The Strengths and Weaknesses of Logic Formalisms to Support Mishap Analysis

C.W. Johnson; Dept of Computing Science, Univ. of Glasgow, Scotland, G12 9QQ.

C.M. Holloway NASA Langley Research Center, Hampton, VA 23681-2199, USA

Keywords: Causal Analysis, Accident Investigation, Accident Analysis.

Abstract

The increasing complexity of many safety critical systems poses new problems for mishap analysis. Techniques developed in the sixties and seventies cannot easily scale-up to analyze incidents involving tightly integrated software and hardware components. Similarly, the realization that many failures have systemic causes has widened the scope of many mishap investigations. Organizations, including NASA and the NTSB, have responded by starting research and training initiatives to ensure that their personnel are well equipped to meet these challenges. One strand of research has identified a range of mathematically based techniques that can be used to reason about the causes of complex, adverse events. The proponents of these techniques have argued that they can be used to formally prove that certain events created the necessary and sufficient causes for a mishap to occur. Mathematical proofs can reduce the bias that is often perceived to effect the interpretation of adverse events. Others have opposed the introduction of these techniques by identifying social and political aspects to incident investigation that cannot easily be reconciled with a logic-based approach. Traditional theorem proving mechanisms cannot accurately capture the wealth of inductive, deductive and statistical forms of inference that investigators routinely use in their analysis of adverse events. This paper summarizes some of the benefits that logics provide, describes their weaknesses, and proposes a number of directions for future research.

Introduction

The last decade has seen the development of a new generation of formal, mathematically based analysis techniques that can be applied to support mishap investigation. These approaches provide grammars (syntax) with well-defined meanings (semantics) so that investigators can interpret models of adverse events without the potential ambiguity that often affects natural language reports. Formal notations also typically provide proof procedures that determine what can and what cannot be inferred from an incident. These procedures can also be used to check the consistency of any analysis prior to the publication of a mishap report. There are also more speculative benefits that might be obtained from the development of formal approaches to mishap investigation. For instance, mathematical notations help to construct abstract representations of the events leading to an adverse event. The same abstract representations that are amenable to deductive reasoning tools might also be used inductively to identify common patterns of failure amongst large-scale collections of mishap models (ref. 1). The potential benefits of formal techniques must be balanced by a number of concerns about the use of 'mishap logics'. In particular, it is important that investigators should not have to make the mishap 'fit the notation'. Further concerns focus on the ability to accurately communicate the results of an investigation to non-mathematicians. The remainder of this paper, therefore, provides a survey of the different mishap logics that might support the analysis of adverse events and near misses.

We illustrate the application of these logics using as a case study the loss of the Air France Concorde crash, flight AFR4590. The Bureau d'Enquêtes et d'Analyses pour la Sécurité de l'Aviation Civile (BEA) enquiry into this accident found that the front right tire of the left landing gear ran over a strip of metal shortly before rotation during takeoff from Charles de Gaulle Airport (ref. 2). The strip had fallen from another aircraft. Damage to the tire created debris that was thrown against the wing. The debris ruptured a fuel tank and a major fire broke out under the left wing. Problems appeared on engine 2 and for a brief period on engine 1 but the aircraft took off. The crew shut down engine 2, following an engine fire alarm. They noticed that the landing gear would not retract. The aircraft flew for around a minute but was unable to gain altitude beyond 200 feet or speed beyond 200 knots. Engine 1 lost thrust, the aircraft's angle of attack and bank increased sharply. The thrust on engines 3 and 4 fell suddenly and the aircraft crashed onto a hotel. We have chosen to focus on the loss of flight AFR4590 because it typifies the complex combinations of events that characterise many high-profile accidents. This case study also typifies a growing class of safety-critical, legacy systems that were once considered to embody 'cutting edge' technology.

Classical Logic

Classical logic is composed of propositions and sentences. Propositions represent facts that we know about the domain of discourse, such as the mishap under investigation. Simple propositions represent observations about individual objects. For instance, they might capture the fact that the 'the metallic strip is part of engine 3 on a DC10'. More complex sentences can be formed from the use of connectives. In classical logic, these include 'NOT', 'AND', 'OR', 'IF'. These connectives can be used to analyze sentences such as the following: 'IF all four engines had been operating THEN the serious damage caused by the intensity of the fire to the structure of the wing would have led to the loss of the aircraft'. The meaning of these sentences can be interpreted by examining the truth tables that are associated with each of the logical connectives. Table 1 illustrates this approach. The previous sentence would be valid from the first line of Table 1 if it could be shown that all four engines were operating (X=True) and the intensity of the fire led to the loss of the aircraft (Y=True). However, the argument would not be valid from the second line of Table 1 if it could be shown that the engines were operating (X=True) and the intensity of the fire did not lead to the loss of the aircraft (Y=False).

Table 1: Truth Table for Material Implication in Classical Logic

X	Y	**IF** X **THEN** Y
True	True	True
True	False	False
False	True	True
False	False	True

The previous example provides a simple illustration of truth tables being used to establish the validity of arguments in an accident report. However, a number of problems complicate this approach. In particular, the simple 'IF…THEN' connective of classical logic (called material implication) cannot convey the many different forms of causal reasoning that are used in accident reports. For example, mishap investigators often distinguish between *necessary* and *sufficient* causes. A *necessary cause* is often identified using arguments of the form 'the mishap would not have occurred if this cause(s) had not also occurred'. A *sufficient cause* can be distinguished by arguments of the form 'the mishap could have occurred if this cause(s) had taken place irrespective of any other of the other circumstances surrounding the incident'. Figure 1 illustrates this distinction. We can see that cause C2 is necessary but insufficient to cause the mishap. In contrast, if we have both C1 and C2 then we have sufficient causes for the mishap to occur. However, this combination of causes is not necessary for the incident to occur because there is another combination of potential causal factors. C2 and C3 are also together sufficient to cause the mishap. They are unnecessary because C1 and C2 represent an alternative causal path. As we have seen, the BEA report into flight AFR4590 argued that "if all four engines had been operating, the serious damage caused by the intensity of the fire to the structure of the wing and to some of the flight controls would have led to the rapid loss of the aircraft". Hence the loss of power was not a necessary cause of the accident. The fire was sufficient to cause the loss of the aircraft without this additional problem. In Figure 1, the fire is represented by C4, a sufficient cause for the accident. The loss of engine power can, arguably, be represented by C3. It is insufficient as a cause without some additional failures not mentioned in the previous citation.

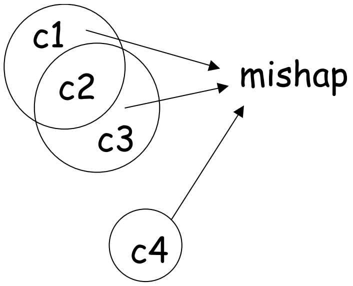

Figure 1: Necessary and Sufficient Causes.

A number of further problems complicate the use of material conditions in classical logic to analyse causal arguments in accident reports. In particular, it is possible to create valid statements when the antecedent and consequent are both false. 'If the grass is blue then the sky is yellow'. This follows from the final line of the truth table for the conditional in table 1. Grice (ref. 3) and Jackson (ref. 4) have addressed this concern and argue that material implication remains a valid form of argument for *indicative* conditionals. Speakers do not say 'If *P*, then *Q*' when they know that *P* is false. It is simpler and more informative to say 'not *P*'. In the previous example, it would be better to assert that 'grass is not blue' rather than construct a more complex argument of the form 'If the grass is blue then the sky is yellow'. Such linguistic arguments might seem to be remote from the practical concerns of accident investigation. However, many reports make use of these rhetorical devices. For example, the BEA report argues, "even if instantaneous ignition (of kerosene) were postulated rapid propagation would require appreciable localized forward airflows". There were no appreciable forward airflows and it is acknowledged that the flame propagation speed of a kerosene fire is not instantaneous, seldom exceeding 6 m/s.

Addressing Limitations of Material Implication: C.I. Lewis and Strict Implication

C.I. Lewis (ref. 5) goes beyond the material implication of classical logic to develop the notion of strict implication. This is based on the idea that a proposition *strictly implies* all others, which are true, in all possible circumstances where it is true. The semantics for this form of strict implication is based around that of modal logics. Each of the ovals in Figure 2 represents a 'possible world' of information. If A strictly implies B then it is impossible for us to reach a world in which A holds but B does not. It is, however, possible for B to hold without A. This is important for mishap investigation because, as we have seen, if A is not a necessary cause of B then there may be alternative sufficient ways in which B might occur. Lewis' strict implication provides a means of avoiding many of the paradoxes that undermine the use of classical logic as a means of reasoning about adverse events and near misses (ref. 1).

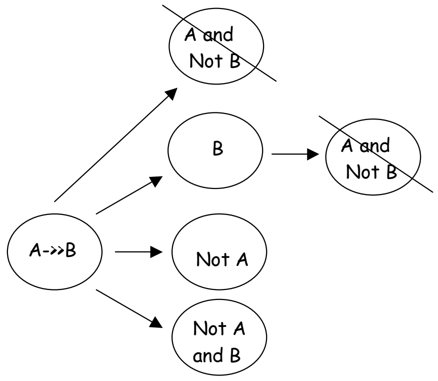

Figure 2: Possible World Semantics for Strict Implication

An investigator might use the Lewis formulation to model the BEA assertion that “If one of the ‘door open’ sensors is destroyed, the information transmitted is ‘the door is not open’ and the gear retraction sequence cannot begin”. The Lewis semantics assert that it must never be the case that one of the door sensors is destroyed without the corresponding transmission to block the retraction sequence. Not only are there considerable practical problems in ensuring that any complex, safety-critical system satisfies this form of strict implication, there are also problems in interpreting the ideas proposed by Lewis. In particular, it is important to be clear about the meaning of the arrows in Figure 2. How do we move between these different worlds of knowledge? One approach would be to associate these transitions with the introduction or revision of information. In our example, no new information should create a situation in which a sensor is destroyed and the information is not transmitted. The additional commitments of strict implication would suggest that the information should be transmitted even if there was damage to associated avionics. As mentioned, this can be difficult to guarantee. Any formal analysis might, therefore, introduce additional caveats into the antecedent of a strict implication to express the degree of damage that might be sustained before the information could not be transmitted. Of course, this ‘limitation’ of the formal technique forces investigators to be precise in the characterisation of those properties that play an important role in a mishap

Further paradoxes affect the Lewis semantics for strict implication. For example, a true consequent allows the introduction of an arbitrary antecedent, such as p in the following p ->> (q v $\neg q$). The engineering objectives of our study might persuade us to overlook these apparent deficiencies; they would arguably have few practical effects on the application of mishap logic. However, philosophers and logicians have used these deficiencies to justify the development of alternative means of representing and reasoning about causal arguments and conditional statements.

D. Lewis, Why-Because Analysis and Counterfactual Reasoning

Even if we overlook the paradoxes of strict implication, a number of further problems prevent the application of this approach to support reasoning about accidents and incidents. In particular, the formalisms discussed thus far can not capture many of the causal arguments that are put forward in accident reports. For example, the BEA investigation argued that ‘if a partial hydraulic failure…had then occurred, only the landing gear located on the side of the rupture would have been affected’. This represents a particular form of subjunctive argument known as a counterfactual. It relies upon an antecedent, which represents a past tense subjunctive sentence of the form "If X *had been the case* …then Y would have happened. These sentences are known as counterfactuals because there is an assumption that the antecedent is false. In other words, X is known not to have been the case. For example, an investigator might argue: If the metal wear strip from the thrust reverser door on the DC-10 had been maintained correctly then it would not have damaged the tires of flight AFR4590. This is counterfactual because the BEA present evidence that the strip was poorly maintained and that it did damage the tires. This example also illustrates the manner in which many counterfactual arguments about the causes of a mishap also embody a false consequent of the general form ‘then the failure would not have occurred’. There are numerous dangers associated with this form of argument. For example, investigators might suggest that ‘if the operators had been more vigilant then this accident would not have happened’. Readers often infer from such statements that the operators were not vigilant even though we have presented no evidence to support this assertion.

The biases that can affect informal counterfactual argument helped to motivate David Lewis’ (ref. 6) development of logics for counter-factual arguments about causation. These formalisms rely on a modal semantics, which is broadly similar to that used in C. Lewis’ work on strict implication. Both depend on accessibility relationships between possible worlds of knowledge. Strict implication ensures that certain properties hold in all possible worlds that are accessible from the world in which an implication is introduced. In contrast, D. Lewis’ logics can be used to state that A is a causal factor of B, if and only if A and B both occurred and in the nearest possible worlds in which A did not happen neither did B. This implies that A is not only a sufficient but also a necessary cause of B. It precludes the observation that other causal factors may have led to B in any of the nearest possible worlds. This does not rule out the existence of alternative causes. It does, however, imply that those causes may only arise in worlds that are remote from the present one that is under consideration. For example, we might use Lewis’ logic to argue that the accident would have been avoided if the DC-10 wear strip had not fallen onto the runway. We can envisage other potential causes of the tire damage but these causes would not be so ‘close’ to the particular incident being modelled in the aftermath of flight AFR 4590.

Lewis’ work on counterfactual arguments is particularly important because it lies at the heart of Ladkin and Loer’s (ref. 7) Why-Because Analysis. This is one of the most influential of the recent generation of formal mishap logics. It has been applied to reason about the causes of a wide range of aviation accidents (ref. 1). WBA begins by a reconstruction phase where a semi-formal graphical notation models the sequences of events leading to a mishap. These sequences can be represented in a form of temporal logic and then iteratively analysed to move towards a causal explanation using counterfactual arguments. Ladkin and Loer introduce the =>> operator which can informally be read as ‘causes’ and the []-> operator to represent a counterfactual relationship. Informally, A []-> B captures the notion that B is true in possible worlds that are close to those in which A is true. The following inference rule can be constructed to relate these connectives. Ladkin and Loer also provide a range of additional proof rules that can be used to ensure both the consistency and sufficiency of arguments about the causes of a mishap:

$$\begin{array}{c} A \ \wedge B \\ \underline{\neg A\ [\,]\text{->}\ \neg B} \\ A\ \text{=>>}\ B \end{array} \qquad (1)$$

The counterfactual approach to causation does not provide a panacea for analyzing conditional arguments. As we have seen, material implication is truth functional. We can determine whether ‘if A then B’ is true by first determining the validity of the antecedent, A, and consequent, B and then by looking at a relevant entry in the table 1. In contrast, it is not possible to use a truth functional style of analysis with counterfactual arguments. By definition, the antecedent of the counterfactual is assumed to be false and so every counterfactual is assumed to be true constrained only by the concepts of nearness or proximity to some agreed notion of the present world. For example, it is valid to argue that the accident would have been avoided if the crew had seen the foreign object from the cockpit. This argument is clearly remote from the ‘real world’ given the size of the strip and the speed of the aircraft. Even Lewis is forced to rely on an appeal to expert judgement in order to identify the bounds of proximity. Such judgements are often contradictory and subject to change over time (ref. 1).

Bayesian Logic

Bayesian logic is build around the conditional probability of a proposition given particular evidence. The value of a conditional probability is, typically, represented by a real number, between zero and one. We can use p(h|e) to represent the probability of some proposition or hypothesis, h, given some evidence, e. Most applications of Bayesian reasoning embody a form of implication or conditional statement in which the observation of some evidence strengthens, or alternatively weakens, the support for particular hypotheses. In other words, if e is observed then this increases the credibility of h. This is significant for accident investigation because we can talk about the manner in which evidence will support our findings about the causes of an adverse event. We can also use Bayesian techniques to support abductive reasoning about the causes of an incident given that we know an accident has occurred. The following formula considers the probability of a given hypotheses, B, in relation to a number of alternative hypotheses, B_i where B and B_i are mutually exclusive and exhaustive:

$$\mathrm{Pr}(B \mid A \wedge C) = \frac{\mathrm{Pr}(A|B \wedge C).\mathrm{Pr}(B|C)}{\mathrm{Pr}(A|B \wedge C).\mathrm{Pr}(B|C) + \Sigma i \ \mathrm{Pr}A|B\_i \wedge C).\mathrm{Pr}(B\_i \,|C)} \qquad (2)$$

This formula can be used to assess the likelihood of a cause B given that a potential effect, A, has been observed. This provides a means of using information about previous incidents to guide the causal analysis of future occurrences (ref. 1). In our case study, investigators might be interested to determine the likelihood that a foreign object on the runway had damaged the front right tire of AFR4590. The analysis begins by assessing the likelihood of runway debris. We might either choose to use subjective estimates or frequencies derived from the analysis of previous incidents, assuming that such data are available and reliable. For demonstration purposes, we assume that the likelihood of finding a foreign object on a particular runway is 0.98; Pr(foreign_object | C) = 0.98. This is a high subjective assessment. It is justified by the BEA observation that the part had fallen from an aircraft five minutes before the Concorde attempted to take off. Without an automated system, they argued it is impossible to guarantee the detection of such objects. As we have seen, there may also be other sufficient causes of the tire damage. For example, the BEA investigation found that the central spacer was absent from the left main landing gear. This could have led to an asymmetrical trajectory and tyre overheating; Pr(spacer_missing | C) = 0.001. They also considered the possibility of tire related defects. Even though a previous incident had led to the strengthening of Qualification Test Program requirements for the tires resistance to twice the normal load compared with 1.5 on other aircraft; Pr(tire_defect | C) = 0.001. The next stage is to determine how likely it is that these potential causes would lead to the deflation of a tire. Further analysis might reveal that 93 incidents involved runway debris out of every 10,000 tire deflations. Recall that Bayesian techniques can be applied both to incident statistics and to subjective estimates. The actual values used here are purely intended to illustrate the application of the approach: Pr(tire destruction | foreign_object ∧ C) = 0.0093, Pr(tire destruction | tire_defect ∧ C) = 0.95, Pr(tire destruction | spacer_missing ∧ C) = 0.0407. We can now integrate these observations into the previous formula to calculate the probability that a foreign object was present on the runway given that a tire failure has been reported. The following calculation suggests that there is about a 90 per cent chance that a failure involved such an object:

$$\mathrm{Pr}(\text{foreign\_object} \mid \text{tire\_burst} \wedge \text{C})\}$$

$$= \frac{\{\mathrm{Pr}(\text{tire\_burst} \mid \text{foreign\_object} \wedge \text{C}).\mathrm{Pr}(\text{foreign\_object} \mid \text{C})\}}{((\mathrm{Pr}(\text{tire\_burst} \mid \text{foreign\_object} \wedge \text{C}).\mathrm{Pr}(\text{foreign\_object} \mid \text{C})) + (\mathrm{Pr}(\text{tire\_burst} \mid \text{tire\_defect} \wedge \text{C}).\mathrm{Pr}(\text{tire\_defect} \mid \text{C})) + (\mathrm{Pr}(\text{tire\_burst} \mid \text{spacer\_missing} \wedge \text{C}).\mathrm{Pr}(\text{spacer\_missing} \mid \text{C}))}$$

$$= \frac{(0.0093).(0.98) \qquad .}{(0.0093).(0.98) + (0.95).(0.001) + (0.0407).(0.001)}$$

$$= 0.901 \tag{3}$$

A number of caveats can be raised against this application of Bayes' theorem. It is difficult to have any confidence in prior probabilities. For instance, estimates of the likelihood of an illness within the general population can be validated by extensive epidemiological studies. Unfortunately, it is difficult to extend these techniques to support numerical assessments about the likely causes of technological failure. The relatively slow growth of the FAA’s Global Aviation Information Network illustrates the difficulty of encouraging commercial and regulatory organizations to exchange incident data. Further technical difficulties complicate the validation of numerical estimates for the likelihood of human ‘error’ and software failure. The provision of subjective probabilities is also subject to systematic biases (ref. 8).

Comparative Probabilities and Partition Models

A further class of techniques enables analysts to talk about the likelihood of particular events without referring to precise, subjective, or quantitative values. Many of these approaches are built around the observation that *a* may cause *b* in a context C if there is a high probability that *b* is true given that *a* is also true in C. In other words, we might require that:

$$\mathrm{Pr}(\text{b}|\text{a} \ \wedge \text{C}) > \text{Very\_likely}. \tag{4}$$

Such observations founder when we attempt to explain what is meant by ‘very likely’. This may again be seen to introduce the subjective, numeric estimates that have been criticised as a weakness of other techniques. In consequence, a number of authors have presented refinements on this initial model (ref. 1). We might require that *a* is causally related to *b* in context C if the probability of A and B in C is not same as the probability of B in C and the probability of A in C. The following formulae adopt the convention of using upper case to denote token types, or general classes of observations; lower case is used to indicate particular instances of these more general events:

$$\mathrm{Pr}(\text{B} \ \wedge \text{A} \mid \text{C}) <> \mathrm{Pr}(\text{B}|\text{C}).\mathrm{Pr}(\text{A}|\text{C}) \tag{5}$$

It can be argued that A is a potential cause of B if an occurrence of A makes B more likely. Conversely, A can be a barrier to B. An occurrence of A, therefore, makes B less likely. We can apply this approach to elements of the BEA investigation. For example, the probability that a foreign object was on the runway and that a tire burst has occurred is greater than the independent probabilities that there is a foreign object multiplied by the probability that a tire burst had occurred:

$$\text{Pr (foreign\_object} \wedge \text{tire\_burst} \mid \text{C)} > \text{Pr (foreign\_object} \mid \text{C).Pr(tire\_burst} \mid \text{C)} \qquad (6)$$

Such formulae form part of a wider research initiative to gradually refine probabilistic models so that they more closely model informal causal concepts. For example, more recent approaches require that the probability of B and A in C is greater than probability of B given that we know that A and C are not true. Informally, knowing A increases the probability of B above a similar situation in which we know ¬ A. Alternatively, we can argue that *a* causes *b* if the probability of B and A in C is greater than probability of B given only C. This deals with a situation in which we do not know about A. In other words, we assume that C in Pr(B|C) contains no information about A:

$$\text{Pr(B} \wedge \text{A} \mid \text{C)} > \text{Pr(B}\mid \neg\ \text{A} \wedge \text{C) v Pr(B|A} \wedge \text{C)} > \text{Pr(B|C)} \qquad (7)$$

The presence of a foreign object on the runway leads to a tire burst if the probability that such an object was present and that a mishap occurred is greater than that associated with mishaps in which tire burst occurred without foreign objects or situations in which nothing is known about the presence of foreign objects:

$$\begin{aligned}&\text{Pr(tire\_burst} \wedge \text{foreign\_object} \mid \text{C)} > \text{Pr(tire\_burst} \mid \neg\ \text{foreign\_object} \wedge \text{C) v}\\ &\qquad\qquad \text{Pr(tire\_burst} \mid \text{foreign\_object} \wedge \text{C)} > \text{Pr(tire\_burst} \mid \text{C)} \qquad (8)\end{aligned}$$

However, *a* and *b* might both be effects of some other common cause. In order to rule out such a situation, investigators must look back in an incident reconstruction to explicitly preclude other causal factors. This raises further complex issues because some of these preceding factors can both promote and confound particular effects. A factor that contributes to the causes of *a* may also have an independent but negative influence on the occurrence of *b*. Partition models provide one approach to the complex relationships that can exist between different causal factors. These models are constructed from a partition, S_j, of all the relevant factors excluding A and C that might contribute to or prevent a mishap. Factors represent negative or positive causal factors, c_1,..., c_m, that must be held fixed to observe the causal effect of *a*. In other words, in order to demonstrate that *a* causes *b*, we have to show that this effect was not caused by another other factor or combination of factors. More formally, any element, *d*, of a subset in S_j is in c_i if and only if it is a cause of *b* or ¬ *b*, other than *a*, and it is not caused by *a*. It can, therefore, be argued that *a*'s cause *b*'s in circumstances C if and only if:

$$\forall \text{ j: Pr(B|A} \wedge \text{S\_j} \wedge \text{C)} > \text{Pr(B|S} \wedge \text{C)} \qquad (9)$$

Each of the factors in c_1,..., c_m must be represented in each subset. This results in 2^m possible combinations of present or absent factors. However, some combinations of causal factors are impossible and can be excluded. Other combinations result in *b* being assigned a probability of 1 or 0 regardless of *a* and can be excluded. All the remaining combinations of causal factors must be considered. In other words, *a*'s must cause *b*'s in every situation described by S_j. Again, this approach can be most easily explained using an example from the BEA report. Recall that a factor is in c_1,...,c_m if and only if it is a cause of *b* or ¬ *b*, other than *a*, and it is not caused by *a*. For example, the following factors might be considered relevant to the tire burst: c_1 might represent abnormal use of brakes, c_2 a tire defect, c_3 might represent a missing central spacer, c_4 might represent correct inflation of the tire and so on. It then remains to be shown that a foreign object on the runway would result in the mishap, under all of the combinations of other factors represented in the partition.

As with all of the techniques assessed in the paper, caveats can be raised about the utility of any causal, mishap analysis that might be performed using such partition models. A particular problem here is the requirement that the partitions, S_j, should consider all of the possible factors that might contribute to, or prevent, the effect that is being studied. In complex mishaps, it can be difficult to identify all of these potential factors. For instance, an initial analysis of the BEA report has identified more than thirty such factors that might be considered relevant to the tire burst. This is a conservative estimate. Clearly, the extent of any partition must be affected by the stopping rule that helps to determine the bounds of any incident investigation. It might, therefore, be argued that this apparent limitation of partition models is no different from the requirement to scope the bounds of a mishap analysis and that this requirement applies to all investigation techniques.

Conclusions

The increasing costs and complexity of accident investigation are imposing new demands on previous generations of analysis techniques (ref. 1). This has led to the development of mishap logics that can be used to represent and reason about the causes of adverse events. We have taken pains to make the different approaches accessible to those with a greater interest in the application of mishap logics than in their philosophical underpinnings. The previous pages have, therefore, summarized the strengths and weaknesses of these techniques. Pearl (ref. 9) argues that logic formalisms cannot be used to prove causes in the same way that one might prove propositions or theorems. Causal expressions in natural language often allow for numerous exceptions that create problems when attempts are made to codify these expressions in the deterministic forms of classical logic. Stochastic techniques create further problems because analysts must validate numeric assessments of likelihood. Partition models avoid some of these problems but they again create the need to demonstrate, or prove, that a potential cause will result in an incident under the various circumstances within any particular partition.

An important aim of this paper has been to convey the complexity of causal analysis. This complexity arises partly from the difficulty of capturing the many informal concepts that relate to causation. Increasingly, however, it also stems from the complex interactions that characterise the engineering of safety-critical systems. For instance, the BEA report could not accurately identify the indirect forces that would have been necessary for a projectile, such as a piece of the tire, to damage the tank in the manner that was observed. They also remark that 'the ignition of the kerosene leak, the possible forward propagation of the flame, its retention and stabilisation occurred through complex, phenomena, which are still not fully understood' (ref. 2). In other words, we do not fully understand the precise causal mechanisms that led to the loss of AFR4590. This situation was compounded by disagreement between the French judicial investigation, the BEA enquiry and the UK Air Accident Investigation Branch analysis. For example, the AAIB strongly favoured arcing from damaged wheel-brake fan power supply cables in the left main landing gear bay as the most probable source of ignition for the fire. In contrast, the BEA also considered it possible that ignition occurred by the forward propagation of flames that were ignited by reheat surfaces. The AAIB discounted this cause and the BEA responded by arguing 'aviation safety can only gain through taking into account the various causes considered as possible by the experts'. Several of the logics in this paper can be extended to represent these different viewpoints on the likely causes of an adverse event. For example, the *C* context parameter can be used in partition models. Alternatively epistemic logics distinguish between the different forms of knowledge and belief that are available to individual agents, such as investigators from different regulatory and investigatory organisations. It remains to be seen whether such techniques have a useful role to play in helping us to understand the increasingly complex interactions among the causes of failure in safety-critical systems.

References

1. C.W. Johnson, The Failure of Safety-Critical Systems: A Handbook of Accident and Incident Reporting, Springer Verlag, London, in press and to appear 2003.

2. Bureau d'Enquêtes et d'Analyses pour la Sécurité de l'Aviation Civile (BEA), Accident on 25 July 2000 at La Patte d.Oie in Gonesse (95) to the Concorde registered F-BTSC operated by Air France, Report f-sc000725a. http://www.bea-fr.org/anglaise/actualite/concorde-en.htm.

3. H.P. Grice, Studies in the Way of Words. Harvard University Press, Cambridge, 1989.

4. F. Jackson, On Assertion and Indicative Conditionals. Philosophical Review (88):565-589, 1979.

5. C.I. Lewis and C.H. Langford (1932), Symbolic Logic, The Century Co. New York, 1932.

6. D. Lewis, Counterfactuals, Oxford University Press, Oxford, UK, 1973.

7. P. Ladkin and K. Loer, Why-Because Analysis: Formal Reasoning About Incidents, Bielefeld, Germany, RVS-Bk-98-01, Technischen Fakultat der Universitat Bielefeld, Germany, 1998.

8. C. Puppe, Distorted Probabilities And Choice Under Risk, Springer Verlag, Lecture Notes In Economics And Mathematical Systems, No 363, Berlin, Germany, 1991.

9. J. Pearl, Causality; Models, Reasoning, and Inference, Cambridge University Press, UK, 2000

Biographies

Prof. C.W. Johnson, MA, MSc, DPhil, CEng, Glasgow Accident Analysis Group, Dept. of Computing Science, Univ. of Glasgow, G12 9QQ, Scotland, tel. +44 141 330 6053, e-mail – johnson@dcs.gla.ac.uk, http://www.dcs.gla.ac.uk/~johnson. Prof. Chris Johnson heads a research team specifically devoted to new generations of accident analysis techniques. He helped to author European guidelines for mishap reporting in Air Traffic Management. He has developed an incident analysis scheme for the UK Health and Safety Executive to support the investigation of adverse events involving programmable systems across the process industries.

C.M. Holloway, NASA Langley Research Center, MS 130 / 100 NASA Road, Hampton, VA 23681-2199, USA. c.m.holloway@larc.nasa.gov, http://shemesh.larc.nasa.gov/people/cmh/. C. Michael Holloway is a senior research engineer at the NASA Langley Research Center. He is interested in accident analysis, software system safety and high-integrity software development. Mr. Holloway has a B.S. in computer science from the Univ. of Virginia, and completed all-but-dissertation towards a Ph.D. from the Univ. of Illinois. He is a member of the IEEE, the IEEE Computer Society, and the System Safety Society. He is chair of the 2003 workshop on the Investigation and Reporting of Incidents and Accidents (IRIA 2003).