The Epistemics Of Accidents

C.W. Johnson
Glasgow Accident Analysis Group,
Department of Computing Science, University of Glasgow,
Glasgow, United Kingdom, G12 8QQ.
E-mail: johnson@dcs.glasgow.ac.uk
WWW: http://www.dcs.gla.ac.uk/~johnson

ABSTRACT
Human intervention has played a critical role in the causes of many major accidents. At Three Mile Island, the operators isolated a reactor from its heat sink. The pilots of the Boeing 737 at Kegworth managed to shut down their one functioning engine. The staff of the London Underground continued to allow trains to deposit passengers in Kings' Cross after the fire had started. In retrospect, it seems impossible to predict all of the ways that human intervention might threaten safety. It is, therefore, important that we learn as much as possible from those accidents that do occur. This task is complicated because conventional accident reports contain many hundreds of pages of prose. They present findings drawn from many different disciplines: metallurgy; systems engineering; human factors; meterology. These findings are, typically, separated into a number of distinct chapters. Each section focuses upon a different aspect of the accident. This makes it difficult for designers to recognise the ways in which operator intervention and system failure combine during major failures. The following pages use formal methods to address these concerns. It is argued that mathematical specification techniques can represent human factors and system failures. Unfortunately, formal methods have not previously been used to analyse the factors that motivate operator intervention. This paper, therefore, argues that epistemic extensions of mathematical notations must be recruited in order to support the formal analysis of major accidents.

Keywords: formal methods, accident analysis; epistemic logic; operator error; human factors.

1 Introduction

This paper addresses two weaknesses in the way that accidents are currently documented and reported. Firstly, accident reports are frequently inconsistent and incomplete. These documents contain many hundreds of pages of natural language analysis that are prepared by many different experts. This makes it difficult to obtain a clear and coherent view of the many different events that lead to major accidents. The second problem is that there is relatively little integration between systems engineering and human factors recommendations. These findings are usually separated into different chapters. As a result, the events that play a central role in one view of an accident are often ignored or omitted in other sections of a report. This makes it difficult for

readers to understand the complex ways in which human and system failures combine during major accidents.

1.1 The Problems of Accident Analysis

Figure 1 illustrates the way that accident reports help to communicate the findings of accident investigations. After an accident occurs, forensic scientists, engineers and human factors analysts conduct an enquiry into the probable causes of any failure (AAIB, 1989). These causes are then documented in a final report which is distributed to commercial, governmental and regulatory organisations. These documents are intended to influence the design and implementation of similar systems. They are published with the aim of preventing future accidents. Unfortunately, if companies cannot use these reports in an effective manner then previous weaknesses may continue to be a latent cause of future failures.

A number of factors make it difficult to produce clear and consistent accident reports. Accident reports contain the work of many different experts. Forensic scientists, metallurgists, meterologists as well as software engineers and human factors experts all present their findings through these documents. Their conclusions are, typically, separated into a number of different chapters. This has important consequences for the 'usability' of the resulting documents. It can be difficult to trace the ways in which system 'failure' and operator 'error' interact to produce the conditions for major accidents. Readers must trace the ways in which the events described in one chapter interact with those mentioned in other sections of the report. For example, the Air Accident Investigation Branch's report into the Kegworth accident contains a section on the `Engine Failure Analysis` and another on the `Crew Actions` (AAIB, 1989). Many of the events that are mentioned in the first chapter are not presented in the second and vice versa. This makes it difficult to trace the exact ways in which equipment failures complicated the task of controlling the aircraft (Johnson, 1994, 1995).

The poor level of integration between human factors and systems analysis can partly be explained by the lack of any integrated tools. For example, the United States' Department of Energy Standard for Hazard Categorisation and Accident Analysis (DOE-STD-1027-92) identifies a number of techniques that can be used to support accident analysis. These include Hazard and Operability Studies (HAZOPS), fault trees, probabilistic risk assessments and Failure Modes, Effects and Criticality Analysis (FMECA). None of these approaches provides explicit means of representing human factors 'failures' (Reason, 1990). This paper, therefore, identifies more integrated means of analysing operator 'error' and system 'failure' during major accidents.

1.2 The Kegworth Accident

The events leading to the Kegworth disaster are used to illustrate the argument that is presented in this paper. The Air Accident Investigation Branch's investigation concluded that the accident was caused by a complex interaction between system and human failures. The main events leading to the accident can be summarised as follows. A British Midlands Boeing 737-400 took off from Gatwick on course for Belfast. Thirteen minutes later the turbofan in the number one (left) engine fractured. This caused a series of compressor stalls in that engine. The pilots were unsure about what had happened. There were excessive vibrations throughout the airframe. Loud noises and a smell of fire could also be detected in the cockpit. There were transitory fluctuations on the primary instrument displays and a maximum reading on the airborne vibration meter (AVM) for the left engine. The crew's initial reaction was to diagnose a fire. The First Officer commented:

"It's a fire Kevin. It's a fire coming through."

In response the commander dis-engaged the auto throttle and said:

"Which one is it though?"

The First officer replied:

"Its the le, its the right one."

The commander then asked the First Officer to throttle the right engine back. This was a mistake because it was the left hand engine that was faulty. The healthy right engine was throttled back. Surprisingly, when the First Officer did this all indications of trouble in the left engine went away except for a maximum reading on the AVM for the left engine. The pilots did not notice this warning. The smoke cleared and so the commander ordered the right engine to be shut-down. During this period the AVM reading for the left engine dropped from a maximum of five to two units, slightly above normal. This suggested that the pilots had solved the problem.

It was decided to divert the aircraft to East Midlands Airport. This activity engaged the First Officer for some time since he had difficulty in re-programming the flight management system with the diversion route. The aircraft appeared to be flying normally and there was no reason to suspect that the wrong engine had been shut down. Passengers and cabin crew had seen smoke and sparks emanating from the left engine and smoke entered the passenger cabins. The cabin steward went forward to the flight deck and reported that the passengers were panicky. The commander then used the intercom to reassure the passengers. He said that they had experienced trouble with their right hand engine but that the problem had been solved by shutting it down. Survivors of the accident reported noticing the discrepancy between what they had witnessed the problems with the left engine and the Captain's solution which involved the right engine. Approximately fifteen minutes later, as the aircraft was being manouvered for its final approach to runway twenty-seven at East Midlands Airport, there was a sudden fall in power in the left engine. Attempts to re-start the right engine failed. The commander then tried to stretch the glide of the aircraft. He failed and the aircraft crashed on the embankment of the M1 motorway two hundred yards short of the start of the runway. A total of forty-six passengers died and a further seventy-four suffered serious injury.

1.3 Outline Of The Paper

Section 2 describes the limitations of existing approaches to accident analysis. Fault trees cannot easily be used to represent complex synchronisation properties. Petri Nets provide limited means of representing and reasoning about human 'error'. Failure Modes, Effects and Criticality Analysis cannot easily be used to examine the knock-on effects that characterise the events leading to many major accidents. Section 3 builds on this analysis and explains how temporal extensions to first order logic can provide an integrated account of the Kegworth accident. Section 4 then illustrates an important benefit of this approach; formal reasoning techniques can be applied to the logic as a means of validating the conclusions that are presented in an accident report. Section 5 builds on this and argues that the same notation can also be used to describe requirements that are intended to prevent accidents from recurring in future applications. Unfortunately, this approach cannot easily be used to represent and reason about the motivations that lie behind operator 'errors'. Section 6 demonstrates that epistemic extensions to our first order, temporal logic can be used to overcome this problem. This provides a re-interpretation of

the findings that were presented in the Air Accident Investigation Branch's report. A number of cognitive scientists have extended the relatively simple epistemic notation that is advocated in this paper. Section 7, therefore, describes how further work might determine whether the additional complexity of these notations can be justified in terms of additional expressiveness when investigating the causes of major accidents. Finally, Section 8 presents the conclusions that can be drawn from this research.

2 The Limitations of Existing Techniques

This section reviews a number of accident analysis techniques that might be used to represent and reason about human 'error' and system 'failure'. Two principle weaknesses are observed. Firstly, they do not capture the temporal properties that have a profound impact upon the course of major accidents. Secondly, they fail to provide adequate means of reasoning about operator behaviour.

2.1 Failure Modes, Effects And Criticality Analysis

Table 1 illustrates a product of Failure Modes, Effects And Criticality Analysis (FMECA). Operators must schedule increased maintenance in order to avoid the fractured turbofan that caused the engine failure. Better training must be provided if the First Officer is to avoid the verbal slip that led to the wrong engine being throttled back. Unfortunately, FMECA

Component	Failure Mode	Failure Cause	Criticality	Improvement
Turbofan	Fracture	Fatigue	High	Maintenance
First Officer	Verbal slip	High workload	Medium	Better training

Table 1: A Failure Mode, Effects And Criticality Analysis Table.

tables provide extremely poor support as a means of representing and reasoning about operator intervention. Terms such as More inspection do not capture the detailed tasks that must be performed if future accidents are to be avoided. The second line of analysis in the table can be used to re-iterate this point. There is no means of representing the precise conditions that led to the High workload. This, in turn, makes it difficult for analysts to identify means of avoiding the problem in future systems. Generic 'improvements' such as Better training are not precise enough. A related technical criticism about these tables is that they lack any structuring mechanism. This makes it difficult to identify the knock-on effects that faults have upon other components in the system. For example, Table 1 does not describe the impact that increased maintenance will have upon the productivity of the aircraft. FMECA might be extended to capture these relationships but this would reduce the tractability of the tabular representation.

2.2 Fault Trees

Fault-trees are in widespread use amongst systems engineers (Vesely, 1981). For example, Figure 2 was produced by the Eindhoven Safety Management Group in response to the Kegworth accident. The leaves of the tree represent primitive causes. The AND gates, represent the conjunction of factors leading to the accident. For example, the bottom-right branch of this fault-tree shows that the pilots' failure to consider the impact of the auto-throttle helped to mask the symptoms of the accident. There are, however, a number of limitations that restrict

the utility of this notation for the analysis of interactive systems. The European Federation of Chemical Engineering's International Study Group On Risk Analysis concludes:

> "Fault-trees have difficulties with event sequences... parts of systems where sequence is important are, therefore, usually modelled using techniques more adept at incorporating such considerations" (European Federation of Chemical Engineering, 1985).

Gate conditions are assumed to hold simultaneously or at some time after their branch conjunctions and disjunctions. The leaves of a tree are assumed to hold simultaneously or at some time before their parents. This lack of temporal information is a significant limitation because event sequences affect the quality of interaction between an operator and their system. Determining the order in which faults occur and commands are issued forms an important stage in any accident analysis.

2.3 Petri Nets

Petri nets have been explicitly developed to capture concurrent, temporal properties of complex systems. Figure 3 illustrates how this notation can also be used to represent human factors and system 'failures' (Johnson, McCarthy and Wright, 1995). This diagram is constructed from places and transitions. Places are represented by unfilled circles. They denote the facts that are known during various stages of an accident. Transitions are represented by the boxes in Figure 3. They denote the events that lead to those facts being true. For example, if the transition labelled `Fan blade fracture` takes place, or 'fires', then the two places that lead from it would be true. This would mean that `There is smore` and `There is vibration in engine number 1`. A transition can only fire if all of the places that lead to it are true. In our example, this would enable the transitions labelled `Smoke enters ventillation` and `Sensors begin to detect vibration`. The process of marking and firing helps to capture the sequential and concurrent properties that cannot be represented using fault trees. However, a number of problems limit the utility of Petri Nets for accident analysis. In particular, it is difficult to identify appropriate labels for the places and transitions. On the one hand, they must be concise and precise if the networks are to remain relatively tractable. On the other hand, they must provide necessary contextual information about the events leading to the accident. For example, the label `1st Officer is liable to make a verbal slip` provides little idea of the justifications that might support such an analysis.

3 Representing an Accident

This section demonstrates that logic can be used to represent and reason about human 'error' and system 'failure' during major accidents. It is not our intention to formalise all aspects of an accident. Rather, the notation is being used to explicitly represent those factors that are assumed to have directly contributed to the failure.

3.1 Critical Entity Tables

It is important to identify the different agents or entities that contribute to system 'failure' and operator 'error'. The following list presents a number of categories that can be used to group these 'entities'. It is important to emphasise that the list provides a heuristic guide. It

is the product of experience in applying this technique (Johnson, 1994, Johnson and Telford, 1996), rather than the consequence of theoretical analysis. It is entirely possible that subsequent analysis might identify as number of additional categories:

- *Operators.* It is important to identify the key individuals that were involved in an accident. This helps analysts to understand the ways in which operator intervention affects the course of system failures. For example, the Commander and First Officer played a key role in the Kegworth accident;
- *Tasks.* Analysts must understand the tasks that operators were or should have performed if they are to trace the ways in which human intervention safeguarded the system or exacerbated any failures. For instance, the task of re-programming the navigation equipment occupied much of the First Officer's time during the immediate aftermath of the engine failure;
- *Components.* Analysts must identify the critical process components that are assumed to have failed during an accident. For example, the fan blade that fractured in the left engine is a focal point for any engineering account of the accident;
- *Displays.* It is important to identify the information that was presented to operators during an accident if analysts are to understand the human reaction to emergency situations. Much of the AAIB report centered around the presentation of information about the faulty engine through the Airborne Vibration Monitors (AVM);
- *Commands.* An operator's response to an accident is frequently determined by the range of commands that they are provided with by their systems. In our case study, the crew had the option of idling their engines or of shutting them down. The distinctions between these different commands played a critical role in the final tragedy;
- *Utterances.* It is important for analysts to identify any communication between the operators that are involved in an accident. Misunderstandings have a profound impact upon safety. Previous sections have already described the First Officer's comment, "Its the le, its the right one", that may have sown doubt in the Commander's mind about the exact location of the fault.

The following table shows the critical entities that can be identified in the opening sections of the AAIB report. AVM represent the Airborn vibration monitor. FMS represents the flight management system. The important point is that analysts should explicitly represent the various 'entities' that contribute to an accident. These entities are frequently hidden within the labels of FMECA or Petri Nets. The development of such a table also helps to avoid some of the ambiguity that can arise during accident analysis and reporting. For instance, the AAIB report refers to the number 1 and 2 engines, it also refers to the left and right engines. Here we simply refer to the *no_1_engine* and the *no_2_engine*. The development of a critical entity table, therefore, helps to establish a common reference scheme for accident investigation. It is important to note that in formal terms, the elements of Table 2 define the types that will be used to analyse the Kegworth accident.

3.2 The Events Leading To Accident

First order logic can be used to represent and reason about the events that lead to an accident. In order to do this, analysts must first provide the source material for the formalisation process.

Operators	*Tasks*	*Components*
commander	reprogram_fms	no_1_engine
first_officer	radio_communication	no_2_engine
cabin_steward		blade_dampers
passengers		fan_blade_17
		shroud
		compressor
Displays	*Commands*	*Utterances*
avm	shut_down	'It's a fire Kevin. Its a fire coming through'
fms	throttle_back	'Which one is it though?'
	disengage	'Its the le, its the right one'

Table 2: Critical Entity Table for the Kegworth Accident.

In this case, we are using the AAIB report for the Kegworth disaster. Elsewhere we have used the preliminary documents that are produced before a final report is drafted. For example, the following quotation describes the primary cause of the Kegworth accident:

> "The No.1 engine suffered fatigue of one of its fan blades which caused detachment of the blade outer panel. This led to a series of compressor stalls, over a period of 22 seconds until the engine autothrottle was disengaged." (Finding 19, page 144)

From this it is possible to identify two systems problems; a fan-blade fractured and there were compressor stalls in the number one engine. Unfortunately, it is not possible to identify which of the fan-blades actually failed from the previous quotation. Readers must search through earlier sections of the report to find this information:

> "Since blade No.17 was the only fan blade to have suffered serious fatigue, it was concluded that the initiating event for Event 2 must have been the ingestion by the fan of a foreign object." (page 115).

Logic provides a means of combining the facts that are embodied within these natural language statements. The intention is to provide a common framework that can be used to describe the events that led to the failure. This reduces the burdens that are associated with cross-referencing the many different sections of a report:

$$fracture(no_1_engine, fan_blade_17). \quad (1)$$

$$stall(no_1_engine, compressor). \quad (2)$$

The previous clauses illustrate how logic can be used to represent information that may be distributed throughout a natural language report. A related problem is that in some cases, it may be impossible to provide readers with the information that they are searching for. Given the limitations of flight data recorders and sensing technology there can be genuine uncertainty about the course of events during an accident. In such circumstances, it is important to explicitly represent the absense of information. For example, Finding 19 does not mention who issued the command to disengage the autothrottle. This lack of information can be explicitly represented

using the $\exists$ quantifier (read as 'there exists') and a variable, c, to stands for the unspecified crew member. The important point here is that any ambiguity should reflect a lack of evidence rather than any vagueness in the use of language:

$$\exists c : input(c, disengage, auto_throttle). \tag{3}$$

Unfortunately, previous clauses suffer from the same problems that limit the utility of FMECA and fault trees. They abstract away from the timing information that is essential to an understanding of most major accidents. This limitation can be avoided by introducing time-points into the previous formalisation. For example, an analysis of the Flight Data Recorder revealed that the probable starting point for the Kegworth accident was 20:04 hrs:

$$fracture(no_1_engine, fan_blade_17, 20_04_00). \tag{4}$$

The previous clauses illustrates the problems of constructing an unambiguous time-line for the events that contribute to an accident. Finding 19 provided a number of relative timings that could only be calculated in relation to the fan-blade fracture. This fixed starting point was documented in Appendix 4 of the main report. Such relative timings introduce considerable burdens for the readers. They must notice and remember the 20:04:00 starting point in order to calculate the real-time at which the autothrottle was disengaged. Finding 19 suggests this took place some 22 seconds after the fracture:

$$\exists p : input(p, disengage, auto_throttle, 20_04_22). \tag{5}$$

It is less easy to specify exactly when the compressor stalls took place. Finding 19 suggests that they occurred between the initial fan-blade fracture and the disengagement of the auto-throttle. We can again use the $\exists$ quantifier to represent this uncertainty. The compressor must have stalled at some time, t, between 20:04:00 and 20:04:22:

$$\begin{aligned} &\exists t : stall(no_1_engine, compressor, t) \wedge \\ &before(20_04_00, t) \wedge before(20_40_22, t). \end{aligned} \tag{6}$$

The problems of constructing a time-line for the events in an accident are not simply limited to system failures, such as those described by previous clauses. The evidence that is necessary to reconstruct operator behaviour may also be scattered throughout natural language documents:

> "The flight crew did not assimilate the readings on the engine instruments before they decided to throttle-back the No. 2 engine. After throttling back the No. 2 engine, they did not assimilate the maximum vibration indication apparent on the No. 1 engine before they shut down the number 2 engine 2 minutes after the onset of vibration, and 5nm south of EMA (East Midlands Airport). The aircraft checklist gave separate drills for high vibration and for smoke, but gave no drill for a combination of both." (Finding 7, page 142)

In order to formalise this finding, it is first necessary to know when the crew decided to throttle-back the number 2 engine. This information was introduced some 140 pages earlier:

> "The commander's instruction to throttle back was given some 19 seconds after the onset of the vibrations when according to the FDR, the No. 2 engine was operating with steady engine indications" (page 3)

Again, however, in order to provide a timing for the decision to throttle-back the number 2 engine, we must first know when the onset of the vibrations occurred. This final piece in the jigsaw can be found earlier on page 3:

> "At 2005.05 hrs, as the aircraft was climbing through FL283 some 20nm south-south-east of East Midlands Airport, the crew experienced moderate to severe vibration and a smell of fire." (page 3)

We know from the previous two quotations that the command to throttle back the number 2 engine was issued nineteen seconds after the onset of vibration. This places the action at 20:05:24. From Finding 7 we know that the decision to shut-down the number one engine was taken some two minutes and seven seconds after that:

$$\exists c : input(c, throttle_back, no_2_engine, 20_05_24). \tag{7}$$

$$\exists c : input(c, shut_down, no_2_engine, 20_07_31). \tag{8}$$

Given clauses (7) and (8) we can now formalise the events described in Finding 7. It is important to note that the $\forall$ (read as 'for all') quantifier provides a means of specifying that at all times before 20:05:24 it was not the case that the commander or the First Officer observed the engine information system for the number 1 and number 2 engines. Here we assume that t ranges from the start of the accident which we take to be the time of the fan blade fracture, that is to say 20:04:00. The term *pei* stands for primary engine instrumentation:

$$\begin{aligned}&\forall t : \neg(observe(commander, pei(no_1_engine), t) \vee \\ &observe(commander, pei(no_2_engine), t)) \wedge before(t, 20_05_24).\end{aligned} \tag{9}$$

$$\begin{aligned}&\forall t : \neg(observe(first_officer, pei(no_1_engine), t) \vee \\ &observe(first_officer, pei(no_2_engine), t)) \wedge before(t, 20_05_24).\end{aligned} \tag{10}$$

The same approach can also be used to formalise other observations from Finding 7. In particular, that at all times before 20:07:31 and after 20:05:24 neither the commander nor the First Officer observed the airborne-vibration monitor for the number 1 and number 2 engines:

$$\begin{aligned}&\forall t : \neg(observe(commander, avm(no_1_engine), t) \vee \\ &observe(commander, avm(no_2_engine), t)) \wedge \\ &before(20_05_24, t) \wedge before(t, 20_07_31).\end{aligned} \tag{11}$$

$$\begin{aligned}&\forall t : \neg(observe(first_officer, avm(no_1_engine), t) \vee \\ &observe(first_officer, avm(no_2_engine), t)) \wedge \\ &before(20_05_24, t) \wedge before(t, 20_07_31).\end{aligned} \tag{12}$$

The complexity of the previous argument illustrates the very real difficulties that people face when attempting to form an unambiguous account of human 'error' and system 'failure' from an accident report. Unless readers can construct such time-lines then it is difficult, if not impossible, to form a coherent view of the events leading to major accidents.

4 Explaining the Causes of an Accident

The previous section has argued that first order logic can be used to represent and reason about the events that lead to major accidents. This is of little benefit if analysts cannot use the formalisation to achieve a better understanding of the causes of system 'failures' and operator 'error'. The following sections, therefore, apply formal reasoning techniques to identify causes of an accident that might otherwise be lost amidst the mass of contextual data in most accident reports.

4.1 System Failures

The AAIB report draws upon a wide range of disciplines to explain the causes of system 'failures' during the Kegworth accident. For example, a metallurgical examination of fan blade 17 established that:

> "... the fracture had propagated initially by fatigue, the origin of which appeared to be on the pressure face of the blade about 1.25mm aft of the true leading edge." (page 117)

The following clause illustrates how the conclusions of an accident analysis can be formalised. This is qualitatively different from the previous use of logic in the paper. Rather than asserting facts about the events that occurred, the following clause uses the $\Leftrightarrow$ operator (read as 'if and only if') to represent an inferred relationship. In other words, the metallurgical analysis suggests that the fan blade fractures if and only if it suffered from fatigue on its pressure face at the time of the failure. The use of the $\Leftrightarrow$ formalises a strong theory about the cause of the accident in that fatigue in the pressure face now becomes a necessary condition for the failure to occur:

$$\forall t : fracture(no_1_engine, fan_blade_17, t) \Leftrightarrow fatigue(no_1_engine, pressure_face(fan_blade_17), t). \quad (13)$$

Formal reasoning provides a means of linking such inferred relationships to the events that are known to have occurred during the accident. For instance, we can use the following inference rules to deduce the timing for the fatigue on the pressure face:

$$\forall t : P(t) \Leftrightarrow Q(t), \exists t' : P(t') \vdash \exists t' : Q(t'). \quad (14)$$

$$\forall t : P(t) \Leftrightarrow Q(t), \exists t' : Q(t') \vdash \exists t' : P(t'). \quad (15)$$

The first of these rules can be applied to clauses (4) and (13) to identify the moment at which there must have been fatigue on the pressure face of the faulty fan-blade. This might seem like a trivial finding based on the previous analysis from page 117. It is important to stress, however, that we have no evidence for any earlier fatigue based upon the extracts from the AAIB report:

$$fatigue(no_1_engine, pressure_face(fan_blade_17), 20_04_00). \quad [By\ applying\ 14\ to\ 4\ and\ 13] \quad (16)$$

Clause (13) provides a high-level explanation of the systems engineering problems that led to the Kegworth accident. It does not consider the possible causes of the fatigue that led to the fracture. This evidence is presented in the matallurgical examination of the engines:

"...it became clear that there was a generic problem affecting the -3C-1 variant of the CFM56... Examination of the blade vibration modes established during certification revealed that the nodal line indicated by the plane of the three fractures matched, amongst others, that to be expected from a second order vibration mode" (page 117-118)

The previous quotation had to be extracted from several pages of the AAIB account. This illustrates how readers must select critical causal information from the mass of contextual detail that is, typically, presented in an accident report. Unfortunately, this filtering process is typically both implicit and subjective. As a result, different readers will often identify a number of different causal factors from the same report. This makes it difficult to achieve concensus when implementing the recommendations from an accident investigation (Johnson, 1994). This problem can be addressed by using logic abstractions to explicitly represent the causal factors behind system failure. For example, the following clause formalises the assertion that fatigue in the pressure face of fan-blade 17 was caused by the fact that it was a -3C-1 variant of CFM56 and that at some time before the fatigue the engine had experienced second order vibrations. It does not represent the similarities between fan blade 17 and the nodal lines of previous fractures. This contextual information might have been included if analysts considered them critical to an account of the accident:

$$\forall t, \exists t' : fatigue(no_1_engine, pressure_face(fan_blade_17), t) \Leftarrow cfm56_3c_1(fan_blade_17, t) \wedge second_order_vibration(no_1_engine, t') \wedge before(t', t) \quad (17)$$

This clause formalises the assertion that was made in the quotation taken from pages 117 to 118 of the AAIB report. Proof techniques can again be used to determine whether this analysis is supported by the evidence that is presented in other chapters of the report. It is important to note, however, that we cannot simply apply the previous inference rules, (14) and (15), to conclude that there were second order vibrations, given that fatigue occurred. The implication, $\Leftarrow$, is weaker than the bi-implication, $\Leftrightarrow$. In particular, the conclusion of an implication may be true even though the antecedents are false. In effect, the two inference rules of $\Leftrightarrow$ are reduced to the following:

$$\forall t : P(t) \Leftarrow Q(t), \exists t' : Q(t') \vdash \exists t' : P(t'). \quad (18)$$

The practical consequences of this are that in order for (17) to hold we must establish that fan blade 17 was indeed a -3C-1 variant of CFM56 and that the engine experienced second order vibrations. These facts can be introduced as axioms providing that we have some supporting evidence. The type of the fan-blade is a matter of record. The fact that second-order vibrations occurred is a matter of debate. However, flying test-beds can be used to determine whether such vibrations could have occurred:

"The tests performed, after the two later incidents, on strain-guaged engines in the flying test-bed showed that the previously undetected system mode vibration of the fan was consistently excited when using -3C rated climb power above 10,000ft." (page 120)

This illustrates the way in which conventional accident analysis techniques can be used to provide the evidence that is necessary to support formal argumentation. It is possible to establish that fan-blade 17 was a -3C-1 variant of the CFM56. Results from the flying test-bed show that

second order vibration occurred in the No. 1 engine of the aircraft at some time prior to the fan-blade fracture:

$$cfm56_3c_1(fan_blade_17, 20_04_00). \tag{19}$$

$$\exists t : second_order_vibration(no_1_engine, t) \wedge before(t, 20_04_00). \tag{20}$$

This in turn enables us to form an argument in which the type of the fan blade and the second order vibration are used as evidence to prove that the fan blade fractured:

$$fatigue(no_1_engine, pressure_face(fan_blade_17), 20_04_00).$$
$$[By\ applying\ (18)\ to\ (17),\ given\ (19)\ and\ (20)] \tag{21}$$

The previous clauses have shown how formal argumentation and proof techniques can be used to identify the necessary evidence that is required to support particular theories about the causes of an accident. The intention is to provide a more precise and concise alternative to the many pages of natural language argumentation that are, typically, found in accident reports. For instance, the test-bed evidence is never explicitly linked to the probable causes of the fracture in the AAIB investigation. The reader is left to construct this inferential chain as they study the report. The use of formal proof techniques, forces the analyst to make these links explicit.

4.2 Operator Error

First order logic not only provides a means of representing the system failures that lead to major accidents. It can also be used to represent the human factors problems that exacerbate those failures. For example, the fan blade fracture might not have had disastrous consequences if the crew had correctly diagnosed the source of the problem. For example, the AAIB report found that:

> "Throughout the period of compressor surging, the No. 2 engine showed no parameter variations but because the first officer was unable to recall what he saw on the instruments, it has not been possible to determine why he made the mistake of believing that the fault lay with the No. 2 engine. When asked which engine was at fault he half-formed the word 'left' before saying 'right'. His hesitation may have arisen from genuine difficulty in interpreting the readings on the engine instruments or it may have been that he observed the instruments only during the 6 second period of relative stability between the second and third stages." (page 97)

This quotation reveals how accident reports may not provide a single, definitive analysis of the events leading to a major failure. Either the First Officer observed the engine information systems and experienced 'genuine difficulty in interpreting the readings' or 'he observed the instruments only during the 6 second period of relative stability'. The following timings for the period of quiecense in clause (23) were obtained from a trace of the critical engine parameters in Appendix 4 to the main report and could not be directly deduced from the pages surrounding the quotation, given above:

$$\exists t, \forall t' : input(first_officer, throttle_back, no_2_engine, t) \Leftarrow$$
$$\neg(observe(first_officer, pei(no_1_engine), t') \vee$$
$$observe(first_officer, pei(no_2_engine), t')) \wedge before(t', t). \tag{22}$$

$$\exists t, \forall t' : input(first_officer, throttle_back, no_2_engine, t) \Leftarrow$$
$$observe(first_officer, pei(no_1_engine), t') \vee$$
$$observe(first_officer, pei(no_2_engine), t') \wedge$$
$$before(20_05_10, t') \wedge before(t', 20_05_16) \wedge before(t', t). \quad (23)$$

It is possible to use formal reasoning as a means of demonstrating that these hypotheses are supported by the description of the accident that is presented elsewhere in the report. For example, we have already cited Finding 7 which states that:

> "The flight crew did not assimilate the readings on the engine instruments before they decided to throttle back the No. 2 engine." (Finding 7, page 142)

This was formalised in clause (10) which can, in turn, be used to support the previous theorem about the causes of the First Officer's 'error'. Informally, given that the First Office failed to observe his primary engine instruments before 20:05:24 and that the First Officer's decision to throttle back the working engine is explained by his failure to correctly monitor the instrumentation, we can conclude that the engine was throttled back:

$$\forall t' : input(first_officer, throttle_back, no_2_engine, 20_05_24) \Leftarrow$$
$$\neg(observe(first_officer, pei(no_1_engine), t') \vee$$
$$observe(first_officer, pei(no_2_engine), t')) \wedge before(t', 20_05_24).$$
$$[By\ instantiating\ 20_05_24\ for\ t'\ in\ (22)\ from\ (7)]. \quad (24)$$

$$input(first_officer, throttle_back, no_2_engine, 20_05_24).$$
$$[By\ applying\ (18)\ to\ (24),\ given(10)] \quad (25)$$

This proof effectively reconstructs the chain of evidence that is implicit in the previous quotation. It also helps to identify an important conflict between the human factors chapter and the rest of the report. The human factors account admits the possibility that the First Officer did observe the engine information systems but that he did so during the seconds of quiecence, see clause (23). In contrast the overall findings of the report, used in the previous proof, do not consider this possibility. The First Officer's error was explained in terms of his failure to monitor the systems rather than the failure of the systems to provide information about the engine problem throughout the failure. Discrepancies between different strands of analysis, such as those described above, can have profound consequences for accident reports. By ignoring the period of quiecense in the engine information system, it is possible to overlook other failings in the aircrew's information systems. In contrast, the human factors account in the AAIB report did consider the possibility that the First Officer correctly polled his displays. This led them to consider further problems with the presentation of information to the aircrew:

> "Unlike the transient fluctuations that would have appeared on the primary engine instruments, the reading on the No. 1 engine vibration indicator rose to maximum and remained there for about three minutes On the EIS (Electronic Engine Information Systems) however not only is the pointer of the vibration indicator much less conspicuous than a mechanical pointer but, when at maximum deflection, it can

be rendered even less conspicuous by the close proximity of the No. 1 engine oil quantity display, which is the same colour as the pointer and is the dominant symbology in that region of the display." (page 100)

This quotation raises a number of points. Firstly, terms such as 'for about three minutes' are unnecessarily vague when Finding 7 states that the maximum reading persisted for 2 minutes and 7 seconds after the onset of the vibrations. Secondly, this quotation illustrates the point that by assuming a blanket cause for operator 'error', as was presented in Finding 7, it is possible to ignore the more detailed problems of interface design and presentation that actually led to those failures. For example, the previous quotation might be formalised to state that the First Officer's fails to observe the airborne vibration monitor (AVM) at some point during the flight if the AVM is the same colour and is close to the oil quantity indicator:

$$\begin{aligned}\exists t : \neg observe(first_officer, avm(no_1_engine), t) \Leftarrow \\ close(avm(no_1_engine), oil_quantity(no_1_engine), t) \wedge \\ same_colour(avm(no_1_engine), oil_quantity(no_1_engine), t).\end{aligned} \quad (26)$$

It is important to emphasise, however, that this formalisation does not ensure that the analysis is 'correct'. The link between the oil quantity indicator and the First Officer's failure must be proven by conventional accident analysis techniques. In other words, given that we know about the detailed presentation and layout of the EIS, is it safe to conclude that at some time the commander will fail to monitor the AVM? The AAIB have addressed this issue by conducting a survey:

"The overall results, however, showed that a large majority of pilots considered that the EIS displayed engine parameters clearly and showed the rate of change of parameters clearly. Fewer that 10% of pilots reported any difficulty in converting to the EIS from the earlier hybrid instruments and only a few reported minor difficulty in alternating between the EIS and hybrid displays. 64% of pilots stated that they preferred instruments with electromechanical pointers to the EIS. (page 69).

The results of this survey are inconclusive; they do not directly support the argument presented in (26). Further evidence might be gathered through the experimental analysis of pilot performance with the AVMs. In terms of the reasoning process, this approach is quite different from the investigation that was used to support clause (17). In the previous case, airborne test-beds were used to gather evidence in support of a causal relationship that was assumed to be true. In this instance, we have all of the necessary evidence about the situation itself. We know the layout and location of the AVM and oil quantity indicator. In contrast, evidence is being sought to support the hypothesised relationship between that layout and the First Officer's error.

5 Formalising The Recommendations In Accident Reports

It is important to emphasise that accident reports are intended to perform a dual role, as shown in Figure 1. Firstly, they record the events that contribute to accidents. Secondly, they help to identify ways in which accidents might be avoided in other systems. An important benefit from the use of a formal notation is that the same language can be used to describe the critical events that contribute to major accidents *and* to represent requirements for future systems. For example, the AAIB report recommended that:

"The Civil Aviation Authority in conjunction with the engine manufacturer, consider instituting inspection procedures for the examination of the fan stage of CFM56 engines to ensure the early detection of damage that could lead to the failure of a blade (Made 10 February, 1989)" (Recommendation 4.2, page 149).

The following clause provides a formalisation of this recommendation. Fatigue does not occur in a fan-blade, f, of an engine, e, at any time, t, if that fan-blade has been inspected within some $time_limit$. This assumes that $fatigue$ is a specific form of 'damage that could lead to the failure of a blade' and that the $time_limit$ represents a requirement for the 'early detection' of the fault:

$$\forall e, f, t, \exists t' : \neg fatigue(e, pressure_face(f), t) \Leftarrow \\ inspected(e, f, t') \wedge before(t' + time_limit, t). \tag{27}$$

Formal proof techniques can be used to demonstrate that such approaches would have interrupted the events that led to the accident. For example, assuming that the fan-blade, f, of an engine, e, were regularly inspected then we might conclude that no fatigue occurs for that fan-blade. Clause (28) provides the necessary inference rule that supports this argument:

$$\forall t : P(t) \Rightarrow Q(t), P(t) \vdash Q(t). \tag{28}$$

$$\forall t, \exists e, f : \neg fatigue(e, pressure_face(f), t). \\ [By\ applying\ (28)\ to\ (27)\ given\ evidence\ of\ inspection] \tag{29}$$

This analysis depends upon an accurate inspection having taken place. Evidence for this can only be gathered by the normal techniques of safety analysis and quality assurance. It is, however, possible to use logic to specify the requirements for an accurate inspection. This approach is illustrated in Johnson (1995) and builds upon the constructive use of logic illustrated by clause (27). In contrast, the following clauses show how the previous safety requirement might have affected the Kegworth accident. The first stage in this analysis is to assume that the aircraft had been subjected to the inspections mentioned above:

$$\forall t : \neg fatigue(no_1_engine, pressure_face(fan_blade_17), t). \\ [Instantiation\ of\ e\ and\ f\ in\ (29)] \tag{30}$$

The second step is to introduce a proof rule which states that if we know that some fact Q is not true and that P is true if and only if Q is true, then we can conclude that P is not true:

$$\forall t : P(t) \Leftrightarrow Q(t), \exists t' : \neg Q(t') \vdash \exists t' : \neg P(t'). \tag{31}$$

Finally, given that regular inspection helps to prevent fatigue, from clause (30), and that fatigue led to the fracture of fan-blade 17, from clause (13), we can apply (31) to argue that Safety-Recommendation 4.3 would have prevented this cause of the accident:

$$\forall t : \neg fracture(no_1_engine, fan_blade_17, t). \\ [By\ applying\ (31)\ to\ (30)\ given\ (13)] \tag{32}$$

This analysis is open to many different forms of attack. For instance, clause (13) stated that fan-blade 17 fractured if and only if it suffered from fatigue. If this interpretation were incorrect

then there is no guarantee that detailed inspection of the component would have prevented the failure. The fan-blade might have fractured for other reasons. Similarly, the previous proof relies upon the effective implementation of Safety Recommendation 4.2, formalised in clause (29). The proof does not describe means of implementing or verifying the necessary inspections; neither does the AAIB report. The important point here is that none of these objections relate to the role of formalism in the investigation process. They relate to the informal safety arguments that are presented in the accident report. The role of the formal proof is to provide a structure for this argumentation and to expose the implicit analysis that is, typically, scattered throughout the many pages of conventional reports.

Human factors recommendations are amenable to the same sort of formal analysis as systems engineering requirements. For instance, the AAIB report argued that:

> "The Civil Aviation Authority should require that the engine instrument system on the Boeing 737-400, and other applicable public transport aircraft, be modified to include an attention-getting facility to draw attention to each vibration indicator when it indicates a maximum vibration (Made 30 March 1990)" (Recommendation 4.9, page 150).

Rather than directly formalising this safety requirement, as was done in (27), the following clause represents a technique that might be used to satisfy the recommendation. For example, an audio alarm might be used to re-enforce the visual representation of the AVM on the engine instrumentation system. The intention is that even if the AVM is close to other critical indicators then the attention grabbing mechanism should still ensure that the crew member, c, observes the warning for an engine e at any time, t:

$$\begin{aligned}\forall c, t, e : observe(c, avm(e), t) \Leftarrow \\ close(avm(e), oil_quantity(e), t) \wedge \\ same_colour(avm(e), oil_quantity(e), t) \wedge \\ audio_alarm(avm(no_1_engine), t).\end{aligned} \tag{33}$$

In order for such clauses to support the development of safety-critical applications, it is important that designers have some means of reasoning about the detailed characteristics of human-machine interfaces. For example, the following clauses state that an *audio_alarm* for the AVM has the timbre of a chime, the amplitude of the warning is 60dB and its pitch is 260Hz. By indexing the alarm by the *no_1_engine*, it is possible to specify that a different pitch or timbre is used to warn of an AVM problem on the number 2 engine:

$$\exists t : timbre(audio_alarm(avm(no_1_engine)), chime, t). \tag{34}$$

$$\exists t : amplitude(audio_alarm(avm(no_1_engine)), 60dBA, t). \tag{35}$$

$$\exists t : pitch(audio_alarm(avm(no_1_engine)), 260Hz, t). \tag{36}$$

One of the benefits of using a temporal notation is that designers can represent dynamic behaviours. The timbre of the audible alarm can be changed over time to represent the increasing urgency of the warning. It is important to note, however, that formal analysis does not guarantee the success of a particular presentation technique. Pilots may not be able to discriminate between *chimes*, *bells* and *horns* against the noise of a commercial cockpit (Patterson, 1990). Experimental analysis must provide the necessary evidence to support the requirement described

in (37):

$$\begin{aligned}\forall t, e : maximum(avm(no_1_engine), t) \Rightarrow \\ timbre(audio_alarm(avm(no_1_engine)), chime, t) \wedge \\ timbre(audio_alarm(avm(no_1_engine)), bell, t + time_out_1) \wedge \\ timbre(audio_alarm(avm(no_1_engine)), horn, t + time_out_2). \end{aligned} \quad (37)$$

This section has provided a stepping stone between accident analysis and the constructive use of formal methods in design (Bowen and Stavridou, 1993). We have shown that the recommendations in accident reports can be directly translated into requirements for future systems. We have also shown that logic can be used to represent and reason about the presentation techniques that satisfy those requirements. There are, however, a number of weaknesses in our approach. In particular, we are concerned to show that recent advances in the field of cognitive psychology can be used to guide a formal approach to accident analysis.

6 Epistemics and the Psychological Precursors of Error

This section argues that epistemic logics, or logics of belief, provide a bridge between user modelling and formal accident analysis. Such approaches are essential if analysts are to understand and combat the sources of human 'error' during major accidents.

6.1 Linking User Modelling and Accident Analysis

Reason (1990) argues that unsafe acts can be reduced by analysing the 'psychological precursors' of operator error. These precursors include the cognitive factors that influence interaction with complex systems. Unfortunately, previous sections have not demonstrated that logic can be used to capture these causes of major accident. Formulae, such as (1) and (7), have focussed upon the observable consequences of system and operator behaviour. Other clauses, such as (22) and (26), describe observable interaction problems, such as the layout of the AVM displays. None of these formulae represent the cognitive problems that frequently exacerbate system failures. For example, the AAIB report contains the following finding:

> "Flight crew workload during the descent remained high as they informed their company at EMA (East Midlands Airport) of their problems and intentions, responded to ATC (Air Traffic Control) height and heading instructions, obtained weather information for EMA and the first officer attempted to re-programme the flight management system to display the landing pattern for EMA." (Finding 10, page 143).

The following clause formalises this conclusion in a similar fashion to that of previous sections. The Commander experiences high workload if at any time they are controlling the aircraft's flight, conducting radio communications and are reviewing the symptoms associated with a particular failure:

$$\begin{aligned}\forall t : high_workload(first_officer, t) \Leftarrow \\ perform(first_officer, radio_communication, t) \wedge \\ perform(first_officer, reprogram_fms, t). \end{aligned} \quad (38)$$

Such an analysis is unsatisfactory. By simply listing a number of competing tasks, the formula ignores the wealth of human factors research into the causes of high workload (Kantowitz and Casper, 1988). This is dangerous because it does not automatically follow that the Commander's workload would be reduced to a safe level by simply reallocating one of his tasks. Similarly, the previous clause provides little constructive guidance for designers who might be faced with similar situations in the future. In Reason's terms, clause (38) cannot easily be used to identify the 'psychological precursors' of the accident.

An alternative approach to that shown above would be to formalise one of the many user models that have been developed over the last decade. Elements of such a model might then be used to represent the 'psychological precursors' of an accident in a formal notation. For example, Barnard's Interacting Cognitive Subsystems (ICS) model provides a complex but thorough account of cognitive problems such as high workload (Barnard and Harrison, 1989). Duke, Barnard and Duce (1995) have recently used a logic-based formalism to represent elements of ICS. Unfortunately, a number of problems limit the application of user models to support accident analysis. There are theoretical limitations. For instance, Rich (1983) catalogues the individual factors which bias the insights that may be obtained from these representations. This is a particular problem in the context of accident reports. The limited evidence that is available after a major disaster can prevent analysts from developing detailed user models of the individuals that were involved in an accident. The AAIB only provides brief biographies and limited psychological 'profiles' of individual crew members. Given this lack of detail, it is almost impossible for analysts to accurately identify the precursors of problems such as high workload. It is, therefore, important to identify some 'coarse grained' means of representing and reasoning about these cognitive factors that are described in accident reports. A further requirement is that this technique should be easily integrated with more advanced user modelling techniques as accident analysts gain a greater awareness of the psychological precursors to major accidents.

Epistemology, or the study of knowledge, has a history stretching back to the Ancient Greeks (Halpern, 1995). The modern development stems from von Wright (1951) to Hintikka (1962) and then Barwise and Perry (1983). This work has produced a number of logics that can be used to represent changes in an individual's knowledge over time. Epistemic logics have a number of attractions for accident analysis. In particular, it is possible to build more complex user models out of epistemic formulae. For example, each of the cognitive systems in the ICS approach contain an epistemic subsystem (Duke, Barnard and Duce, 1995). The rest of this paper, therefore, demonstrates that epistemic logics can capture the cognitive observations that are embedded within the AAIB report:

> "It is also likely that, if the No. 2 engine had not been shut-down, the accident would not have happened and some explanation must be sought for the commander's decision to shut it down. It is now known that the engine was operating normally but, because the decision to shut it down was made after its throttle had been close, having failed to recognise its normal operating parameters before closing the throttle, the crew could no longer confirm its normal operation by comparison with the No.1 engine instruments." (page 102)

The commander thought that there was a problem with the number two engine when we know that there was no problem with that engine:

$$\exists t : know(commander, problem(no_2_engine, t). \tag{39}$$

$$\forall t : \neg problem(no_2_engine, t). \tag{40}$$

It is important to emphasise the difference between epistemic clauses, such as *knows*, and other temporal formulae. In the former case, the logic represents an individual's beliefs about an accident as it progresses. In the latter case, formulae represent the analysts' knowledge about an accident after it has taken place. The use of the $\exists$ quantifier in (39) captures the observation that at some point during the accident, the Commander thought that there was a problem in the number 2 engine. The use of this quantification, rather than $\forall$ (read as 'forall'), enables designers to specify that there exists a subsequent time in the accident when the commander realised his mistake:

> "Fifty three seconds before ground impact ... there was an abrupt decrease in power from the No. 1 engine. The commander immediately called for the first officer to relight the No. 2 engine." (Findings 12 and 13, page 143)

The Commander knew that there was a problem with the Number 1 engine some time after 20:23:50. This timing was calculated from the observation that the decrease in power occurred fifty three seconds before the impact and from the fact that Appendix 4 of the main report records the moment of impact as 20:24:43. In the following clause, the temporal variable t' indicates that the Commander realised that the fault had developed at some time before he noticed the problem in the number one engine:

$$\exists t, t' : know(commander, problem(no_1_engine, t'), t) \wedge \\ before(20_23_50, t) \wedge before(t', 20_23_50). \quad (41)$$

It is possible to build upon this analysis. The change in the Commander's actions indicate that he did not know that there was a fire in the number one engine until fifty-three seconds prior to the impact:

$$\forall t : \neg \ (know(commander, problem(no_1_engine, t), t) \wedge before(t, 20_23_50)). \quad (42)$$

Having constructed this initial time-line for the commanders beliefs about the state of his system, it is possible to use formal proof rules to introduce evidence about the true state of the underlying systems engineering. This again demonstrates that formal notations can be used to integrate information that is, typically, distributed throughout the many different chapters of an accident report. The first stage in this process is to introduce the systems observation. In this case, we simply link the commander's knowledge to the underlying problem in engine number 1:

$$\forall t : \neg \ (know(commander, problem(no_1_engine, t), t) \wedge before(t, 20_23_50)) \wedge \\ fracture(no_1_engine, fan_blade_17, 20_04_22). \\ [Introduction\ of\ \wedge,\ given\ (42)\ and\ (4).] \quad (43)$$

The previous clause states that the commander did not know of the problem in the number 1 engine at any time before 20_33_50, the moment of the decrease in power, and that there was a fan-blade fracture in that engine at 20_04_22.

6.2 Epistemics and the Reasons for Belief

The use of an epistemic formalism helps analysts to identify the operator's perspective upon critical events during an accident. This is extremely important. In retrospect, it might seem obvious that the Commander should have considered the possibility of a problem with the

number one engine. At the time of the accident, things were not so clear cut. In particular, the Commander thought that the number two engine was failing and this may have prevented him from correctly diagnosing the failure in the number one engine. This is significant because the crew might not have decided to shut down their healthy engine had they entertained any doubts about their original diagnosis. The human factors analysis in the AAIB report found that:

> "Once the No. 2 engine had been shut down, it would appear that that the apparent absense of any manifestation of abnormality other than the No. 1 engine vibration indication, which they did not notice, persuaded both pilots that, in the commander's own words, 'the emergency had been successfully concluded and the left engine was operating normally'." (page 102).

In other words, the Commander's belief that engine number two was failing, prevented him from believing that there might have been a problem with the number one engine. The importance of this distinction cannot be under-emphasised. It was the decision to shut down the healthy engine rather than the failure to act on the faulty power unit that led to the crash. The following clause states that after the shut-down of the number two engine at 20:07:13 and before 20:23:50, when the Commander realised his mistake, he thought that there was a problem with the number two engine. He also thought that there was no problem with the number one engine. The timing for the decision to shut down engine number two is derived from Appendix 4 of the main report. The time at which the Commander realised his mistake is taken from clause (41):

$$\begin{aligned}\exists t, t' : & \mathit{know}(\mathit{commander}, \mathit{problem}(\mathit{no_2_engine}, t'), t) \wedge \\ & \mathit{know}(\mathit{commander}, \neg\ \mathit{problem}(\mathit{no_1_engine}, t'), t) \wedge \\ & \mathit{before}(t, 20_23_50) \wedge \mathit{before}(20_07_13, t) \wedge \mathit{before}(t', 20_27_13). \end{aligned} \quad (44)$$

It is important to compare the previous clause with (43). The initial formalisation stated that the Commander did not know about the problem in the engine; $\neg\ \mathit{know}(\mathit{commander}, \mathit{problem}(\mathit{no_1_engine}, t), t)$. In the previous clause, this is strengthened so that the Commander now knows that there is not a problem with the engine; $\mathit{know}(\mathit{commander}, \neg\ \mathit{problem}(\mathit{no_1_engine}, t'), t)$. In other words, the absense of knowledge has been replaced by a positive belief that something is not true. Both of these clauses describe the *commander*'s view of the accident and are independent of the actual state of the engine.

It is possible to generalise the findings of the previous analysis. The Commander thought that the number one engine was not on fire because he thought that the number two engine was on fire:

$$\begin{aligned}\forall t, t' : & \mathit{know}(\mathit{commander}, \neg\ \mathit{problem}(\mathit{no_1_engine}, t), t') \Leftarrow \\ & \mathit{know}(\mathit{commander}, \mathit{problem}(\mathit{no_2_engine}, t), t'). \end{aligned} \quad (45)$$

This clause illustrates some of the dangers that are inherent in the formalisation of accident reports. It is tempting to suggest that the positive diagnosis of a problem in one engine, actually interfered with the diagnosis of a problem in the other engine. This is implied but never explicitly stated in the accident report. There is a danger, therefore, that clauses such as (45) read too much into the analysis provided by previous quotations from the AAIB report. A second problem with the previous formalisation is that it does not state the reasons why the Commander believed that the number two engine was failing. One explanation is that it was difficult for the crew to observe the AVM warnings. The problems with this display were represented in clause (26). Other reasons were also cited:

> "After the accident, he (the commander) stated that he had judged the No. 2 engine to be at fault from his knowledge of the aircraft air conditioning system. His reasoning was that he thought the smoke and fumes were coming forward from the passenger cabin; the air for the cabin came mostly from the No. 2 engine; therefore the trouble lay in that engine." (page 98).

The following clause links the Commander's diagnosis to his knowledge about the air conditioning system. It is important to note that the moment when he believed there to have been an engine problem, t', occurred before the moment at which his belief holds, t. This accurately describes the situation in which the Commander is providing evidence about his knowledge at some time after the accident has occurred:

$$\begin{aligned}\forall t, t' : know(commander, problem(no_2_engine, t'), t) \Leftarrow \\ know(commander, air_feed(no_2_engine, passenger_cabin), t') \wedge \\ know(commander, smoke(passenger_cabin), t') \wedge before(t', t).\end{aligned} \tag{46}$$

The importance of this formalisation is that it enables analysts to explicitly represent and reason about the chains of belief that can influence an operator's intervention during system 'failures'. For instance, the following proof rule describes how inferences can be drawn from a series of implications:

$$\forall t : P(t) \Leftarrow Q(t), Q(t) \Leftarrow R(t), R(t). \vdash \exists P(t) \tag{47}$$

This rule can be applied to reason about the Commander's reliance on the number 1 engine during the Kegworth accident. Informally, the belief that smoke from the passenger cabin was passed through an air-feed from the number 2 engine helped the Commander to infer that there was a problem in that engine. The belief that there was a problem in that engine helped the Commander to infer that there was no problem in the number 1 engine:

$$\begin{aligned}&\forall t, t' : know(commander, \neg problem(no_1, engine, t), t'). \\ &[(47)\ \textit{applied to}\ (46)\ \textit{and}\ (45)\ \textit{given knowledge of air system, see below}]\end{aligned} \tag{48}$$

In order for this chain of inference to stand, analysts must prove that the Commander really did use information about the air-conditioning to influence his diagnosis of the problem in the number two engine. The following quotation casts doubt on the analysis embodied in clause (46). This again illustrates the point that formal reasoning techniques are not intended to replace the traditional skills of accident investigation:

> "While this reasoning (about the air conditioning system) might have applied fairly well to other aircraft he had flown, it was flawed in this case because some of the air conditioning for the passenger cabin of the Boeing 737-400 comes from the No. 1 engine... It seems unlikely that in the short time before he took action his thoughts about the air conditioning system could have had much influence on his decision. It is considered to be more likely that, believing the first officer had seen positive indications on the engine instruments, he provisionally accepted the first officer's assessment" (page 98).

The important point about the use of epistemic logic in this section is that it provides analysts with a means of representing and reasoning about the consequences of particular hypotheses,

such as that stated in the previous quotation. By arguing that the Commander did not use his knowledge of the air condition systems, the analyst breaks a chain of inference that explains the Commander's mistaken belief about the state of engine number 2. This, in turn, removes support for his diagnosis about the state of the number 1 engine. The consequences of the analysis on page 98 are never explicitly stated in the accident report.

6.3 The Epistemics of Group Work

The previous analysis forces analysts to identify other reasons why the Commander believed that there was a problem in the number 2 engine. The quotation cited above suggests that the Commander based his diagnosis upon an inference about the First Officer's beliefs. This is a form of recursive epistemics; the user's belief is based upon what he believes about another person's beliefs. The following clause formalises this analysis. The temporal variable t' represents the moment at which the First Officer is assumed to have diagnosed the failure in the number two engine. The variable t'' represents the moment at which the First Officer is assumed to have believed that the failure actually occurred:

$$\begin{aligned}\forall t, t', t'' : know(commander, problem(no_2_engine, t''), t) \Leftarrow \\ know(commander, know(first_officer, problem(no_2_engine, t''), t'), t) \wedge \\ before(t', t) \wedge before(t'', t').\end{aligned} \tag{49}$$

This formalisation does little to explain why the Commander inferred that the First Officer had accurately diagnosed a problem with the number two engine. The justification for this comes from the crew's verbal exchanges immediately before the decision to throttle back the working engine. The AAIB's transcript of the cockpit voice recorder provides the following details:

> "From the CVR it was apparent that the first indication of any problem with the aircraft was as it approached its cleared flight level when, for a brief period, sounds of 'vibration' or 'rattling' could be heard on the flight deck. There was an exclamation and the first officer commented that they had 'GOT A FIRE'. The autopilot disconnect warning was then heard, and the first officer stated 'ITS A FIRE COMING THROUGH', to which the first officer replied, 'ITS THE LE...ITS THE RIGHT ONE'. The commander then said 'OKAY, THROTTLE IT BACK'." (page 28)

From this evidence, it was the First Officer's initial reply to the Commander's question that led to the Commander's belief that the fault lay with the number two rather than with the number one engine. The following clause provides a means of explicitly representing this relationship which is implicit within the AAIB account:

$$\begin{aligned}\forall t, t' : know(commander, know(first_officer, problem(no_2_engine, t'), t'), t) \Leftarrow \\ speech(first_officer, 'Its\ \ the\ \ le,\ \ its\ \ the\ \ right\ \ one', t).\end{aligned} \tag{50}$$

An initial analysis of the interaction between the crew might indicate that the First Officer had made a simple verbal slip. However, previous clauses illustrate how epistemic logic can be used to explore the complex consequences of this 'error' for the Commander's beliefs about his system. Similar comments can be made about the First Officer. For example, the human factors analysis of the accident found that:

> "However, any uncertainty that he (the First Officer) may initially have experienced appears to have been quickly resolved, because, when the commander ordered him to 'THROTTLE IT BACK', without specifying which engine was to be throttled back, the first officer closed the No. 2 throttle." (page 97)

This indicates that these is a mutual dependency between the beliefs of the First Officer and the Commander. Both based their diagnosis upon what they believed the other person knew about the state of the system. These beliefs were, in turn, based upon their interpretation of their colleague's utterances:

$$\forall t, t' : know(first_officer, know(commander, problem(no_2_engine, t'), t'), t) \Leftarrow speech(commander, 'Throttle\ it\ back', t). \quad (51)$$

We have argued that epistemic logic can be used to capture critical observations about the human factors of major accidents. In our case study, it has been possible to distinguish between the Commander's lack of knowledge about the state of the number one engine and knowledge that the engine was not on fire. It has also been possible to express situations in which a positive belief that the fire was in the number two engine also prevented the Commander from believing that there might be a fire in the number one engine. We have shown that the Commander's diagnosis was influenced by what he believed the First Officer knew about the failure. Finally, we have shown that the First Officer's diagnosis was, in turn, influenced by what he knew about the Commander's beliefs. We have not, however, shown that epistemics can be used to help the design of future systems.

6.4 Epistemics And Design

One of the key recommendations from the AAIB in response to the Kegworth accident was that air crew must receive additional training in decision making under abnormal situations. The implicit intention behind this recommendation was that training would reduce the chances of communications problems, such as those described in (50) and (51), from jeopardising the safety of future flights:

> "The Civil Aviation Authority should review current airline transport pilot training requirements with a view towards considering the need to restore the balance in flight crew technical appreciation of aircraft systems, including systems response under abnormal conditions, and to evaluate the potential of additional simulator training in flight deck decision making.' (Recommendation 4.15, page 150).

This recommendation is pitched at an extremely high level of abstraction. This causes problems for commercial organisations because there are many different aspects of flight crew decision making that might be supported through additional training. Epistemic logics can, however, be used to reason about the specific requirements that satisfy recommendations such as that cited above. For example, simulator training might be conducted to ensure that the First Officer does not make assumptions about the Commander's diagnosis of a problem in a particular engine, e, at all times, t, without his direct verbal confirmation, s:

$$\forall e, t, s : know(first_officer, \neg\ know(commander, problem(e, t), t), t) \Leftarrow speech(commander, s, t) \wedge \neg confirmation(s, e, t). \quad (52)$$

An important benefit of this formalisation is that is can be used to analyse the impact which particular requirements might have had upon previous failures. For example, the AAIB report contains the observation that:

> "...the commander ordered him (the First Officer) to 'THROTTLE IT BACK', without specifying which engine was to be throttled back" (page 97).

In other words, the commander failed to adequately confirm his diagnosis with his command to throttle back the engine. The timings for this utterance can be obtained from the cockpit voice recorder and are recorded in Appendix 4 of the AAIB report:

$$\forall e, t : \neg confirmation('Throttle\ it\ back', e, t). \tag{53}$$

$$speech(commander, 'Throttle\ it\ back', 20_05_24). \tag{54}$$

We can now prove that the First Officer should **not** have assumed that the Commander knew about the problem in the number two engine given his command at 20:05:24. Informally the reasoning proceeds as follows, (52) states that the First Officer should not infer that the Commander knows which engine that the fault is in if they do not provide specific confirmation of that engine. We know that the Commander failed to provide this confirmation from clause (53). Therefore, the First Officer should not have concluded that the Commander know which engine was failing:

$$know(first_officer, \neg know(commander, problem(no_2_engine, 20_05_24), 20_05_24), 20_05_24) \Leftarrow$$
$$speech(commander, 'Throttle\ it\ back', 20_05_24) \wedge$$
$$\neg confirmation('Throttle\ it\ back', no_2_engine, 20_05_24) \wedge$$
$$before(20_05_24, t) \wedge before(t', t).$$
$$[Instantiating\ number_2_engine\ for\ e,\ 20_05_24\ for\ t'',\ 'Throttle\ it\ back'\ for\ s\ in\ (52).] \tag{55}$$

$$know(first_officer, \neg know(commander, problem(no_2_engine, 20_05_24), 20_05_24), 20_05_24).$$
$$[Applying\ (18)\ to\ (55),\ given\ (7)\ and\ (54).] \tag{56}$$

The previous clauses clearly represent the relationship between analysis and recommendations in accident reports. This is an important benefit for epistemic logics. Previous work has shown that some accident reports have made recommendations that cannot easily be justified in terms of the events that caused human 'error' and systems 'failure' (Johnson, 1997). It is important to emphasise, however, that the previous inferences depend upon the First Officer understanding the requirement described in (52). In other words, company regulations must ensure that aircrew do not make inferences about their colleague's diagnoses without explicit verbal confirmation. Unfortunately, neither natural language reports nor formal requirements can guarantee that operators will *always* behave in a particular fashion. The AAIB report recognises this when it argues that the crew":

> "...reacted to the initial engine problem prematurely and in a way that was contrary to their training" (page 148)

The following section, therefore, demonstrates that epistemic logic can be used to reason about the systems engineering that can be recruited to support the training of system operators.

7 Further Work: Belief Sets and Cognitive Science

The Air Accident Investigation Branch argued that the introduction of video cameras might have helped the first officer to review his beliefs about the state of the number 2 engine:

> "(this would) provide flight deck crews with views of cargo bays to aid in decision making following an associated fire warning" (page 95).

The intended effect of such changes to the systems engineering can be represented as follows:

$$\forall e, t : know(first_officer, \neg problem(e,t), t) \Leftarrow display(video, \neg\ problem(e,t), t). \quad (57)$$

The problem with this clause is that it fails to capture the impact that previous knowledge might have upon the interpretation of new sources of information. For example, the First Officer might not believe the video 'footage' if they already thought that the engine was failing. Wickens (1984) argues that such prior beliefs can have a profound impact upon subsequent operator interaction:

$$\forall e, t, \exists t' : \neg know(first_officer, problem(e,t), t) \Leftarrow know(first_officer, problem(e,t'), t') \wedge display(video, \neg problem(e,t), t) \wedge before(t', t). \quad (58)$$

The Civil Aviation Authority recognise this problem when they argue that:

> "The integration of video information into the routine and emergency procedures used by flight deck crews will need careful consideration. This aspect should be investigated in parallel with the equipment development, to ensure that the final system is compatible with the prescribed flight procedures"(page 96)

Relatively simple epistemic clauses, such as (58), help to identify particular problems that frustrate the introduction of novel information systems. Cognitive scientists have, however, developed a range of more complex epistemic notations. Further work intends to determine whether these logics can also be recruited ti support accident analysis. For instance, Alchourron and Makinson (1985) have developed belief sets to explicitly represent learning over time. Informally, these sets can be thought of collections of the facts that an individual knows about their environment. More formally, this approach introduces a language L of propositions whose elements include a number of belief sets, K_i, where the i subscript is used to index the set to a particular individual. A single belief set can be associated with every operator in the system. The advantage of this more complex epistemic notation is that it provides a framework for belief revision. When a new fact, A, becomes known to a user then there is a unique belief set, $K_i * A$, that represents the revision of K_i with respect to A. This is distinguished from the simple addition of information $K_i + A$ which leads to a contradiction if $\neg A \in K_i$. In other words, the revision operator, $*$, holds whether or not a fact was already held as one of the operator's beliefs. Conversely, when some item of information, B, is no longer known to be true by an operator then there is a unique belief set, $K_i \doteq B$ that represents a contraction of their earlier beliefs with respect to B. It is hypothesised that these $*, +$ and $\doteq$ operators provide a concise means of reasoning about the changes in knowledge that characterise an individual's view of an accident as it develops. It remains an open question as to whether these enhanced epistemic notations will actually deliver benefits that analysts can use to support the pragmatics of accident investigation.

Further work also intends to identify means of validating the techniques that are described in this paper. The UK Engineering and Physical Sciences Research Council have recently funded a major project to investigate the effectiveness of this approach compared to more conventional forms of accident analysis (GR/L27800). This raises a large number of methodological issues. In particular, it is a non-trivial task to recruit domain experts for the periods of time that are necessary to learn and apply our techniques. This is not such a severe limitation as it might appear. Mathematicians are already widely employed to develop continuous models of physical processes, such as fire, during accident investigations. However, abstract formulae provide non-formalists with an extremely poor impression of the events leading to an accident. Simulations provide a much better idea of the course of an accident. Elsewhere we describe how the Prelog tool can be used to develop interactive prototypes from formal descriptions (Johnson, 1993). Human factors analysts and systems engineers can animate the clauses that have been presented in this paper. The tool provides facilities for 'running' the traces of events that lead to system failures. This helps to 'debug' the description of an accident. If causal sequences are omitted then the simulation will not follow the course of the events that are described by eye-witnesses and monitoring applications. An important benefit of Prelog is that it can also be used to pose queries about changes in an operator's knowledge. This provides a means of monitoring the impact that warnings and displays may have had upon their assessment of the accident. Unfortunately, the existing interface is entirely textual. Further work intends to address this issue by developing Prelog's presentation facilities. One means of doing this would be to produce graphical time-lines from the logic clauses that have been presented in this paper.

8 Conclusions

Accident reports contain the findings of many different experts: systems engineers; forensic scientists; human factors specialists; meteorologists. These documents, typically, separate the work of each discipline into distinct chapters. This makes it difficult to form a coherent view of the way in which human factors and systems failures contribute to major accidents. This paper has used a mathematical notation to address this problem. In particular, we have used a temporal form of first order logic to represent system problems, such as a fan blade fracture, and operator 'errors', including a failure to observe critical warnings on primary engine instrumentation. We have also argued that it is important for analysts to reason about the causes of such 'failure'. In terms of the systems engineering; second order vibrations led to fatigue in a -3C-1 variant of a CFM56 fan-blade. Unfortunately, it is less easy to explain the motivating factors behind operator 'error'. Recent work has suggested that syndetic models might be used to analyse the way in which an operator's knowledge can affect their behaviour. This approach integrates formal specifications of system requirements with observations about a users' cognitive characteristics. We have extended this work by showing that it can support accident analysis. We have not, however, adopted the detailed propositional, implication, morphonolexical subsystems that have been a feature of previous approaches (Barnard and Harrison, 1989). This is justified by the limited evidence that is available about the cognitive and perceptual characteristics of system operators after major accidents. We have, therefore, argued that relatively simple epistemic notations can be used to represent and reason about changes to an operator's beliefs during the course of an accident.

Acknowlegement

Thanks go to the members of the Glasgow Accident Analysis Group, Glasgow Interactive Systems Group (GIST) and to the Formal Methods and Theory Group in Glasgow University. This work is supported by the UK Engineering and Physical Sciences Research Council, grants GR/JO7686, GR/K69148, GR/K55040 and GR/L27800.

References

Air Accidents Investigations Branch, Department of Transport (1990). *Report On The Accident To Boeing 737-400 G-OBME Near Kegworth, Leicestershire on 8th January 1989*, number 4/90, London, United Kingdom, 1990. Her Majesty's Stationery Office.

C. Alchourron and D. Makinson (1985). On the logic of theory change: Safe contraction. *Studia Logica*, 44:405–422.

P. Barnard and M.D. Harrison (1989). Integrating cognitive and system models in human computer interaction. In A.G. Sutcliffe and L.A. Macaulay, editors, *People And Computers V*, pages 87 – 104. Cambridge University Press.

J. Barwise and J. Perry (1983). *Situations And Attitudes*. Bradford Books, Cambridge, United States of America.

J. Bowen and V. Stavridou (1993). Safety-Critical Systems, Formal Methods And Standards. Software Engineering Journal,189-209, July 1993.

D. Duke, P. Barnard, and D. Duce (1995). Systematic development of the human interface. In *APSEC'95- Second Asia-Pacific Software Engineering Conference*.

R. Fagin, J. Halpern, Y. Moses, and M. Vardi (1995). *Reasoning About Knowledge*. MIT Press, Boston, United States of America.

J. Halpern (1995). Reasoning about knowledge: A survey. In *Handbook Of Logic And Artificial Inteligence: Volume 4 - Epistemic And Temporal Reasoning*, pages 1–34. Clarendon Press, Oxford, United Kingdom.

J. Hintikka (1962). *Knowledge And Belief*. Cornell University Press, Ithica, United States of America.

C.W. Johnson (1993). A probabilistic logic for the development of safety-critical interactive systems. *International Journal Of Man-Machine Studies*, 39(2):333–351.

C.W. Johnson (1993a). Specifying and prototyping dynamic human-computer interfaces for stochastic applications. In J.L. Alty, D. Diaper, and S. Guest, editors, *People And Computers VIII*, pages 233–248. Cambridge University Press, Cambridge, United Kingdom.

C.W. Johnson (1994). The formal analysis of human-computer interaction during accidents investigations. In G. Cockton, S.W. Draper, and G.R.S. Weir, editors, *People And Computers IX*. Cambridge University Press, Cambridge, United Kingdom.

C.W. Johnson (1995). The application of Petri Nets to represent and reason about human factors problems during accident analyses. In P. Palanque and R. Bastide, editors, *The Second Eurographics Workshop On The Design, Specification And Verification of Interactive Systems*, pages 345–357. Springer Verlag, Berlin, Germany, 1995.

C.W. Johnson (1997). Proving Properties of Accidents. Submitted to *Ergonomics*.

C.W. Johnson, J.C. McCarthy, and P.C. Wright (1995). Using a formal language to support natural language in accident reports. *Ergonomics*, 38(6):1265–1283.

C.W. Johnson and A. Telford (1996). Using Formal Methods to Analyse Human 'Error' and System Failure during Accident Investigations. *Software Engineering Journal*, 11(6).

B.H. Kantowitz and P.A. Casper (1988). Human workload in aviation. In E.L. Wiener and D.C. Nagel, editors, *Human Factors In Aviation*, pages 157–187. Academic Press, London, United Kingdom.

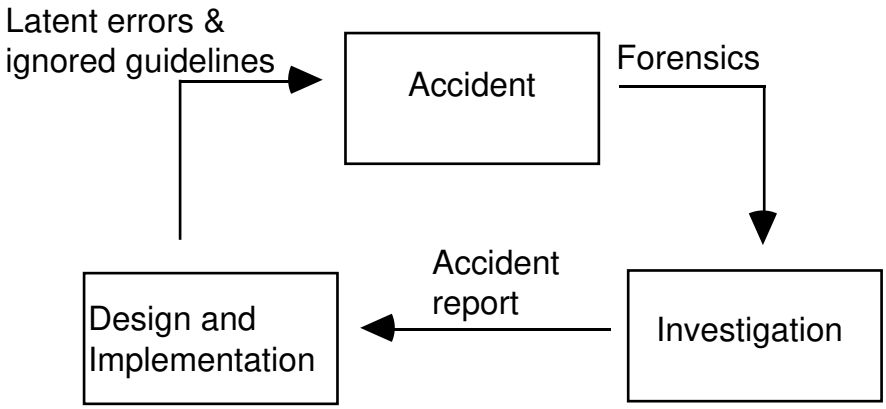
R.D. Patterson (1990). Auditory Warning Sounds In The Work Environment. In D.E. Broadbent, J. Reason and A. Baddeley, editors, *Human Factors In Hazardous Situations*, pages 37-44. Clarendon Press, Oxford, United Kingdom.

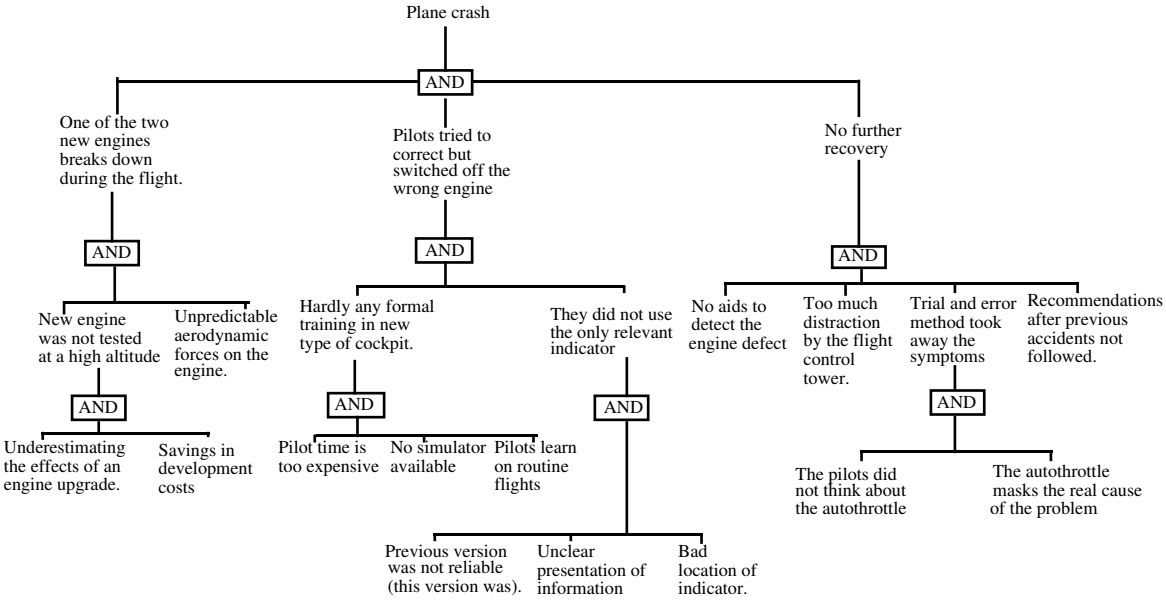
J. Reason (1990). *Human Error*. Cambridge University Press, Cambridge, United Kingdom.

E. Rich (1983). Users are individuals: Individualising user models. *International Journal Of Man-Machine Studies*, 18(3):199–214.

G.H. von Wright (1951). *An Essay In Modal Logic*. Elsevier, North Holland, Netherlands.

C.D. Wickens (1984). *Engineering Psychology And Human Performance*. C.E. Merrill Publishing Company, London, United Kingdom.


Latent errors &
ignored guidelines
Accident
Forensics
Accident
report
Design and
Implementation
Investigation


Plane crash
AND
One of the two new engines breaks down during the flight.
Pilots tried to correct but switched off the wrong engine
No further recovery
AND
AND
AND
New engine was not tested at a high altitude
Unpredictable aerodynamic forces on the engine.
Hardly any formal training in new type of cockpit.
They did not use the only relevant indicator
No aids to detect the engine defect
Too much distraction by the flight control tower.
Trial and error method took away the symptoms
Recommendations after previous accidents not followed.
AND
AND
AND
AND
Underestimating the effects of an engine upgrade.
Savings in development costs
Pilot time is too expensive
No simulator available
Pilots learn on routine flights
The pilots did not think about the autothrottle
The autothrottle masks the real cause of the problem
Previous version was not reliable (this version was).
Unclear presentation of information
Bad location of indicator.

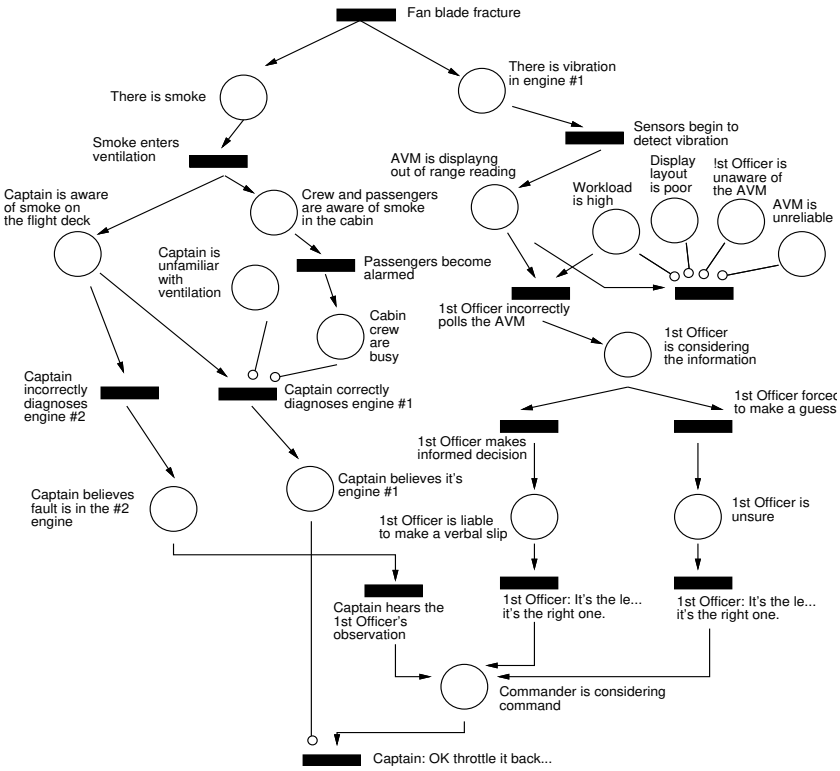

Fan blade fracture
There is smoke
There is vibration in engine #1
Smoke enters ventilation
Sensors begin to detect vibration
AVM is displayng out of range reading
Display layout is poor
!st Officer is unaware of the AVM
Workload is high
AVM is unreliable
Captain is aware of smoke on the flight deck
Crew and passengers are aware of smoke in the cabin
Captain is unfamiliar with ventilation
Passengers become alarmed
Cabin crew are busy
1st Officer incorrectly polls the AVM
1st Officer is considering the information
Captain incorrectly diagnoses engine #2
Captain correctly diagnoses engine #1
1st Officer forced to make a guess
1st Officer makes informed decision
Captain believes it's engine #1
Captain believes fault is in the #2 engine
1st Officer is liable to make a verbal slip
1st Officer is unsure
Captain hears the 1st Officer's observation
1st Officer: It's the le... it's the right one.
1st Officer: It's the le... it's the right one.
Commander is considering command
Captain: OK throttle it back...

Figure 1: The Accident-Reporting Cycle.

Figure 2: Fault-tree Representing the Kegworth Accident

Figure 3: Petri Net Description of the Kegworth Accident