee q)$, and from this $?q$. So we shall first transform the above criterion to a criterion that is equivalent with respect to **CL**, viz. that $?A$ is derivable from $\Gamma \cup \{?B\}$ iff one of the possible answers to $?A$ is derivable from one of the possible answers to $?B$ together with Γ. Next, we eliminate irrelevant questions by requiring a $\mathbf{CL}^a$-derivation, rather than a **CL**-derivation. However, this transformation introduces another problem. From $p, q \vdash q$ one should not conclude that $?q$ is derivable from $\{?p, q\}$. The obvious way out is to require that no answer to the derived question should be derivable from the descriptive premises alone. This leads to a nice definition from a systematic point of view, but is problematic from a pragmatic point of view because **CL** is undecidable and there is no positive test for $B_1, \ldots, B_n \nvdash_{\mathbf{CL}} A$.

[20] The consequence relation is non-monotonic, the consequences are a selection of conditions of Qa rather than formulas derived in the proof, etc. All this is highly non-standard, but nothing much follows from this fact in itself.

[21] Some logicians will balk at this and, selectively quoting Aristotle, claim that the derivation of Ra on the condition Ta is sound because it is truth preserving (in that it comes to unconditionally deriving $Ta \supset Ra$). Our reply to this is double. On the one hand, logic is a goal-directed activity. Thus, if the premise is p and the goal is $p \vee q$, the derivation of $p \vee r$, $(p \vee r) \vee r$, etc., is certainly truth preserving but nevertheless a candid example of bad (and stupid) reasoning. On the other hand, the whole truth-preservation argument works only in terms of a semantics. Getting acquainted with the semantics of some adaptive logics might liberate truth-preservationists from a couple of prejudices.

There is a sensible approach to this problem in terms of an adaptive logic—see a forthcoming paper by Kristof De Clercq and Joke Meheus (that solves the problem within the context of Wiśniewski's approach). However, given the aim of the present paper, the right approach seems to require a *pragmatic* notion of derivability, rather than a systematic one. Whether a person or group X is justified in deriving a question at some point in time, depends on the knowledge of X at that time. Suppose that X faces the question $?A$, that B is derivable from A together with some descriptive statements (that are $\mathbf{CL}^a$-derivable from X's knowledge), and X has no proof, at time t, that either B or $\sim B$ is derivable from the descriptive statements alone. In such a case, X is clearly justified in deriving the question $?B$ at t. It may turn out later that either B or $\sim B$ is derivable from the descriptive statements, but this does not undermine the justification for X to derive $?B$ at time t. We shall write

$$B_1, \ldots, B_n \nvdash^X_{\mathbf{CL}} A$$

to denote that no proof available to X is a proof of $B_1, \ldots, B_n \vdash_{\mathbf{CL}} A$.[22]

Remark that $B_1, \ldots, B_n \nvdash^X_{\mathbf{CL}} A$ contains an implicit reference to time—logical insights in $B_1, \ldots, B_n$ may change over time (and usually change in the course of searching a proof from $B_1, \ldots, B_n$). However, we shall not further complicate the notation because it will always be clear from the context which point in time is referred to. The same holds for the logic of questions to which we now proceed.

We upgrade $\mathbf{CL}^a$ to the logic of questions $\mathbf{CL}^q$. The language will be that of $\mathbf{CL}$ extended with the question operator ?, but without operations on questions. $\mathbf{CL}^q$ is obtained by extending $\mathbf{CL}^a$ with the following two rules. QD enables one to derive questions, QA to answer questions by non-analysing steps.

QD Where $n \geq 0$,

if $A, B_1, \ldots, B_n \vdash_{\mathbf{CL}^a} C$ or
$\sim A, B_1, \ldots, B_n \vdash_{\mathbf{CL}^a} C$ or
$A, B_1, \ldots, B_n \vdash_{\mathbf{CL}^a} \sim C$ or
$\sim A, B_1, \ldots, B_n \vdash_{\mathbf{CL}^a} \sim C$,
and $B_1, \ldots, B_n \nvdash^X_{\mathbf{CL}} C$ and
$B_1, \ldots, B_n \nvdash^X_{\mathbf{CL}} \sim C$,
then $?A, B_1, \ldots, B_n \vdash^X_{\mathbf{CL}^q} ?C$.

QA Where $n \geq 0$,

if $?C$ occurs in the proof
and $B_1, \ldots, B_n \vdash_{\mathbf{CL}} C$ (respectively $B_1, \ldots, B_n \vdash_{\mathbf{CL}} \sim C$),
then C (respectively $\sim C$) may be added to the proof.

[22] One may strengthen this, for example by requiring that no known decision method—remember that many fragments of **CL** are decidable—answers $B_1, \ldots, B_n \vdash^X_{\mathbf{CL}} A$ in the negative. In our view, however, our characterization in the text is the appropriate one. On this, we definitely agree with Hintikka. If a question can be (easily) answered by observation, to do so may be preferable over attempting to derive the answer, especially as the attempt may be longwinded and unsuccessful.

Let us start with some general comments. QD concerns yes-no questions. We could introduce a similar rule for wh-questions, but will not do so in the present paper. Next, the consequence relation $\vdash^X_{\mathbf{CL}^q}$ is a pragmatic one. What we are interested in, in the present context, is whether it is justified for the explanation seeking person or group X to derive a question in view of X's relevant logical insights. Finally, as $\vdash_{\mathbf{CL}^a}$ and $\nvdash^X_{\mathbf{CL}}$ are defined only for formulas that belong to the language of **CL**, no "?" occurs in A, $B_1, \ldots, B_n$ or C in the above rule. Remark also that all rules of $\mathbf{CL}^a$ are valid in $\mathbf{CL}^q$; the rule QD extends $\mathbf{CL}^a$ to handle questions.

Rule QA simply enables one to derive answers to questions (that occur in the proof) even if the answers cannot be obtained from the descriptive premises by analysing rules. Thus if $?(p \vee q)$, then $p \vee q$ may be derived from p, and $\sim(p \vee q)$ may be derived from $\sim p$ and $\sim q$.[23]

We now move to some properties of $\mathbf{CL}^q$. We have

$$?(A \vee B) \vdash^X_{\mathbf{CL}^q} ?A$$

because

$$\sim(A \vee B) \vdash_{\mathbf{CL}^a} \sim A\,.$$

However, $?p \nvdash^X_{\mathbf{CL}^q} ?(p \vee q)$ because there is no $\mathbf{CL}^a$-proof of either $p \vee q$ or $\sim(p \vee q)$ from either p or $\sim p$. Similarly, $?p \nvdash^X_{\mathbf{CL}^q} ?(p \& q)$. Remark also that $?A$ and $?{\sim}A$ are variants of the same question. A positive answer to $?A$ means that A is the case, a negative one that $\sim A$ is the case. A positive answer to $?{\sim}A$ means that $\sim A$ is the case, a negative one that $\sim\sim A$ and hence A is the case.

Quite different comments concern the intuitive justification of our logic of questions. First, $\mathbf{CL}^q$ concerns only deriving questions from other questions together with descriptive premises. Actually, this problem is an extremely important one. If one takes serious the idea that science is problem driven, then one has to explicate the reasoning about and from problems. Accepted theories and data play an essential role in this. Consider the question $?p$. If we know that $q \supset p$, then the question $?q$ is useful for answering $?p$—if the answer to $?q$ is positive, so is the answer to $?p$. If we know that $p \supset q$, then again the question $?q$ is useful for answering $?p$—if the answer to $?q$ is negative, so is the answer to $?p$.

Remark that the logic $\mathbf{CL}^q$ presupposes that a question is useful as soon as one of its answers is. To see this, suppose that $?A$ is derived from $?(A \vee B)$. If the answer to $?A$ is positive, we have an answer to $?(A \vee B)$. If the answer to $?A$ is negative, we still need an answer to $?B$ in order to have an answer to $?(A \vee B)$. Here, even a negative answer to $?A$ takes us a step forwards in that $?(A \vee B)$ is reduced to $?B$. But this is not always the case. If $?A$ is derived from $?B$ and $A \supset B$, a positive answer to $?A$ provides an answer to $?B$, but a negative answer to $?A$ is useless for answering $?B$. Nevertheless, if we cannot

[23]It can be shown (by relying on QD) that the derivation of answers by non-analysing rules cannot lead to the derivation of irrelevant questions (see below) because the presence of $p \vee q$ or $\sim(p \vee q)$ cannot lead to questions that were not derivable in the presence of $?(p \vee q)$.

obtain an answer to $?B$ in a direct way, and we can obtain one to $?A$, then it is clearly sensible to derive the latter question.

Summarizing what we have so far, $\mathbf{CL}^q$ enables one to 'split up' questions into relevant subquestions, and to derive questions from other questions in view of empirical knowledge. Remark, however, that the relevance of a question to another question is not transitive. Thus from $?A$, $B \supset A$, and $B \supset C$, one may first derive $?B$ and next $?C$. But no answer to $?C$ provides (for all the premises reveal) an answer to $?A$, which was the question we started from. This may seem problematic and hence deserves some more attention.

In the example, one of the answers to $?C$, viz. $\sim C$ provides an answer to $?B$, viz. $\sim B$. So, although $?C$ is not useful for directly answering $?A$, it is indirectly useful in this connection. It reveals that one possible way for obtaining a positive answer to $?A$, viz. from $?B$, fails, and hence that we have to look for another way to obtain an answer to $?A$. The use of this insight becomes even more clear if the descriptive premises are somewhat more complex. Consider the premises $?A$, $B \supset A$, $D \supset B$, $E \supset D$, and $B \supset C$. From this we may derive the questions $?B$, $?D$ and $?E$, and a positive answer to any of them provides a positive answer to $?A$. But if one cannot obtain an answer to any of these derived questions in a direct way, but obtains a negative answer to the derived question $?C$, then we know at once that, in order to answer $?A$, it is useless to derive further questions from $?B$, $?D$ or $?E$. So, a whole branch of possible questions (in other cases a whole subtree) is 'cut away' by a negative answer to $?C$. This suggests part of the strategy that we shall apply in the next section. Given a question for which no answer can be obtained in a direct way, we shall first try to derive a question for which an answer can be obtained in a direct way. If that is impossible, we shall derive questions from which in turn may be derived questions for which an answer can be obtained directly, etc. In general, we shall look for the shortest paths that lead to questions for which we can obtain an answer directly.

$\mathbf{CL}^q$ enables one to derive questions that are relevant for answering derived questions, but are not necessarily relevant for answering a starting question. Precisely for this reason, it is essential that we impose heavy restrictions on the deductive closure of the descriptive premises, viz. close them under $\mathbf{CL}^a$ rather than $\mathbf{CL}$. Let us see why this is so. Suppose first that we replace $\mathbf{CL}^a$ by $\mathbf{CL}$ in QD. Then, from $?A$ would follow $?(A \vee B)$, and from this $?B$. In other words, we could derive any question from any other question in two steps. Similar problems arise if we leave QD unchanged, but add it to $\mathbf{CL}$ rather than to $\mathbf{CL}^a$.

Some will argue that $\mathbf{CL}^a$ does not allow one to derive all relevant questions. It may be impossible to obtain an answer to $?A$ directly, and possible to obtain an answer to $?(A \vee B)$ directly; and a negative answer to $?(A \vee B)$ entails a negative answer to $?A$. By this reasoning, it seems that $?(A \vee B)$ should be derivable from $?A$ (at least for some B). We think that this argument does not hold water. If (first case) a negative answer to $?(A \vee B)$ is derivable from the descriptive premises, then so is a negative answer to $?A$. And if (case 2) there are observational means that provide a negative answer to $?(A \vee B)$ directly (and not to $?A$), then there is something one can observe, or some measurement that leads to a certain state of a measuring instrument, or some test the outcome of which can be observed, and this something or state or outcome will have a name in any language that is not utterly artificial. So, there will be a C from which one may conclude $\sim(A \vee B)$ and hence $C \supset \sim A$ will be among the descriptive premises or will be derivable from them by $\mathbf{CL}^a$.

5 Deriving the Right Questions

In Section 3, we have presented the means to arrive at possible initial conditions. These provide one with the starting questions, which one will try to answer by relying on one's full knowledge system. If one of the starting questions is answered positively, one has located an initial condition that holds true, and hence has found an explanation. From the starting questions (and one's knowledge) one will derive further questions, aiming at a set of questions that can be answered by observation and experiment, or by deduction from one's knowledge. This typical pragmatic process is often neglected in the literature. This is a pity, especially as the process is in general radically different from seeking possible initial conditions. Indeed, it almost never proceeds in terms of the explaining theory (premises 1–4 in the simple example from Section 3). Nearly always, different available knowledge has to be invoked, and *all* available knowledge *may* be invoked. To answer a starting question, one may have to use instruments (and hence rely on the theories describing their use), one may have to set up experiments, relying on theories that may have nothing in common with the theory used in the explanation, etc.

So, in finding out whether a possible initial condition is true, one will rely on one's whole knowledge system. There is one proviso. It is well-known and already occurs in [16]: we should obtain the answers in a way that is independent of establishing the truth of the *explanandum*. This is a simple and obvious matter and we shall not comment on it any further.[24]

Let us move on to our logic, which (for reasons that will soon appear) will be an adaptive logic. It extends $\mathbf{CL}^q$ but is very different from it. Clearly, we need a special premise-like rule for starting questions. If, for example, Pa is a possible initial condition for explaining Qa, we introduce the starting question as follows:

i	$?Pa$	SQ	$\langle Pa, Qa\rangle$

The far right of this line contains its *condition*. It is a couple of formulas (in real life, statements) and suggests the following reading: an answer to the question (here $?Pa$) is relevant (and in this case exhaustive) for establishing the

[24] Actually, the proviso has a weak and a strong version. The weak version requires that the *explanandum* is not used to establish the truth of the initial condition. The strong version requires that there is a set of empirical results by which the truth of the *explanandum* is established, and that no element in this set is used to establish the truth of the initial condition.

truth or falsehood of Pa, which is a possible initial condition for explaining Qa. Conditions as the one of line i should not be confused with the conditions we met in Section 3. The present conditions will play a central role for *marking* certain lines of the proof. That a line is marked at a stage of the proof will mean that the question derived at that line is not any more useful for answering the connected why-question. The why-question may be read off from the second element of the condition of the line—for line i above: why-Qa.

There is no need to remember, for the sake of the present proofs, the theory that should be combined with (in our example) Pa in order to obtain an explanation of Qa. This theory is identified in the proofs introduced in Section 3. The present endeavour is directed at obtaining an answer to $?Pa$. We do, however, have to remember that we are looking for an explanation of Qa. One reason is that, in establishing the truth or falsehood of Pa, we have to take into account the proviso mentioned in the second paragraph of this section, viz. that one should not rely on the truth of Qa in order to establish the truth of Pa.

There is a further reason for explicitly mentioning the *explanandum* in the condition. Our aim, in the present section, is not merely to apply $\mathbf{CL}^q$, but rather to define an adaptive logic (that extends $\mathbf{CL}^q$ and) that serves a specific purpose, viz. to derive only those questions that are useful for arriving at an explanation. Thus if it is found out, either by deductive means or from new premises (answers to questions obtained from observation and experiment) that Pa is false, then the above listed starting question proves useless, at that point in time, for explaining Qa. The same applies to all questions derived from $?Pa$ (in any number of steps). Once Pa is established to be false, no answer to any derived question will change this. By way of bookkeeping, we shall mark all those lines, viz. all lines that have $\langle Pa, Qa\rangle$ as their condition. As we shall need several marks, these ones will be called E-marks.

Suppose that there are several possible initial conditions for Qa (as is the case in the example from Section 3), and hence several starting questions. As soon as one of the starting questions is answered positively, there is an explanation of Qa, and hence it is not useful to continue searching for answers to the other starting questions (that concern the *explanandum* Qa).[25] Even the starting question that has been answered in the positive ceased, by this very fact, to be a useful question for finding an explanation Qa. It should not provoke any empirical research or any attempt to derive one of its answers by logic. In other words, as soon as an explanation is available, *all* questions relating to the same *explanandum* should be marked. The marks indicating this will be called A-marks.

Clearly, whether a question is useful with respect to a specific explanatory aim, depends on one's logical insights and factual knowledge. Both of these change over time, the first by reasoning, the second by observation and experiment. This is why the use of a question should be judged at a *stage* of our inquiry. The latter can safely be identified with the stage of our proof,

[25]There is nothing wrong, of course, with searching for several explanations for Qa, but this is not the standardly pursued aim and we shall stick to the latter.

because the proof contains all relevant deductions as well as all new empirical data (which are introduced as new premises). This shows that we need a third kind of marks, which we shall call Q-marks. As soon as a question is answered, it should not be pursued any more. So, lines containing answered questions will be Q-marked.[26] We shall see later that our proof format is very useful for deciding whether a question is Q-marked or not.

We should add a final warning before presenting the promised logic. This logic does not improve on the knowledge system (of a person or group X). If this knowledge systems contains $(\forall x)((Rx \& Sx) \supset Px)$ as well as $(\forall x)((Rx \& {\sim}Sx) \supset Px)$, then $(\forall x)(Rx \supset Px)$ is derivable from it, and hence, while $?Ra$ is relevant for $?Pa$, $?Sa$ actually is not. But if the knowledge system is given in this clumsy way, our logic will not prevent one from deriving $?Sa$ (in two steps) from $?Pa$. The only way to prevent this (in an absolute sense), is by referring to derivability (from the knowledge system). But as **CL** is undecidable, any such reference will lead to a nice systematic definition, but not to a useful pragmatic tool.[27]

By now the reader should be sufficiently prepared to have a look at the (generic) rules of the adaptive logic **Q**. We have a Premise rule, a (conditional) rule for introducing starting questions, and an unconditional rule for other derivations.

PREM If $A \in \Gamma$, then one may add to the proof a line consisting of

(i) the appropriate line number,
(ii) A,
(iii) "–",
(iv) "PREM", and
(v) $\emptyset$.

SQ Given a possible initial condition B for an explanation of A, as located in an explanation seeking deduction, one may add to the proof a line consisting of

(i) the appropriate line number,
(ii) B,
(iii) "–",
(iv) "SQ", and
(v) $\langle B, A \rangle$.

RU If $A_1, \ldots, A_n \vdash_{\mathbf{CL}^q} B$ $(n \geq 0)$, and $A_1, \ldots, A_n$ occur in the proof on the conditions $\Delta_1, \ldots, \Delta_n$ respectively,[28] then one may add to the proof a line consisting of

(i) the appropriate line number,

[26] Listing the question in the condition (which would then be a triple) would involve a rather useless repetition (even if it would agree with the standard proof format of adaptive logics).

[27] It is not difficult to devise an adaptive logic that eliminates irrelevant 'connections' from the knowledge system. We fear, however, that another complication might scare off the reader.

[28] If no A_i is a question, all Δ_i will be empty. Moreover, it is obvious in view of rule QD that it is useless to include several questions among the A_i. So, if one proceeds sensibly, at most one of the Δ_i will be non-empty.

(ii) B,
(iii) the line numbers of the A_i,
(iv) "RU", and
(v) $\Delta_1 \cup \ldots \cup \Delta_n$.

As explained above, the marking of lines in a proof is governed by a Marking Definition. We shall, as promised, introduce three such definitions. The first eliminates starting questions that received a negative answer. It ensures that a line with condition $\langle B, A \rangle$ is marked if it turns out that B is not a possible explanation of A. The second eliminates all questions aiming at finding an explanation of A as soon as such explanation is available. The third Marking Definition eliminates all questions that have been answered (whether positively or negatively).

Definition 1 *A line is E-marked at stage s iff, where $\langle B, A \rangle$ is its fifth element, $\sim B$ is unconditionally derived in the proof at stage s.*

Definition 2 *A line i is A-marked at stage s iff, where $\langle B, A \rangle$ is its fifth element, some line j in the proof has $\langle C, A \rangle$ as its fifth element, and C is unconditionally derived in the proof at stage s.*

Definition 3 *A line i is Q-marked at stage s iff, where $?A$ is its second element, either A or $\sim A$ is unconditionally derived in the proof at stage s.*

In view of these definitions, we define two forms of derivability as is usual for dynamic proofs. A line will be called marked iff it is E-marked, A-marked or Q-marked. A formula is derived at stage s of a proof from Γ iff it is the second element of a line of that proof and the line is non-marked at stage s. A formula is *finally derived* at line i of a proof at a stage from Γ iff it is the second element of line i, line i is non-marked at that stage, and will not be marked in any extension of the proof. A formula is *finally derivable* from Γ iff it is finally derived at some line in a proof at a stage from Γ.

Definition 4 $\Gamma \vdash_{\mathbf{Q}} A$ *iff A is finally derivable from Γ.*

All descriptive statements derived in the proof are finally derived. The final derivability of a question indicates that none of its answers is derivable from the premises[29] that were introduced, that its starting question was not answered, and that no explanation was obtained for the *explanandum* it is related to. Whether more data (some of them answers to questions) are obtained later is a different matter.

In most adaptive logics, we are interested in the finally derivable consequences. This is the reason for inserting the usual definition above. But here the matter is rather different: if the search for an explanation (of one or more

[29] We mean the original premises (one's knowledge system at the start) as well as the 'new' premises (answers to questions).

explananda) is successful (in that, for each of them, one of the starting questions was answered in the positive), all lines at which questions were derived are marked. This indicates that no empirical inquiry and no deductive job is left. Of course, the marked lines are still there. From them one may read off the explanations.

As soon as a non-starting question $?A$ is answered, it will be clear from the proof whether this has an effect on the question from which $?A$ was derived. Thus, if the answer is A, and $?A$ was derived from $?(A \vee B)$, then one simply derives $A \vee B$ from A (by rule QA). If the answer is A and $?A$ was derived from $?B$ and $A \supset B$, then one derives B (by Modus Ponens, which is an analysing rule). In other words, if an answered question is derived with the aim to answer another question, then it will immediately be clear from the proof whether an answer to the 'premise' question is derivable from the answer of the derived question.

The natural strategy[30] for **Q**-proofs was already discussed in the previous section. If no direct observational answer for a question is available, we shall either engage in the attempt to derive the answer from the descriptive premises, or look for the shortest paths that bring us to a question for which an answer can be obtained directly from observation. Roughly, the method for doing so is similar to the strategy described in Section 3. Apart from this, the reader will remember what was said in the previous section on 'cutting off' subtrees from the tree of questions.

6 Putting Things Together

That the explanation seeking deduction was separated from the derivation of questions is rather artificial, and actually was only introduced here for reasons of exposition. Putting both kinds of proofs together is simple enough, provided one makes clear that the theory from which one tries to obtain one or more initial conditions, is clearly distinguished from the rest of our knowledge. Thus, the lines of the proofs introduced in Section 3 may be intermingled with the lines of the proofs from Section 5, provided one, for example, attaches a subscript to each formula of the former proof.

We introduce this change to stress that there is no need to derive all possible initial conditions (with respect to the explaining theory) before starting to derive questions. If a possible initial condition is found, one may at once introduce the starting question, and derive further questions from it. Questions will of course not be given a subscript. This would be useless, as their purpose is obvious in view of their condition and of the context.

Moreover, there is nothing wrong with deriving initial conditions that pertain to different explanatory theories, or even to different *explananda*. Quite to the contrary, this will only make it more transparent that answering some derived

[30]Given the complexity and the novel character of the adaptive logic **Q**, we shall not attempt to push the strategy in the proof itself this time.

questions is necessary (or sufficient) for explaining a phenomenon in one or several ways, or for explaining several phenomena.

There is no danger that logics get mixed up by deriving both initial conditions and questions within the same proof. The conditional lines that contain or lead to possible initial conditions, are **CL**-lines in disguise—we have sufficiently stressed that above. They are all clearly distinguished by a subscript, as are all other consequences of the explaining theories. Moreover, it is easily verified that all descriptive formulas derived in those proofs are $\mathbf{CL}^a$-consequences of the premises. Hence, they may be invoked for deriving questions. The conditional lines that contain questions bear a somewhat looser relation to **CL** (that we shall not try to characterize here), but then their form and their function in the proofs is sufficiently distinct to avoid all confusion.

7 Some Further Pragmatic Aspects

In the two previous sections, we have tacitly supposed that, if an answer to a question can be obtained by observation or experiment, then this answer is obtained and added to the proof. However, this supposition unjustifiably reduces important pragmatic aspects of explanation to unproblematic side-aspects of the search for explanations. Given a set of questions that can be answered by observation or experiment, one faces a serious decision problem: Which observations and experiments will be carried out? Clearly, this cannot be settled by logic. Nevertheless, the logic **Q** provides an important tool for reaching a decision.

The central considerations for making the decision concern available means. Some questions may be answered by observation, provided one is in the right spot at the right time, or one is able to get there. Other questions may be answered by the outcome of an experiment, provided one has (or can afford to provide oneself with) the required materials, instruments, competence to handle the instruments, etc. It is not difficult to see that economic considerations are essential here: buying objects, hiring people, and getting the funding for it.

Sometimes, the means to answer the questions are standard. Often, and especially for the most interesting questions, they are not. Galileo's experiment to 'discover' the laws of the free fall is an excellent example.[31] There were no stopwatches and no electronic eyes. The ingenuity by which he managed, relying on heavy theoretical presuppositions, to rephrase the problem of the free fall in such a way that the inclined plane experiment would answer it, and next managed to combine ears and eyes and hand and a flow of water to measure time, can only provoke awe. The *reasoning* that is behind this ingenuity is heavily relying on logic, but unfortunately is wholly beyond the scope of the present paper. Nevertheless, it is quite clear that generating questions is not the toughest problem.

[31] This question is not so different from the ones we have been dealing with in the present paper—see Section 8.

Consider a scientist X and his available (economic and intellectual) means. Given that X has derived all (or some of) the right questions, which decision should X take with respect to observation and experiment? The answer is less simple than one might expect. Suppose that $Pa \vee Ra$ and Sa are possible explanations of Qa, that the questions $?Pa \vee Ra$ and $?Sa$ were introduced, and that the questions $?Pa$ and $?Ra$ were derived from $?Pa \vee Ra$. Suppose moreover that there is no direct test for answering $?Pa \vee Ra$, that the test for $?Pa$ is inexpensive, the one for $?Sa$ expensive, and the one for $?Ra$ extremely expensive. If the inexpensive test for $?Pa$ delivers a positive answer, the problem is solved. However, if its answer is negative, one will presumably opt for a test for $?Sa$. If that one is positive, the first test was, with hindsight, useless.

As the decision problem may be complex, it is necessary to have a clear view on its different aspects, and precisely this may be provided by the logic **Q**. As we have seen before, the proof makes explicit (in as far as this is not explicit from the structure of the questions themselves), the logical relations that hold between the answers to the questions. This may easily be turned into a search tree, from which one may read off the use of a positive or negative answer to a derived question for answering a starting question (including cutting off paths for answering it). This illustrates the importance of deriving questions from the starting questions. The search path (or paths) for answering the why-question may contain several bifurcations. Each of these requires that one or more derived questions are answered. Let us stress again that this logical information does not in itself enable one to make the best decision with respect to observation and experiment. Nevertheless, it provides the required paths on which economic and similar considerations may be applied.

We briefly comment on the advantage of pursuing several why-questions at the same time (possibly even in the same proof). With a few (sometimes notorious) exceptions, scientists work on clusters of connected problems rather than on a single problem. This is to some extend a matter of taking opportunities—a result that is useless for one problem may be useful for another problem. Often, however, it is also more efficient to work on several problems at the same time (it makes one more alert, causes one to have more ideas, etc.) The matter is not different for why-questions. From a survey of the questions that pertain to different explanations, it may appear that some question is central for several explanations. Even if answering it requires an expensive experiment, the fact that its outcome is crucial for several explanations may influence the decision to stage the experiment.

8 Conclusion and Open Problems

It goes without saying that the logics we have presented require a decent metatheoretic study and need to be extended in several respects. And we hope to have convinced the reader that the logic deserves that metatheoretic study and those extensions. As the first will not interest most of the readers of the present *Festschrift*, we briefly comment on the extensions, especially on extensions that are necessary in view of the claims made by Matti Sintonen in writing.

Our logic of explanation allows only for singular questions and for singular new premises (answers to questions). But clearly, the search for explanations may lead to the formulation (and acceptance) of new theories. Sintonen points this out in [27]. There is no need, of course, for having a single logic that takes care of both kinds of problems. This is fully in line with Hintikka and Halonen's warning, in [19, p. 196], for a mixture of theorizing and explanation. Moreover, one should distinguish between obtaining some relevant empirical generalizations, and obtaining theories that may be highly abstract and theoretical in nature. It seems sensible to require that the search for an explanation should, in simple cases, provide one both with an initial condition and with the explanatory generalizations. But to devise a full blown theory is clearly a matter of theorizing alone. Of course, it is related to why-questions. However, it never concerns the explanation of some singular statement, but the explanation of whole sets of connected singular statements, and the explanation of generalizations.

So, while real theorizing should remain outside the picture, it is not difficult to see how one may extend our logic in such a way that both an initial condition and some required generalizations are obtained. Indeed, an adaptive logic of induction is available—see [9]. It handles background knowledge, and generates and rejects new generalizations, and it does all this exactly as it should.

This logic of induction does not allow one to *withdraw* premises (theories or data), except when they are falsified by the data. This clearly is a disadvantage. Of course, one might introduce a bracketing rule, where bracketed premises are those for which one has a reason to (were it temporarily) withdraw them. We think, however, that such reasons must be made explicit, and that it often is the result of (possibly complex) reasoning, which also must be made explicit. The simplest reason to withdraw A is that, after A was introduced, one obtains direct empirical evidence for $\sim A$. (We refer to situations in which $\sim A$ is a direct conclusion of experience.) On the adaptive approach, $\sim A$ will be added to the premises, and an index will be attached to each premise, where the indices indicate the weight of the evidence for a premise. The adaptive logic should make sure that A and its consequences are neither derivable nor have any effect on the conclusions that still are derivable. If there is no direct empirical evidence for $\sim A$, A is withdrawn because other data or theories were obtained and either contradict A or make it doubtful. The adaptive logic should again prevent A from having any effect on the consequence set. Moreover, in this case it should make explicit the reasoning that leads to the rejection of A. Such adaptive logics are available: see [31] for some inadequate solutions of the problem and [30] and [7] for some adequate ones.

Two further complications concern (i) the case in which a theory turns out to be inconsistent, but no consistent and empirically satisfying alternative is available, and (ii) the case in which a theory is falsified but no empirically satisfying alternative is available. The first problem is amply documented in the literature (for example [14], [25], [26], [29], [24], [22] and [23]) and the second receives growing attention (for example [21]). To extend our logic in such a way that it can cope with those complications is rather simple in view of the results on inconsistency-adaptive logics (actually, the best studied type of adaptive logics). The extension may be carried out along the lines of [8], which extends Hintikka's theory of the process of explanation in such a way that it can handle inconsistent data and theories.[32]

References

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

[2] Diderik Batens. Natural heuristics for proof construction. Part I: Classical propositional logic. *Logique et Analyse*, 127–128:337–363, 1989. Appeared 1992.

[3] Diderik Batens. *Logicaboek. Praktijk en theorie van het redeneren.* Garant, Leuven/Apeldoorn, 1992.

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

[5] Diderik Batens. Dynamic semantics applied to inconsistency-adaptive logics. In *Logical Investigations*, volume 5, pages 74–85. Moscow, "NAUKA", 1998.

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

[7] Diderik Batens. Aspects of the dynamics of discussions and logics handling them. *Logical Studies*, in print.

[8] Diderik Batens. The theory of the process of explanation generalized to include the inconsistent case. *Synthese*, in print.

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

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

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

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

[12] Diderik Batens and Joke Meheus. Shortcuts and dynamic marking in the tableau method for adaptive logics. *Studia Logica*, in print.

[13] George S. Boolos and Richard J. Jeffrey. *Computability and Logic*. Cambridge University Press, 1989. (Third edition).

[14] Bryson Brown. How to be realistic about inconsistency in science. *Studies in History and Philosophy of Science*, 21:281–294, 1990.

[15] Ilpo Halonen and Jaakko Hintikka. Toward a theory of the process of explanation. *Synthese*, this volume.

[16] Carl G. Hempel. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. The Free Press, New York, 1965.

[17] Jaakko Hintikka. *Inquiry as Inquiry: A Logic of Scientific Discovery*. Kluwer, Dordrecht, 1999.

[18] Jaakko Hintikka. Is logic the key to all good reasoning? In *Inquiry as Inquiry: A Logic of Scientific Discovery* [17], pages 1–24.

[19] Jaakko Hintikka and Ilpo Halonen. Semantics and pragmatics why-questions. In *Inquiry as Inquiry: A Logic of Scientific Discovery* [17], chapter 9, pages 183–204.

[20] Jaakko Hintikka, Ilpo Halonen, and Arto Mutanen. Interrogitive logic as a general theory of reasoning. In *Inquiry as Inquiry: A Logic of Scientific Discovery* [17], pages 47–90.

[21] Theo A. F. Kuipers. *From Instrumentalism to Constructive Realism*, volume 287 of *Synthese Library*. Kluwer, Dordrecht, 2000.

[22] Joke Meheus. Adaptive logic in scientific discovery: the case of Clausius. *Logique et Analyse*, 143–144:359–389, 1993. Appeared 1996.

[23] Joke Meheus. Inconsistencies in scientific discovery. Clausius's remarkable derivation of Carnot's theorem. In Geert Van Paemel and *et al.*, editors, *Acta of the XXth International Congress of History of Science*. Brepols, in print.

[24] Nancy Nersessian. Inconsistency, generic modeling, and conceptual change in science. In Joke Meheus, editor, *Inconsistency in Science*. Kluwer, Dordrecht, in print.

[25] John Norton. The logical inconsistency of the old quantum theory of black body radiation. *Philosophy of Science*, 54:327–350, 1987.

[26] John Norton. A paradox in Newtonian gravitation theory. *PSA 1992*, 2:421–420, 1993.

[27] Matti Sintonen. In search of explanations: From why-questions to Shakespearian questions. *Philosophica*, 51:55–81, 1993.

[28] Matti Sintonen. Why questions, and why just why-questions? *Synthese*, 120:125–135, 1999.

[29] Joel Smith. Inconsistency and scientific reasoning. *Studies in History and Philosophy of Science*, 19:429–445, 1988.

[30] Liza Verhoeven. Changing one's position in discussions. some adaptive approaches. *Logic and Logical Philosophy*, in print.

[31] Liza Verhoeven. Proof theories for some prioritized consequence relations. To appear.

[32] Andrzej Wiśniewski. *The Posing of Questions. Logical Foundations of Erotetic Inferences.* Kluwer, Dordrecht, 1995.