KAPITEL 3

Unfall-Analyse

We show how objective, rigorous causal reasoning in the analysis of air transportation accidents can improve our understanding of the factors involved in those accidents, by considering two high-profile digital-automation-related air transport accidents.

3.1 Why Investigate Accidents?

Let us consider *safety* as *freedom from accidents*, where an accident is *an unwanted (but not necessarily unexpected) event that results in a specified level of loss* [44]. Suppose one wants to improve safety. Then one must increase the relative freedom from accidents. One cannot undo accidents that have already happened, so one cannot improve safety by attempting to undo past accidents. Yet detailed accident investigation is widely regarded as a significant tool for improving safety. Why? Why not just say „*Oh dear, we regret very much ... but we must move on with life*“, and ignore the whole event?

When one is trying to ensure safety, one is oriented to the future. Future events have not happened yet; one is trying to avoid those that would be accidents. We must think about the system we have, and we must attempt to assess what could happen and what could not, and if necessary reconfigure the system or its environment of operation or both in order to change what we believe to be the behavioral possibilities.

An accident is a concrete, irrefutable example of system and environment behavior. It is thus a guide to the possibilities. By comparing what we think we knew about the system with what we know from a detailed investigation of the accident, we may

be able to correct and improve our reasoning about and our knowledge of possible system behavior.

Further, suppose one makes a general presumption that system and subsystem behaviors have some statistical distribution. We won't know what that distribution might be. However, the presumption entails that, in normal system use, specific states and events occur with a particular although unknown expected frequency. Events about which we may be very concerned are those events which are or can be involved in accidents. By investigating accidents in detail, one obtains information about which events and states are involved, and may focus on these events and states in this and other recorded instances to obtain information about their actual frequency of occurrence. One may then consider mitigating measures.

There can be no guarantee that one has thereby enumerated *all* events or states that may be involved in accidents. However, if all have some expected frequency, then some of those expected frequencies will be higher than others, and those events are those which we are likely to see — or to have seen — more often. In particular, when we mitigate accident contributors with high expected frequency of occurrence, we attempt to reduce their frequency of occurrence or eliminate it altogether. By mitigating the occurrence of contributing events and states that one has seen in accidents, one can expect to reduce the frequency of occurrence of the most frequent contributors, thereby reducing the overall frequency of likely occurrence of all accident contributors taken together, even if one does not know them all.

These, then, I take to be the general reasons for investigating accidents. Investigation is the art of discovering facts. Some of these discoveries are made "in the field" by finding things, by reading data recorders and listening to cockpit conversation. Others are discovered by reasoning, by inference from facts one has already determined, and enumerating behavior possibilities constrained by the facts one has already determined. Both sharp eyes and sharp minds are essential components of investigation. Both can be improved by methods: methodical ways of searching rubble fields, and methodical reasoning.

3.2 What-If Reasoning

I want to focus on the reasoning. General procedures have been known for over a century for how to add method to reasoning, and to check for one's mistakes. This is

the science of formal logic. One way to become more methodical is to look closely at the features of the reasoning as practiced, identify general principles, justify these principles, and build them in to a formal logic. Then anyone can check whether the reasoning is sound by reproducing it — or failing to — in the formal logic.

What kinds of reasoning are involved in safety, and in accident investigation? One is reasoning about system behavior, and because one is trying to avoid certain kinds of behavior deemed to be accidents, one must engage in so-called *what-if* reasoning. What if this-and-this were to occur in a behavior? What if that-and-that were to occur? HAZOP is an example of this kind of reasoning. Other kinds of reasoning attempt to reason from problem behaviors of the system to contributory problem behaviors of subsystems by using the architecture of the system. Suppose this-and-this were to happen. It would happen if and only if that-and-that were to happen with that part. Fault tree analysis is an example of this kind of part-whole reasoning.

When investigating accidents, one engages also in what-if reasoning. This is what the U.S. Air Force says about accident explanations [70]:

> **3-11. Findings, Causes, and Recommendations.** The most important part of mishap investigation is developing findings, causes and recommendations. The goal is to decide on the best preventive actions to preclude mishap recurrence. To accomplish this purpose, the investigator must list the significant events and circumstances of the mishap sequence (findings). Then they [*sic*] must select from among these the events and conditions that were causal (causes). Finally, they suggest courses of action to prevent recurrence (recommendations).
>
> **3-12. Findings:**
>
> a. Definition. The findings ... are statements of significant events of conditions leading to the mishap. They are arranged in the order in which they occurred. Though each finding is an essential step in the mishap sequence, each is not necessaily a cause factor ...
>
> **3-13. Causes:**
>
> a. Definition. Causes are those findings which, singly or in combination with other causes, resulted in the damage or injury that occurred. A cause is a deficiency the correction, elimination, or avoidance of which would likely have prevented or mitigated the mishap damage or significant

> injuries. A cause is an act, an omission, a condition, or a circumstance, and it either starts or sustains the mishap sequence ...

The phrase „... *would have prevented* ....“ talks about something that could have happened, but did in fact not. The correction, elimination or avoidance of feature X would have prevented the accident. But in fact X occurred, and so did the accident. The supposition, that had X not occurred as it did, the accident would not have happened, is known as a *counterfactual*. So reasoning about causes of accidents in the USAF is reasoning with counterfactuals.

The USAF was not the first to think this way. David Hume gave two definitions of causality over 200 years ago.

> ... we may define a cause to be *an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second*. Or, in other words *where, if the first object had not been, the second never had existed*.

[24, Section VII, Part II, paragraph 60].

We may consider the word *object* to refer also to events, maybe states, as noted in the work of John Stuart Mill [53].

David Lewis notes [45] that of the two definitions given by Hume, over the course of the intervening couple of hundred years, the second has been more neglected by Humean commentators. Hume's second definition is *counterfactual*. Like the U.S. Air Force, it talks of what might have been but was not.

Lewis's Formal Definition of Causal Factor

In *op. cit.*, Lewis gives a formal definition of *necessary causal factor*, based on the counterfactual definition of Hume. Suppose A and B are state descriptions or events. Then A *is a (necessary) causal factor of* B just in case, had A not occurred, B would not have occurred either. This definition is obviously counterfactual. Lewis [46] had already defined a formal semantics, and a complete logic, for counterfactuals, based on the formal-semantical notion of possible worlds, used ubiquitously by formal logicians, with an additional notion of *comparative nearness*: a behavior, or a history, is said to be *nearer* to a reference behavior than another behavior is to that reference behavior. Comparative nearness is a ternary relation - it has three arguments – and

Lewis also required that it have certain formal mathematical properties for whose reasonableness he argued (for those interested in more detail, the properties are listed in [33]).

An Example

Consider a system in which there is a programmable digital component which contains a bit, stored in a variable named X. With systematic ambiguity, we shall refer to this bit as X. Suppose the electronics is wired such that, when X is set, a mechanism (say, an interlock) is thereby set in motion. Suppose the interlock has been well enough designed so that it can only be set in motion by setting X. Then X is a causal factor in any setting in motion of the interlock according to the Lewis definition: *had X not been set, the interlock would not have moved*. Furthermore, let us suppose that the digital component is well-designed, so that X can only be set by a specific operation O of a processor to set it, and that this operation is performed by executing a specific program instruction I. Then,

- *had the operation O not been performed, X would not have been set*, and
- *had the instruction I not been executed, the operation O would not have been performed*.

It follows that

- Performance of O is a necessary causal factor in setting X, and
- Executing I is a necessary causal factor in performing O

The Meaning of A Counterfactual

Lewis's formal meaning for a counterfactual proceeds as follows. We interpret the counterfactual *had A not occurred, B would not have occurred*. The real world history is some behavior. We have a relation of comparative nearness amongst behaviors. In the real world, B occurred, as did A. But we want to know about behaviors in which A did not occur. Did B occur in them? We do not consider all these counterfactual behaviors – Lewis proposes we consider only the *very nearest behaviors* to the real world in which A did not occur. The counterfactual *had A not occurred, B would not have occurred* is defined to be true (in the real world) just in case, in all these nearest behaviors in which A did not occur, B did not occur either. Lewis's formal

requirements on the notion of comparative nearness ensure that there are always very nearest behaviors.

The Semantics Applied to the Example

We can consider behaviors near enough to the real world such that I was not executed. We may presume that the more properties of the system and environment that are the same, the nearer the states of the alternative behavior are to the real world. It follows that in the nearest behaviors the design and intended operation of the system can be assumed to be identical to its design and intended operation in the real world. For these behaviors, then, in which I was not executed, O was not performed. And in these behaviors in which O was not performed, X was not set. And in these behaviors in which X was not set, the interlock was not set in motion. So consideration of the nearest behaviors shows that the counterfactuals are to be evaluated as true. Consequently, the assertions of causality (or, rather, *causal-factorality*) are true.

Causal-Factorality and Causality

It turns out that Lewis's formal notion of causal factor is not transitive, that is

- If A is a causal factor of B, and B is a causal factor of C, this does not necessarily mean that A is a causal factor of C.

Since the intuitive idea of a cause is something that propagates through a "chain" of causal factors, Lewis proposes to define "cause" as the "transitive closure" of the relation of causal factor. The *transitive closure* of a relation R is the smallest (or "tightest", most narrowly defined) relation R^* which, roughly speaking, is transitive and contains R.

An Aside on Causality and Computers

Relation Between Instruction and Execution is Causal

This example also illustrates that, according to the formal definition, the design of a digital system ensures that the relation between the form of an instruction and and its execution is causal. The instruction I says to increment register R. I is executed; R is incremented. Had the instruction not been to increment register R, then R would not

have been incremented. Therefore, the form of I, that I is an instruction to increment R, is a causal factor in incrementing R when the instruction is executed.

Debugging is Causal Analysis

This observation entails that debugging computer programs is a form of causal analysis. One can consider it akin to 'debugging' complex systems. Not only by analogy, but formally.

3.3 Where Does This Get Us?

So the first observation is that counterfactual, or what-if, reasoning is essential not only for reasoning about safety but also for reasoning about causes of accidents. The second observation is that there is a mathematically satisfactory formalisation of counterfactual reasoning. In principle, we can check our safety reasoning and our reasoning about the causes of accidents against objective, rigorous criteria.

In practice, however, one has to put it all together. Karsten Loer and I took a formal logic sufficient for describing formal properties of distributed systems, the temporal logic TLA [41], and combined it with the causal/counterfactual logic of Lewis, adding in some inference rules which we observed were commonly used when arguing for sufficiency of causal explanations. The resulting logic, *Explanatory Logic* or EL, could be used for formal causal reasoning about complex system behavior. We developed a method, *Why-Because Analysis* or WBA, for causally analysing complex system accidents and applying EL to check the reasoning. WBA is described in [34], along with applications to a number of high-profile aviation accidents.

Do we really need all this machinery to help us analyse systems and design safer ones? Or is this just an exercise for academics? I don't want to introduce the details of WBA here. For one thing, there are a lot of technical details, and for another thing, readers might prefer to use a different formalism. My goal here is to persuade that rigorous, counterfactual reasoning is needed for accident analysis.

Thus I would like to provide two examples to persuade readers of the necessity for objective, rigorous reasoning such as proposed in WBA. These examples employ the preliminary part of a WB-Analysis, which we call the WB-Graph method.

Our approach is very simple. For the 1993 Lufthansa Warsaw accident and the

1988 Air France Habsheim accident, Michael Höhl and I took the factual findings in the official accident reports at face value. We listed them all, and then for each pair of facts, say A and B, we applied Lewis's possible world semantical reasoning informally to determine whether A was a causal factor in B or not. We drew the results in a graph, called the *Why-Because Graph* or WB-Graph. I want to comment on what the graphs show.

3.4 The Warsaw Lufthansa A320 Accident [73]

On 14 September 1993, a Lufthansa Airbus A320 landed at Warsaw airport in a thunderstorm. Upon landing, none of the braking systems (air brakes, thrust reverse, wheel brakes) functioned for about nine seconds: the wheel brakes only started to function after about thirteen seconds. The aircraft ran off the end of the runway, collided with an earth bank and started to burn. Primarily because of the superb behavior of the crew, only two people died: one pilot, who died when the aircraft hit the bank, and one passenger, who was unconscious in the front corner and unnoticed in the evacuation as the cabin filled with smoke, and was asphyxiated. It became clear that the logic of the braking systems was indeed a reason why the braking systems hadn't functioned as expected. However, many commentators focused upon this factor as *the* main cause of the accident, which as we shall see is probably incorrect. There were, as is usually the case, many other necessary causal factors.

The WB-Graph

Figure 3.1 shows the WB-Graph derived from the report by considering all the mentioned states and events and assessing their causal relations to each other using the Lewis semantics. An edge passing from a lower node N to a higher node M means that N is a necessary causal factor in M. No attempt was made to identify features of the accident that were not explicitly mentioned somewhere in the report. It is not easy to read all the node labels, so I divide the graph into three parts: the lower part in Figure 3.2, the middle part in Figure 3.3, and the upper part in Figure 3.4. This division also coheres with the statement of probable cause in the final report, and emphasises a missing feature.

The statement of probable cause from the report is as follows:

> Cause of the accident were incorrect decisions and actions of the flight crew taken in situation when the information about windshear at the approach to the runway was received. Windshear was produced by the front just passing the aerodrome; the front was accompanied by intensive variation of wind parameters as well as by heavy rain on the aerodrome itself.
>
> Actions of the flight crew were also affected by design features of the aircraft which limited the feasibility of applying available braking systems as well as by insufficient information in the aircraft operations manual (AOM) relating to the increase of the landing distance.

Decisions and Actions of the Flight Crew

The first sentence of the probable cause statement coheres with what one sees in the lower portion of the graph in Figure 3.2. The events and states in this portion contribute to the "key" node *Decisions and actions of the flight crew in anticipation of wind shear*.

Weather

The weather phenomenon plays a role in the middle portion of the WB-Graph, as may be seen in Figure 3.3. Also in this portion appear the "*design features of the aircraft*" adduced in the second paragraph of the statement of probable cause.

The Destruction Sequence

Most of the upper portion of the graph, in Figure 3.4, enumerates the parameters of the accident. In order to be classified as an accident, people must be killed or severely injured, and/or the aircraft must be significantly damaged. Both occurred in this accident (although, thankfully, only two people lost their lives and other injuries were minor). One can see these factors appearing in this portion of the graph. But what caused all this?

Focusing In on Factors

Let us now focus on the upper portion of the graph where it narrows down to one node. It is rare that a WBA of an accident results in a graph with a width of one. What is this single node?

```
AC hits earth bank
```

Take away this node, and you've avoided the accident. What are its immediate precursors?

```
AC overruns RWY
Earth bank in overrun path
```

The report's attribution of probable cause focused entirely on causal factors contributing to the first of these two events. What about the second? Why was there an earth bank in the overrun path? Because

```
Bank built by airport authority for radio equipment
```

Prophylaxis: Don't Overrun Or Don't Build

So there is clearly something to consider. Don't build earth banks for radio equipment at the ends of runways in the overrun area. Or don't overrun runways. Well, measures are taken to minimise cases of the latter, but most authorities consider that no matter what one does, aircraft will still overrun runways once in a while. So if you want to prevent or minimise such catastrophic overrun accidents, one had better take the other option and not build in the overrun area.

In fact, leaving a clear overrun area at the end of runways is regarded not only as good practice but as essential practice by most Western European and US authorities and by practically all pilots.

Rigorous Causal Reasoning Helps

The report's conclusions about probable cause and contributing factors said nothing about building earth banks in overrun areas.

The WBA of the accident shows clearly that this omission is a mistake in causal reasoning that the report made. The information necessary to infer that it was a contributing cause was contained in the body of the report - that is where we obtained the factors in the WB-Graph in Figure 3.1. The WBA shows it to be a causal factor.

This is not the only causal reasoning mistake in the Warsaw report, neither is it the only report in which significant causal reasoning mistakes may be demonstrated by WBA. Another, the report on the 1995 American Airlines B757 accident on approach to Cali, Colombia is one, which also omits demonstrably causal factors in its statement of probable cause. The omitted factors in that report were, however, taken into account by the U.S. National Transportation Safety Board in their letter to the U.S. Federal Aviation Administration containing their safety recommendations based on their analysis.

Using rigorous methods of causal reasoning such as WBA would thus help considerably in ensuring correctness of these important reports. Prophylactic measures are based on the reports' analyses. It is important to reduce future accidents that resources be pointed in the appropriate directions, and one can only do this if a report's reasoning is correct.

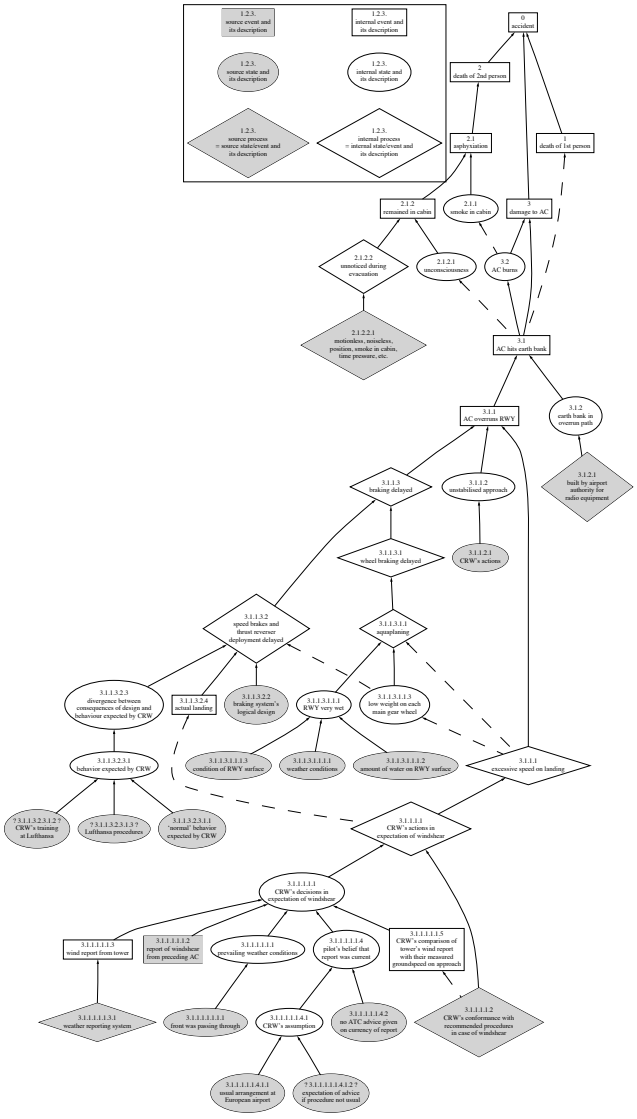

Abbildung 3.1: The Warsaw WB-Graph: overall pattern

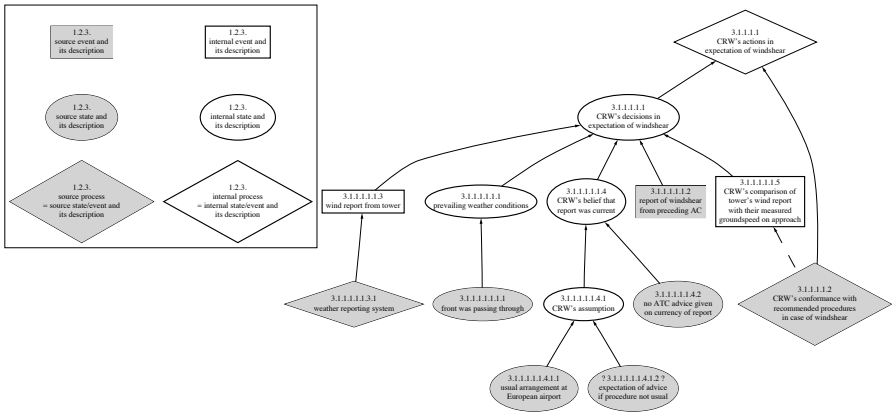

Abbildung 3.2: The Warsaw WB-Graph, lower part

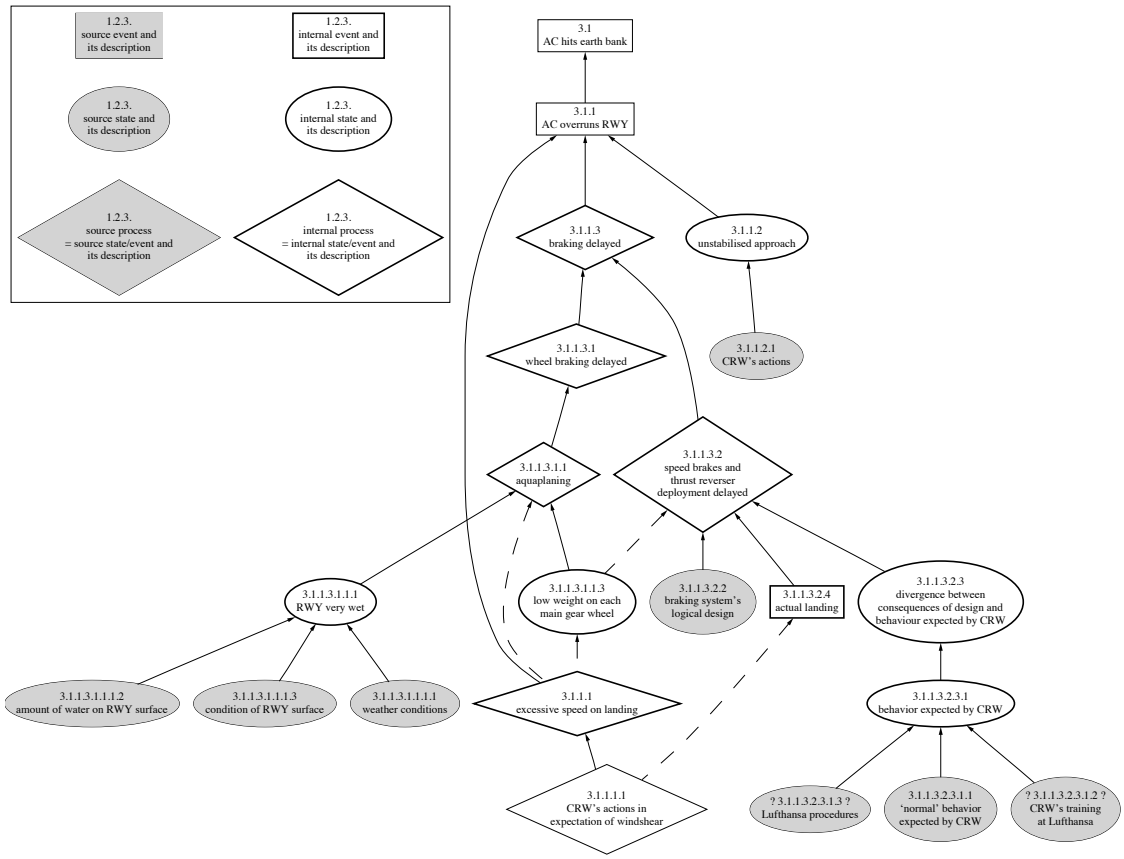

Abbildung 3.3: The Warsaw WB-Graph, middle part

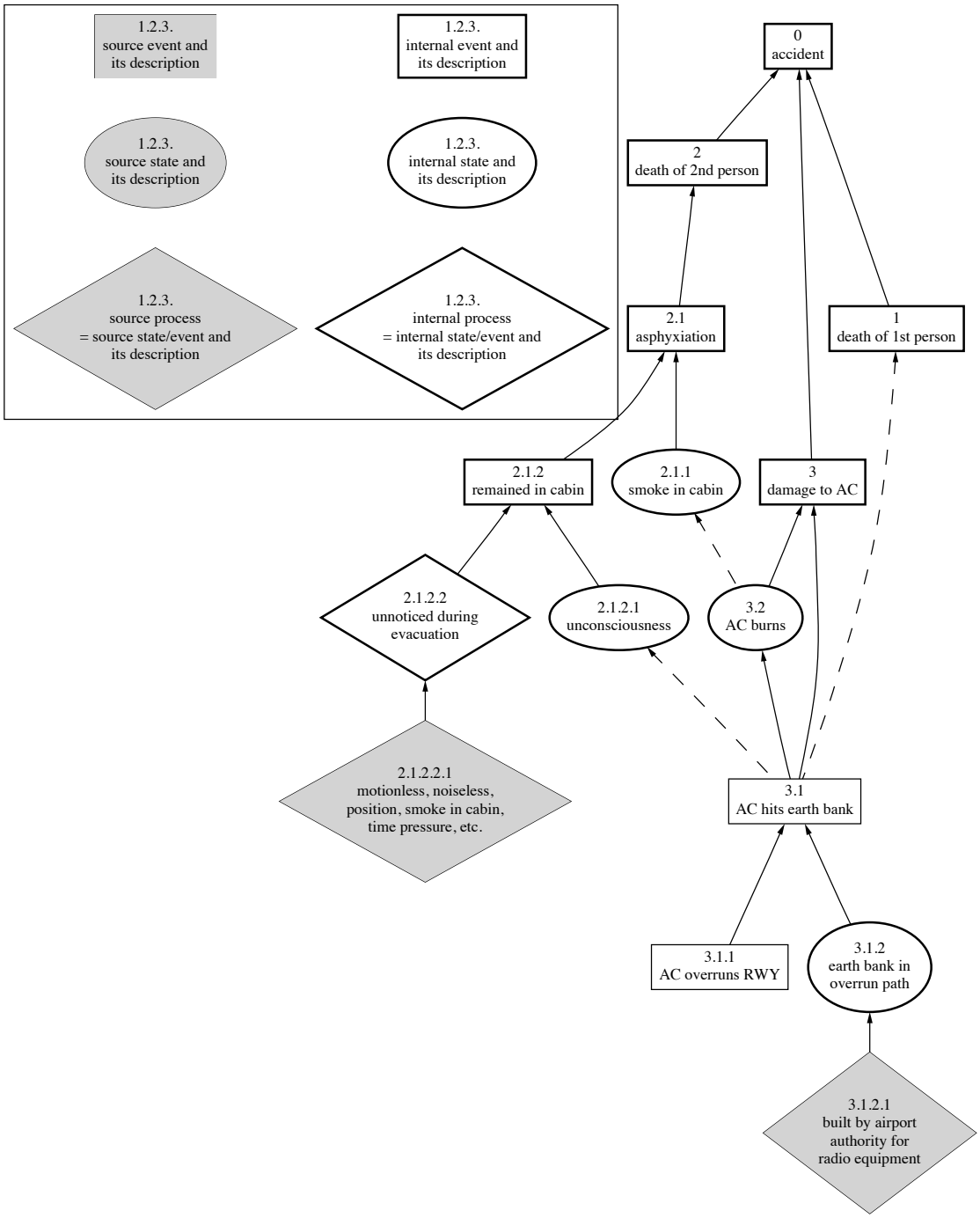

Abbildung 3.4: The Warsaw WB-Graph, upper part

3.5 The 1988 Habsheim Accident [16]

On 26 June, 1988, an Air France A320, new into service with the airline, took off from Basle-Mulhouse airport with sightseeing passengers, intending enroute to put in an appearance at an airshow at the small airport Mulhouse-Habsheim, just a few miles and minutes flying time away. The pilot had planned for a “low-speed pass”, a manoeuver in which the aircraft is configured for landing, flies low along the line of the runway very slowly without landing, and then accelerates up and away. This manoeuver was believed to show off the automatic slow-speed flight protection capabilities of the autopilot, and thereby how the performance of the airplane is enhanced. The manoeuver had been practiced at altitude by the pilot, from a more-or-less level entry.

The pilots had not surveyed the display airport before appearing, and had submitted incomplete flight planning to the Air France administration on Friday. The incomplete planning was approved, although some of it contravened French aviation-legal restrictions on airshow performances by commercial aircraft.

Upon takeoff, the aircraft climbed to an intermediate altitude of 1000 feet above the ground while the pilots identified the airshow airport, which should have been visible almost immediately upon takeoff. A descent was commenced towards the Habsheim airport, which reached a rate of 600 feet per minute with the engines in flight idle. The power setting at flight idle is 29% N1 (a measure correlating with the thrust produced) although the Commission noted that the manoeuver been planned starting from a high power setting.

As the aircraft approached for the low pass and passed through 100 feet above ground level (the planned fly-by altitude), the aircraft was still descending at a rate of 600 feet per minute with the engines in flight idle. The aircraft reached a low altitude of about 30 feet above the runway while attempting to perform the manoeuver. Beyond the end of the runway was a forest, with tree tops considerably higher. “Take-off/go-around” (TOGA) thrust was applied, but the aircraft continued level as the engines accelerated up to TOGA thrust, and the aircraft settled into the trees as the engines ingested tree parts.

Despite a jammed exit door, most passengers were able to leave the aircraft before it was consumed by fire from the burning fuel. Two young children and an adult (presumed to have gone back to help) died from smoke inhalation.

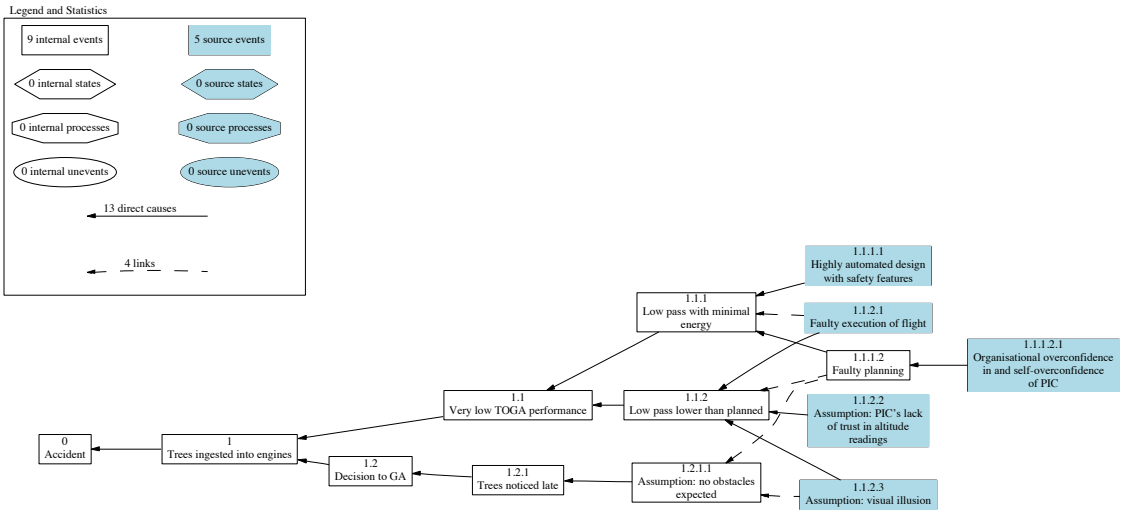

Abbildung 3.5: The Habsheim WB-Graph

Figure 3.5 shows a WB-Graph causally relating the major features of the accident flight, including preparation, from the official report.

Controversy

The accident became controversial when the captain, who was piloting the aircraft during the accident flight, publically asserted

- that the engines did not respond as designed to his TOGA thrust request;
- that about 4 seconds of recording data were missing from the flight data recorder (FDR) trace;
- that there were at least two different FDR boxes presented to the public as “the” FDR, and/or visible at the accident site
- that some of the data ostensibly from the FDR did not fit some of the facts about the flight;
- that required legal procedures for securing the FDR and taking it for analysis were not followed; insecure procedures were followed.

The captain wrote a book containing his version of the events, published a short while after the accident, and other books suggesting official miscreance have appeared. A

decade later, another book about the events is planned to be published.

We may take it as uncontroversial that, had the engines reached TOGA thrust, say, some two seconds earlier, the aircraft would likely have avoided settling into the trees, and thus avoided the crash altogether.

Further Evidence

There was a private video made of the accident fly-by by a spectator at the airshow. This video corroborated the altitude at various points of the fly-by, the timing of events, including (through sound-spectral analysis) the % N1 levels of the engines, the start of thrust increase on the engines, and the settling into trees.

The engines as certificated require up to about 8 seconds to increase from 29% N1 up to TOGA thrust. The official FDR data showed that they performed better than their certification parameters.

Evaluation of the Two Versions

Our concern in evaluating the accident is to identify causes and other contributing factors in order to increase knowledge about safety-related aircraft and crew performance and to mitigate undesirable or unsafe features in future operations.

Thus the sole significant assertion for our purposes amongst those made by the captain is that the engines did not perform according to specification when TOGA thrust was commanded.

What difference would this make to the WB-Graph in Figure 3.5? Indeed, none at all. At the level of detail at which the major factors are stated, the only factor under dispute would be Factor 1.1, "*Very low TOGA performance*. Both versions agree this was so, although for different reasons. Both versions agree that the manoeuver was commenced at commanded thrust equivalent to 29% N1, and that the manoeuver had been practiced, and was usually conducted, commencing at much higher N1 levels. Both versions agree on the descent profile, and that the flight-idle power setting was a result of that. Both versions agree that the aircraft was piloted to within 30 feet of the runway, although the captain planned to overfly at 100 feet. The incomplete and partially legally unsuitable planning, and the lack of oversight, are likewise uncontroversial.

The Political Controversy

As far as our interest goes, then, any dispute is about the exact level of TOGA performance, which disappears into the details when we are looking at the major factors contributing to the accident.

However, the high-visibility political controversy at the time was concerned not just with how the authorities may or may not have acted in the aftermath of the accident, but whether this "wonder aircraft", the A320, in fact could perform according to its manufacturer's and operator's claims. We can see clearly from the WB-Graph that this latter dispute is a matter of mere technical detail as far as the accident is concerned; it does not affect the causal relations of the major factors at all. The asserted performance difference, while passing the Lewis semantic test for a causal factor, is a question of a finer difference that is subsumed within one of the major factors: it is undisputed that the TOGA performance of the aircraft did not suffice to avoid the trees. According to the official evaluation, it could not have been better. The captain thinks it could have been. That is all.

Had the status of this technical dispute been available and appreciated at the time, we can speculate that the major political controversy over the introduction of the A320 into service, following the accident, might have taken a much different form.

3.6 Conclusions

The two examples show that objective reasoning methods, had they been used during the investigation and ensuing controversy in these two cases, might have cast a very different light on things. If the methods of reasoning are not generally accepted and open to independent checking, then it is open to anyone to criticise and query for any reason they wish, and if two parties to a discussion reach significantly different conclusions, then there are no further ways of deciding the issues than deciding whom one believes. This is a highly unsatisfactory situation, and gives grounds for introducing objective reasoning methods. If reasoning methods are agreed to be rigorous and objective, then all parties to a discussion are bound to abide by the results.

Two questions: Do such methods exist, and how severe are the problems that stem from lack of rigor? Our use of the Lewis semantics for causality, and the related

method WBA, show that the answer to the first question is yes.

The second question can be answered by considering what might have happened had a WB-Graph been available.

In the case of Warsaw, had a WB-Graph been constructed by the report writers based on the content of their report, they would have identified omissions in their statement of probable cause, and attention would have been brought to bear on the presence of an airport construction which adversely affected safety. Anecdotes say the mound is still there.

In the case of Habsheim, the heated political debate about the safety of the design of a new aircraft, and its consequences for public acceptance of the aircraft, might have evaporated, in favor of a technical performance debate and review of the sort which goes on every day at aircraft design and manufacturing plants.

Two anecdotes cannot prove a general hypothesis, but they may persuade. My purpose has been to persuade that objective methods of reasoning in accident evaluations are not just an exercise for academics. I believe they would have significant benefits, not only for accident investigation and the safety of air travel, but also for public debate as a whole.

There is another point worth remarking, again while taking care not to draw general conclusions from two individual cases. Both were publically high-profile accidents in which the digital automation on the aircraft was considered by many to have played a major contributory role. It is interesting to observe, when the causal reasoning is finally laid out, how few of the many factors involved in either of these accidents directly concerned the digital automation.