Searching for Singular Causal Explanations: a formal analysis

Leen De Vreese and Erik Weber
Centre for Logic and Philosophy of Science
Ghent University, Belgium
Leen.DeVreese@UGent.be, Erik.Weber@UGent.be

August 29, 2003

1 Introduction

In order to clarify the aim of this paper, we must distinguish two types of causal reasoning. The first includes reasoning on the basis of statistical information and aims at discovering causal relations between variables or between certain types of events. The second type is the search for *singular* causal explanations, e.g. of singular happenings in history such as disasters, accidents, once-only political events such as putsches, ... In this case, the aim is to find an explanation of a particular event. The general causal knowledge (i.e. knowledge about causal relations between types of events or variables) is supposed to be present at the start of reasoning processes of the second type. The task is to determine which potential causes were present in the case at hand.

Van Dyck has developed the adaptive logic for causal discovery **ALCD** [5], trying to capture the first type of causal reasoning. Following among others Judea Pearl [7], he took the causal Markov condition as his most important assumption, leading to the derivation of direct causal relations between variables/types of events. In [9], we applied this logic to a case-study, testing the usefulness of **ALCD** in capturing the dynamics in scientific reasoning towards an explanation on the basis of statistical information. The logic **ALCD** made it indeed possible to execute a thorough analysis of the involved scientific survey concerning the relationship between parental separation and smoking initiation in adolescents [6].

The dynamics in the reasoning towards a singular causal explanation is inclined to be still bigger, because the knowledge on which people base themselves to come to conclusions is often less firm in comparison with the statistical information at the basis of **ALCD**. This means that it is not obvious to capture this kind of reasoning in a logic which is general enough to be applicable to a wide range of cases. The more because a lot of non-formal

elements need to be taken in consideration. However, we are convinced that the big lines and basic strategies are uniform enough to be put in a logic. Nonetheless, our purpose here is not to present a completely developed and well-defined logic. That's the long-term purpose, in the build-up to which we are working on the formal analysis of the reasoning process in particular case-studies.
We will analyse here in more detail a couple of parallel and linked search processes for a singular causal explanation, namely the search for an explanation of a series of similar shipdisasters during the first World War. From this, we hope to deduce some logical rules and general patterns, used in the reasoning towards singular causal explanations.
In what follows, we will first give a short introduction to adaptive logics in section 2. Then we start the analysis of the reasoning towards an explanation of the ship disasters. In section 3, we analyse the case in which only one possible explanation fills people's mind. In section 4, we turn to the case in which several possible explanations are considered. In the final section, we summarize the results we obtained.

2 Basic Features of Adaptive Logics

Adaptive logics [1] were proposed by Diderik Batens as a solution for the problems brought about by inconsistencies. Instead of carrying us to triviality — as inconsistencies do in CL by the property of *ex falso quodlibet* —, or instead of pushing us toward the disqualification of some classical inference rules, — as happens in the rather weak paraconsistent logics —, inconsistent premisses can be interpreted as consistent as possible by applying the adaptive logics. The idea is that not the inference rules of classical logic themselves should be dropped, but those applications thereof that lead to triviality. Each adaptive logic consists of three basic components: a *lower limit logic* (henceforth **LLL**), an *upper limit logic* (**ULL**), and a *marking strategy*. The **LLL** contains all the unproblematic rules, and can be applied unconditionally. The **ULL** is made up of the rules of the **LLL** *plus* the problematic rules. The important feature of the **ULL** rules is that they are applied *conditionally* in an adaptive proof. The conditions always consist of a set of formulas that have to behave "normally" for the specific **ULL** rules to be applicable. Any application of an **ULL** rule can be retracted *whenever* it turns out that the condition is violated and thus that the specific set of formulas do not behave normal. When this turns out to be the case, the sentence that was derived on the condition of normality is *marked* as invalid in accordance with the marking strategy, stating when marking has to occur. Of course, if the sentences to which we applied unconditional rules

[1] For a thorough description of these logics, we refer to [1] and [3]. This summary is based on [5]

were derived on a condition, then the sentence that we introduce at a new line in the proof will also contain a condition: the union of all the conditions of the earlier sentences of which it is a consequence (if an earlier line has to be considered invalid at a certain point of the proof, all its consequences obviously also have to be). All these elements make up for a non-monotonic logic with a dynamic character (some lines that were derived in a proof can be invalidated in a later stage of the proof; the other way round marked lines can even become unmarked later on in the proof). Each line of an adaptive proof consists of five elements:

1. a line number,
2. the sentence derived,
3. the line numbers of the sentences from which (ii) is derived,
4. the rule of inference that justifies the derivation,
5. the set of sentences on the normal behaviour of which we rely in order for (ii) to be derivable by (iv) from the sentences of the lines enumerated in (iii).

Besides a structural rule by which you can introduce premises in a proof (always with an empty fifth element), there are two kinds of inference rules: unconditional ones (being all valid applications of the **LLL** rules), and the conditional ones (all the applications of the extra **ULL** rules). The fifth element of a line, together with the marking strategy, determines when that line will have to be marked in view of the other lines written down in the proof.

It is clear that the search for a singular causal explanation implies a dynamical way of reasoning whereby explanations are considered, afterwards rejected, and/or the reverse. The more information we get, the better we will be able to discern between possible explanations. Furthermore, it seems almost evident that we will be confronted from time to time with inconsistencies during our search. Adaptive logics are developed precisely for this kind of patterns in human reasoning. Therefore, we will take the adaptive logic principles as the basis for our formal analysis.

It is possible to make a pragmatical distinction between corrective adaptive logics and ampliative adaptive logics. Corrective logics allow for logical abnormalities, but interpret a set of premises as much as possible in agreement with the intended standard of deduction. This standard of deduction is usually **CL**. The first adaptive logics developed, e.g. the inconsistency-adaptive ones, were corrective. The logical rules proposed in this paper are to the contrary of an ampliative kind. Ampliative adaptive logics assign to a set of premises a consequence set that extends the consequence set assigned to the premises by the standard of deduction (CL). However, it does not concern

a systematic distinction here, but one based on a non-logical criterion: the choice of a standard of deduction. What's more, some adaptive logics are both corrective and ampliative.

3 What happened to the Bulwark? Reasoning with one possible explanation

Novembre the 26th 1914, early in the morning. In the mouth of the Medway river, a fleet of impressive battle ships is ready to oppose maritime attacks on Great Britain. At a couple of minutes to eight, one of the ships is suddenly surrounded by flames, without the slightest previous warning. Next, a big explosion was heard for ten miles around. Immediately, people began speculating about what had happened to the battleship named the Bulwark. Was it an accident or the result of sabotage? Till know, the cause(s) are unclear.[4]

3.1 Basic, but uncritical patterns of thinking

3.1.1 General question for an explanation

Confronted with this disaster, the search for an explanation started immediately. This spontaneously arising why question will form the starting point of our analysis. For the analysis of the reasoning patterns, we use Diderik Batens *Notes on problem solving* [2], which is on his turn based on goal-directed proofs [8].
The initial question is introduced in our formal analysis by means of the Main Target Rule. This Main Target will be of the form $?E(X_\alpha)$. This formalizes the yes/no-question if their can be found an explanation for the event involving α. In our example we have:

1 $?E(X_b)$ – $Main$ $E(X_b)$
(b=Bulwark)

This is an application of the first rule that should be included in our logic:
Main: At the very beginning of every proof we put a first line consisting of:

(**1**) Line number 1,

(**2**) $?E(X_\alpha)$,

(**3**) "-",

(**4**) $Main$,

(**5**) $E(X_\alpha)$.

A word about the fifth element of the line, namely the condition: If the condition holds, the question is answered. In other words, once an explanation has been found, the question of line 1 is answered and can be marked (we will explain more about the marking rules later in this paper).

3.1.2 Guessing for an explanation

At the first moment, men could only guess at the cause of the explosion of the Bulwark.[4] In such cases, when nothing about (the) possible cause(s) is known yet, people will search in an undirected way. People will guess for an explanation. For example, they conjecture it was a fire on board that caused the loss of the ship, or an attack from outside. The only thing we can rely on for such conjectures is our general background knowledge. Because in this situation, no information is added, we will be left with the open questions. This situation can be expressed as follows.

$\underline{\mathbf{F} \rightarrow \mathbf{E}}$
$?F_b$
(F= fire, b=Bulwark, E=explosion)

The $\rightarrow$ points to a direct causal relation. We use the capital letters in bold to refer to types of events. So, $\mathbf{F} \rightarrow \mathbf{E}$ expresses a potential causal relationship: the supposed type of cause is able to bring about an event of the type to which our target event belongs. (More specifically: the background supposition that a fire is able to cause an explosion similar to the explosion of the Bulwark).
In everyday reasoning (for example about an explanation for the event: "Kim has had a traffic accident last night") people will most of the time reason in this undirected way, just guessing about some possible causes. Sooner or later a direct answer to this sort of questions can be given (e.g. somebody affirming you that Kim was seen drunk before the accident happened). An initial indication confirming one of the guesses, is in everyday reasoning often enough for people to accept it as the explanation. This line of reasoning can be represented by means of the following patterns.

$\underline{\mathbf{D} \rightarrow \mathbf{E}}$
$?D_k$

$\mathbf{D} \rightarrow \mathbf{E}$
$\uparrow D_k$

$D_k \rightarrow E_k$

(D= drunk, k= Kim)

We use $\uparrow X$ to express that there is an indication that the involved possible cause was present. $D_k \rightarrow E_k$ is the formalization of the conclusion that the drunkenness must form the explanation for the accident.
Formalized in general:

$\underline{\mathbf{\Gamma \rightarrow E}}$
$?\Gamma_\alpha$

$$\frac{\mathbf{\Gamma \rightarrow E} \quad \uparrow \Gamma_\alpha}{\Gamma_\alpha \rightarrow E_\alpha}$$

But in fact, this way of reasoning is neither very realistic, nor very interesting. Reasoning this way would be an example of uncritical thinking. (Although Kim was drunk, the cause of the accident could for example have been a ghost-driver.) Further, most of the times, people have at least some *specific* background knowledge from which they reason. For example about the night in which the car accident happened. (e.g. it was bad weather, Kim went to a party...). In the Bulwark case, the wild guessing soon was given up for reasoning that started from specific background knowledge.

3.2 Background Knowledge

In the case of the Bulwark, people are conscious of the risk for (submarine) attacks, because of the state of war. That is clear from the following fact: Before mooring, every ship already hang out the torpedo nets. [4]
The specific background knowledge consists here in the knowledge that the Bulwark is operative during war, which together with the general background knowledge implies there are some reasons to have the initial assumption that a submarine attack has caused the explosion.

$$\frac{W_b \quad \mathbf{W} \supset\Uparrow \mathbf{SA}}{\Uparrow SA_b\{SA_b\}}$$

(W= wartime, SA=submarine attack)

$$\frac{\mathbf{SA \rightarrow E} \quad \Uparrow SA_b\{SA_b\}}{?SA_b\{SA_b\}}$$

The formulas between brackets are conditions. Once the negation of these formulas are derived, these lines need to be marked and are out of the proof. We like to come back once again on the use of the arrows. Those symbols need to be distinguished from $\neg$ and the absence of $\neg$. Here we've been using $\Uparrow$ instead of $\uparrow$ in the previous section. $\Uparrow$ points to an initial assumption which in a critical reasoning process has to be confirmed by supporting evidence. The opposite $\Downarrow$ will later in this paper be used to express that a certain explanation seems on the first sight not applicable. We will use the $\uparrow$ and $\downarrow$ to introduce the supporting positive or negative evidence, as will be demonstrated in the next section. Since the arrows point to initial assumptions ($\Uparrow$ and $\Downarrow$) and indications ($\uparrow$ and $\downarrow$) rather than certainties, it remains possible that the opposite assertion turns out to be true in the end. For example: notwithstanding an initial assumption and an important indication for a submarine attack on the Bulwark, this explanation has afterwards been totally rejected on the basis of additional information.
$\Uparrow$- and $\Downarrow$ - lines will be derived from the background knowledge. It is clear by now, that in critical patterns of thinking, this initial assumption alone will not be enough to derive final conclusions from the background knowledge. This in opposition to the uncritical patterns of thinking typical for daily life reasoning and as described in the previous section.
Let's summarize in a more formal way the till know informally introduced rules. First some rules for the introduction of premisses:

GBK: At any point in the proof, one can add a line introducing some relevant ***g***eneral ***b***ackground ***k***nowledge involving **E** as an effect-variable:

(**1**) suitable line number,

(**2**) $\mathbf{X} \rightarrow \mathbf{E}$,

(**3**) "-",

(**4**) GBK,

(**5**) "-".

SBK: At any point in the proof, one can add a line introducing some relevant ***s***pecific ***b***ackground ***k***nowledge concerning the involved α:

(**1**) suitable line number,

(**2**) Z_α,

(**3**) "-",

(**4**) SBK,

(**5**) "-".

IE: At any point in the proof, one can add some background knowledge introducing some relevant ***i***nitial ***e***vidence:

(**1**) suitable line number,

(**2**) $\mathbf{Z} \supset\Uparrow \mathbf{X}$,

(**3**) "-",

(**4**) IE,

(**5**) "-".

The following is a first derivation rule:

Assume $\dfrac{\begin{array}{l} Z_\alpha \\ \mathbf{Z} \supset\Uparrow \mathbf{X} \end{array}}{\Uparrow X_\alpha\{X_\alpha\}}$

We introduced further one target rule which leads to the initiation of a possible explanation for which we need to find supporting evidence:

Pursue $\dfrac{\begin{array}{l} \mathbf{X} \rightarrow \mathbf{E} \\ \Uparrow X_\alpha\{X_\alpha\} \end{array}}{?X_\alpha\{X_\alpha\}}$

The conclusion of the assume-rule is derived conditionally. The derivation of the negation of X_α would imply the marking of this conclusion and of all lines derived on the basis of $\Uparrow X_\alpha$.

3.3 Testing of Hypotheses

In order to test the possible explanations (as derived by the pursue-rule), people will search for confirming or refuting information. They will examine *indications* for the possible causes and necessary *consequences* of these causes. On the one hand there can be found evidence confirming or denying the possibility of an initial supposition. For example in the Bulwark case, when there was made mention of the observation of a periscope. This was accepted as the confirmation of the initial assumption that a submarine attack had happened.[4]

$\Uparrow SA_b\{SA_b\}$
P_b
$P_b \supset \uparrow SA_b$

$\uparrow SA_b\{SA_b\}$
(P=periscope in the neighbourhood)

As the reader can see, we use $\uparrow$ to indicate that the periscope forms supporting positive evidence for the possible explanation submarine attack. This leads to the following lines:

SA $\rightarrow$ **E**
$\uparrow SA_b\{SA_b\}$

$\uparrow (SA_b \rightarrow E_b)\{SA_b\}$

This last line forms a possible answer to our very first question. Nevertheless, in addition to this conclusion one can take into consideration the necessary consequences of the initial supposition, which can be denied or confirmed by some evidence.
E.g., sabotage by a submarine attack implies an explosion from outside. Opposing this, Winston Churchill announced still the same day of the explosion of the Bulwark, on the basis of the reports of the vice- and rear admiral, that it concerned an internal explosion. Other witnesses confirmed this assertion afterwards. Because of this, the possibility of a submarine attack was at once rejected.

$\Uparrow SA_b\{SA_b\}$
$SA_b \supset EE_b$
$\neg EE_b$

$\neg SA_b$
(EE= external explosion)

Here we do use the negation, because the internal explosion is not just some supporting evidence against a submarine attack which is on its own unable to rule out the opposite. It concerns a well-established and decisive fact which makes the explanation submarine attack totally impossible. This can clarify the difference we make with "arrow-lines" and why we need the up- and downwards arrows in addition to the negation symbol. Further, it is clear that we arrive at inconsistency here. On the one side, the periscope confirms

the possibility of a submarine attack, on the other side, this possibility is totally ruled out by the negation of an external explosion. This clarifies why $\uparrow (X_\alpha \rightarrow E_\alpha)$ will only be derivable on the condition of consistency. Once the strong negation of X_α is derived in the same proof, the conditional lines need to be marked, which means they are out of the proof. With regard to the Bulwark, this means we fall in the starting position again, in which no possible cause is known.
Let's formally summarize the rules introduced in this section. First two introduction-rules:

IND: Indication rule: At any point in the proof following on the introduction of a certain target by the Pursue-rule, one can add a line introducing indications (cf SBK) for or against this pursued target.

(**1**) suitable line number,

(**2**) $Z_\alpha \supset\uparrow X_\alpha$, or $Z_\alpha \supset\downarrow X_\alpha$

(**3**) "-",

(**4**) IND,

(**5**) "-".

CON: Consequence rule: At any point in the proof following on the introduction of a certain target by the Pursue-rule, one can add a line introducing necessary consequences (cf SBK) for this pursued target itself, or for the refutation of the pursued target.

(**1**) suitable line number,

(**2**) $\uparrow X_\alpha \supset Z_\alpha$, or $\downarrow X_\alpha \supset Z_\alpha$

(**3**) "-",

(**4**) CON,

(**5**) "-".

Those introduction-rules go together with a couple of derivation-rules, which form an answer to the pursue-line (so the pursue-line can be marked, once on of these rules has been applied):

D1
$Z_\alpha \supset\uparrow X_\alpha$
$\Uparrow X_\alpha$
Z_α

$\Uparrow\uparrow X_\alpha\{X_\alpha\}$

D2
$\uparrow X_\alpha \supset Z_\alpha$
$\Uparrow X_\alpha\{X_\alpha\}$ $\surd$
$\neg Z_\alpha$

$\neg X_\alpha$

Those are the basic patterns which will be applied in different forms. Depending on the form of the indication- or consequence-lines (introducing positive or negative evidence), and on the available number of those lines, the direction and number of arrows will variate. By consequence, e.g. D1 will also refer to complex variants of the derivation, such as for example:

D1
$Y_\alpha \supset\uparrow X_\alpha$
$W_\alpha \supset\downarrow X_\alpha$
$Z_\alpha \supset\downarrow X_\alpha$
$\Uparrow X_\alpha$
Y_α
W_α
Z_α

$\downarrow\downarrow\uparrow X_\alpha\{X_\alpha\}$

If we have at least an initial supposition with one positive indication (which makes up for two upward arrows), and in addition less downward arrows than upward ones, than we can confirm that X_α forms a possible explanation, in accordance with the following rule:

HEX
$\mathbf{X \rightarrow E}$
$\uparrow X_\alpha\{X_\alpha\}$

$\uparrow (X_\alpha \rightarrow E_\alpha)\{X_\alpha\}$

For reasons which will become more clear in a further section, we will derive this last line only as a ***h***ypothetical ***ex***planation. This means a final answer has not yet been given to our main question at this point in the reasoning

process. Besides, the special background knowledge was given in our example. However, we can imagine similar reasoning processes in which this is not the case, but in which the question for it arises on the basis of an indication- or consequence-rule. This implies two target-rules:

T1
$$\frac{Z_\alpha \supset\uparrow X_\alpha \quad \Uparrow X_\alpha\{X_\alpha\}}{?Z_\alpha\{X_\alpha\}}$$

T2
$$\frac{\uparrow X_\alpha \supset Z_\alpha \quad \Uparrow X_\alpha\{X_\alpha\}}{?Z_\alpha\{X_\alpha\}}$$

What we have been analysing till this point, is a first kind of reasoning process towards a singular causal explanation. It is this kind of reasoning whereby people have only one possible cause in their mind which they are examining without being concerned about other possibilities. In the case of the loss of the Princess Irene, a similar reasoning process has been followed. In 3.4 we briefly consider that case as full illustration of the rules introduced till now.

3.4 Princess Irene

In May 1915, the Princess Irene was destroyed by a similar explosion. In this case, it was immediately clear that it concerned an internal explosion. What's more, there were immediately testimonies by a lot of witnesses about careless conduct on board of this ship. This stories were pointing in the direction of an accident: not conscious of possible defects, an inexperienced sailor who was exercising in the focussing of mines, has almost certainly been putting a defective igniter directly into a mine. The positive evidence has been strong enough to eliminate the consideration of alternative possible causes. Using the logical rules defined above, we can formalize this in the following proof:

(i=Princess Irene, EE=external explosion, BC=bad conduct of sailors on board, AC= accident due to carelesness, WT=witnesses testimonies)

1	$?E(X_i)$	–	**MAIN**	$E(X_i)$
2	$\neg EE_i$	–	**SBK**	–
3	WT_i	–	**SBK**	–
4	**WT** $\supset\Uparrow$ **AC**	–	**IE**	–

5	$\Uparrow AC_i$	3,4	**Assume**	$\{AC_i\}$
6	$\mathbf{AC \rightarrow E}$	–	**GBK**	–
7	$?AC_i$	5,6	**Pursue**	$\{AC_i\}$ $\surd_{10}$
8	BC_i	–	**SBK**	–
9	$BC_i \supset\uparrow AC_i$	–	**IND**	–
10	$\uparrow AC_i$	5,8,9	**D1**	$\{AC_i\}$
11	$\uparrow (AC_i \rightarrow E_i)$	6,10	**HEX**	$\{AC_i\}$

4 And what happened to the Natal, the Glatton, the Vanguard? Reasoning with several possible explanations

Before we go on, some words about cordite as a possible explanation for the shipdisasters. Cordite is a kind of gunpowder, often aboard of the warships at that time, which deteriorates in quality when exposed to high temperatures. Abnormal heat can be the cause of a spontaneous ignition of cordite, which can then result in the explosion of the munition depot on board of the ships. Once the possibility of a submarine attack on the Bulwark had been refuted, a cordite explosion was the first alternative possibility which came into the researchers' mind. It's obvious the same explanation was considered for a couple of other shipdisasters. For example, the Natal.[4]

4.1 Natal

An explosion similar to the one of the Bulwark and the Princess Irene happened in December 1915, on the Natal. As with the Bulwark, what could be discovered was limited. The research team suspected the bad cordite. Part of the cordite on board of the Natal was already on board since 1904. In their report, the researchers wrote: "The council assumes that the loss of the ship has been caused by an explosion of the munition inside the ship, and not by an explosion from the outside. (...) The loss of the ship is not due to the design, to carelessness or shortcomings of people on board." [4] We see a different way of reasoning towards a singular causal explanation here. It applies to people who have different possible explanations in their mind. Or it was a cordite explosion, or it was sabotage, or it was an accident due to carelessness. When people make this kind of list, the implicite supposition is there are no other possible explanations left (cf. the symbol $\leftrightarrow$ in the formalization below). Then it comes to deciding which one is right. In the Natal case, the choice for the cordite as explanation has clearly been made by the negation of the other possibilities. The possibility of an accident due to carelessness is explicitly eliminated on the basis of testimonies,

the possibility of sabotage rather implicitly from the fact that it concerned an internal explosion (which is actually not enough to eliminate this explanation, e.g. a bomb could have been brought aboard). As a consequence, the cordite has been choosen as explanation, although no positive evidence for the spontaneous ignition has been offered. This forms a first possible way of reasoning towards an explanation on the basis of different possibilities: by elimination.

(n=Natal, CE=cordite explosion, S=sabotage, A=accident due to carelessness, EE= external explosion, WT= witnesses testimonies, OC= old cordite, E= target event)

1	$?E(X_n)$	–	**MAIN**	$E(X_n)$
2	$CE_n \vee S_n \vee A_n \leftrightarrow E_n$	–	**PEX**	–
3	$\neg EE_n$	–	**SBK**	–
4	$\neg EE_n \supset\downarrow S_n$	–	**IND**	–
5	WT_n	–	**SBK**	–
6	$WT_n \supset\downarrow A_n$	–	**IND**	–
7	OC_n	–	**SBK**	–
8	**OC** $\supset\Uparrow$ **CE**	–	**IE**	–
9	**CE** $\rightarrow$ **E**	–	**GBK**	–
10	$\Uparrow CE_n$	7,8	**Assume**	$\{CE_n\}$
11	$?CE_n$	9,10	**Pursue**	$\{CE_n\}$
12	$\neg(S_n \rightarrow E_n)$	2,3,4	**ELIM**	–
13	$\neg(A_n \rightarrow E_n)$	2,5,6	**ELIM**	–
14	$CE_n \rightarrow E_n$	2,10,12, 13	**HEX**	$\{CE_n\}$
15	$!E(CEn)$	12	**EX**!	

As becomes clear from the formalization, we have indications for the elimination of two out of three possibilities. There is no evidence for the third one, but this possibility strokes with the background knowledge of line 7, 8 and 9. All this information taken together, gives people reasons to take a cordite explosion as the real explanation.

Some new rules have been added. First, one new introduction-rule:

PEX: At any point in the proof, one can add a line introducing all possible explanations for the involved target event:

(**1**) suitable line number,

(**2**) $C1_\alpha \vee C2_\alpha \vee ... \vee Cn_\alpha \leftrightarrow E_\alpha$,

(**3**) "-",

(**4**) PEX,

(**5**) "-".

Further, three new derivation-rules:

ELIM
$$\frac{C1_\alpha \vee C2_\alpha \vee ... \vee CN_\alpha \leftrightarrow E_\alpha \quad Z_\alpha \quad Z_\alpha \supset\downarrow C1_\alpha}{\neg(C1_\alpha \rightarrow E_\alpha)}$$

Confirm
$$\frac{C1_\alpha \vee C2_\alpha \vee ... \vee CN_\alpha \leftrightarrow E_\alpha \quad \neg(C1_\alpha \rightarrow E_\alpha) \quad \neg(C2_\alpha \rightarrow E_\alpha) \quad \Uparrow C3_\alpha\{C3_\alpha\}}{C3_\alpha \rightarrow E_\alpha\{C3_\alpha\}}$$

EX!
$$\frac{C3_\alpha \rightarrow E_\alpha\{C3_\alpha\}}{!E(C3_\alpha)}$$

As was the case with the D-derivation lines in the previous section, some variation is also possible concerning the application of the Confirm-rule. We suppose the basic idea is clear: enough information need to be present to eliminate all but one possibility, and further at least an initial assumption or some supporting evidence to accept the not eliminated one. What concerns the EX!- rule: this can only be applied if only one possible explanation is left.
However, there will not always be clear indications to eliminate all but one possible explanations. The positive evidence for one of them can however be strong enough to be convincing. This was for example the case when people were searching for an explanation for the explosion of the Glatton.

4.2 Glatton

In September, 1918 the Glatton perished as well due to an internal explosion. This tragedy opened in a certain degree the possibility to slightly justify the conclusions of the various researches concerning the previous disasters. In this case, it was thought that all possible precautions were taken, so that an ignition of the cordite was impossible.

Later on it was nonetheless discovered that the firers had the habit to put red-hot ashes and cinders against the ships' side which separates the munition depot from the stokehold. Maybe this caused a substantial rise of the temperature? An experiment on the sister ship the Gorgon demonstrated to the contrary that this increase in temperature was not big enough to cause a spontaneous ignition.
But in 1919, the old isolation material from the Gorgon has been removed. This led to the discovery of an open space between the boiler room and the munition depot, which was filled with newspapers instead of crumbled cork. Because the Gorgon and the Glatton had been build by the same group of workers, it is possible that also in the Glatton newspapers had been put in the open space between the boiler room and the munition depot. The red-hot ashes and cinders could have caused the ignition of the newspapers, and by consequence of the wooden panels, which on their turn could have caused an increase of the temperature leading to a cordite explosion.
All this information gives us reasons to accept that in the case of the Glatton, a cordite explosion forms indeed the best explanation. The other possible explanations, gathered on the basis of previous explosions, were not anymore considered ...

(g=Glatton, CE=cordite explosion, S=sabotage, A=accident due to carelessness, EE= external explosion, OC= old cordite, PC=precautions taken, HF= habit of firers aboard, EX= experiment on sistership, DS= design of sistership, E= target event)

1	$?E(X_g)$	–	**MAIN**	$E(X_g)$ $\surd_{22}$
2	$CE_g \vee S_g \vee A_g \leftrightarrow E_g$	–	**PEX**	–
3	$\neg EE_g$	–	**SBK**	–
4	$\neg EE_g \supset\downarrow S_g$	–	**IND**	–
5	PC_g	–	**SBK**	–
6	**PC** $\supset\Downarrow$ **CE**	–	**IE**	–
7	$\Downarrow CE_g$	5,6	**Assume**	$\{\downarrow CE_g\}$
8	OC_g	–	**SBK**	–
9	**OC** $\supset\Uparrow$ **CE**	–	**IE**	–
10	$\Uparrow CE_g$	5,6	**Assume**	$\{\uparrow CE_g\}$
11	**CE** $\rightarrow$ **E**	–	**GBK**	–
12	$?CE_g$	10,11	**Pursue**	$\{CE_g\}$ $\surd_{19}$
13	HF_g	–	**SBK**	–
14	$\neg HF_g \supset\uparrow CE_g$	–	**IND**	–
15	EX_g	–	**SBK**	–
16	$\neg EX_g \supset\downarrow CE_g$	–	**IND**	–
17	DS_g	–	**SBK**	–
18	$\neg DS_g \supset\uparrow\uparrow CE_g$	–	**IND**	–

19	$\uparrow\downarrow\uparrow\uparrow CE_g$	10-17	**D1**	$\{CE_g\}$
20	$\neg(S_n \rightarrow E_n)$	2,3,4	**ELIM**	–

First some explication concerning line 16: because the positive evidence for a cordite explosion, derived from the design of the sistership the Gorgon, outweighs the other arguments, we gave it two upward arrows. We are convinced this strokes with human reasoning towards a singular causal explanation: some arguments will outweigh some others... Although there were at the beginning as well a positive as a negative assumption towards the cordite explanation, it is clear the positive evidence outweighs the negative ones. Now we need one more derivation rule: if we have convincing positive evidence for one of a number of possible explanations, and for some or none of the others some negative evidence, and additionally, no information at all for those left, then we will conclude to accept those explanation for which we have the most and convincing evidence.

Confirm2
$$\frac{\begin{array}{l} C1_\alpha \vee C2_\alpha \vee C3_\alpha \leftrightarrow E_\alpha \\ \neg(C1_\alpha \rightarrow E_\alpha) \\ \uparrow\downarrow\uparrow C2_\alpha \end{array}}{C2_\alpha \rightarrow E_\alpha\{CE_\alpha\}}$$

Again, some variation in the number and direction of the arrows is possible. The important thing is that all evidence taken together confirms the elected explanation. The end of our proof will subsequently be as follows:

21	$CE_g \rightarrow E_g$	2,17,18	**Confirm2**	$\{CE_g\}$
22	$!E(CEg)$	19	**EX!**	

4.3 Vanguard

A third possibility when reasoning with several possible explanations is that there can't be made a choice at all for only one possible explanation. Or there is not enough information to arrive at the right single conclusion, or two possible explanations need to be taken together to come to the right conclusion. In other words, we arrive here at a problem of overdetermination. This was for example the case when people were searching an explanantion for the explosion of another ship in the series, the Vanguard.
The loss of the Vanguard in July 1917 pointed to a new possible explanation: one person called John Harston could have caused sabotage by bringing aboard a time bomb. He had been working on both the Natal and the Vanguard, in both cases shortly before their loss by explosion, and working in

the harbour of Chatham, he could easily gain access to the Bulwark too. Nonetheless, nothing suspect or accusatory has been found. For example, the man didn't have German connections. Nevertheless, during interrogation Harston took on some more suspicion. But finally, he has never been confined. In the end, the research team rejected again the possibility of sabotage, and started to suspect once again the cordite. [4]
But in fact, taking all the evidence together, there are no substantial reasons for that choice. We will never be able to know the real cause: it can be one of both or their combination. Internal sabotage brought about by Harston can have led to a fire which caused a cordite explosion.
This doesn't really form a problem when reasoning following the patterns analysed in this section. In our formalization, two possibilities will stay open for consideration, which makes the undecidable situation automatically clear. To the contrary, this example points to a problem with the first way of reasoning analysed by us in section 3. If people focus on one possible explanation, for which they find supporting evidence leading to the acceptance of that explanation, they will never be able to conclude that there are still other possible explanations, perhaps working together. This is why we derived an explanation only hypothetically by the rule **HEX**. When reasoning in a critical way, we need at least to examine other possibilities. Only when those latter can explicitly be eliminated, we have thorough reasons to accept our preconceived explanation, and by consequence to mark the main target-line, which means a final explanation has been found.
All this points further to a frequently occuring human reasoning pattern leading to faulty conclusions. Possible explanations are in the first place created on the basis of certain indications in the direction of the concerning possible causes. By consequence, the absence of certain indications make people overlook possible explanations. For example, when there are not many survivors who are able to give testimony of careless conduct by the sailors on board, the possibility of an accident due to carelessness will almost automatically be implicitly eliminated. This can be expressed by $\neg WT \supset\!\!\downarrow AC$. This shows at the same time the usefulness of the distinction made by us between $\neg$ and $\downarrow$ as a tool to make a distinction between explicit and implicit, strong and weak evidence. In the just mentioned case for example, it would obviously be wrong to state $\neg WT \supset \neg AC$. Accepting this would be very uncritical.

5 Concluding Remarks

In this paper, we made a thorough analysis of the search process for a series of singular causal explanations. The patterns and rules derived, must be seen as elements of a logic which has to be further developed. We like to stress again it was not our goal to present a complete and well-developed logic.

Which general patterns can be deduced from the previous analysis?

1. Possible explanations are in many cases created on the basis of indications pointing at one or more possible causes. Background knowledge is activated on the ground of such indications.
2. There are two important ways of reasoning towards a singular causal explanation. Or one focuses on one possible explanation (cf. section 3), or one reasons from several possible explanations (cf.section 4). The first way of reasoning is rather uncritical, and the explanations we obtain are very sensitive to collapse if further evidence becomes available.
3. The choice for one of a set of possible explanations sometimes happens through elimination of explanations (due to a lack of positive evidence and/or the introduction of negative evidence; e.g. Natal), and sometimes through the confirmation of one alternative by means of positive evidence (e.g. Glatton).
4. Though one aims at singular causal explanations, the comparison with similar events can form an important basis to enlarge the background knowledge to rely on to derive conclusions.
5. Overdetermination forms a problem, especially when reasoning following the patterns of section 3. There is no way to be sure that there is no overdetermination of causes. This problem characterizes on the other hand a frequently occurring fault in human reasoning, in which focussing on one possible explanation easily leads to the wrong conclusions.

We introduced three kind of rules:
First, quite a lot of rules to introduce information: GBK (to introduce general background knowledge), SBK (to introduce special background knowledge), IE (to introduce intitial evidence), IND (to introduce indications), CON (to introduce necessary consequences), PEX (to introduce all possible explanations for the target event). That we need so much different premise rules, is due to the fact that a lot of non-formal information needs to be taken in consideration, as announced in the introduction.
Second, a lot of derivation rules have been abstracted from the examples. They make it possible to derive further information and conclusions from the premisses: Assume, D1, D2, HEX, Elim, Confirm, EX!, Confirm2.
Finally, some target rules (which can be seen as a special kind of derivation rule): the most central rule MAIN (to introduce the main target), pursue (to find evidence supporting an initial assumption), T1 and T2 (to find special background knowledge).
All these rules have been explicitly defined. The rules for the marking of lines on the other hand, have only implicitly been introduced. There are three reasons to mark a line. In the ideal case, one can mark the main target line at the end of the proof. This means a single causal explanation has been found, and by consequence no undecidabilities or overdetermination. Second, marking occurs when one of the other target-lines is answered (in

exactum, when found an answer to a pursue-line by means of D1 or D2, or when found an SBK-line giving an answer to T1 or T2). A last reason for marking is correlated to the conditionally derived lines. If a condition is contradicted, the conditionally derived line is marked and by consequence out of the proof. For example in the Bulwark case, when the possible explanation submarine attack was contradicted by the hard evidence that it concerned an internal explosion.
In the future, we like to examine the applicability of these reasoning patterns and logical rules to the formal analysis of other case-studies. That way, we hope in the long run to be able to bring all these results together to become an adaptive logic for singular causal explanation which is applicable to a wide range of cases.

References

[1] Diderik Batens. Inconsistency-adaptive logics. In Ewa Orłowska, editor, *Logic at Work. Essays Dedicated to the Memory of Helena Rasiowa*, pages 445–472. Physica Verlag (Springer), Heidelberg, New York, 1999.

[2] Diderik Batens. Notes on problem solving. *Research Report*, 2002.

[3] Diderik Batens. A general characterization of adaptive logics. *Logique et Analyse*, in print.

[4] Philip Mac Dougall, editor. *Schepen die verdwenen: over mysterieuze verdwijningen en onverklaarbare rampen ter zee.* De Boer Maritiem, Baarn, 1994.

[5] Maarten Van Dyck. Causal discovery using adaptive logics. 2002.

[6] James B. Kirby. The influence of parental separation on smoking initiation in adolescents. *Journal of Health and Social Behavior*, 43:56–71, 2002.

[7] Judea Pearl, editor. *Causality. Models, Reasoning, and Inference.* Cambridge University Press, Cambridge, 2000.

[8] Diderik Batens & Dagmar Provijn. Pushing the search paths in the proofs. a study in proof heuristics. *To appear.*

[9] Leen De Vreese & Erik Weber. Application of the adaptive logic for causal discovery. 2003.