A Survey of Logic Formalisms to Support Mishap Analysis

Chris Johnson (a), C.M. Holloway (b)

(a) Department of Computing Science, University of Glasgow, Scotland.
johnson@dcs.gla.ac.uk

(b) NASA Langley Research Center, MS 130 / 100 NASA Road, Hampton, VA 23681-2199, USA
c.m.holloway@larc.nasa.gov

Abstract: Mishap investigations provide important information about adverse events and near miss incidents. They are intended to help avoid any recurrence of previous failures. Over time, they can also yield statistical information about incident frequencies that helps to detect patterns of failure and can validate risk assessments. However, the increasing complexity of many safety critical systems is posing new challenges for mishap analysis. Similarly, the recognition that many failures have complex, systemic causes has helped to widen the scope of many mishap investigations. These two factors have combined to pose new challenges for the analysis of adverse events. A new generation of formal and semi-formal techniques have been proposed to help investigators address these problems. We introduce the term 'mishap logics' to collectively describe these notation that might be applied to support the analysis of mishaps. The proponents of these notations have argued that they can be used to formally prove that certain events created the necessary and sufficient causes for a mishap to occur. These proofs can be used to reduce the bias that is often perceived to effect the interpretation of adverse events. Others have argued that one cannot use logic formalisms to prove causes in the same way that one might prove propositions or theorems. Such mechanisms cannot accurately capture the wealth of inductive, deductive and statistical forms of inference that investigators must use in their analysis of adverse events. This paper provides an overview of these mishap logics. It also identifies several additional classes of logic that might also be used to support mishap analysis.

Keywords: Mishap logics, Causation, Implication, Accident Investigation, Mishap Investigation.

Introduction

Mishap reports yield information about the hazards that threaten safety-critical applications. They, therefore, provide important means of validating risk assessments and safety cases. As a result, a growing number of international standards require that mishap-reporting systems be integrated into safety management schemes. For example, IEC 61508 is widely used as a standard for the development of safety-critical applications that incorporate computer systems. This includes the recommendation that manufacturers should: "…implement procedures which ensure that hazardous Mishaps (or Mishaps with potential to create hazards) are analysed, and that recommendations are made to minimise the probability of a repeat occurrence." (IEC, paragraph 6.2.1). There are also industry specific standards, such as those governing the US and European development of medical devices (cite), that require reporting systems be used to monitor adverse events.

NASA provides an example of an organisation that has responded to these international and national initiatives by developing sophisticated support for mishap reporting. Each centre operates a safety related reporting system. These local systems ensure that safety concerns are addressed at a local level where they are first raised. This can ensure a prompt response from those individuals who are best placed to understand the context in which a mishap occurs. Figure 1 provides screen shots from one of the Johnson Space Centre's (http://www.jsc.nasa.gov/Safety/Alert) web-based information systems. This provides staff with feedback on the outcome of local mishap investigations. The format and content of this information is distinct from the web-based mishap reporting system exploited by, for example, the Langley Research Centre (http://safety.larc.nasa.gov).

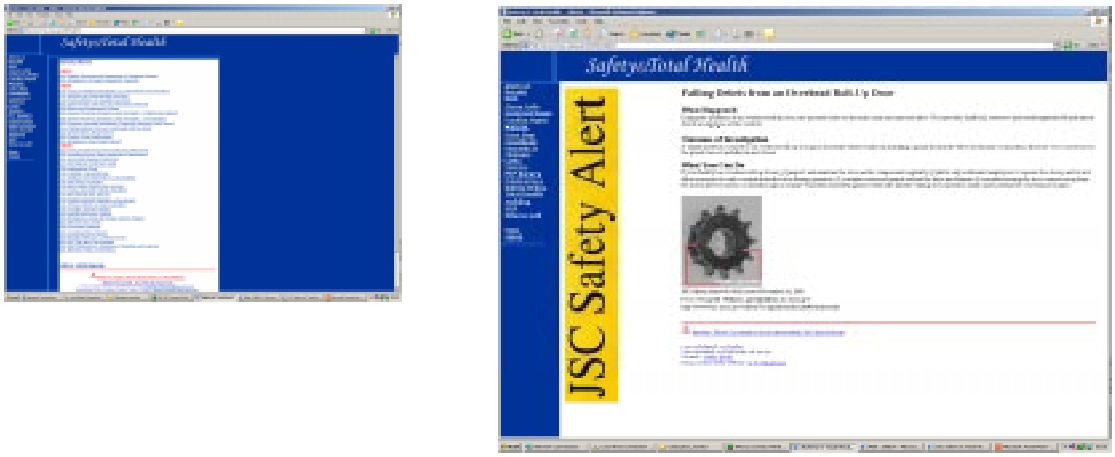

Figure 1: Examples of NASA Mishap Reports

Local mishap reporting systems, such as that illustrated in Figure 1, are supported by more centralised systems. For instance, the NASA Safety Reporting System provides staff with the means of reporting concerns that they feel may not have been adequately addressed at a local level (http://www.hq.nasa.gov/office/codeq/narsindx.htm). Similarly, the NASA lessons learned system (http://lls.nasa.gov) provides a centralised mechanism for exchanging both safety and operational information between individual sites.

Overview of the Mishap Investigation and Reporting Process

The amount of resources that are devoted to the investigation and analysis of mishaps is, typically, determined by an initial assessment of the risks associated with a recurrence of an adverse event. It is, however, possible to identify a number of stages that are common to most mishap investigations. For instance, Figure 2 provides an overview of the reporting process recommended in the EUROCONTROL guidelines for reporting occurrences in European Air Traffic Management (Johnson, Le Galo, Blaize, 2000). Mishaps must first be detected and reported, for instance either by members of staff, the public or by automated monitoring systems.

After a mishap has been reported, it is important to gather data about what happened during the incident or near miss. For example, investigators must gather any physical components that may have failed. They must also safeguard the data logs that are compiled by monitoring equipment. This data gathering stage can also involve eyewitness interviews. Once evidence has been gathered, investigators can use the data to reconstruct what happened during a mishap. This reconstruction stage must consider the latent, or long-term, failures that created the conditions in which an incident occurred. It should also consider the more catalytic events that triggered the mishap. Investigators must also consider the mitigating events that might have helped to reduce the consequences of an adverse event or near miss. This analysis is important because it can help to identify the successful barriers that might protect against similar incidents in the future.

After investigators have determined what happened during any mishap, it is then important to establish 'why' it occurred. Reconstruction and causal analysis can be considered as distinct but interdependent phases of a mishap investigation. Reconstructions, typically, consider a mass of contextual details that are important to understanding the course of an adverse event but which are not considered to have an impact on the causes of a mishap. In brief, reconstruction focuses on what happened during an incident while analysis helps to determine why the incident occurred. As can be seen from Figure 2, however, there are a series of feedback looks between this and previous stages of the investigation process. Analysis can reveal new causal hypotheses that may, in turn, identify the need to gather further evidence. For instance, an investigation might initially focus on the possibility of human error. Subsequent investigations might reveal the need to consider organisation and managerial factors in greater detail.

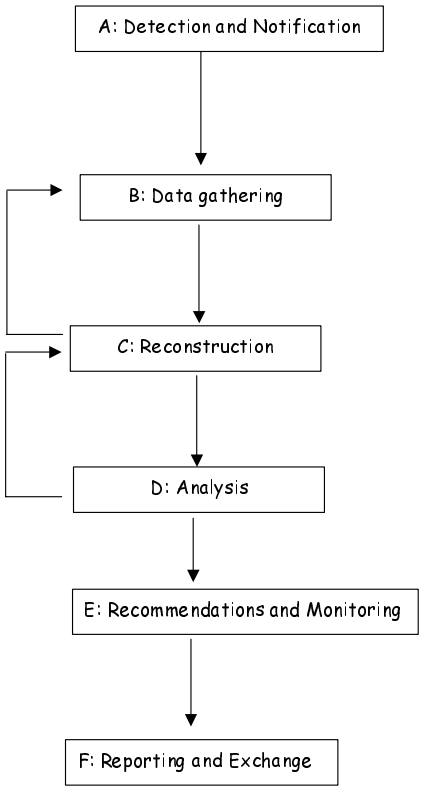

Figure 2: An Overview of the Mishap Reporting Process

Mishap investigation does not end once the causes of an event have been identified. Recommendations must be identified and implemented if the future safety of the system is to be protected. It is particularly important to monitor the success or failure of any changes that are made in the aftermath of a mishap because there is a danger that recommendations may fail to address the underlying the causes of adverse events. For example, previous studies of mishap reporting systems have shown a tendency to recommend low cost remedies, such as additional training, rather than recommend longer-term changes in equipment provision or management practices (Busse and Wright, 2000). Finally, it is important to ensure that the insights gained from a mishap investigation are disseminated both within and between organisations. The NASA lessons learned system, introduced in the previous section, provides an example of how this requirement can be satisfied using web-based technology. This information exchange is important because other groups within an organisation may react to protect themselves against similar mishaps. It can also help to elicit information about similar failures that may not previously have been reported.

It is important to stress that several alternative schemes guide mishap investigations. Some of these reflect the use of particular tools. For instance, many NASA mishap teams use fault tree analysis. Boolean algebra can structure conjunctions and disjunctions of events. The resulting models can then be analysed to identify minimal cutsets. These describe the minimum conjunction of events that are necessary for a mishap occur. There may be several alternate cutsets, each of which describes the possible causes of a mishap. A necessary cause is any event that is common to every cutest. If this form of analysis is used then the investigation process involves sorting and ranking cutsets, developing a case for specific cutsets, communicating the findings, and then taking corrective and preventative actions with follow-up evaluations (Sampino, 2001).

Barriers to the Investigation Process

Safety managers and incident investigators face a number of problems that complicate the task of establishing and maintaining mishap reporting systems. It is often difficult to elicit reports about adverse events and near-miss incidents. Potential contributors often have a concern that any report might result in disciplinary action. There is a reluctance to provide reports especially if they concern the behaviour of colleagues or more senior staff. In consequence, many organisations operate confidential or anonymous systems. Further problems complicate the reconstruction activities, described above. It can be difficult to build up complete time-lines or identify the many different factors involved in an incident. Evidence is often missing and eyewitnesses frequently contradict each other.

Further problems arise even if investigators can piece together what happened before, during and after a mishap. It is important that they resist a number of biases and influences that can affect their interpretation of event. Investigators often find themselves exposed to the many subtle and implicit pressures that affect all complex organisations. For example, hindsight bias occurs when investigators use data that is unavailable to the participants in an adverse event. It is relatively easy to argue that operators should have responded in a particular way with the benefit of hindsight. Political bias occurs when an analyst places greater weight on the analysis of an individual because of their status rather than the judgement itself. It is difficult to avoid the implicit effects of this form of bias given that status is often an accurate indicator of expertise. However, this need not always be the case. Similarly, sponsor bias occurs when an investigator's analysis will affect the reputation of a group or individual that the investigator is responsible for. Supervisors and managers are often required to perform the initial analysis of mishaps that occur within their teams. This can create conflicts of interest if a mishap reveals problems in the management of their employees and sub-contractors. Professional bias occurs when colleagues favour particular outcomes from the causal analysis of a mishap or 'near miss'. In its most explicit form 'whistle blowers' have been expelled from organisations because they have brought their profession into disrepute. It can also affect the analysis of adverse events if investigators continue to identify an individual's failure to observe professional codes rather than the inadequacies of those codes as a means of ensuring professional conduct.

Johnson (2002) provides a more complete analysis of these many different influences that can affect an investigators interpretation of a mishap. The important point is that these influences can, in turn, affect the recommendations that are made in the aftermath of an adverse event or 'near-miss'. For instance, frequency bias occurs when investigators continue to identify the most frequently occurring causes even though they may not have contributed to the mishap currently being investigated. This implies that any recommendations will continue to address a restricted vocabulary of causal factors. It is, therefore, important to monitor both the frequency with which investigators identify particular causal factors and the repertoire of recommendations that they propose in the aftermath of adverse events. This can help to ensure that similar incidents elicit consistent recommendations and that those recommendations are implemented across complex, safety-critical organisations. These tasks are complicated by many different problems. For example, local and individual factors can affect the way in which different groups within an organisation interpret a particular recommendation. In some situations this is necessary to ensure the future safety of an application. In other situations, different interpretations of the same recommendation can satisfy the preconditions for further mishaps (Johnson, 2002).

Potential Solutions: Formal Methods and Mishap Logics

A number of techniques have been developed to help safety managers and mishap investigators address the problems that were identified in the previous section. Many of these techniques stem from the innovative work conducted by the US Department of Energy and the National Transportation Safety Board during the 1970's and 1980's. Techniques such as Multilinear Event Sequencing and Events and Causal Factors Charting provide means of first reconstructing an adverse event and then analysing the resulting graphs to distinguish root causes from contributory factors. Other techniques, such as Management Oversight and Risk Trees support the classification of root causes according to a number of carefully predefined categories such as 'less than adequate risk assessment'. These techniques have been widely applied to support the investigation of incidents in industries ranging from oil production and transportation through to healthcare and pharmaceutical manufacture (Johnson, 2002). It is also important to distinguish between inductive techniques, such as Event Sequence Analysis, and deductive techniques, including the accident Fault Trees mentioned in the previous section. Inductive techniques help analysts to piece

together the ways in which individual events combine during the course of an adverse event or near miss. The available evidence about these events drives the analysis. The success of an inductive approach, therefore, depends upon investigators gathering sufficient evidence to begin the analysis. Conversely, deductive techniques work back from the adverse event or potential outcome. Investigators must then trace back a causal sequence to identify initial events. These approaches can be used in situations where evidence is missing or hard to gather. However, these techniques can lead to biased results if causal hypotheses overly determine the gathering of evidence (Johnson, 2002).

A number of important caveats can be made about this previous generation of semi-formal approaches. In particular, they were intended to support the analysis of the types of incidents that commonly occurred when they were initially developed. These did not commonly involve the types of tightly integrated computer-controlled systems that typify many current safety-critical applications. There are further caveats. For example, many of these techniques do not reflect the current focus on the systemic causes of failure. In consequence, they can be used to identify causal sequences without necessarily helping investigators to identify the mass of different contextual, causal factors that create the preconditions for failure. These criticisms cannot easily be levelled at techniques that emerged or were revised during the 1990's, including Reason's (1997) Tripod and van der Schaaf's (1992) PRISMA method. These approaches explicitly encourage investigators to look beyond immediate system failures and operator 'errors' to consider the latent causes of adverse events and near misses. However, they provide little support for the analysis of software related failures and it is questionable whether they can scale-up to the complex, technological mishaps that can affect an organisation such as NASA.

The last decade has seen the development of a new generation of formal, mathematically based analysis techniques that can be applied to support mishap analysis. Arguably the most notable of these approaches has been Ladkin and Loer's (1998) Why-Because Analysis (WBA). This approach uses a modal logic to provide a clear syntax and semantics for the languages that are used both to reconstruct a mishap and then to identify its causes. It also provides proof techniques that can be used to establish that a causal analysis is well justified by a reconstruction. This is not the only approach. Burns (2000) has exploited deontic logic to identify the ways in which mishaps often stem from the violation of 'normal' working practices. Johnson (2002) has used epistemic logics to model the different perspectives and views that investigators can often hold about the same incident. He has also used temporal logics to represent and reason about the ways in which an incident can change over time. In earlier work, he used the diagrammatic form of the Petri Net notation to support a similar analysis of mishaps that involve complex communications between many different concurrent systems (Johnson, 2002). Kjellen (2000) and others have extended the application of Fault Trees from risk analysis and design to support mishap investigation using the gates that are associated with Boolean algebra.

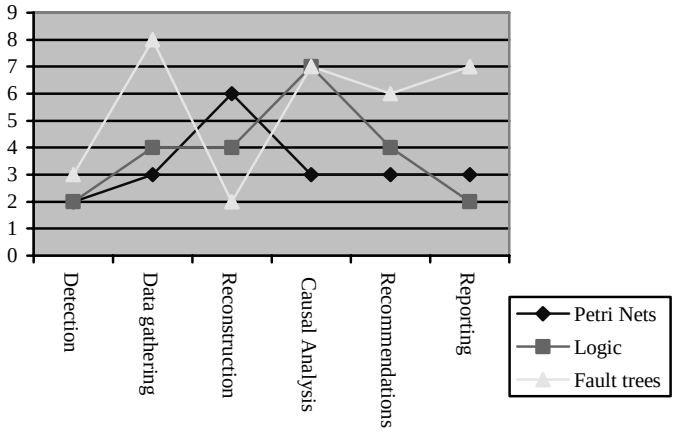

Figure 3: Different Levels of Support for the Mishap Investigation Process

It is important to stress that these formal techniques, typically, focus the support that they offer within particular stages of the investigation process. In related work, we have conducted a survey to elicit opinions about where these approaches might provide the greatest benefits (Johnson and Holloway, 2002). Figure 3 provides an excerpt from the results of this work. As can be seen, logic is argued to provide greatest support for the causal analysis of mishaps but there remain doubts about whether it can or should be used to support data gathering or the dissemination of incident reports. Graphical representations, such as Fault Trees, arguably offer more support for the presentation of any causal analysis. If formal notations are used then they must be integrated into the wider investigation process. Brevity prevents a more detailed discussion of this issue and the interested reader is directed to Johnson (2002).

Formal techniques, such as Why-Because Analysis, offer a number of potential benefits for mishap investigation. They provide a clear syntax and semantics so that it should be possible for other investigators to interpret reconstructions and the products of a causal analysis without the potential ambiguity that often affects the interpretation of natural language reports. This is important when mishap reports can trigger the re-design of safety-critical applications, often involving considerable investment. Formal notations also, typically, provide proof procedures that determine what can and what cannot be inferred from an incident reconstruction or a causal analysis. These proof procedures can be used to reason about the consistency of any analysis. This provides a means of checking the work of an investigator or investigation team prior to the publication of a mishap report. In particular, proof procedures can be used to check for inconsistencies and omissions in formal models of adverse events and near misses. The ability to provide a clear syntax and semantics for formal notations and to agree upon proof procedures can also support the development of automated tools, such as theorem provers and model checkers. These can be used to avoid some of the errors that can affect the manual construction of formal proofs about complex systems. 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 (Johnson, 2002).

The potential benefits of formal techniques must be balanced by a number of concerns about the use of ‘mishap logics’. Pearl (2000) argues that one cannot use logic formalisms to prove causes in the same way that one might prove propositions or theorems. For example, 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. It is also possible to question the utility of any approach that encourages the consistent analysis of adverse events. Forcing investigators to consider a limited range of potential causes often attains inter-analyst agreement. The use of’mishap logic’ or any other formal technique should not sacrifice the analysts ability to explore a wide range of plausible causes. In particular, it is important that investigators should not have to make the mishap ‘fit the notation’. Further concerns focus on the ability of any formalist to accurately communicate the results of an investigation to non-mathematicians. This is particularly important because domain experts often cannot read or construct the abstract models that are developed using mishap logics. This can make it difficult to validate the results of any formal analysis. With these more general caveats in mind, the remainder of this paper provides a survey of the different mishap logics that might support the analysis of adverse events and near misses. In order to do this, we first introduce a complex, safety-critical mishap that will be used to compare the utility of these different approaches.

NASA maintains a wide range of wind tunnels that are used during the development and testing of aerospace applications including the Space Shuttle. In 1999, one of these tunnels was included within a facility improvement programme, which formed part of a wider initiative to modernise NASA facilities. One of the aims of this work was to improve the performance characteristics of the tunnel by replacing the turbine blades that generate the forces within the tunnel. The intention was that the existing aluminium construction would be replaced with blades made from a composite material. A NASA model shop was, therefore, commissioned to develop a number of prototypes. Two of the resulting blades were installed and tested in the tunnel for approximately 1,300 hours. This included testing under supersonic conditions with no apparent problems. It was, therefore, decided to proceed with the procurement of a set of blades based upon the prototype design. This involved the fabrication of some 200 blades. As the procurement progressed, a number of significant changes were made to the original construction of the prototype blades.

The delivered components passed an initial acceptance test and were then installed within the tunnel. Initial Integrated Systems Tests to assess the performance of the new composite blades progressed without incident. However, discrepancies were first observed during a subsonic test in May 1999. A blade had separated and raised approximately 1/8th of an inch from its 'rootblock'. Subsequent tests confirmed that it was unlikely that a blade could escape from the rootblock and so NASA staff determined that this behaviour was an anomaly and that the Integrated Systems Tests could continue. Supersonic tests were conducted over a three-day period from Wednesday 13th October-Friday 15th October. During this time, personnel were busy achieving and holding the Mach speed of the tunnel and in correlating the throat setting or aperture of the tunnel with the Mach speed. On the Monday after these tests, an initial inspection of the tunnel discovered rootblock material in the section of the Wide Angle Diffuser that lies downstream of the compressor section were the blades are located. Some of the composite blades showed evidence of delamination and of carbon fibre damage. They also discovered that there were sections missing from the trailing edges of some of the new composite blades. There was also evidence that some had separated from the rootblock beyond the tolerances that had been established during the investigation of the initial incident in May 1999. In August 2000, a Test Discrepancy Review Board identified the following causes of this mishap:

> "…(there was) a combination of poor material selection in the rootblock, inadequate quality provision in the supply contract, latent and patent discrepancies between the developed prototype and the as-manufactured blades, failures in NASA's quality program and an abrupt geometric transition in the rootblock to blade intersection, as key contributors to the failure of the new composite blades…" (NASA TDRB, 2000, page 3)

Classical Logic

This paper focuses on 'mishap logics'. The survey reviews a broad range of textual notations that might be used to support the analysis of adverse events and near misses. A companion paper provides an overview of the broader range of graphical and semi-formal techniques that have been proposed to support mishap investigations see Johnson and Holloway (2002) and Johnson (2002). The present focus on textual notations is also justified by the observation that many of these diagrammatic forms are extensions of the logics that are discussed in this paper. For example, there is a correspondence between Petri Nets, Fault Tree components and classical logic (Biljon, 1988, Hura and Attwood, 1988). This should not be surprising given that classical logic is, typically, used to provide the semantics for these diagrammatic notations. It is, therefore, appropriate to ask whether classical logic might itself provide a suitable means of analysing the causes of mishaps such as the composite blade test discrepancy, introduced in the previous section.

Classical logic is composed of propositions and sentences. Propositions represent facts that we know about the domain of discourse. In the context of this paper, the domain of discourse is the mishap under investigation. 'One place' propositions represent simple observations about individual objects. For instance, they might capture the fact that the 'as-manufactured casting resin is unsuitable'. Propositions can also be used to form relational sentences between several objects. For instance, they might be used to capture the fact that 'the prototype 4415 casting resin is more pliable than the as-manufactured 828 epoxy'. 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 analyse sentences such as the following:

> 'Static AND monitored dynamic stress data…were below those experienced during Phase 4' (p.9)
>
> 'The initial blade failure was NOT thoroughly understood' (p. 4)
>
> '…The blade had risen in the rootblock OR pulled out of the block' (p. 9)
>
> 'Thus, IF the rootblock as-fabricated conditions are such that the casting resin must carry substantial load, (THEN) the reduced material properties at compressor operational temperatures will produce a blade that appears very soft…' (p.14)

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. For instance, the first of the previous sentences describes how 'static AND monitored dynamic stress data…were below those experienced during Phase 4' (p.9). If both the static data and the dynamic data can be shown to be below the Phase 4 levels then this sentence is true. The first line of the first truth table in Table 1 illustrates this. However, if the static data (X) were below the Phase 4 level but the dynamic data (Y) was not then this sentence would be false. The second line of the first table in Table 1 illustrates this interpretation. The key point here is that investigators can use the truth tables associated with each connective to identify conditions or criteria that must be established in order to prove sentences that form their analysis. This technique is known as determining the 'truth functionality' of a sentence.

X	Y	X **AND** Y
True	True	True
True	False	False
False	True	False
False	False	False

X	Y	X **OR** Y
True	True	True
True	False	True
False	True	True
False	False	False

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

X	**NOT** X
True	False
False	True

Table 1: Truth Tables for Connectives of Classical Logic

Classical Logic and Material Conditions

This section identifies a number of problems that complicate the interpretation of conditional sentences within classical logic. These sentences are, typically, composed from an antecedent and a consequent. In the sentence "If *A* then *B*", *A* is known as the antecedent of the conditional. *B* is the consequent. Similarly, it is possible to identify and antecedent and a consequent in the sentence "If blade aerodynamic shape is altered to try and gain some aero performance improvements, (then) this approach needs further assessment…" (P. 26). Here the antecedent is that the 'blade aerodynamic shape is altered to try and gain some aero performance improvements' and the consequent is that 'the approach needs further assessment'. The truth table for conditional sentences in classical logic suggests that if the antecedent holds then the consequent is true. If the blade shape is altered then the approach does need further assessment. The truth table also reveals that it is possible for these to be other situations in which an assessment is required and that the blade shape has not been altered. The third line of the truth table illustrates this. The interpretation of the conditional illustrated by the truth table is known as 'material implication' or the 'material conditional'. Put another way, to state "If *A* then *B*" asserts that it is not the case that *A* is true and *B* is false. It is not the case that the blade aerodynamic shape is altered and no further assessment of the design is needed.

As mentioned, a number of technical problems complicate the application of classical logic to support the analysis of adverse events and near miss incidents. In particular, several paradoxes characterise valid statements in classical logic but which often seem to be counter-intuitive. The key point about these paradoxes is that they can help to confuse non-logicians and serve to undermine the credibility of results that are obtained using classical forms of mishap logic. For example, the following formulae is valid in classical logic:

$$p \rightarrow (q \rightarrow p).$$

This represents a circular form of argument. For instance, it might be suggested that 'If no other blades exhibit this failure then (if this failure is considered an anomaly then no other blades exhibit this failure)' which can be rephrased as 'since no other blades exhibit this failure, this failure is considered an anomaly because no other blades exhibit this failure'. It seems unlikely that any investigator would construct a sentence that explicitly exhibits this circularity. However, studies of the rhetoric that is used in accident and mishap reports have revealed precisely this form of argument spread across the many different pages of natural language accounts (Johnson, 2002, Snowdon, 2002). A number of further paradoxes affect the material implication in classical logic. For example, the following is also a valid formula:

$$\neg p \rightarrow (p \rightarrow q).$$

An example of this form of argument is 'if other blades do not exhibit this failure then (if other blades do exhibit this failure then the failure is considered an anomaly)'. This paradox seems to contain a contradiction. Either the blades do or do not exhibit the failure. It is, however, a valid formula. Similarly, the following paradox illustrates how it is possible to introduce arbitrary propositions into some forms of the material implication:

$$(p \rightarrow q) \text{ v } (q \rightarrow r).$$

An example of this form of arbitrary introduction is that 'if today is Tuesday then the failure is considered an anomaly or if the failure is considered an anomaly then today is Friday'. As before, the key point is not that such extreme examples are likely to occur within the individual sentences of a mishap report but that they might be embedded within the wider argument that is constructed about a near miss or adverse event. The inability of classical proof procedures to identify and, therefore, avoid such paradoxes helps to undermine confidence in the use of such analytical techniques. These doubts are exacerbated by the way in which many investigators use conditional sentences to represent causal information. A number of further technical problems make classical logic particularly unsuited for the representation and analysis of causal relationships. Take the sentence 'IF the rootblock as-fabricated conditions are such that the casting resin must carry substantial load, (THEN) the reduced material properties at compressor operational temperatures will produce a blade that appears very soft…' (p.14). This can be interpreted as meaning that operating load caused the blade softness. Problems arise because the simple 'IF…THEN' connective of classical logic cannot convey the many different and varied interpretations of causal information. For example, mishap investigators often distinguish between necessary and sufficient causes. Analysis may also distinguish between necessary and sufficient causes. A necessary cause is often identified using counter-factual 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 4 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 sufficient to cause the mishap. They too are unnecessary because C1 and C2 represent an alternative causal path. Figure 5 illustrates the sufficient conditions identified in the NASA test discrepancy report. The question marks denote that there may be other combinations of conditions that would be sufficient to create similar mishaps in the future. These may include subtle combinations of necessary conditions that were identified in the case study report. It is the focus of any investigation to remove these necessary conditions so that future mishaps can be avoided. As we shall see, the truth-functional interpretation of conditional 'IF…THEN…' sentences, illustrated by Table 1, fails to capture these causal distinctions.

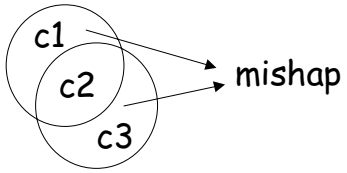

Figure 4: Necessary and Sufficient Causes.

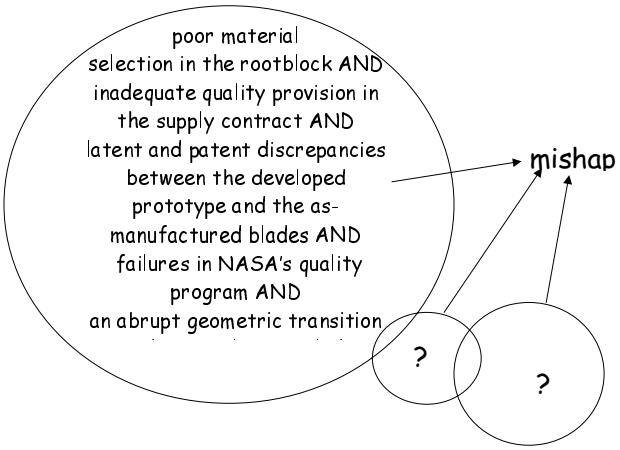

Figure 5: Necessary and Sufficient Causes

.

Although the material condition might not seem to capture the distinctions between necessary and sufficient causes, illustrated in Figure 4. It can be argued that this need not prevent the use classical logic to support the analysis of adverse events. There are, however, a number of additional technical problems that complicate the use of material conditions in mishap logics. In particular, classical logic does not require that there be any meaningful connection between the propositions that form the antecedent and consequent in a material implication. For instance, it is perfectly valid to state that 'If grass is green then the sky is blue'. This is clearly a concern in mishap analysis. For instance, an investigator might assert that 'if the final Integrated Systems Test was run on a Friday then the blades failed'. We know that the final stage of the IST was run on a Friday and that the blades suffered some form of anomaly hence the conditions required by the truth table are satisfied. The implication, therefore, holds even though there may be no connection between the day on which the final IST was conducted and the outcome of the mishap. All that the material condition of classical logic requires is that we can determine the truth or falsity of the antecedent and consequent and that they satisfy the conditions in the truth table illustrated in the previous section. In contrast, investigators typically express a causal relationship between the antecedent and the consequent in the conditions within a mishap report:

> 'Thus, IF the rootblock as-fabricated conditions are such that the casting resin must carry substantial load, (THEN) the reduced material properties at compressor operational temperatures will produce a blade that appears very soft…' (p.14)

The limitations of material implication in representing causal arguments become apparent when one analyses the different ways in which investigators use conditional statements. These go well beyond the simple forms that have been presented in the previous sections of this paper. For instance, subjunctive conditionals assert hypothetical claims of the form 'If P were to happen then Q would be the case'. This can be illustrated by the following sentence:

> If accepted industry practice was followed then the temperature limit for the material would be 140 degrees… (see p. 18).

Many causal arguments are constructed using a form of subjunctive conditional that is not characterised by material implication. In particular, counterfactual conditionals rely 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 that X is known not to have been the case. For example, an investigator might assert that 'If he had been further away, then he would not have been hurt'. There is an implication that he was NOT further away and also that he was, in fact, hurt. In the context of our case study, it can be argued that:

If the as-manufactured blades had thicker and integral blade end caps then its construction may have provided more support and helped prevent any rootblock material from tearing out (see p. 8).

It is difficult to underestimate the importance of counterfactual arguments in mishap investigation. Most mishap investigations identify root causes, X and Y, using informal arguments to support assertions that 'if X and Y had not occurred then the mishap would have been avoided'. A number of incident investigation guidelines explicitly recommend that investigators use this form of argument as a heuristic to guide their analysis (Johnson, 2002). Counterfactual arguments are also a central feature of recent formal analysis techniques for mishap investigation, including Why-Because Analysis. Consequently, the lack of support that classical logic provides for such counterfactual arguments remains a major limitation for their use in mishap investigations.

It is possible to summarise the previous arguments about the representation of causation in material implication in the following way. Assume that an event A causes another event *B*. Using, the material condition if *A* really caused *B* then *A* and *B* must have actually occurred. However, as we have seen it is not necessary for there to be any actual causal connection between *A*'s occurrence and *B*'s occurrence. Implications can hold between any two arbitrary known events or propositions. In contrast, the counterfactual condition can be expressed as 'if, all else being equal, *A* had not occurred, then *B* would not have occurred'. The previous paragraph has argued that this form of sentence can be used to distinguish root causes from contextual factors in mishap investigations. If a root cause had been prevented then the mishap would have been avoided. It is, however, possible to go beyond the counterfactual condition to look at a more general form of subjunctive argument in which event A causes B only if an instance of this causal relationship forms part of a wider, more regularity occurring pattern; 'If, all else equal, another *A* were to occur, then another *B* would occur'. It is important to construct such arguments in order to identify the measures that might prevent future accidents based on observations about previous mishaps. The paradoxes of material implication and the counterfactual, subjunctive conditions in causal arguments, therefore, justify attempt to look beyond classical logic as a tool for mishap analysis.

Addressing Limitations of Material Implication (1): Indicative Conditionals

A number of logicians, philosophers and linguists have recognised the limitations of strict implication and have responded by constructing alternative logics, which avoid the problems of classical logic, or by analysing the ways in which people construct implicational statements using material conditions. Grice (1989) and Jackson (1979) have exploited this latter approach. They argue that material implication remains a valid form of argument for *indicative* conditionals. In particular, Grice and Jackson observe that most people use arguments to communicate information in the most 'cost effective' means possible. They are anxious to avoid the costly repair actions that are necessary whenever misunderstandings occur. One consequence of this is that people will not assert weaker forms of a proposition when they can assert a strong form. In particular, 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*'.

Grice and Jackson's analysis is important because it can be used to avoid some of the problems that arise from material implication between two arbitrary false statements. Recall that material implication would allow a statement of the form 'If snow is black, then grass is red' to be true. Grice and Jackson argue that people do not reject such statements because they believe them to be 'false'. Instead, they argue that our reservations stem from the impression that such arguments would misleadingly suggest that we are unsure about the colour of snow. By analogy, our analysis of the test discrepancy might suggest that: if testing is useless, then the blades were reliable. Such arguments illustrate a common rhetorical device in which investigators use an obviously false antecedent to emphasise the falsity of the consequent. We know testing is not useless and this encourages the reader to question the reliability of the blades. However, Grice and Jackson's analysis questions the utility of such rhetorical devices, or tropes. In contrast, they encourage investigators to state the stronger proposition that testing is not useless.

Addressing Limitations of Material Implication (2): C.I. Lewis and Strict Implication

Grice and Jackson's study of linguistics preserves material implication for indicative conditionals by urging investigators to avoid implications where stronger propositions are known. In contrast, C.I. Lewis (see

Lewis and Langford, 1932) goes beyond the material implication of classical logic to develop the notion of strict implication. This is based upon the idea that a proposition *strictly implies* all others, which are true, in all possible circumstances where it is true. Syntactically, we can introduced the ->> connective to denote strict implication:

p ->> q

This can be read as 'Necessarily, if p then q'. The semantics for this form of strict implication is based around that of modal logics. Hence, we have that *A*->> *B* is true at world *w* if and only if for all *w'* such that *w'* is accessible to *w*, either *A* fails in *w'* or *B* obtains there. The sketch shown in Figure 5 can informally illustrate this.

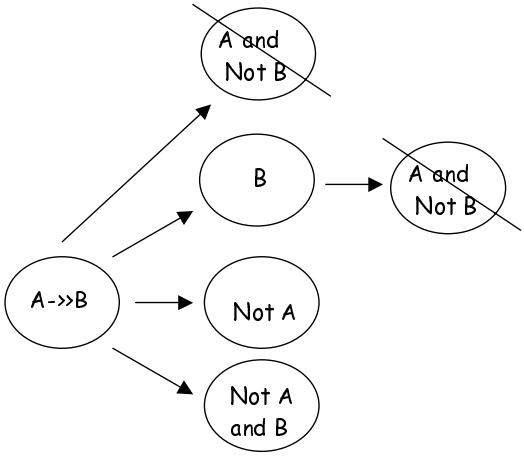

Figure 5: Possible World Semantics for Strict Implication

Each of the ovals in Figure 5 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. In particular, we no longer have A ->> (¬A ->> B) as a valid statement. Informally, unfolding this statement across possible worlds we can reach a situation in which we have both A and not A. Similarly, strict implication avoids the paradox in which B ->>(A ->> B). Again, informally this can be rejected because we can envisage a state in which B is true without A being true.

One of the apparent limitations with the Lewis semantics for strict implication is that it is conceptually more difficult to comprehend. This is a significant problem given the need to validate the results of any analysis with domain experts and incident investigators. It should be stressed that we have had to stretch the precise meaning of strict implication in order to present the underlying concepts in a manner that might be accessible to non-formalists. Having raised these caveats it is possible to suggest ways in which strict implication might be used to analyse properties of the case study. As mentioned, strict implication assumes that: *A*->>*B* is true at world *w* if and only if for all *w'* such that *w'* is accessible to *w*, either *A* fails in *w'* or *B* obtains there. For instance, an investigator might use the Lewis formulation to assert that:

> If the casting resin elastic properties degrade with increasing temperature then there will be a downwards shifting in the blade resonance frequencies (see p.4)

The Lewis semantics assert that it must never be the case that elastic properties degrade without a downward shift in resonance. This might seem like a reasonable statement. However, there are considerable practical problems in applying this form of strict implication. These can be explained by

attempting to define what the accessibility relation between states means. In other words, we must determine what the arrows indicated in Figure 5 mean. 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. If this is the case then no new information should create a situation in which elastic properties degrade without a downward shift in resonance. Another approach is to associate these transitions with the passage of time, as is the case in temporal logics. In this case, the previous implication must *always* hold throughout time. Whatever the interpretation for the possible world semantics, there are very few properties that can be expressed using strict implication without many complex antecedents. For instance, consider the previous example. If we assume a temporal semantics then it seems unreasonable to argue that it is always the case if elastic properties degrade with increasing temperature then there will be a downwards shift in blade resonance. There is likely to eventually be a stopping point beyond which any further degradation has no impact on resonance. Hence any formal analysis of this effect on the mishap would have to introduce additional caveats into the antecedent of a strict implication to express the temperature ranges of which it might hold. Of course, this need not be seen as a limitation of the formal technique because it forces the investigator to be precise in the characterisation of those properties that played an important role in the eventual mishap.

More serious concerns focus on a number of remaining paradoxes that affect Lewis' semantics for strict implication within a mishap logic. For example, this approach allows the following as a valid statement

$$(p \wedge \neg p) \text{ -}\!\!>\!\!> q$$

This would enable an investigator to create arguments of the form 'if the contract did and did not require fatigue testing, then the blades were reliable'. The initial contradiction enables us to introduce an arbitrary proposition, *q*, and still yield a valid statement. Johnson (2002a) provides an example of this form of argument in an accident report, which provided contradictory evidence for and against the conclusion that a software firm followed an industry defined management process. If such statements were formalised using the Lewis form of strict implication then it would be possible to introduce an arbitrary consequent and still have a valid statement. A number of further paradoxes affect the application of this non-classical condition:

$$p \text{ -}\!\!>\!\!> (q \text{ -}\!\!>\!\!> q)$$

This form of reasoning might be used to argue that 'if the blades were reliable then, if the contract required fatigue testing then the contract required fatigue testing'. This is similar to one of the previous paradoxes of material implication. Again, we can introduce arbitrary antecedent propositions into the implication because we know that the consequent will always be true. The strict implication is true because trivially the consequent will be true whenever the antecedent is true. There are further variants of this paradox, for instance:

$$p \text{ -}\!\!>\!\!> (q \vee \neg q)$$

This can be illustrated by the argument that 'if the blades were reliable then the contract did or did not require fatigue testing'. Again the truth of the consequent allows the introduction of an arbitrary antecedent. It seems prudent to ask whether these paradoxes introduce any practical constraints upon the use of Lewis' strict implication within mishap logic. Are investigators likely to use this counter-intuitive style of argument? If so, would the validity of these paradoxes serve to undermine confidence in any conclusions that were reached using such formalism? In the past, these paradoxes have inspired philosophers and logicians to develop alternative semantics for implication. The engineering objectives of our study might, however, permit us to live with these apparent deficiencies in the knowledge that they are unlikely to have any practical effects on the application of mishap logic. Previous pages have, however, cited examples where these paradoxes have been introduced into existing incident reports either explicitly as rhetorical devices or accidentally as part of a more complex causal argument.

Relevance Logics

Many of the paradoxes that affect the Lewis semantics for strict implication arise because the antecedent is irrelevant to consequent. A number of logicians have responded to this apparent problem by developing what are known as relevant or relevance logics (Anderson, Belnap and Dunn, 1992). One approach is to construct logics that explicitly forbid or reject the paradoxes of material and strict implication. Another approach develops an associated proof theory that provides a notion of 'relevant' proof. Anderson and Belnap (1975) have developed two variations on this proof technique. The first requires that premises and conclusion must share a variable in valid conditionals. This requirement can help to ensure that the antecedent and consequent refer to the same object in an assertion. Alternatively, the proof theory of relevance logics can require that conclusions can be directly derived from a premise without the introduction of arbitrary antecedents and consequents. This is intended to ensure that any premises really are used to obtain a valid conclusion.

As we have seen, a contradiction in classical logic can be used to entail any proposition. It, therefore, follows that implications can be derived for which the antecedent and consequent do not share any variable. Many of the other paradoxes of material implication arise because any true proposition is derivable from any other proposition. The proof theory of relevance logic excludes these approaches. This can be illustrated by variants of the natural deduction systems that have been developed for relevance logic. Indices are used to keep track of the premises and assumptions that are used in the proof. This helps to avoid the arbitrary introductions that lead to the paradoxes of classical logic. The indices at each step of the proof provide a reminder of those premises that support that stage. All premises must be used, as indicated by the presence of their associated index in the appropriate column. This approach is illustrated by the following natural deduction of the permutation of implication.

Permutation, $A \rightarrow B \rightarrow C \mid\!- B \rightarrow A \rightarrow C$		
Sentence	**Indices**	**Proof Step**
(1) $A \rightarrow B \rightarrow C$	1	Premise
(2) B	2	Assumption
(3) A	3	Assumption
(4) $B \rightarrow C$	1+3	1,3, $\rightarrow$ O
(5) C	1+3+2	2,4, $\rightarrow$ O
…	= 1+2+3	Commutivity of addition
(6) $A \rightarrow C$	1+2	3-5, $\rightarrow$ I

The proof techniques of relevance logic might also be used to ensure that investigators do not introduce arbitrary assertions into their reasoning about a mishap. For example, one of the most common means of analysing an adverse event or near miss is in terms of a chain of assertions. We are, therefore, concerned to show the manner in which the relevance logic style of natural deduction might be used to support this form of reasoning. Consider the following conditional; 'if the blade construction had provided more support then it would have prevented rootblock material from tearing out' (see p. 8). We can denote this implication as follows, assuming that B represents that the observation that the blade construction provides more support and C is used to denote the observation that the rootblock material was prevented from tearing out:

$$B \rightarrow C$$

Similarly, we might represent the assertion that 'if the as-manufactured blades had been thicker (A) then its construction would have provided more support (B)' by the following proposition:

$$A \rightarrow B$$

As mentioned, mishap investigations often use such assertions to construct inference chains to represent the successive causes of an adverse event or near miss. In order to do this, we might use the natural deduction style of relevance logics to demonstrate that 'if the as-manufactured blades had been thicker (A) then it would have prevented rootblock material from tearing out (C)'. In other words, we would like to perform a formal manipulation of the previous premises to demonstrate that:

$A \rightarrow C$

The following proof provides an example of the way in which relevance logic can be used to construct such as transitive proof, again using the indices at each step to ensure that all premises are used and that none are spuriously introduced in a manner that might lead to the paradoxes identified in the previous paragraphs:

$B \rightarrow C$; $A \rightarrow B$ |- $A \rightarrow C$		
Sentence	**Indices**	**Proof Step**
(1) $B \rightarrow C$	1	Premise
(2) $A \rightarrow B$	2	Premise
(3) A	3	Assumption
(4) B	2+3	2,3, $\rightarrow$ O
(5) C	1+(2+3)	1,4, $\rightarrow$ O
….	= (1+2)+3	Associativity of +
(6) $A \rightarrow C$	1+2*	3-5, $\rightarrow$ I

Brevity prevents a more sustained analysis of the ways in which relevance logics might be applied to support reasoning about adverse events. The interested reader is directed to Dunn (1986). The meta-level point is that these techniques provide well-established means of avoiding the paradoxes of material implication. They have not, however, been widely used to support reasoning about near misses and adverse events. A number of further caveats also affect this potential application of relevance logic. In particular, the elimination or avoidance of paradoxes does little to capture the subjunctive and counterfactual forms of implication that we have identified as key components of the causal analysis in many mishap reports.

D. Lewis, Why-Because Analysis and Counterfactual Reasoning

David Lewis (1973, 1973a) developed a number of logics that can be used to capture counter-factual arguments that capture important aspects of causation. Recall from the previous discussion that a conditional statement takes the general form of 'if A then B' where A is the antecedent and B is the consequent. A *counterfactual* is a conditional where the antecedent is false. For example, an investigator might argue:

> If there had been fewer differences between the prototype blade design and the as-fabricated components then the test discrepancies would not have arisen.

This is a counterfactual argument because, in fact, there were considerable differences between the prototypes and the delivered components. 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' or as in this case, the discrepancies would not have arisen. D. Lewis' logics for counterfactual reasoning rely upon a modal semantics, which is broadly similar to that used in C. Lewis' work on strict implication. Both depend upon accessibility relationships between possible works 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 focuses more on the relationships between possible worlds rather than creating constraints within particular accessible worlds. In particular, his 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 is a stronger form than many of the conditionals that we have met in previous sections because it 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.

Lewis' work on counterfactual arguments about causation is particularly important in the context of this survey because it lies at the heart of Ladkin and Loer's (1998) Why-Because Analysis. This, in turn, is

one of the most influential of the recent generation of formal mishap logics. The technique begins by a reconstruction phase when a semi-formal graphical notation can be used to identify and construct sequences of events leading to a mishap. These sequences can be represented in a form of temporal logic and then iteratively analysed to move from a formal reconstruction to a causal explanation of why the incident occurred 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:

$$\frac{A \wedge B \qquad \neg A \; []\text{->} \; \neg B}{A \; \text{=>>} \; B}$$

In other words, if we know that A and B occurred and that if A had not occurred then B would not have occurred then we can conclude that A causes B. These formalisations together with the Lewis semantics help to provide important tools for the analysis of adverse events. 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. As with previous techniques, brevity prevents a complete introduction to this approach. However, it is possible to illustrate the ways in which WBA provides a framework for the formal analysis of our case study. The mishap report includes the following observation:

> "It was also revealed that one of the original prototype blades, designed and fabricated in-house, had experienced a separation at this same interface. Based on the design details of the prototype blade rootblock, it had been judged at that time to be inconsequential and not an indication of concern. This assessment had been based on engineering judgement and perceived understanding of the prototype design with no engineering analytical support" (p. 11).

The previous inference rule provides a structure for representing the causal relationship that is implicit within this natural language statement. In particular, it can be used to represent the observation that the lack of analytical support directly led to the conclusion that the prototype's separation was inconsequential had there been such support then this conclusion could not have been reached:

$$\frac{\text{No engineering, analytical support} \wedge \text{prototype separation judged inconsequential} \qquad \neg \text{ No engineering, analytical support } []\text{->} \neg \text{ prototype separation judged inconsequential}}{\text{No engineering, analytical support =>> prototype separation judged inconsequential}}$$

In practice, it would be important to revise the language used in this example to avoid the confusion that can arise where logical negation, denoted by ¬, appears in the same sentence as natural language terms that also denote negation, such as 'No engineering, analytical support'. We have not done this here so that readers can trace the connections between the language used in the mishap report and the components of the inference rule provided by Ladkin and Loer. As mentioned, WBA provides a range of additional inference rules that can be recruited to structure the formal analysis of adverse events and near misses. In particular, the []=> relation can be used to denote necessary and sufficient causes. This usually takes the following form where the overall conjunction is sufficient and yet each individual A_1 is a necessary cause of B:

$$A\_1 \wedge A\_2 \wedge \ldots A\_n \; []\text{=>} \; B$$

The following inference rule can be used to establish the necessary and sufficient causes of an adverse event. Recall that =>> denotes a causal relationship:

$$\frac{C \qquad \neg C \; \text{=>>} \; \neg B \qquad \neg B \; \text{=>>} \; \neg C}{C \; []\text{=>} \; B}$$

As before, this inference rule provides a framework for the individual proof steps that support mishap investigation within Why-Because Analysis. The following example illustrates the application of the framework to our case study. Here it is assumed that the lack of engineering, analytical support led to the separation of the prototype not being judged as an important incident and conversely that the fact that the separation was not recognised as being important also helped to cause the lack of analytical support allocated to the project:

No engineering, analytical support
¬ No engineering, analytical support =>> ¬ Prototype separation judged inconsequential
¬ Prototype separation judged inconsequential =>> ¬ No engineering, analytical support
No engineering, analytical support []=>> prototype separation judged inconsequential

It is apparent that the reasoning to support such inferences relies as much upon information arguments about the validity of the component propositions as it does on the more formal inference rules provided by Ladkin and Loer. The full form of Why-Because Analysis provides means of increasing the level of formality involved in supporting such inferences. These include techniques for reasoning about operator behaviour as well as component failures. However, the proponents of this approach argue that analysts should only recruit the level of formality that is appropriate to their needs. Increasing the level of detail in a supporting proof can lead to a corresponding if not a proportionately greater increase in the resources that are required by any analysis.

Why-Because Analysis provides one of the most complete techniques for the formal analysis of adverse events. However, this does not mean that it the counterfactual approach to causation removes all of the limitations that affect other forms of conditional argument. As we have seen, material implication is truth functional. That is, the validity of an implication relies only on the interpretation of the component propositions. We determine whether 'if A then B' is true by determining the validity of the antecedent and consequent. The previous section on relevance logics identified the lack of any connection between the antecedent and consequent as a limitation of this truth functional approach. 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. Even Lewis is forced to rely upon an appeal to expert judgement in order to identify the bounds of proximity. These caveats enable us to construct intuitively appealing counterfactual arguments that have little basis in fact. For instance, we might argue that:

'If the new blades had not damaged the tunnel, something else would have'

This property of counterfactual argumentation also exposes a number of further problems. It is possible to create apparently contradictory statements. For example, it is possible to derive the following counterfactual arguments:

'If the new blades had not damaged the tunnel, something else would have'

'If the new blades had not damaged the tunnel, nothing else would have'

Any system that holds both statements to be true can appear to be both contradictory and, arguably, flawed. The significance of these observations is that they may undermine an investigator's confidence in the surrounding proof system that is intended to support their arguments about the causes of adverse events and near miss incidents. As before, it is important to offset such caveats against the wider range of methodological support that is provided by techniques such as Why-Because Analysis. It remains to be seen whether such objections substantially affect the application of the counterfactual approach or whether they remain purely philosophical objections.

Disposition Logics

The remainder of this paper surveys a range of techniques that have been developed to capture aspects of causation that do not stem so directly from a dissatisfaction with the paradoxes of material implication.

They are, nevertheless, suitable candidates for the future development of mishap logics because they often capture aspects of causation that are not directly addressed by the techniques that we have reviewed. For instance, disposition logics start from the premise that certain objects particular properties within a particular set of circumstances. These properties are termed dispositions. For instance, salt has a disposition to be soluble in water. In terms of our case study, it might be observed that:

4415 casting resin is pliable. (see p.8)

The statement of a disposition is intended to go well beyond the reported results of small number of observations or experiments. They are in tended to capture very general properties and hence have a strong predictive power. In consequence, it might be argued from the previous disposition that *whenever* salt is placed in water, *then* it will dissolve. Or in terms of our example:

whenever casting resin 4415 is subject to a transverse pressure, *then* it will bend.

As can be seen from these examples, the natural language statement of dispositional properties, typically, involves a 'whenever… then' construction. From this it can be argued that 'whenever' is a generalisation from the more conventional conditionals that we have seen in previous sections. Disposition logics differ from the approaches that we have examined in previous sections because they distinguish between deterministic and probabilistic causal relationships. For instance, it can be argued that the solubility of salt in water is a 'sure fire' disposition. In contrast, other dispositions are probabilistic from one moment to the next. For instance, the disposition of a radioactive nucleus to decay can be expressed as a *probability* that decay will occur within a particular time interval. Similarly, a mishap investigator might argue that probabilities are the only way to express whether or not the disposition of 4415 casting resin will continue to resist a particular amount transverse pressure from one interval to the next. There may be a very good chance that blades made from the resin will eventually break under a particular loading. Also, dispositions will change over time even while a disposition is not being manifested. For instance, a component made from 4415 casting resin may lose its ability to resist pressure as it ages even if no pressure is applied to it.

These general comments ignore the underlying problem of how to interpret a disposition. We have not provided any semantics for the concept. In other words, we have not explained what it *means* to say that 4415 casting resin is pliable. In order to do this we must provide a meaning for the conditional phrases that characterise this class of properties. It is clear from the previous sections that this form of conditional cannot refer to material implication. For example, assume that P(x) means `*x* is pliable'; S(x,t) means `*x* is stressed at time *t;* B(x,t) means `*x* breaks at time *t*'. We might use the material condition to construct a definition of this disposition in the following manner:

$$P(X) = (\forall t)S(x,t) \rightarrow B(x, t);$$

Unfortunately, one consequence of this would be that every object that was never stressed would have a disposition to be pliable. Alternatively, we might assume that P(x,t) to mean *x* is pliable at time t and construct the following formalisation of the disposition:

$$P(x,t)=S(x,t) \rightarrow B(x,t)$$

This would result in every object that was not stressed at time *t* have a disposition to being pliable at time *t*. Mackie has attempted to resolve this apparent impasse by arguing that dispositional statements are constructed from non-material conditional statements. The implications can only defined in terms of modal structures similar to those used by C. Lewis' account of strict implication and D. Lewis' development of counterfactual argumentation. However, Mackie (1973) shows considerable flexibility in his interpretation of the dispositional condition arguing that it can take an 'open' or equivocal form, a subjunctive or a counterfactual form depending on circumstances. For instance, if a particular sample of the resin was known not to be stressed at time *t*, then a counterfactual form is appropriate: 'If the resin was stressed at *t* then it would have broken'. Alternatively, if it was known that a sample of the resin was stressed at time *t*, then Mackie exploit the simple construct $S(x,t) \equiv B(x,t)$. In other words, 'if the resin is stressed then it breaks'. If it is not known whether a sample was stressed or not, then the subjunctive

conditional should be used. That is to say 'if the resin were to be stressed, then it would break'. Mackie calls these different forms *minimal dispositions*. He argues that the identification of a minimal disposition involves the identification of a suitable non-material condition. This analysis can be seen in one of two ways. Either Mackie's pragmatism reflects the many subtle distinctions that are embodied in natural language accounts of causal phenomena or his catholic use of different non-material conditions introduces a host of paradoxes and problems associated with each of the individual forms.

Disposition logics can be used to describe how an object, S, has a disposition, P, to do action A in some circumstance C IF there is a non-zero likelihood of S doing A in that circumstance. An important potential benefit of this approach is that is can be used to describe the subtle and changing influences that management and organisational structures have upon the course of a mishap. For example, the test discrepancy report from the wind tunnel incident contains the following observation:

> "Lacking the callout for a structural adhesive between the rootblock ribs and the blade flare, the specified tolerance did not "raise a flag" to the contractor as a crucial fit-up item. Further, the contract did not specify specific quality provisions for this fit-up. As a result, the contractors' attention to this fit-up tolerance was not exacting. The over-tolerance gap between the ribs and the blade flare exacerbated the structural situation for the casting resin." (p. 4)

An investigator might interpret this statement using the concepts of disposition logic. A contractor (S) has a disposition (P) to ignore a fit-up tolerance (A) if the contract fails to flag this tolerance (C) as being significant. We can use this example to further illustrate the different non-material conditionals that were highlighted by Mackie's analysis of minimal dispositions. For instance, a counterfactual approach would yield an argument that:

> If the contract had flagged tolerance then the contractor would not have ignored the fit-up tolerance…

Alternatively, the weaker form of subjunctive implication would yield the following from of argument:

> If a contract does not flag the fit-up tolerance as being important then the contractor is likely to ignore it.

It can be argued that disposition logic creates as many problems as it answers. Mackie shows how the semantics of these techniques must, typically, be constructed from the non-material, model forms of implication introduced in previous sections. As we have seen, these can often introduce additional caveats and concerns. Similarly, the extension of disposition logics beyond 'sure fire' properties can raise questions about what exactly is meant by phrases such as 'non-zero likelihood'? It is important to emphasise that such concerns are not confined to the proponents of disposition logic. Many logicians and philosophers have sought to address the paradoxes of material implication by recruiting probabilistic concepts. As we have seen, material implication makes no connection between the antecedent and consequent. Hence we have valid arguments of the form 'if snow is white, then grass is green'. Edgington (1995) has, therefore, argued that a probabilistic connection is implicit within conditional statements. These statements reflect the subjective probability that speaker assigns to conditional. Evidence that snow is white increases the likelihood of grass being green. The consequent is therefore connected to the antecedent by an implicit form of conditional probability. It is, therefore, appropriate to extend the scope of this survey to consider Bayesian techniques that can represent subjective rather than objective views of probability. In this view, the observation of a cause makes an effect more likely. Only in a relatively small number of situations will it make it completely certain that an effect will occur. It can be argued that this high-level framework provides a more accurate analysis of many incidents that the more deterministic approaches that we have reviewed in previous sections.

Bayesian Logic

Bayesian analysis provides analytical techniques for representing and reasoning about subjective views of probability and cause (Jayne, . It can be contrasted with the statistical, frequentist or objective views that will be the focus of the next section. These objective views focus on determining the probability of events

that we can count, such as coin tosses or the likelihood of selecting a particular card from a deck. In contrast, subjective approaches to probability capture the beliefs of an idealised rational agent. They focus on the likelihood of events that it may be difficult to directly observe or count, such as 'what is the chance of nuclear war in next ten years?'. Such applications are particularly relevant for mishap analysis. By analogy, analysts may determine the likelihood of rare adverse events that may not yet have occurred within a particular facility.

Conditional probabilities are at the centre of the Bayesian approach. The conditional probability of a proposition given particular evidence represents an agent's belief in that proposition, given the 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. An example of the application of this approach is to consider the probability that a coin will land showing the side with a head on it. The likelihood of this outcome can depend on a number of factors, such as the manner in which the coin was tossed or the way in which it is caught. In particular, it might depend upon the fairness of the coin. If we have evidence of previous tosses we might use this information to revise our assessment of a particular outcome, especially for instance if we saw ten heads come up in a row. In such circumstances, our subjective estimate of the probability of the hypothesis that heads would occur next, h, would be dependent on the evidence, e, that we had previously seen heads appear in the ten previous throws.

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. It does not, typically, imply the necessity of h given that we are dealing with probabilistic inferences. Bayesian techniques can also be used abductively. If we have a conditional probability of the form pr(h|e) and instead of observing the evidence we rather observe that our hypothesis is true then we might also make a number of inferences about the likelihood of e also being true. These different forms of reasoning can best be illustrated through an example. Assume that if the grass is wet then the likelihood that it has been raining is 0.8 (A). Similarly, if the sprinkler system is on then the probability that we observe wet grass is 0.95 (B). Finally, we might argue that if it has been raining then the likelihood that we will observe a wet lawn is 0.93 (C). If we then observe that the grass is wet then we can conclude that it is likely to have been raining. The probability for this is given as 0.8 by rule (A). Notice that we have not given a rule of the form 'if wet grass is observed then the sprinkler has been on'. The addition of such a rule would enable us to compare the relative probabilities of the two different explanations but we would also have to consider the likelihood that, for instance, it had both been raining and the sprinkler had been operating. Given the previous rules, all we can conclude is that it is likely to have been raining without making any statement about the state of the sprinkler system. However, if we observe that it has been raining we cannot use rule A to reason that the lawn will be wet. This form of abduction relies upon rule C.

We can adapt the previous example to show how forms of Bayesian reasoning might be used to support mishap analysis. For instance, the following statements might be used to link the likelihood of a failure in the bond used in the blade rootblock shell bottom butt joint and the different forms of adhesive that were used. In the following, it is important to stress that the precise values need not be derived from failure frequencies but can be derived from subjective expert judgements. In this case the numeric values are purely indicative of the way in which this technique might be used:

- if there was bond separation then 5-minute epoxy on butt joint (0.08)
- if 5-minute epoxy was used on butt joint then bond separation (0.95)
- if EA9394 structural adhesive was used on the butt joint then bond separation (0.0093) (see page 8).

As in the previous example, if bond separation is observed then an investigator might use the first rule to conclude that the use of a 5-minute epoxy-bonding agent was the likely cause. Conversely, if EA9394 was observed then we might need to search for other contributory causes given the relatively small likelihood of any bond separation. A key point here is that this approach encourages investigators to explicitly state the subjective probabilities that they associate with particular causal factors. This can be used early in an

investigation to help uncover any potential biases that might unreasonably exclude certain causal factors that the investigator considers to be unlikely to have contributed to a particular mishap. This is illustrated in the previous example by the relatively small subjective probability associated with the EA9394 structural adhesive that was used in the prototype assembly contributing to a subsequent bond separation. Conversely, if an investigation team concurs with a relatively low probability estimate then investigation resources can be channelled to eliciting the necessary evidence that might establish high probability causes.

Bayesian techniques enable mishap investigators to reason about the manner in which the observation of evidence affects our belief in causal hypotheses. For instance, 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:

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

This formula can be used to assess the likelihood of a cause B given that a potential effect, A, has been observed. This has clear applications in the causal analysis of adverse events and near misses. In particular, it provides a means of using information about previous incidents to guide the causal analysis of future occurrences (Johnson, 2002). In our case study, investigators might be interested to determine the likelihood that reported different bonding agents had caused damage to a bond. The analysis begins by assessing the likelihood that different agents are used in the assembly process. 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. In this case, the likelihood of finding that EA9394 was used in a context C is given as 0.98. The likelihood associated with the other bonding agents are described in a similar fashion and again, in this example, the actual values are purely intended to illustrate the application of the approach:

$\Pr(\text{EA9394} \mid C) = 0.98$
$\Pr(\text{5-minute epoxy} \mid C) = 0.01$
$\Pr(\text{other bond} \mid C) = 0.01$

The next stage is to determine how likely it is that the use of one of these different agents would lead a bond to separate. Further analysis might reveal a likelihood of 0.0093 that bond separation stemmed from the use of EA9394. Alternatively, this can be interpreted as 93 incidents involving EA9394 out of every 10,000 reported bond separations:

$\Pr(\text{bond separation} \mid \text{EA9394} \wedge C) = 0.0093$
$\Pr(\text{bond separation} \mid \text{5-minute epoxy} \wedge C) = 0.95$
$\Pr(\text{bond separation} \mid \text{other bond} \wedge C) = 0.0407$

We can now integrate these observations into the previous formula to calculate the probability that a bond used EA9394 given that a bond failure has been reported. This following calculation suggests that there is almost a 50 per cent chance that a failure involved this material. This may seem counterintuitive, however, it reflects the relatively low probability of other agents being used:

$$\Pr(\text{EA9394} \mid \text{bond separation} \wedge C)\}$$

$$= \frac{\{\Pr(\text{bond separation} \mid \text{EA9394} \wedge C).\Pr(\text{EA9394} \mid C)\}}{((\Pr(\text{bond separation} \mid \text{EA9394} \wedge C).\Pr(\text{EA9394} \mid C)) + (\Pr(\text{bond separation} \mid \text{5-minute epoxy} \wedge C).\Pr(\text{5-minute epoxy} \mid C)) + (\Pr(\text{bond separation} \mid \text{other bond} \wedge C).\Pr(\text{other bond} \mid C))}$$

$$= \frac{(0.0093).(0.98)}{(0.0093).(0.98) + (0.95).(0.01) + (0.0407).(0.01)}.$$

$= 0.48$

A number of caveats can be raised against this application of Bayes' theorem. Dembski (1998) argues that it is seldom possible to have any confidence in prior probabilities. Such figures can only be trusted in a limited number of application domains. For instance, estimates of the likelihood of an illness within the general population can be validated by extensive epidemiological studies. It is difficult to conduct similar studies into the causes of safety-critical mishaps. Especially given that such incidents are, typically, rare occurrences. For example, it seems unlikely that there would be any means of validating the subjective assessment that 93 incidents involved EA9394 out of every 10,000 reported bond separations. Given the specialised nature of the wind tunnels systems in this case study, any design that suffered this number of failures would be modified or closed down before such statistics could be compiled.

There are a number of extensions to the Bayesian approach described in the previous paragraphs. None of these techniques have been explicitly developed to support mishap investigation. However, some have considerable potential to support both the causal analysis of adverse events and the application of frequency information derived from the statistical analysis of mishap data. For instance, Bayesian Belief Networks have been developed to represent particular forms of causal relationships. They can be used to distinguish between marginally independent and conditionally dependent observations. For example, rain has no impact on whether or not a sprinkler is on. These two concepts are marginally independent. However, if we observe that the grass is wet then because this can be explained away by rain or the sprinkler then there concepts are conditional dependent. The conditional dependency is created by the observation of a wet lawn. By analogy, this is similar to a situation in which a mishap had been observed and creates a conditional dependency between the different causal hypotheses that are considered by the investigator. Bayesian Belief Networks provide mechanisms and procedures for analysing and updating information about particular hypotheses as more evidence is obtained. Brevity prevents a more sustained analysis of this approach and the interested reader is directed to Johnson (2002).

Comparative Probabilities and Partition Models

The main concern about the use of Bayesian statistics to support mishap analysis is that they rely upon the provision of subjective probabilities. These are subject to systematic biases (Puppe, 1991). Given the low frequency of many mishaps and the relatively slow development of international systems for the exchange of incident statistics, it seems unlikely that it will be possible to develop quantitative means of validation these estimates. There are further complexities. For instance, it is unclear how one might account for the role of human behaviour and software failure within many adverse events. There are a wide number of technical objections to those techniques that have been developed to support the derivation of numerical reliability assessments for these system 'components' (Leveson, 2002). There are technical solutions to some of these problems. For example, Bayes factors can be used to quantify the degree of support for a hypothesis in an existing data set and hence can be used to avoid stating subjective prior probabilities. However, most of these techniques remain the subject of considerable debate. For example, Lavine and Schervish (1997) identify problems with Bayes Factors that do not arise when using other measures such as the posterior odds ratio. However, this approach does not avoid the problems of subjectivity that motivated the use of these factors in the first place. 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:

$$Pr(b|a \wedge C) > Very\_likely$$

This is little more than an extension of the Bayesian analysis that was presented in the previous section. However, such an observation can be seen to founder when we attempt to explain what is meant by the phrase '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 (Hausman, 1998). 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.

$$Pr(B \wedge A \mid C) <> Pr(B|C).Pr(A|C)$$

In other words, the probability of a and b are dependent if there is a causal connection between A and B. This can work in one of two ways. 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 such arguments to the test discrepancy report that provides the case study for this paper. For example, the probability that a 5-minute epoxy has been used and that a bond separation has occurred is greater than the independent probabilities that the epoxy is used multiplied by the probability that a bond separation had occurred. This reflects the causal relationship between the bond separation and the use of this agent:

$$Pr(\text{5-minute epoxy} \wedge \text{bond separation} \mid C) > Pr(\text{5-minute epoxy} \mid C).Pr(\text{bond separation} \mid C)$$

Conversely, it can be argued that the likelihood of EA9394 being used in a situation where a bond separation has occurred is less that the probability of that medium being used multiplied by the probability of a bond separation occurring:

$$Pr(EA9394 \wedge \text{bond separation} \mid C) < Pr(EA9394 \mid C) .Pr(\text{bond separation} \mid C)$$

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 a before b in any causal sequence. A causal connection between a and b might also require that the probability of B and A in C is greater than probability of B given that we know $\neg$ A and C. Informally, knowing A increases the probability of B above a similar situation in which we know $\neg$ 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:

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

Again, we can apply this form of 'non-quantitative' causal reasoning using probabilities to the case study mishap report. The use of 5-minute epoxy leads to bond separation if the probability that this material was used and a mishap occurred is greater than that associated with mishaps in which bond separation occurred without the use of this material or situations in which nothing is known about the use of this bonding agent:

$$Pr(\text{bond separation} \wedge \text{5-minute epoxy} \mid C) > Pr(\text{bond separation} \mid \neg \text{ 5-minute epoxy} \wedge C) \text{ v}$$
$$Pr(\text{bond separation} \mid \text{5-minute epoxy} \wedge C) > Pr(\text{bond separation} \mid C)$$

A number of further factors complicate this analysis. For instance, 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 $\neg$ 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:

$\forall$ j: Pr(B|A $\wedge$ S_j $\wedge$ C) > Pr(B|S $\wedge$ C)

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. For example, if there is no fuel then the presence of an ignition source cannot start a fire. 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 NASA test discrepancy report. Recall that a factor is in c_1,...,c_m if and only if it is a cause of b or $\neg$ b, other than a, and it is not caused by a. For example, we might focus on the following factors in the blade mishap:

c_1 represents '5-minute epoxy',
c_2 represents `EA9394',
c_3 represents `blade end caps not integral to rootblock shell'.

This yields the expected eight combinations of possible causal factors:

{{c_1, c_2, c_3}, {c_1, c_2}, {c_1, c_3}, {c_2, c_3}, {c_1}, {c_2}, {c_3}, {}}.

We can then begin the process of pruning the number of test conditions that we must consider in order to demonstrate that the use of 5-minute epoxy was a cause of the observed anomaly. An investigator might exclude the conditions represented by {c1, c2} and {c1,c2,c3} because this would suggest the use of two different structural adhesives within the same assembly. Of course, co-workers might argue against such an assumption. The key point, however, is that the technique associated with partition models can help investigators to exclude alternative causal hypotheses during the analysis of a mishap. Similarly, it might be argued that we cannot have the situations represented by {c3} and {} because this might imply that no structural adhesives were used. Recall that c_1,..., c_m is intended to be an exhaustive characterisation of possible causes. Similarly, {c1} and {c2} might be rejected as they imply that there were no end caps. This process continues until investigators are satisfied that the remaining combinations of causal factors cannot be excluded from consideration. It then remains to be shown that a causes b, or that 5-minute epoxy would result in the mishap, under all of the combinations of other factors represented in the partition. As in previous sections, it is important to note that the partition formulae are parameterised with the context, denoted by C. This is intended to show that any partition is relative to the context in which a mishap occurs. Changes in that context can yield a new and revised set of potential causal factors that would then have to be considered in a revised causal analysis. Similarly, arguments that previously excluded certain combinations of factors might need to be revised to reflect the changed operating practices.

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 then 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 Test Discrepancy Report case study has identified more than twenty such factors that might be considered relevant to the blade anomaly. 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 and Further Work

Mishap investigations provide important information about adverse events and near miss incidents. They are intended to help avoid any recurrence of previous failures. Over time, they can also yield statistical information about incident frequencies that helps to detect patterns of failure. In theory, this information can also be used to validate the risk assessments and safety cases that, typically, guide the development and operation of safety-critical systems. However, the increasing complexity of many safety critical systems is

posing new challenges for mishap analysis. The recognition that many failures have complex, systemic causes has helped to widen the scope of many mishap investigations. A new generation of formal and semi-formal techniques have been proposed to help investigators address these problems. In particular, a number of mathematically based 'mishap logics' have been developed. These can be used to formally prove that certain events created the necessary and sufficient causes for a mishap to occur. These proofs can be used to reduce the bias that is often perceived to effect the interpretation of adverse events. This paper has provided an overview of these mishap logics. It has also identified several additional classes of logic that might also be used to support mishap analysis.

The previous discussion has extended the application of many of these mishap logics well beyond the original intentions of the philosophers and logicians who original constructed them. In most cases, their intention was to improve our understanding of the many different ways in which we perform causal reasoning. Our extension of their work creates the risk that we have misinterpreted aspects of the underlying models or that we have extended the assumptions that support the logics to a point at which they hold few similarities with the initial formalism. Further work is required to verify that many of the results, which hold for the initial logics, also hold for our extensions.

It is also important to stress that this paper has provided only a partial review of the approaches that might be used to support mishap analysis. For instance, we have argued that a number of paradoxes undermine the use of the Lewis semantics for counterfactual reasoning. As we have seen, many these paradoxes stem from the same problems of relevance between the antecedent and consequent of a conditional that also affect classical logic. Mares and Fuhrmann (1995), therefore, extend a variant of the relevance logics that we have presented in this paper to support the counterfactual reasoning that might support the causal analysis of adverse events and near miss incidents. Further work is required to determine whether the benefits that such hybrid techniques provide, for example in removing the relevance paradoxes, can justify increasingly complex semantics.

We have only sketched ways in which these approaches might be recruited to address the problems of system complexity and integration that are exacerbating the problems of mishap analysis. Further work is required to validate our studies and to determine whether the proposed benefits hold for a range of different mishaps beyond the test discrepancy report that has formed our case study. Conversely, further analysis is required to determine whether they may be further benefits from the use of mishap logics that we have not identified in this paper. For instance, it might be that the use of mishap logics might support the development and use of formal specifications in systems development. Techniques that enable investigators to reason about what caused a system to fail should also enable them to identify what needs to be done in order to avoid future failures.

Acknowledgements
This research has been funded by a NASA contract NAS1-97046, Task 212. UK Engineering and Physical Sciences Council grant (EPSRC GR/M98302) has provided additional support.

References

A.R. Anderson and N.D. Belnap (1975), *Entailment: The Logic of Relevance and Necessity*, Princeton, Princeton University Press, Volume I.

A.R. Anderson, N.D. Belnap and J.M. Dunn (1992) *Entailment*, Princeton, Princeton University Press, Volume II.

W.R. Van Biljon, (1988), Extending Petri Nets For Specifying Man-Machine Dialogues, International Journal of Man-Machine Studies, (28)4:437-455.

C. Burns, (2000), Analysing Accidents Using Structured and Formal Methods, PhD Thesis, Department of Computing Science, University of Glasgow, Scotland, UK.

D. Busse and D. Wright (2000), Classification and Analysis of Incidents in Complex, Medical Environments, Topics in Health Information Management, (20)4:1-11.

W.A. Dembski (1998), The Design Inference: Eliminating Chance Through Small Probabilities, Cambridge University Press, Cambridge, U.K.

M. Dunn, (1986), Relevance Logic and Entailment. In D. Gabbay and F. Guenther (eds.), Handbook of Philosophical Logic, Vol III, 117-224, D. Reidel Publishing.

Edgington (1995), On Conditionals. Mind, (104):235-329.

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

D.M. Hausman, (1998), Causal Asymmetries, Cambridge University Press, Cambridge, U.K.

G.S. Hura and J.W. Attwood, (1988) The Use Of Petri Nets To Analyse Coherent Fault Trees, IEEE Transactions On Reliability, (37)5:469-473.

IEC 61508, (2000) Functional safety of electrical/electronic/programmable electronic safety-related systems, International Electrotechnical Commission. See http://www.iec.ch/61508 for further details.

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

C.W. Johnson (2002 in press), A Handbook for the Reporting of Incidents and Accidents, Springer Verlag, London, UK.

C.W. Johnson (2002a), The London Ambulance Service, Computer Aided Dispatch System: A Case Study in the Integration of Accident Reports and the Constructive Design of Safety-Critical Computer Systems, Reliability Engineering and Systems Safety.*** details

C.W. Johnson, G. Le Galo, M. Blaize (2000), Guidelines for the Development of Occurrence Reporting Systems in European Air Traffic Control, European Organisation for Air Traffic Control (EUROCONTROL), Brussels, Belgium.

C.W. Johnson, C.M. Holloway and B. Strauch (2002), Subjective Evaluations of Formal Notations for Mishap Analysis, Technical Report.

U. Kjellen (2000), Prevention of Accidents Through Experience Feedback, Taylor and Francis, London, United Kingdom.

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

N. Leveson, (2002), A Systems Model of Accidents. In J.H. Wiggins and S. Thomason (eds) Proceedings of the 20th International System Safety Conference, 476-486, International Systems Safety Society, Unionville, USA.

M. Lavine and M. J. Schervish (1997), Bayes Factors: What they are and what they are not. Technical Report 652 1/97. Department of Statistics, Carnegie Mellon University.

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

D. Lewis, (1973), Causation, Journal of Philosophy,70: 556-567.

D. Lewis, (1973a), Counterfactuals, Oxford University Press, Oxford, UK.

J.L. Mackie, (1973), Truth, Probability and Paradox, Oxford University Press, Oxford, U.K.

E.D. Mares and A. Fuhrmann, (1995), A Relevant Theory of Conditionals. Journal of Philosophical Logic. (24)6:645-665.

NASA TDRB, (2000) Investigation of Unitary 3-Stage Compressor Composite Blade Test Discrepancy, NASA Ames, Test Discrepancy Review Board, August 22.

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

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

J. Reason (1997), Managing the Risks of Organizational Accidents, Ashgate Publishing, Aldershot, UK.

E.J. Sampino (2001), Applications of Fault Tree Analysis to Troubleshooting the NASA GRC Icing Research Tunnel. Proceedings of the 2001 IEEE Annual Reliability and Maintainability Symposium, 16-21, IEEE Press, New York.

P. Snowdon, (2002), Analysing Argumentation Structures in Accident Reports, PhD Thesis, Department of Computing Science, University of Glasgow, Scotland, UK.

T.W. van der Schaaf, (1992), Near Miss Reporting in the Chemical Process Industry, PhD thesis, Technical University of Eindhoven, Eindhoven, The Netherlands.