Chapter 11

Alternative Causal Analysis Techniques

The previous chapter showed how a range of existing techniques can be applied to identify the root causes and causal factors that lead to failures in high-technology systems. In particular, we have shown how Events and Causal Factor (ECF) charts can be derived from the findings of primary and secondary investigations. The scope of these diagrams can be both broadened and deepened using barrier analysis and change analysis. Counterfactual reasoning can then be applied to distinguish causal factors from other contextual influences on an incident. Finally, tier analysis and non-compliance analysis can be used to distinguish the root causes that threaten future safety from the causal factors that characterise individual incidents. The intention was to provide a relatively detailed case study in the application of these particular analytical techniques. The choice of approach was motivated by the recommendations of the US Department of Energy and of NASA NPG 8621.1.

The following pages build on this analysis by introducing a range of alternative techniques. The intention is to provide a broader perspective on causal analysis. As we shall see, some of these techniques can be integrated into the approach that was described in the previous chapter. For instance, ECF charts can be replaced by Sequentially Timed and Events Plotting or by Multilinear Events Sequencing [72, 348]. The justification for broadening the scope of the previous chapters is that there have been few investigations into the comparative utility of causal analysis techniques. There are some notable exceptions. For instance, Benner [73] provides a rating of accident models and investigative methods. Munson has more recently presented a comparative analysis of accident modelling techniques applied to Wildland Fire Fighting Incidents [553]. It is important to note that both of these studies are more concerned with the range of factors that are captured by particular modelling notations and their integration into investigatory processes. Neither directly studies the ultimate application of these models to support causal analysis. In the absence of such comparative studies, it is important that investigators have a clear idea of the alternative approaches that might be used to support the causal analysis of safety-critical incidents.

A fire on-board the bulk carrier Nanticoke provides a case study for the remainder of this chapter [623] This is appropriate because it provides a further contrast to the Pipeline expolosion that was modelled in Chapter 8.3 and the Mars case studies that were analysed in Chapter 9.3. The Nanticoke departed Camden, New Jersey, USA, on 19 July 1999. It was carrying 29,000 tons of petroleum coke. Between 12:00-16:00 on the 20th July, an engineer cleaned the forward fuel filter on the Nanticoke's port generator as part of a preventive maintenance routine. The engineer started the generator and tested the filter for leaks around 15:00. At 15:15 the chief engineer entered the engine-room and inspected the generators. He found that all temperatures and pressures were normal and, therefore, continued on to the control room. Shortly after this, a fire drill was started. The chief engineer relieved the duty engineer who had to go to an assigned fire station. During this time, the chief engineer and a mechanical assistant maintained their watch from the engine control room where they could not directly observe the state of the generator. The fire drill ended at 16:00. Shortly after this, the chief engineer noted a high cooling water temperature alarm from port generator cylinder No. 1 from the engine control room displays. He left the control room and discovered that the engine-room was full of smoke.

The chief engineer returned to the control room and sounded the general alarm. He then called the bridge and informed them of the fire. He shut down the port generator, isolated its fuel supply and then put on a smoke hood so that he could find the mechanical assistant. The mechanical assistant had already left the engine-room and so the chief engineer returned to the control room. The control room was not equipped with an emergency exit and so he was forced to follow handrails to the engine-room exit door on the main deck. The starboard generator was left running to supply power to the rest of the vessel.

On the bridge, the master sent a security call that was acknowledged by the United States Coast Guard in New York City. They then transmitted a Mayday as the extent of the fire became more apparent. The fire parties were standing down from the drill and were in the process of removing their protective fire suits when the alarm sounded. Two crew members entered the engine-room using air packs and protective suits that were already to-hand following the fire drill. They initially used carbon dioxide extinguishers to fight the fire but were driven back by the heat. A second team then repeated the attempt using a fire hose but this also failed to completely extinguish the fire. The chief engineer then performed a headcount and ensured that the engine-room vents were closed. He then discharged the Halon extinguishing system around 16:40. The fire was fully extinguished by 17:22. Shortly after this time, the gangway doors were opened to ventilate the engine-room.

The remaining pages use this incident as a case study to illustrate a number of alternative causal analysis techniques. This provides investigators with an overview of the rival approaches to the ECF and Causal Analysis techniques that were presented in Chapter 9.3. The following pages also introduce complementary techniques that can be used to supplement or replace the method that was described in the previous chapter.

11.1 Event-Based Approaches

ECF charts can be used to analyse the way in which various chains of events and conditions contribute to safety-critical incidents. Failure sequences can be sketched, edited and extended as other techniques, such as barrier analysis, drive further insights into an incident. Unfortunately, a number of limitations reduce the utility of this approach. For instance, Munson argues this method is labour intensive and often requires considerable amounts of time to complete even a preliminary analysis [553]. It also requires a considerable range of domain knowledge, in additional to the technical knowledge required to perform the analysis [292]. For instance, tier analysis relies upon a knowledge of the managerial structure of the many organisations that are involved in an incident. As we have seen in the previous chapter, commercial barriers and the complexity of some management organisations can frustrate attempts to elicit this information even in cases where serious failures have occurred. Further limitations stem from the manner in which temporal information is included within individual events and conditions. There is an implicit assumption that time flows from the left to the right in an ECF chart. An event or condition is assumed to occur after events or conditions that are placed to their left. There is, however, no time scale associated with ECF charts. In consequence, investigators must manually search through dozens of nodes in these diagrams to determine what might have happened at any particular moment in time.

11.1.1 Multilinear Events Sequencing (MES)

Multilinear Events Sequencing (MES) provides an alternative to the ECF charts in Chapter 9.3. It is different from the more general modeling techniques introduced in Chapter 8.3, such as Petri Nets and Fault Trees, because it was specifically developed to support accident and incident analysis [72, 348]. It is intended to help investigators model and analyse an incident as an investigation progresses [706]. This implies that the approach avoids some of the overheads associated with the more elaborate techniques that are presented in previous chapters.

The basic premise that underlies MES analysis is that both successful operations and failures are the result of processes that are comprised of interactions between events. Rimson and Benner go on to argue that incidents occur "when changes during a planned process initiate an unplanned process which ends in an undesired outcome" [706]. Such comments must be balanced against situations in which two planned processes interact to produce an undesired outcome [449]. The underlying assumption, however, is that by analysing changes in a planned process it is possible to identify the potential causal factors that lead to adverse events. Processes are described in terms of a relationship between events. This is very similar to the approach adopted in ECF charts.

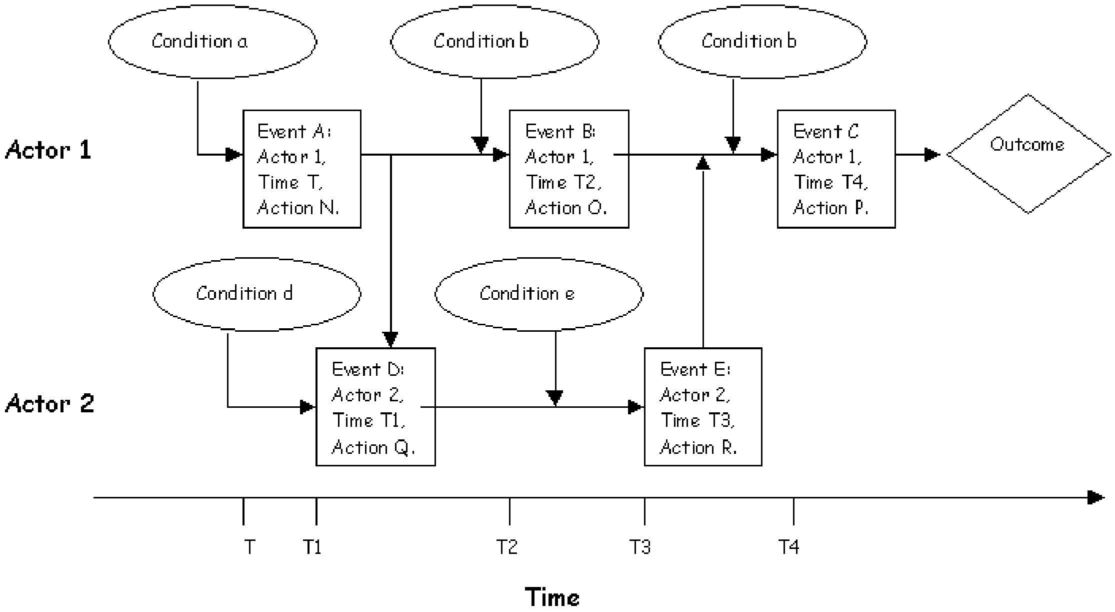

Figure 11.1: Abstract View of A Multilinear Events Sequence (MES) Diagram

Figure 11.1 presents the high-level form of MES flowchart. Each of the events in Figure 11.1 is described in terms of a block of information. These represent an actor performing an action at a particular time. A time-line is also included at the bottom of MES charts. This is used to explicitly represent the timing of events. It is important to note, however, that the relative position of a condition does not explicitly convey any temporal information. As can be seen, there is a deliberate attempt to help investigators identify situations in which simultaneous events contribute to an incident. The intention is to to "discover possible unknown linking events, causes, and contributing factors" [553]. The resulting diagrams resemble annotated flowcharts. This should not be surprising. The developers of MES argue that "if you can't depict a process in a flowchart, you don't understand it!" [706]. Such statements should be interpreted with care. The underlying importance of constructing accident models that are easily understood by a number of different investigators cannot, however, be denied. The MES methodology can be summarised as follows:

1. *Identify the boundaries of an incident.* A key objective behind the development of MES was to construct a method that could be used to delineate the beginning and the end of an accident sequence. Peturbation Theory (or P-Theory) was proposed to support these objectives. This starts from the assumption that the "dynamic equilibrium of successive events progresses in a state of homeostasis requiring adaptive behaviour or adaptive learning by the actors involved in maintaining the stable flow of events" [72]. Incident sequences begin with a perturbation that disturbs this dynamic equilibrium. If the system adapt to these changes then homeostasis can be maintained. If the system fails to adapt then an accident or incident sequence begins. Initial

peturbations can initiate cascading sequences of events that, in turn, place further pressures on other system components. These components can either fail or they can adapt to changing circumstances. P-Theory defines a 'near miss' incident to occur if system components adapt to any perturbations before an injury or other form of loss occurs. A number of caveats can be applied to this aspect of the MES technique. Some authors have proposed that the search for peturbations should end when "the final damaging event" is identified [553]. As we have seen, however, any analysis should ideally also consider the immediate response to any adverse occurrence given that this can either exacerbate or mitigate the consequences of any initial failure. Secondly, there are some incidents in which it is difficult ever to identify homeostasis. For instance, the relationship between LMA and JPL continued to evolve throughout the Mars Surveyor'98 missions. It is, therefore, very difficult to apply P-Theory as a means of identifying any single external event that triggered the failures. It is important to reiterate, however, that the intention behind P-Theory is simply to establish the boundaries of an incident so that investigators can begin to delineate the more detailed flow of events that contribute to a failure.

2. *Construct event blocks.* Investigators must construct a 'block' of information about each event that leads to an incident. This information must identify the actor that is associated with each event. It must also identify the action that led to the event. Both the actor and their action must be described as precisely as possible without "qualitative adjectives, adverbs, or phrases" [72]. Finally, investigators should note the time at which the event occurred. These requirements can raise a number of practical difficulties. Previous chapters have described the reliability problems that often frustrate attempts to use advanced automated logging and tracking systems to derive precise timings for critical events in the aftermath of an incident or accident. It can also be difficult to identify the agent that is associated with the ignition of the fire onboard the Nanticoke. The most probable high-temperature sources were identified as the indicator tap that protruded from the generators cylinder head and an uncovered exhaust manifold associated with the engine's turbocharger. Neither of these inanimate objects can easily be interpreted as agents even though the ignition event is critical to an understanding of the incident. One solution is to extend the notion of 'agency' to include systems and subsystems that exhibit particular behaviours in response to environmental changes. The developers of the MES method have an even broader interpretation in which actors include inanimate objects such as tires, machines and even water [72].

3. *Construct an MES flowchart.* An MES flowchart maps each event block onto two axes. The X axis represents the actors involved in an incident. In the Nanticoke case study, the master could be listed above the engineer. The engineer, in turn, might be inserted above the mechanical assistant and so on. The Y axis denotes the passage of time during an incident. The developers of the MES approach argue that because each actor is typically involved in a number of sequential events, their actions will appear as a horizontal line of event blocks across the chart. Again, this raises a number of concerns. Firstly, human factors research has shown that operators often interleave sub-tasks [668]. Interruptions can force individuals to suspend particular actions only to resume them once the immediate situation has been addressed. Similarly, it is a routine occurrence for operators to simultaneously perform multiple control tasks. Further problems stem from the construction of the MES flowchart. The granularity of the time-line must be appropriate to the circumstances that are being considered. As we have seen for time-lines, this can cause problems for incidents that are characterised by distal events that occur many months before a large number of more proximal failures. In consequence, investigators can be forced to exploit differing time-scales over the period under consideration. Each event block is then inserted into the two-dimensional array at the position determined both by the agent responsible for the event and the time at which the event is assumed to occur.

4. *Identify Conditions* The construction of an MES flowchart provides investigators with an overview of the events leading to an incident. This, in turn, can help to identify those condi-

tions that make particular events more likely to occur. Each condition is linked to at least one event using an arrow. Each condition can itself be the outcome of other external peturbations. These events can also be introduced into an MES flowchart, providing analysts with a further means of expanding the scope of any investigation. This process helps analyst to explore the underlying conditions that might trigger future perturbations and, hence, could trigger any recurrence of an incident. Experience in applying the MES approach persuaded its developers that conditions ought to be omitted from subsequent versions of the technique. It was argued that the inclusion of conditions in the MES flowcharts is superfluous because conditions are stable until changed by some action. Investigators should, therefore, focus on analysing the events that characterise an incident. This is an important difference between the version of MES that is used in this section, where conditions are included, and the STEP methodology in the following section, where conditions are omitted.

5. *Validate the assignment of event blocks within the flowchart.* After having constructed an initial flowchart, analysts must ensure that they have a coherent model of the events leading to an incident. This involves two checks. Firstly, they must ensure that the array accurately reflects the ordering for each pair of events performed by any agent. In other words, investigators must ensure that all events to the right of any particular event occur after that event. Secondly, analysts must ask whether the preceding events are both necessary and sufficiency for any following events to occur. Additional analysis must be performed if either of these tests fails. For example, the labels that are used to identify each event can be ambiguous. In such circumstances, investigators may be forced to break them down into more detailed 'sub-events'. Alternatively, events may have been omitted during the early stages of any investigation. Additional evidence can be gathered to identify any missing event blocks.

6. *Identify causal relationships.* The second of the two validation criteria, mentioned above, can be used to identify causal relationships between event blocks. Investigators annotate the flowchart so that it is possible to identify the necessary and sufficient conditions for each event to occur. Arrows can be used to represent a causal relationship between events and conditions. It should be emphasised that this is orthogonal to the temporal relationships that are denoted along the X-axis of the MES flowchart. Once this has been done, it is important that investigators consider whether there are any alternative causal hypotheses that are not reflected by the relationships that have been denoted on the flowchart. For instance, an oil leak from the forward filter cover and the ignition source provided by the turbocharger exhaust together describe sufficient conditions for the Nanticoke fire. Each event is also necessary in order for the incident to occur. There may, however, be other causal explanations. For instance, the ignition source might have been provided by the indicator tap. Either of these hypotheses might provide the necessary conditions for the subsequent mishap. Analysts must, therefore, conduct further investigations including reconstructions and empirical studies to determine which of the hypotheses is most likely. The previous requirements of temporal coherence and causal 'sufficiency' should again be applied if the chart is revised to reflect a new hypothesis. This stage is important because it encourages analysts to consider whether there may be alternative causal complexes that might have resulted in the same consequences [508]. There are further benefits. For instance, it is possible to compare the causal model in the MES flowchart with alternative models of the intended process behaviour. This can be used not only to identify the external peturbations that lead to an incident but also the ways in which internal barriers must fail in order for an incident to progress.

7. *Identify corrective actions.* Investigators can annotate the resulting MES flowchart to denote any events or conditions that should form the focus for future interventions. These potential intervention points must be analysed to identity means of mitigating the undesired outcome or of making any peturbations less likely. Recommendations can then be made to commissioning and regulatory authorities.

Figure 11.2 illustrates the results of an initial MES analysis of the Nanticoke case study. The analysis begins by identifying an event that disturbs the previous homeostasis or equilibrium of the system.

P-Theory suggests that if the system adapts to these changes then homeostasis can be maintained. If the system fails to adapt then an accident or incident sequence begins. Initial peturbations can lead to cascading sequences of events that, in turn, place further pressures on other system components. In the Nanticoke example, modifications to the forward filter removed the seating grooves that helped to maintain a seal between the copper gasket and its securing bolt. This created problems for the watchkeeping engineer when they attempted to achieve such a seal.

Figure 11.2 illustrates further events that contributed to the engineer's problems. Copper gaskets are often deformed by the pressures that they sustain under normal operating conditions in engine filters. The engineer was, however, forced to anneal and re-use the existing component as there were no spares on-board the Nanticoke. Under normal circumstances, this need not have had serious consequences. However, the deformation of the gasket may have contributed to the engineer's difficulties in sealing the filter assembly. As can be seen, the time-line in the MES flowchart provides a reference point for th events that contributed to this incident. The fuel started to escape under pressure at some point after the Chief Engineer's inspection at 15:15. The fuel was ignited by a source on the port generator at some time after it started to spray from the filter.

Figure 11.2 illustrates some of the issues that complicate the development of MES flowcharts. For instance, the ignition event, labelled D2, is associated with the engine as a whole. This diagram could, however, be refined to represent a lower level of detail. The ignition source was either the exposed indicator tap or the exhaust manifold. These two agents could be introduced to replace the generator. Unfortunately, this creates further problems. The proponents of the MES approach argue that investigators must minimise any uncertainty over the events that contribute to incidents and accidents [72]. MES flowcharts do not have any equivalent of an OR gate in a fault tree. In consequence, it is difficult to denote that the ignition source was either the indicator or the manifold. Figure 11.2 therefore refers to the port generator rather than its specific components. Part/sub-part ambiguity is used to avoid the disjunction associated with alternative events. Ideally, such imprecision might be avoided by empirical tests and simulations. As we have seen, however, it is not always easy to obtain the resources that are required to support such investigations even in the aftermath of safety-critical incidents.

Benner's P-Theory suggests that incidents are distinguished from accidents by the manner in which the system regains equilibrium without adverse consequences. This is illustrated by the outcome event in Figure 11.2. This is linked to three other events. D2 described the ignition of the fuel source. C3 describes the escalation of the fire after the O-rings on the filter's main covers were melted. Event E3, in contrast, describes the Chief Engineer's mitigating actions in shutting down the port generator.

Figure 11.3 introduces a number of conditions into the event structure that was shown in Figure 11.2. This follows the general approach that was introduced for ECF charts. The use of events and conditions offers a number of benefits. In particular, it helps to distinguish between an event and its outcome. This is illustrated by the event A1. In Figure 11.2 this was initially used to denote modification to forward filter cover/bolt sealing surface removes groove for copper washer. This captures the event, the maintenance, as well as its outcome, the removal of the seating groove. In Figure 11.3 the event is simplified to Modifies forward filter cover/bolt sealing surface. The outcome is denoted by a condition Grooves for copper washer are removed, sealing surface is uneven and grooved with file marks. These distinctions are important for the subsequent analysis of an incident. By separating the representation of an event from its outcome, analysts are encouraged to think of alternative consequences for key events during any mishap. In this instance, it may not be possible to prevent future modifications to the sealing surface but action could be taken to ensure that the sealing surface is levelled prior to operation.

Further conditions help to explain the reasons why particular events occurred. For instance, we had to explain why the watchkeeping engineer annealed the existing copper gasket, denoted by event B1 in Figure 11.2. In contrast, Figure 11.3 introduces a condition to explain that there were no spare gaskets on board the vessel at the time of the maintenance operations. A condition is also used to explain that the deformation of a gasket, event C1, can make it difficult to obtain a good seal. Finally, Figure 11.3 introduces a condition to explain that the ignition, denoted by event D2, was possible at temperatures below the flash point for the fuel because it was being sprayed under

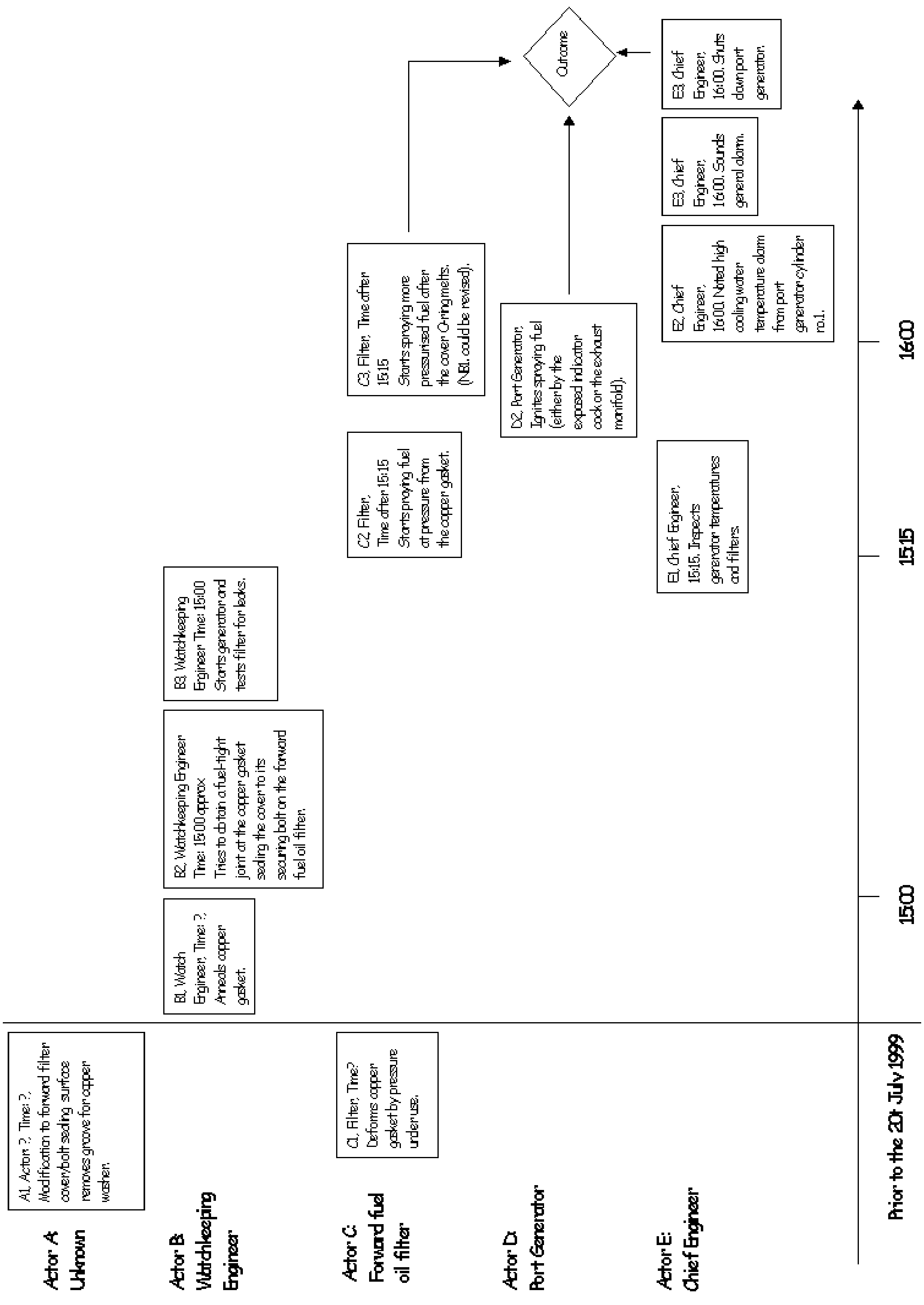

Figure 11.2: An Initial Multilinear Events Sequence (MES) Diagram

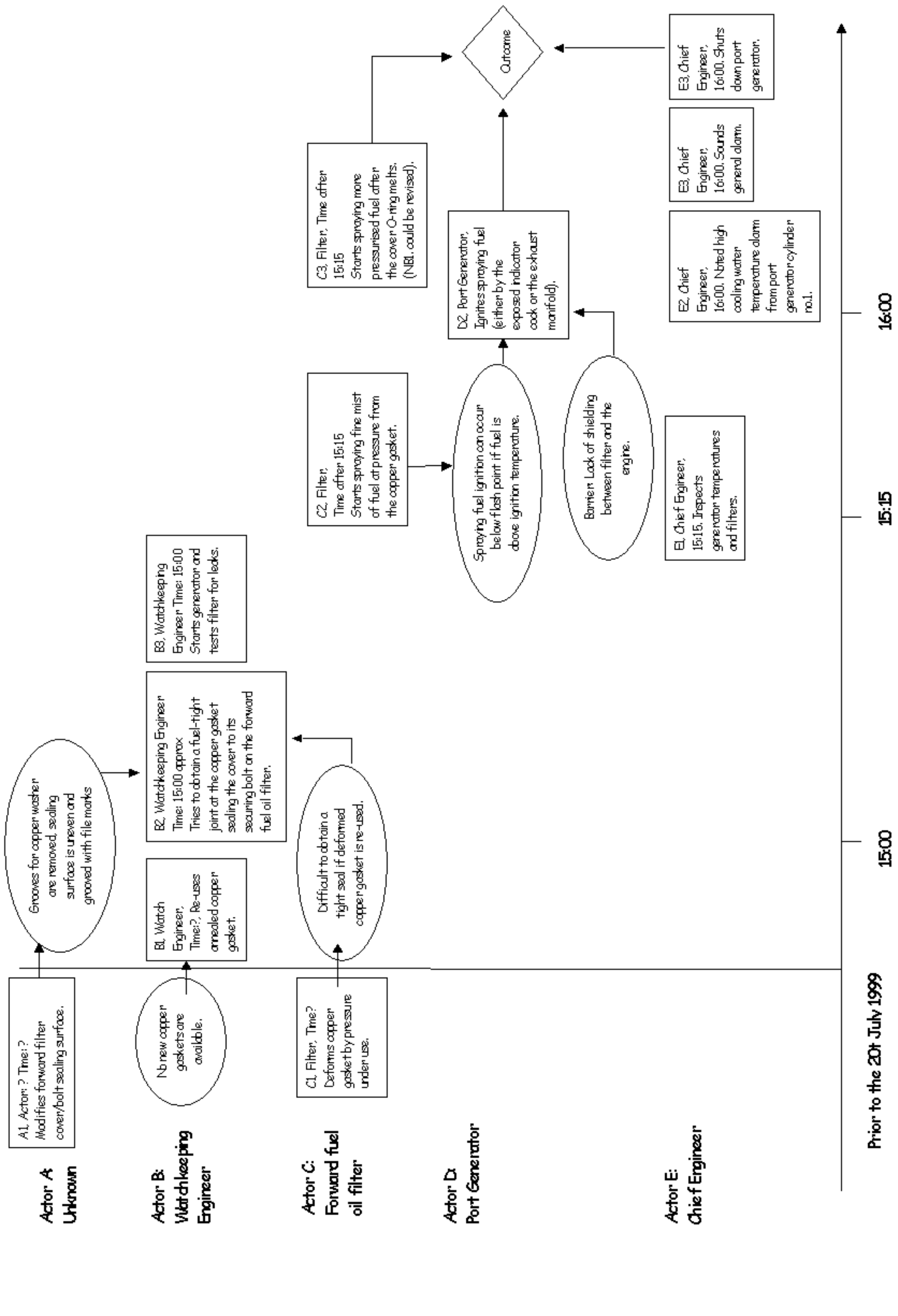

Figure 11.3: A MES Flowchart showing Conditions in the Nanticoke Case Study

pressure, denoted by event C2. These conditions do not simply to separate out information about an event and its consequences. They provide important contextual details that can help the members of a multidisciplinary investigation team to understand the significance of particular events. The importance of this should not be underestimated. Without such explicit annotations, investigators may rely upon inappropriate assumptions about their colleagues' ability to reconstruct the ways in which particular events contribute to the course of an incident.

Figure 11.3 extends the notation described in Benner's original work [72]. A condition represents the absence of a barrier; lack of shielding between the filter and the engine. The initial MES notation makes it difficult for analysts to represent both the absence of barriers and errors of omission. This is entirely deliberate. It can be argued that investigators must focus on what *did* happen during an incident rather than what *might* have happened. By drawing other investigators' attention to the absence of particular protection measures, analysts can potentially obscure information about the performance of those barriers that were available. These objections also argue against our previous use of barrier analysis to drive ECF modelling in Chapter 8.3. A number of arguments support our use of conditions to represent the lack of shielding in Figure 11.3. Firstly, there is no empirical evidence to support either position in this argument. Until such evidence is obtained it is difficult to determine whether or not the introduction of information about missing barriers will bias an investigator's analysis of an incident. Secondly, even if information about errors of omission and absent barriers are excluded from incident models, these details must be explicitly considered during any subsequent analysis.

Figure 11.4 illustrates the results of introducing causal information into Figure 11.3. As mentioned above, this involves a variant of the counterfactual reasoning introduced in previous chapters. Starting with the earliest event or condition on the time-line, analysts must ask whether the next event or condition in time would have happened if this earliest event had not occurred. If the answer is no then they form a causal pair and an arrow can be drawn from the leftmost event or condition to the related event or condition. If the answer is yes then the earliest event or condition is not a necessary cause of the subsequent event or condition. No link is drawn. The analysis continues until the investigators has asked whether all of the subsequent events or conditions were potential causes by the initial event or condition. The entire process is then repeated for each subsequent event or condition in the MES flowchart. In practice, however, a number of 'optimisations' are often made. For instance, transitive links are omitted. If event or condition A causes event of condition B, which in turn, causes event or condition C then arrows need only be drawn between A and B and between B and C. The causal arrow between A and C is implied.

Causal analysis can help investigators to identify potential revisions to an existing MES flowchart. For example, Figure 11.4 introduces an event labelled fuel tight joint at the copper gasket sealing the cover to its securing bolt on the forward fuel oil filter fails. This proved necessary in order to link the previous observations about the watchkeeping engineers difficulty in obtaining a seal to the later events that described the course of the fire. As can be seen, a question mark is used to denote a degree of uncertainty in this causal link. Without it, however, there would have been no explicit means of representing that the maintenance task was a potential cause of the incident.

Figure 11.4 also illustrates the way in which a causal analysis can help to identify events that are otherwise isolated from the causal 'flow' that leads to an incident. In this case, there is an event which denotes that the Chief Engineer inspects generator temperatures and filters at 15:15. This event is important for our understanding of the incident because it helps us to determine that the fire did not take hold before that moment in time. It does not, however, play a direct role in the incident. The proponents of MES analysis, therefore, argue that it ought to be omitted from future diagrams. It is important not to underestimate the pragmatic benefits of such guidelines. It is very rare to find that any modelling or causal analysis technique provides advice about when *not* to introduce additional information that might obscure or otherwise hinder subsequent investigations.

A number of limitations can be identified with the MES techniques described in this section. As mentioned, the developers found that investigators used conditions in an arbitrary and ad hoc manner. Previous sections have argued that this is an important strength of the ECF approach. Investigators can use conditions to denote broad insights into the context in which an incident occurs. In contrast, Benner views this as a dangerous abuse because conditions can introduce superfluous

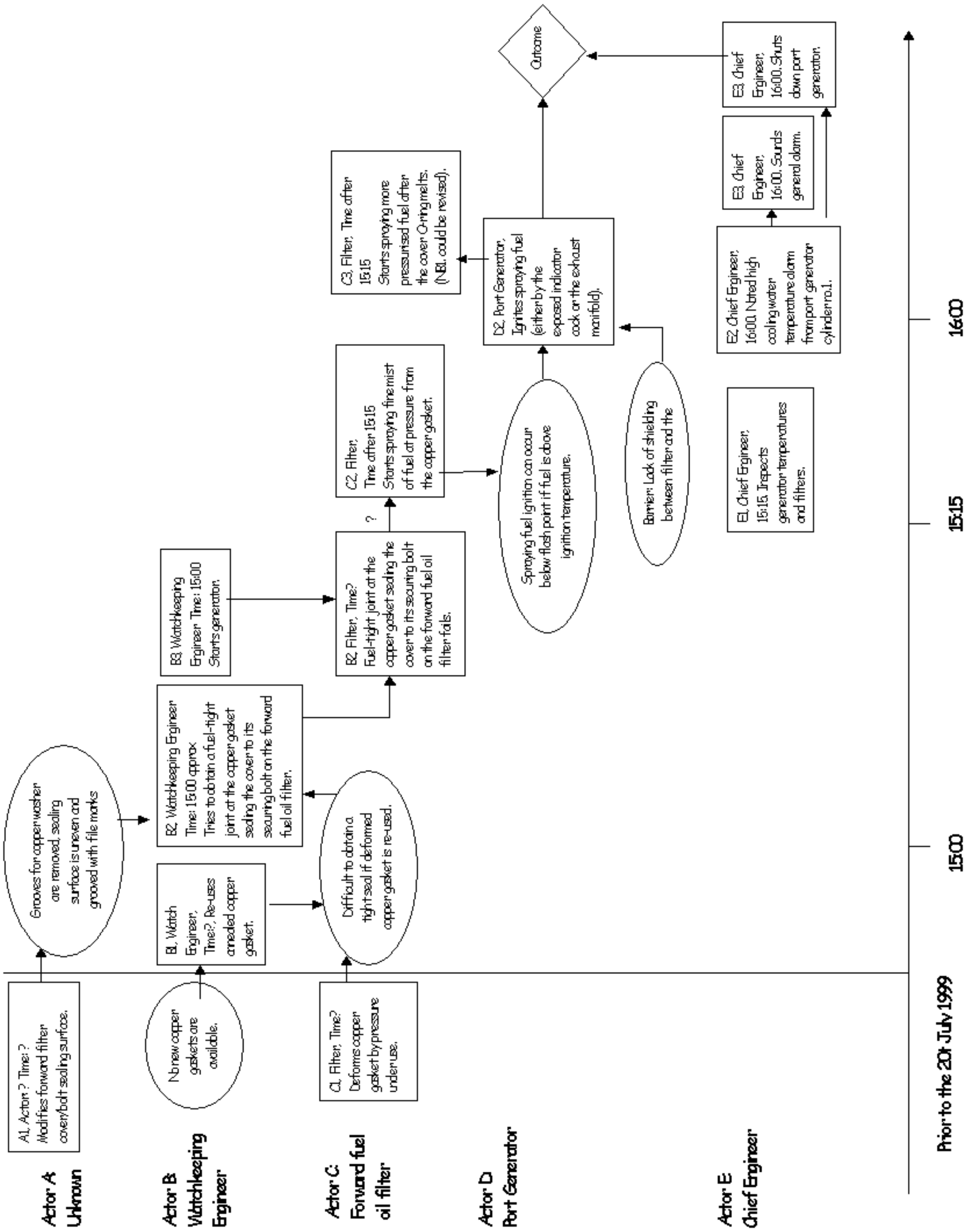

Figure 11.4: A MES Flowchart showing Causation in the Nanticoke Case Study

information that might otherwise be represented more directly by the events that stem from those conditions. He also argues conditions are often used to represent unsubstantiated factors that are difficult to validate after an incident has occurred. Others have argued that the MES approach is limited by the perceived complexity in developing and analysing the flowcharts [553]. As mentioned above, it can be difficult to identify a stable state for many complex technological systems. This, in turn, frustrates attempts to apply the P-Theory that drives MES analysis.

11.1.2 Sequentially Timed and Events Plotting (STEP)

The concerns mentioned at the end of the previous section led Hendrick and Benner to revise the MES approach [348]. Sequentially Timed and Events Plotting (STEP) provides a synthesis of ECF charting and MES [553]. It begins with the compilation of STEP cards. These provide an initial means of recording information about key events that occur during the course of an incident. They can be completed during any stage of a primary or secondary investigation. This reflects the concern that STEP should provide a pragmatic tool for investigators. It, therefore, attempts to avoid some of the notational excesses of the other analytical techniques that we have presented in previous chapters.

Event card identifier:	
Actor:	
Action:	
Time event began:	
Event duration:	
Data source/evidence:	
Event location:	
Description:	

Table 11.1: STEP card used to consolidate event information [348]

Table 11.1 illustrates the format of a STEP card. As can be seen, the information on these cards is closely modelled on the event blocks that support MES analysis. STEP cards do, however, record a number of additional items of information. In addition to the actor, time and action information that is captured by MES, STEP cards also introduce a free-text description of the event. They include information about the event location and its duration. Finally, STEP cards also record a Source identifier. This can be used to refer to the evidence that helped to identify the event. Such information can be useful when considering whether or not to support particular hypotheses about the course of events. The evidence that supports an event can be used to determine whether or not it should be retained in the face of competing explanations about the course of an incident.

Event information again provides the building blocks that are used to reconstruct the course of an incident. STEP relies upon a tabular format rather than the MES flowchart. The abscissa or vertical scale denotes the passage of time during an incident. The beginning and end of the accident sequence are, therefore, represented by the first and last columns in the matrix. Actors are represented on the ordinate, or horizontal, scale in the matrix rather than along the Y-axis in a MES chart. This tabular format offers a number of potential benefits during the initial stages of an investigation. Spreadsheets can be used to reduce the burdens associated with inserting new events and actors into an existing matrix. This might appear to be a trivial issue. As we have seen, however, the overheads involved in constructing graphical diagrams that involve many hundreds of nodes can dissuade investigators from using many of the more 'advanced' techniques that have been proposed to support incident analysis.

As mentioned before, STEP matrices do not include conditions. These were initially included to explain why an event occurred. Experience suggested, however, that investigators used conditions to introduce a range of biases into MES flowcharts. For instance, conditions were used to represent contextual factors that might not have had a direct impact upon the course of an incident. They can also be used to modify events so that they seem to be less significant that they might otherwise appear. The decision to exclude conditions from the STEP methodology was also justified by the observation that conditions are, typically, the result of previous actions. It can, therefore, be argued that they are superfluous to any subsequent investigation. Previous chapters have argued that conditions provide an important means of introducing some of the broader contextual factors that affect the course of many incidents. The decision to omit them from STEP matrices is, therefore, open to debate. It remains a continuing focus for on-going research into accident and investigation analysis. However, the following pages adopt the conventions introduced by the original STEP papers. Conditions are omited from the tabular representations of event sequences.

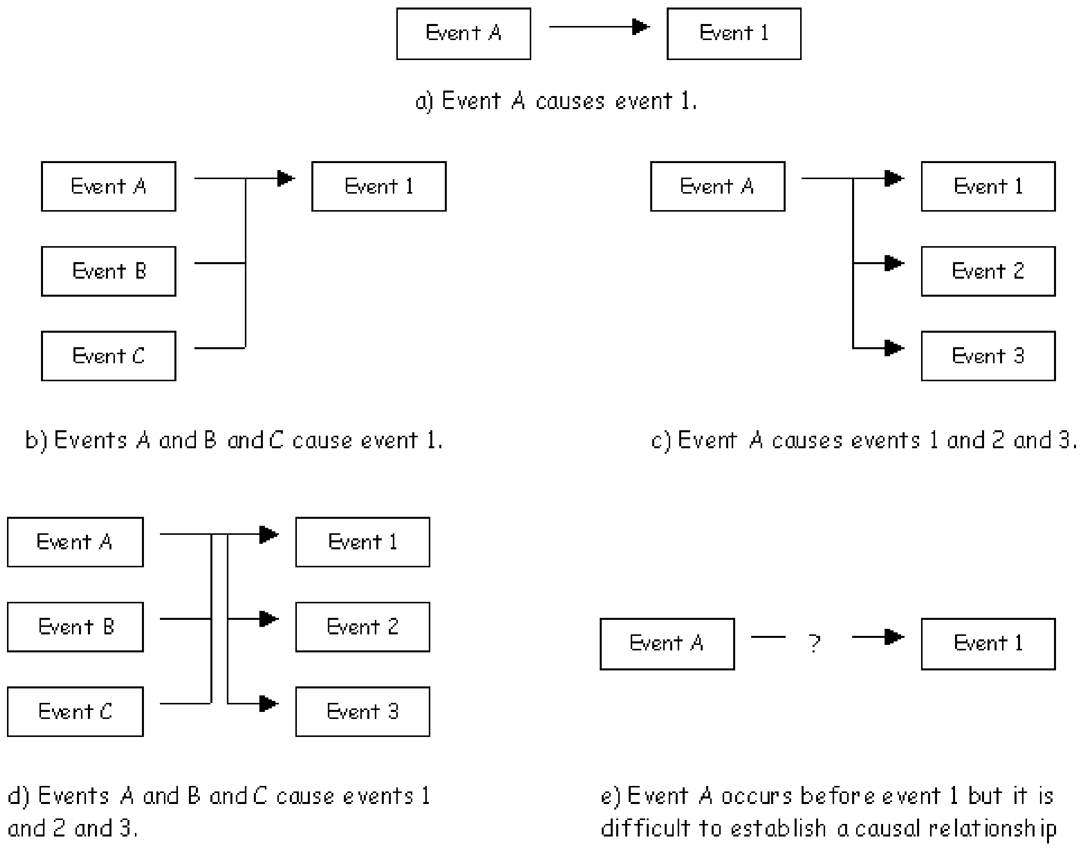

Figure 11.5: Causal Relationships in STEP Matrices

The construction of a STEP matrix follows the P-Theory process outlined for MES analysis. The same consistency checks are also performed to ensure that the resulting worksheet provides a coherent temporal ordering over the events that it presents. The causal analysis of a STEP matrix also follows a procedure that is similar to that described for the MES flowchart. More recent expositions of STEP [74, 75] enumerate a broader range of causal relationships than appeared in the initial MES papers [72]. These are illustrated in Figure 11.5. Diagram a) denotes that event A is a direct cause of event 1. In other words, A is both a necessary and sufficient cause of 1. Diagram b) is similar to an AND gate within a fault tree. Events A, B and C are individually necessary for event 1 to occur. However, none of these events are sufficient for event 1 to occur unless A and B and C all occur at the times denoted by the abscissa. Diagram c) denotes a situation in which event A causes events 1, 2 and 3. This is important if the outcome of an event has an impact upon many other actors throughout a system. Events 1, 2 and 3 might represent these distributed, knock-on consequences. Diagram d) combines elements of diagrams b) and c) to denote that A, B and C are individually necessary and collectively sufficient for 1, 2 and 3 to occur. Finally, diagram e) denotes events that have a clear relationship in time but for which no causal explanation can be established. Such ambiguities form the focus for subsequent investigation of the underlying physical processes that characterise complex applications.

It is important to note that although diagram b) can be thought of as an AND gate, there is no equivalent of an OR gate within STEP matrices. If it is unclear what caused an event then investigators must introduce an event block that is labelled by a question mark. This is intended to avoid indicating "uncertainty about what happened" which is argued to be a weakness of the OR gate approach [75]. Whereas the use of events labelled by a question mark indicates "uncertainties in the description" [75]. It is difficult to interpret such distinctions. There are also pragmatic difficulties. Previous chapters have identified the limitations of current data recording devices. We have also described the problems associated with determining the causes of failure in hostile environments, such as space, where telemetry is strictly limited. This chapter does, however, follow the STEP conventions [75] . Disjunction are avoided.

Event card identifier: A1	
Actor:	?
Action:	Modifies forward filter cover/bolt sealing surface
Time event began:	Prior to 20th July 1999
Event duration:	?
Evidence:	Post incident inspection shows file marks are present on the cover/bolt sealing surface which was flat with no recess, unlike aft filter.
Event location:	Nanticoke forward fuel filter.
Description:	The copper washer gasket grooves are removed and this makes the sealing surface uneven.

Table 11.2: STEP card for the Nanticoke Filter Modification

Having introduced the underlying components of STEP, the following paragraphs apply this technique to analyse the Nanticoke case study. Matrices 11.2 and 11.3 present STEP cards that document information about key events. Investigators are intended to use these cards to help document the investigation progresses. Given the constraints of this case study, these cards were completed post hoc. They do, however, provide an illustration of the range of information that can be captured using these documents. For example, previous sections have explained the reasons why conditions are excluded from STEP matrices. This information can, however, be retained within the STEP card descriptions of key events. The condition labelled grooves for copper washer are removed, sealing surface is uneven and grooved with file marks in Figure 11.4 now forms part of the free-text description in Table 11.2.

STEP and MES are unusual in that they have been specifically intended to help investigators conduct a causal analysis during secondary, and even primary investigations. Other techniques, including ECF analysis and the application of Fault Trees, are far less explicit about when any causal analysis should begin. Many of the publications that propose the application of these approaches seem to make an implicit assumption that investigators have already secured any relevant information. We have argued in previous chapters that this is unrealistic. The identification of a potential causal factor can often lead to further investigation. For instance, if there is only circumstantial evidence that an event actually occurred. There is, therefore, a great deal to be learned from the comparatively simple documentary support offered by STEP cards. They avoid many of the maintenance overheads that are associated with the revision of more complex graphical and text-based analyses when new evidence becomes available.

The example STEP cards in Tables 11.2 and 11.3 illustrate further differences between this

Event card identifier: C2	
Actor:	Forward Fuel Filter
Action:	Starts spraying fine mist of fuel at pressure from the copper gasket.
Time event began:	After Chief Engineer's inspection at 15:15.
Event duration:	Until port generator shut-down at 16:00
Evidence:	When fire burns at high intensity, soot deposited on nearby surfaces is burnt off leaving a 'clean burn'. This is present slightly inboard of port generator valve covers 1 and 2; the general location of the fuel filters. Inspection of lubricating oil under the valve covers and two other starboard upper fuel filters rules out these sources.
Event location:	Nanticoke forward fuel filter.
Description:	If fuel was released under pressure from the copper gasket of the forward fuel filter then ignition could occur below the flash point of the fuel.

Table 11.3: STEP card for the Nanticoke Fuel Release

approach and the techniques that have been introduced in previous chapters. In particular, both include information about the evidence that supports the identification of particular events. The impact of modifications to the filter cover, event A1, is supported by a post incident inspection, which shows file marks are present on the cover/bolt sealing surface which was flat with no recess unlike the aft filter. The escape of fuel under pressure from the forward fuel filter, event C2, is supported by a more complex line of reasoning. When fire burns at high intensity, any soot that is deposited on nearby surfaces is burnt off leaving an area of 'clean burn'. Post incident inspections detected an area of clean burn slightly inboard of the port generator valve covers 1 and 2. This was in the general location of the fuel filters. These inspections also eliminated the possibility of the fire being fueled from three alternative sources. The importance of explicitly documenting such evidence should not be underestimated. The STEP approach encourages analysts to construct a single, 'deterministic' failure scenario. Disjunctions are not allowed when constructing STEP matrices from cards, such as those shown in Tables 11.2 and 11.3. Ambiguities are to be avoided as much as possible. Investigators must justify their analysis if their colleagues are to understand the evidence that supports the particular version of events that, in turn, supports any causal findings.

Figure 11.6 shows how a STEP matrix can be constructed to represent the causal relationships that exist between the various events that are described on STEP cards, such as those shown in Tables 11.2 and 11.3. As can be seen, there are strong similarities between this matrix and the MES flowcharts that were introduced in previous sections. However, there are no conditions. Some causal links have to be re-drawn because conditions are excluded from this form of analysis. For instance, in Figure 11.4 the modification event A1 led to a situation in which the sealing surface was uneven. This condition, in turn, affected the Watchkeeping Engineer's ability to obtain a fuel tight joint. In contrast, Figure 11.6 omits the condition. The modification event A1 might therefore have been shown as a direct causal link to event B2, which represents the Watchkeeping Engineer's attempts to obtain the fuel-tight seal. In contrast, Figure 11.6 shows that the modification event causes the fuel escape. This might seem like a subtle distinction but it reflects important differences between the MES and STEP techniques. In the former case, the initial event caused a condition that affected the Engineer's actions. Hence a causal link could be drawn from A1 to B2 through the mediating condition. In the STEP matrix, it cannot be argued that the modification event directly caused the Engineer to attempt to form a fuel-tight seal. Hence the modification event A1 and the Engineer's efforts, B2 contribute to the fuel release, event C2. The proponents of STEP argue that this clarifies

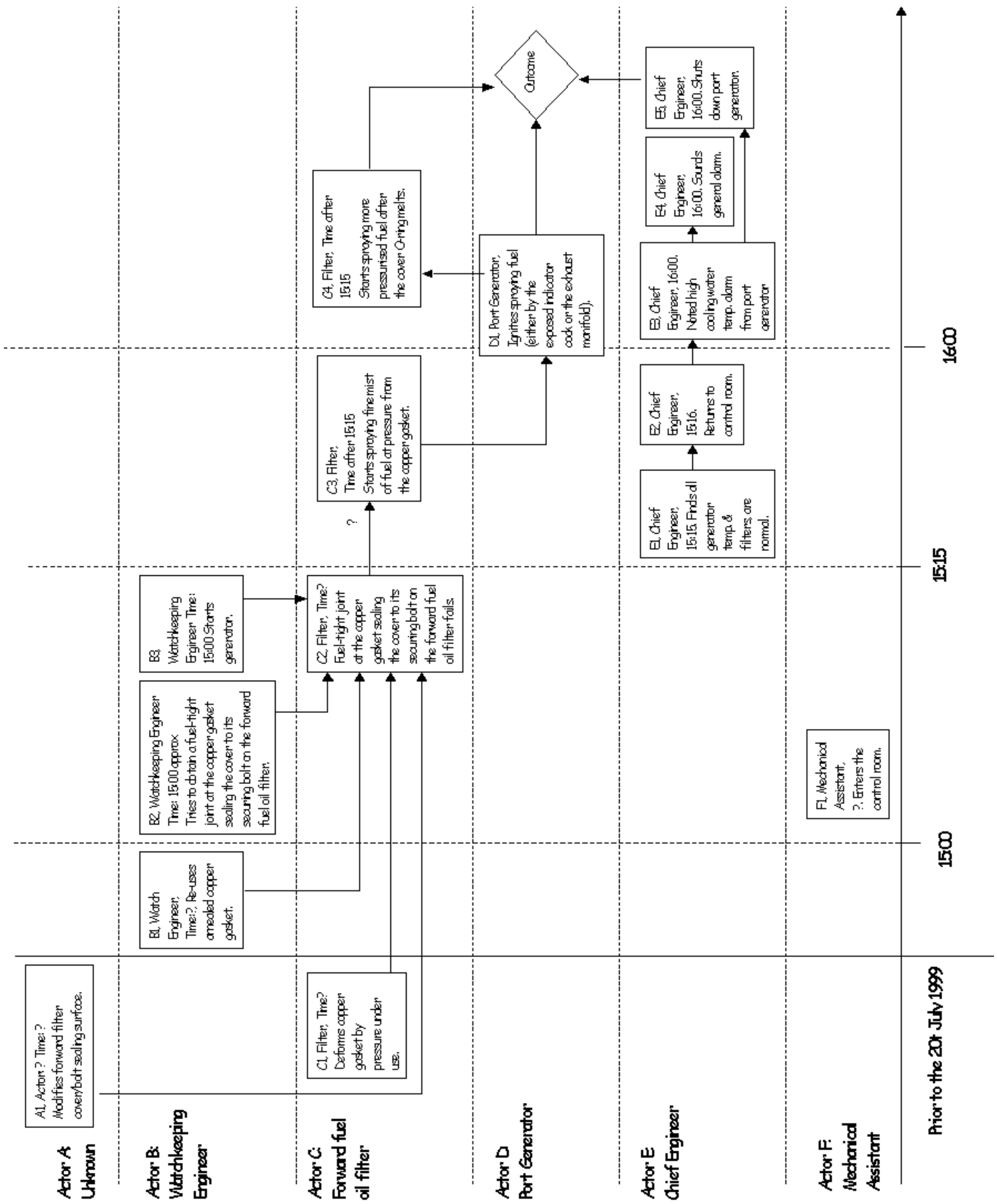

Figure 11.6: STEP Matrix for the Nanticoke Case Study

the causal relationships between events. The condition into the MES diagram introduces a form of indirection between the modification event and the eventual fuel release that is not apparent in the STEP matrix of Figure 11.6. This practical standpoint is support by the more philosophical work of Lipton who argues that only events can be causes [496].

Figure 11.6 extends the previous MES analysis by considering a number of additional causal factors. In particular, the role of the Chief Engineer and the Mechanical Assistant are considered in greater detail. Events are introduced to denote that the Mechanical Assistant Enters the control room and that the Chief Engineer returns to the control room. These are used to explain why the fire was not detected until 16:00. This again illustrates how the application of causal analysis techniques continues to depend on the skill and expertise of the investigator. There is no automatic means of determining that these additional events ought to be introduced into a MES flowchart or STEP matrix. Table 11.4 illustrates how such insights may force investigators to develop additional STEP cards to represent information about a wider range of events. In this case, the card is used to record details about the Chief Engineer's return to the control room after his inspection at 15:15.

Event card identifier: E2	
Actor:	Chief Engineer
Action:	Returns to control room.
Time event began:	Approximately 15:16.
Event duration: Until high cooling water temperature alarm around 16:00.	
Evidence:	Witness evidence (Watchkeeping Engineer, Mechanical Assistant and Chief Engineer).
Event location: Engine control room.	
Description:	The Chief Engineer returned to the control room after observing that the generators and filters appeared to be functioning normally. The significance of this event is that they could not observe the port-side of the engine room from the control room. Neither the chief engineer nor the mechanical assistant made a visual inspection between 15:15 and 16:00 and this gave the fire an opportunity to take hold.

Table 11.4: STEP card for the Chief Engineer's Monitoring Activities

Table 11.4 also illustrates a number of problems that complicate the application of the STEP approach. Firstly, the card explains the significance of the Chief Engineer's decision to return to the control room. He could not observe the port-side of the engine room from the control room and, therefore, was unlikely to directly observe the fire until it had taken hold. This information is not included on the STEP matrix in Figure 11.6. This introduces cross-referencing problems that affect the use of multiple representations for the same events. Investigators must not only understand the causal relationships represented on the matrix but they must also follow the more detailed information that is represented on each of the cards. This might seem to be a relatively trivial demand. It can, however, impose significant burdens when STEP matrices are used to represent complex, safety-critical incidents involving several hundred events.

There are further problems. The STEP card in Table 11.4 records that neither the chief engineer nor the mechanical assistant made a visual inspection between 15:15 and 16:00 and this gave the fire an opportunity to take hold. The previous STEP matrix does not document this temporal information. One solution would be to introduce an additional field into a STEP card. This would distinguish the duration of an event and from the duration of its effects. For example, the Chief Engineer only took a few seconds to enter the control room but they remained there until 16:00. Such additions to the STEP card introduce further problems. Events often trigger a number of different effects. The ignition event created heat and smoke, it also eventually triggered temperature alarms. Each of these effects have different durations. The smoke and heat were eventually countered by ventilating the engine room after Halon gas had been used to extinguish the fire. The alarms continued until the ship had been secured. The tractability of the STEP cards approach would clearly be sacrificed if investigators had to introduce this duration information into the more concise summaries of key events.

These overheads can be avoided by explicitly introducing stopping events into a STEP matrix. The continued presence of the Mechanical Assistant and the Chief Engineer in the control room can, therefore, be inferred by the absence of any event to denote that they left the control room. Such inferences carry a degree of uncertainty. Investigators may forget to introduce these terminating events. Figure 11.6 uses event C4 to denote that pressurised fuel begins to spray at an increased rate when the filter cover O-rings melt after 15:15. We have not, however, specified when this fuel release ended. In constructing the STEP matrix it was assumed that investigators would recognise that the release ended when the engineer shut-down the port engine at 16:00. Unless explicit stop-events are introduced, there is a danger that investigators may rely upon incorrect inferences about the duration of key properties during an incident. These problems should not be surprising. The difficulty of representing events and duration also affected the time-lines introduced in Chapter 8.3 and the ordinate scale of the STEP matrix can be viewed as a time-line.

Previous sections have explained how both MES and STEP derive directed graphs of an incident. Nodes represent events in STEP, or events and conditions in MES. Edges represent the causal relations that hold between nodes in the graph. We have not, however, described how investigators can identify root causes from the various causal factors that are used to construct these graphs. One solution would be to replicate the analytical techniques that were introduced a the end of Chapter 9.3. In addition to the validation steps, which ensure that causal factors are both necessary and sufficient, analysts must distinguish those events that represent more general (root) causes from those that characterise a particular incident. It is important to note, however, that the developers of the STEP and MES techniques have been highly critical of previous attempts to derive methods for root cause analysis. Benner, in particular, has argued that attempt to distinguish root causes from causal factors can misdirect investigators to find a few 'silver bullets' instead of understanding, describing and explaining the entire incident process [76]. He goes on to argue that root causes are often 'judgemental, unverifiable conclusions' that typically cannot be validated by 'objective quality controls'. These comments are consistent with the STEP focus on determining the particular events that contributed to an incident. Conditions that might represent wider causal factors are deliberately excluded from this approach. In contrast, STEP focuses on the evidence that supports the introduction of particular events into the associated matrices.

P-Theory suggests that investigators must focus on the initial perturbation that causes any subsequent failure. Figure 11.6 starts with the initial modifications to the forward filter cover/bolt sealing surface. P-Theory also suggests that investigators should consider causal events that compromise protective barriers. For example, the Watchkeeping Engineer might have reported the problems experienced in fitting the gasket. These events represent missed opportunities for the system to return to an initial 'homeostasis'. This focus on the particular causes of an incident provides important benefits. It is intended to reduce the likelihood that external pressures will 'persuade' investigators to introduce arbitrary contextual factors, or conditions, as a means of explaining particular events [74]. There is, however, a danger that the application of MES and STEP will miss important underlying causes of an incident. For example, previous sections have argued that organisational and managerial failures can jeopardise a number of different barriers. These failures can be analysed and measured, for instance in terms of participation rates in incident reporting schemes or in the number of regulatory sanctions that were previously applied to an organisation. It is unclear how such factors might be represented as causal events within a STEP analysis.

A number of further problems complicate the application of STEP [74, 75]. These include limitations that affect this particular approach. They also include more general issues that affect all forms of causal analysis:

1. *Incomplete chains between the first and last events.* If it is not possible to establish a path through the causal connections in the matrix then investigators must seek additional evidence about events that might not have been identified. This can involve the use of additional techniques, such as Fault Trees or the Change analysis and Barrier Analysis methods that were used in conjunction with ECF charts. Alternatively, the scope of any report might be confined to those events that can be accurately identified from the available evidence. This clearly jeopardises the insights that might have been obtained from any analysis of the incident. Investigators must, typically, take steps to increase the amount of 'diagnostic' information that can be obtained from any potential future incidents.

2. *Unconnected events after the causal analysis.* The developers of the STEP method argue that investigators must avoid unconnected events. For example, events F1 in Figure 11.6 denotes that the Mechanical Assistant enters the control room. It does not, however, have any direct causal relationship with the subsequent events in the Nanticoke case study. It has been argued that, at best, these unconnected events can divert investigator's attention away from more important causal sequences. Scare development resources can be allocated to deal with these extraneous peturbations that need not have affected the course of an incident. At worst, it is argued that they provide "handles for others to grasp to raise irrelevant, unnecessary and invalid questions about the accident" [75]. It is argued that investigators should delete these unconnected events from a STEP matrix because they can mislead rather than enlighten other investigators. The dangers with such a policy are clear. Investigators run the risk of deleting information that might enable their colleagues or other readers to identify important causal relations that might have been overlooked in any previous analysis. If analysts follow this advice then there ought to be some documentary evidence to record their decision so that others can follow the justification for removing information from the matrix.

3. *Inconsistent data requirements.* The increasing inter-connection and functional sophistication of safety-critical systems poses considerable challenges for incident analysis. This complexity has been exacerbated by the increasing recognition that more and more factors ought to be considered during any investigation. The scope of any analysis has broadened beyond individual operator error and component failure to examine more systemic causes of incidents and accidents. It is not surprising, therefore, that analytical techniques such as STEP should yield complex accounts of the mishaps they represent. This can lead to conflict if managers expect 'simple' descriptions of complex failures. Further problems can arise if the products of a STEP analysis do not correspond to the categories expected by a regulators reporting system. As Benner notes, "this very frequent problem often arises after statisticians design forms for data collection, then declare that the statistical elements on the forms are significant investigative data and train investigators to 'fill out the form' rather than investigate the accident" [75]. Later sections will assess these problems in greater detail. For now it is sufficient to recognise that they stem from the wider organisational and regulatory environment that surrounds an incident reporting system. Investigators must clearly be aware of such issues before attempting to pioneer the introduction of new analytical techniques.

This section has identified that changes that have been introduced between earlier version of the MES analysis technique and the more recent STEP approach. MES and STEP can be seen as variants of the same underlying ideas. Both rely upon the notion of event blocks that are associated with actors and can be mapped onto a time-line. These similarities should not be surprising given that STEP extends Benner's earlier work on MES [72]. Some confusion has arisen because these two different terms have been used synonymously. Investigators have referred to MES when applying the tabular forms associated with the techniques in the STEP handbook [348]. We have attempted to make a clear distinction between these techniques, however, readers should be aware of the potential confusion given these strong similarities.

11.2 Check-List Approaches

Previous sections have focussed on event-based techniques that encourage analysts to reconstruct or model the development of an incident over time. A number of alternative techniques have, however, rejected this approach. In contrast, they often assume that analysts develop and maintain more implicit models of the events that contribute to an incident. This arguably reflects a more pragmatic attitude to the partial nature of the evidence that is available in the aftermath of many mishaps. These approaches instead provide checklists that prompt investigators to look for a number of predefined features that are common to a wide range of incidents and accidents. The US National Patient Safety Foundation's (NPSF) report on the 'Scientific Basis for Progress on Patient safety' summarised the strengths and weaknesses of these approaches:

> "Collections of incidents and accidents cry out for classification. The apparent similarities and differences between the events, their outcomes, and the circumstances that precede them encourage us to organise them in categories and rank them in severity. But classification also has its own hazards, especially in complex domains where there are multiple possible paths to any outcome and multiple possible outcomes from any path. Classification involves identifying relevant similarities and differences; their effective use depends on being able to know a priori what relevant means... Classification does involve a type of analysis but a type that greatly constrains the insights that can be obtained from the data. Typically, when classification systems are used as the analysis, a report of an incident is assigned, through a procedure or set of criteria, into one or another fixed category. The category set is thought to capture or exhaust all of the relevant aspects of failures. Once the report is classified the narrative is lost or downplayed. Instead, tabulations are built up and put into statistical comparisons. Put simply, once assigned to a single category, one event is precisely, and indistinguishably like all the others in that category." [182]

The following paragraphs use a number of different causal analysis techniques to illustrate and expand on these observations. In contrast, Chapter 14.5 describes how checklist approaches to causal classification can also be used as the indices in information retrieval systems.

11.2.1 Management Oversight and Risk Tree (MORT)

Figure 11.7 illustrates the Management Oversight and Risk Tree. This is the central component of what has become known as MORT [429]. As can be seen, MORT diagram is constructed using the elements of a fault tree. An undesired event can be either the result of oversights and omissions or it is the result of an assumed risk. Assumed risks "are defined as only those risks that have been analysed and accepted by the proper level of management; unanalysed or unknown risks are not considered to be Assumed Risks" [204]. If an oversight or omission has occurred then it can be categorised as being the result of either a management failure or of a failure in specific technical controls. If there has been a break-down in management then either there was a failure in the implementation of some policy or the policy was flawed or the risk assessment was less than adequate. A failure in management risk assessment can occur if incorrect goals were established for a project or the information systems used to support a risk assessment were less than adequate or the hazard analysis process was flawed or the safety review program was less than adequate. As can be seen, the components of the MORT diagram provide a check-list that can be used to analyse and categorise the potential causes of an incident.

The MORT diagram was intended to provide a template that might guide the causal analysis of incidents and accidents. There is an obvious danger that investigators will force an incident to fit one or more of the categories in the MORT checklist. The proponents of this approach have responded by extending the range of factors that are included in the MORT diagram. For instance, the version of Figure 11.7 includes over 1,500 basic events. This leads to a difficult trade-off. By extending the scope of the MORT diagram, investigators are more likely to find an appropriate causal factors that describes their incident. By extending the scope of the MORT diagram, investigators may

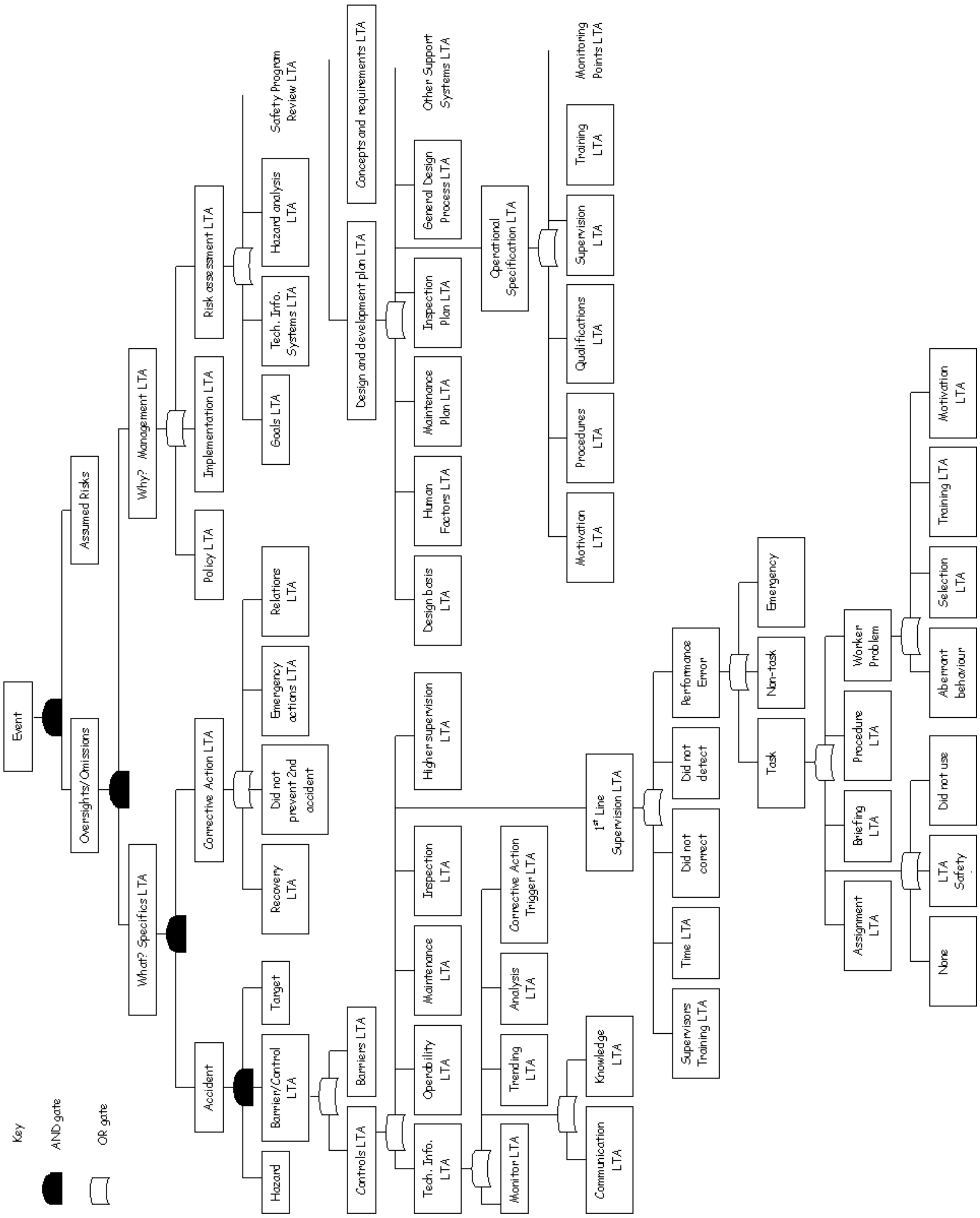

Figure 11.7: The Mini-MORT Diagram

also experience more difficulty in distinguishing between the many different forms of failure that are described by each of the leaf nodes. In consequence, the US Department of Energy advocates the use of a stripped-down version of the full MORT diagram [205]. This mini-MORT provides approximately fifty basic events but each of these denotes a far broader set of causal factors than the more detailed versions of the diagram.

MORT diagrams embody their developers' view of accident causation. The branches of the tree reflect a concern to assess management responsibility. There is also provision for assessing the technical context in which an incident occurs. Human factors issues are also captured, arguably in a rather narrow fashion, by focusing on errors of commission . A further branch traces the failure of barriers. As a result, the barrier analysis introduced in Chapter 9.3 is often used as a precursor to MORT classification. There is also a preoccupation with understanding and assessing the causes of any potential energy release [300]. One consequence of this is that MORT also provides an implicit definition of incidents and accidents. An accident occurs if an unwanted energy flow affects a vulnerable target. An incident occurs if an unwanted energy flow occurs without hitting such a target [457]. This is consistent with the use of barrier analysis and reflects their common origin within the nuclear industry. Johnson developed most of the MORT approach while working for the US National Safety Council and under a contract from the US Atomic Energy Commission [429]. As mentioned, the US Department of Energy continues to advocate this method [205, 204]. The MORT approach, therefore, combines concepts from management and from safety analysis. It captures the notion that management has a profound impact upon the effectiveness of barriers that prevent unplanned energy releases.

MORT analysis consists of two principle stages. Firstly, analysts must consider what happened during an incident. This involves a traversal of the what? sub-tree under the oversights and omissions branch. This is intended to help the analyst identify the barrier or control problems that contributed to the incident. Secondly, the analyst must then identify any management elements on the why branch of the MORT diagram that contributed to these particular problems. It is important to document each of the problems and summarise the findings of the analysis.

This process of iteratively describing what happened and then searching for causal explanations in the why branch is guided by a number of questions that analysts can ask as they inspect each node in the MORT diagram. For example, the US Department of Energy MORT user guide provides the following question that can be asked to determine whether or not any emergency response was adequate. This corresponds to the leaf node with the following pathEvent: Oversights and Omissions: What? : Corrective Actions: Emergency Actions:

> "*Emergency Action (Fire Fighting, Etc.) Less Than Adequate* Was the emergency response prompt and adequate? Which emergency response teams were required? Were they notified and did they respond? [Include local facility fire brigade, health physics team, fire department, bomb squad, and other speciality teams. Be sure to consider delays or problems in both notification and response.] " [204]

These questions appear, at first sight, to be relatively straightforward. Unfortunately, a number of factors complicate this analysis. The use of the term 'less than adequate' implies a value judgement. There can often be considerable disagreement about what does, and what does not, represent an adequate response. Even in countries that publish national guidelines for response times, there can be considerable debate about whether the nature of any response was appropriate given the scale of an incident [218, 211]. Some investigators, including Benner [72], argue that these value judgements are open to political pressure and bias in the aftermath of safety-critical incident.

Even with the additional complications created by the validation of value judgements, the previous question is relatively simple in contrast to some of the other guidelines that are intended to support MORT analysis. This point is illustrated by the following questions. These are intended to guide the analysis of a supervisor's failure to correct a hazard. Each of these questions relates to further basic events that are present in more complete versions of the MORT diagram. They would be shown under Event: Oversights and Omissions: What? : Accident : Barrier/Controls/ Controls/1st Line Supervisor/ Did Not Correct Hazards in Figure 11.7:

> *Did Not Correct Hazards:* Was an effort made to correct the detected hazard?

- Interdepartment Coordination Less Than Adequate: If the accident/incident involved two or more departments, was there sufficient and unambiguous coordination of interdepartment activities? [Interdepartment coordination is a key responsibility of the first line supervisor. It should not be left to work level personnel.]
- Delayed: Was the decision to delay correction of the hazard assumed by the supervisor on behalf of management? Was the level of risk one the supervisor had authority to assume? Was there precedent for the supervisor assuming this level of risk (as then understood by him)? [Note a decision to delay correction of the hazard may or may not transfer to the Assumed Risk branch. It was an assumed risk only if it was a specific named event, analysed, calculated where possible, evaluated, and subsequently accepted by the supervisor who was properly exercising management-delegated, decision-making authority.]
- Was the decision to delay hazard correction made on the basis of limited authority to stop the process?" [204]

The previous two examples have illustrated the questions that can be used to guide the analysis of the what sub-branch in a MORT analysis. Previous paragraphs have, however, argued that investigators must also identify the reasons why these events occurred. This involves an analysis of the why sub-branch under the oversights and omissions node. Questions can again guide this form of analysis. For example, the following guidelines corresponds to the leaf node with the following path Event: Oversights and Omissions: Why? : Management : Risk Assessment : Safety Program Review : Design and Development Plan : Human Factors. They direct an analyst to consider the impact that a managerial failure to consider human factors issues may have had upon the course of an incident:

> "*Human Factors Review Less Than Adequate:* Has consideration been given in design, plan, and procedures to human characteristics as they compete and interface with machine and environmental characteristics?
>
> - Professional Skills Less Than Adequate: Is the minimum level of human factors capability, needed for evaluation of an operation, available and will it be used? (275)
> - Did Not Describe Tasks: For each step of a task, is the operator told: When to act? What to do? When the step is finished? What to do next? (276)
> - Allocation Man-Machine Tasks Less Than Adequate: Has a determination been made (and applied) of tasks that humans excel in versus those tasks at which machines excel?
> - Did Not Establish Man-Task Requirements: Does the review determine special characteristics or capabilities required of operators and machines?
>   - Did Not Define Users: Is available knowledge about would-be users defined and incorporated in design?
>   - Use of Stereotypes Less Than Adequate: Are checklists of stereotypes (typical, normal, expected behaviour) used in design? (e.g., Is a control turned right to move a device to the right?) Are controls coded by size, colour, or shape?
>   - Displays Less Than Adequate: Are displays used which can be interpreted in short time with high reliability?
>   - Mediation Less Than Adequate: Is consideration given to delays and reliability of interpretation/action cycles?
>   - Controls Less Than Adequate: Are controls used which can be operated in short times with high reliability?
> - Did Not Predict Errors: Is there an attempt made to predict all the ways and frequencies with which human errors may occur, and thereby determine corrective action to reduce the overall error rate?

– Incorrect Act: Have all the potential incorrect acts associated with a task been considered and appropriate changes made?
– Act Out of Sequence: Has the consequence of performing steps of a task in the wrong order been considered and has appropriate corrective measures been made?
– Failure to Act: Is there an attempt to reduce the likelihood of operators omitting steps or acts which are required by procedure?
– Act Not Required: Are all the steps that are needed to accomplish a task required in the procedures? Are only those steps in the procedure?
– Malevolence: Are deliberate errors and other acts of malevolence anticipated and steps taken to prevent them or reduce their effect?" [204]

The MORT user guidelines emphasise a number of additional practical observations that have emerged from the application of this technique during incident and accident investigations [204]. The approach works best if it is used to focus discussion and debate. Any figures or forms that are produced during the analysis should be considered as working documents that can be revised and amended as work progresses. MORT, therefore, provides analytical guidance; 'it helps avoid personal hobbies, bias, or the tunnel vision that commonly results from pet theories of accident causation' [204]. It should not be seen as a framework to be imposed upon a final report. It can, however, be used as a quality control mechanism to identify any potential omissions in a final report. Investigators can use the questions to ensure that they have described both what happened and why those events occurred. Finally, experience in applying MORT has shown that even the full version of the diagram cannot cover all aspects of some incidents. If a mishap is not covered in any of the branches then analysts are encouraged to extend the existing diagram using the basic fault tree gates that were introduced in Chapter 8.3.

Having raised these caveats, it is possible to illustrate the application of MORT to the Nanticoke case study that was introduced in previous sections. As mentioned above, MORT analysis begins by determining what happened during an incident or accident. Investigators traverse the what branch of the tree, such as that shown in Figure 11.7, asking whether or not each potential failure contributed to the incident under investigation. MORT assumes that investigators have sufficient evidence to perform such an analysis. It does not provide any explicit guidance on how to go about satisfying this prerequisite, however, others have extended the approach to provide this support [444]. At the highest level, this traversal of the MORT diagram encourages investigators to identify the hazard that threatened potential targets within the system [300]. In our case study, the hazard can be identified as the danger of a fire being started by a pressurised fuel release from a fuel filter onto the adjacent indicator tap or uncovered exhaust manifold. This hazard threatened a number of different targets. Most immediately it posed a danger to the people and systems in the engine room. Ultimately, it threatened everyone on the vessel and even other ships that were operating in the same area as the Nanticoke.

As can be seen from the left sub-branches of Figure 11.7, analysts must also identify the ways in which any barriers or controls were circumvented during an incident. Barriers typically protect or shield a target from a hazard. Controls make it less likely that a hazard will occur in the first place. These terms are, however, often used interchangeably [553]. This imprecision is justified by the practical difficulties of distinguishing between these two different forms of defence. For instance, more regular inspections of the filter assembly might have made the fire less likely. Crew members might have noticed the leak before ignition. More frequent inspections might also have acted as a barrier by raising the alarm as soon as the fire had started. The practical problems of distinguishing between these different forms of protection helps to explain an imbalance in the MINI-Mort tree of Figure 11.7. This diagram provides considerable detail about the potential forms of control failure. This level of detail is not, however, reflected by the portion of the tree that considers inadequate barriers. This imbalance is also justified by the observation that these failures often take similar forms. Inadequate technical information or maintenance procedures can threaten both of these potential defences.

Barriers prevent hazards from having adverse consequence once they occur. They can be thought of as protection devices or shields that guard the target from the hazard. It can be argued that the barriers worked well in the Nanticoke case study because the fire was ultimately extinguished without loss of life or serious injury. Conversely, it can be argued that the barriers failed because the ship suffered considerable damage. The relatively limited fire managed to burn through the common cable tray that contained all of the steering systems. After 1st September 1984, duplicated steering power and control systems had to be routed as widely as possible throughout a vessel so that an isolated fire was unlikely to destroy all of these redundant systems. The Nanticoke was built in 1980 and so lacked the protection offered by the 1984 requirement. In consequence, the vessel was effectively disabled until an alternative power source could be rigged to the steering gear.

As mentioned, controls make it less likely that a hazard will occur. Figure 11.7 documents a number of potential weaknesses that can jeopardise adequate control. For example, the Nanticoke incident was arguably caused by inadequate maintenance. The modifications to the forward filter cover and bolt sealing surface left grooves that made it hard for the watchkeeping engineer to achieve a fuel-tight joint. This analysis shows how the MORT diagram can be used as a check-list to guide the analysis of what happened during an incident. It also illustrates some of the complexity that frustrates the use of checklist techniques. Damage to the seating surface not only suggests inadequate maintenance, it also indicates that there may have been inspection problems. Crew members might have recognised the potential for a fuel leak during previous rounds of preventive maintenance. This illustrates the way in incidents often stem from the failure of several different controls. Problems arise if investigators form different opinions about the salience of these failures. For instance, some analysts might discount the importance of inspection failures by arguing that the true significance of the seating damage could only have been determined with hindsight. Other analysts might stress the importance of inspection failures by arguing that the watchkeeping engineer should have reported their problems in obtaining a fuel tight seal during the maintenance that immediate preceded the incident. Such differences of interpretation make it very important that analysts both document and justify the findings of their MORT analysis. These justifications can then be reviewed and challenged before any subsequent causal analysis.

There are a number of differences that distinguish checklist approaches, such as MORT, from event-based techniques, such as STEP and MES. In particular, checklist approaches often abstract away from the temporal properties that are a central concern of the flowcharts and tabular forms in previous sections. The initial stages of a MORT analysis identify instances of generic failure types. They do not chart the timing of events. This is both a strength and a weakness. The MORT diagram cannot, in isolation, be used to reconstruct the way in which an incident developed over time. There is, therefore, no guarantee that investigators will identify omissions or inconsistencies in the events leading up to an incident. On the other hand, previous sections have criticised event-based techniques that force analysts to model precise event sequences which are unlikely to recur in future incidents. The identification of MORT failure types can, in contrast, generalise from the particular observations that characterise an individual incident. There are further benefits. By abstracting away from temporal properties, the MORT classification process can help investigators to identify similarities between latent and catalytic failures. Such similarities can be difficult to demonstrate with event-based techniques that deliberately separate the presentation of events that occur at different times during an incident. For instance, inadequate inspections may have contributed to the latent conditions behind the Nanticoke incident. Crew members failed to recognise the damage to the seating surface and this ultimately made it difficult for the engineer to achieve a fuel-tight seal. Inspection failures also characterised more immediate events during the incident. The engine room was not inspected between 15:15 and 16:00. Subsequent analysis might determine that these different failures had very different causes. The key point is, however, that the MORT style of analysis can help to identify potential similarities between failures that occur at different times during the same incident.

As with any checklist approach, MORT provides prompts that encourage analysts to consider a broad range of potential failures that might contribute to incidents and accidents. For example, the Nanticoke case study partly stemmed from operability problems. There were no new copper gaskets. Once a used copper gasket has been deformed by use, it is more difficult to obtain a tight seal for subsequent use even if it has been annealed. Other failures can be associated more directly with individual operators. For instance, the Mini-MORT diagram of Figure 11.7 includes a branch that represents inadequate intervention by the first line supervisor. As we have seen, it can be argued that they failed to correct the damage incurred during previous modifications to the filter. It can also be argued that they failed to detect the leak or the fire before it had taken hold.

Figure 11.7 also shows how further branches focus on the response to an incident. For instance, it can be argued that the emergency actions that were taken in response to the incident were complicated by the lack of any emergency exit from the control room. In consequence, the chief engineer had to follow hand rails out of the engine room. The corrective actions branch of the MORT diagram also includes a node Did not prevent 2nd accident. This supports the analysis of incidents in which the same hazard occurs more than once. For example, the fuel might have reignited after the initial fire had been extinguished. More widely, this node can encourage investigators to consider whether an incident forms part of a wider pattern. Chapter 14.5 will stress the importance of such activities. Investigators must look beyond the immediate response to an incident in order to learn from previous attempts to address similar failures. For example, the Transportation Safety Board of Canada identified that four similar engine room fires had occurred on Canadian ships within six months of the Nanticoke incident [623]. Previous ship safety bulletins had not resulted in adequate barriers being placed between potential fuel sources and adjacent exposed, hot surfaces. Subsequent analysis of the reasons why the fire occurred must, therefore, explain this failure to act upon previous safety bulletins.

Sub-Tree: What/Accident	
What?	Rationale
Hazard	Danger of a fire being started by a pressurised fuel release from a fuel filter onto the adjacent indicator tap or uncovered exhaust manifold.
Targets	People and systems in the engine room. Everyone on the vessel. Other ships in the same area as the Nanticoke.
Barriers	
Did not use	More frequent inspections might raised the alarm sooner.
Did not provide	Fire burnt through common cable tray containing all of the steering systems. Nanticoke was disabled until alternative power source was rigged for steering gear.

Table 11.5: MORT (Stage 1) Summary Form for Hazard, Targets and Barriers

Tables 11.5 and 11.6 summarise the results of the first stage in the MORT analysis of the Nanticoke case study. These tables are intended to provide a focus for discussion. Previous paragraphs have argued that considerable disagreements are possible over our interpretation of which nodes best capture the failures that contributed to this incident. It is also important to notice that Table 11.5 extends the Barrier branch from Figure 11.7. The nodes Did not use and Did not provide reflect types of failure that were described as part of the introduction to barrier analysis in Chapter 9.3. This illustrates the way in which analysts may have to extend the pre-defined categories within a Mini-MORT diagram. In this case, however, these additional nodes are consistent with those included in the full MORT diagram.

Previous sections have described how this first stage of identifying *what* happened helps to drive a more detailed causal analysis of *why* those failures occurred. Before making this transition, however, it is possible to make a few observations about the use of MORT to drive an initial assessment of the Nanticoke case study. As we have seen, there is no automatic or semi-automatic procedure for

Sub-Tree: What/Accident	
What?	Rationale
Controls	
Inspection LTA	More regular inspections of filter assembly might reduced likelihood of fire. Crew members might have noticed the leak before ignition.
	Crew members (arguably) might have reported problems in obtaining a fuel tight seal during maintenance immediately before the incident.
	Engine room was not inspected between 15:15 and 16:00.
Maintenance LTA	Modifications to forward filter cover and bolt sealing surface left grooves that made it hard to achieve a fuel-tight joint.
Operability problems	No new copper gaskets. Copper gaskets are deformed by use and pose more problems in obtaining a tight seal even if they have been annealed.
1st Line Supervision LTA	Failure to identify and correct damage incurred during previous modifications to the filter.
	Failed to monitor engines during interval prior to the fire (15:15 to 16:00).
Emergency actions LTA	No emergency exit from the control room. Chief engineer had to follow hand rails out of the engine room.
Did not prevent 2nd accident	Four similar engine room fires occurred on Canadian ships within six months of the Nanticoke incident. Ship safety bulletin (13/85) had not resulted in adequate barriers being placed between potential fuel sources and adjacent exposed, hot surfaces.

Table 11.6: MORT (Stage 1) Summary Form for Controls

identifying the particular failures that characterise an incident or accident. In contrast, investigators must rely on subjective judgement and prior expertise to determine which of the MORT nodes most accurately describe what led to the incident. There are no guarantees that different investigators will derive similar classifications for the same incident. This would seem to be unlikely given that particular conditions, such as the damage to the seating, can be the result of several inadequacies throughout the left-hand branch of the MORT diagram. The proponents of this approach have argued, however, that MORT provides a focus for discussion rather than a method for deriving a definitive analysis or single interpretation of events. This is an important observation given that there can be considerable disagreement not simply about the course of an incident but also about the precise meaning of each category within the MORT diagram. As we have seen, investigators often experience considerable practical difficulties in distinguishing between a barrier and a control. Some organisations have responded to these potential problems by developing considerable in-house documentation to support the use of MORT [204]. This material includes training material, case studies and style guides that reflect a particular approach to the MORT technique. Others have gone further. For instance, Kjellén has extended MORT to develop SMORT (Safety Management and Organisation Review Technique) [444]. This provides explicit support for data collection during incident investigations. As we have seen, this support was not part of the initial MORT approach. Such elaborations combined with explicit encouragement to extend the MORT diagram if it does not capture key aspects of an incident have resulted in a situation in which the term MORT is often used to describe a very varied collection of subtly different techniques. These techniques vary both in the checklists that are used and in the supplementary methodological support that is provided to guide their application.

A number of further observations can be made about the Nanticoke case study. The MORT diagram illustrated in Figure 11.7 captures the emphasis that this technique places upon failure. The diagram prompts investigators to identify what went wrong by systematically considering the ways in which various aspects of performance were less than adequate. Previous chapters have, however, argued that near-miss incidents often provide vital information about those barriers and controls that worked effectively to prevent an accident from occurring. For example, the Halon system on the Nanticoke provided an effective final resort after the crew made two unsuccessful attempts to fight the fire themselves. It can, therefore, be argued that investigators ought to repeat their analysis of a MORT diagram to identify these mitigating factors whose performance was At or Beyond Expectation (ABE) and not Less Than Adequate (LTA).

The second stage of MORT analysis helps investigators to determine the causes of an incident. This is done by identifying those elements in the why branch that contributed to each of the failures that were summarised in Tables 11.5 and 11.6. At the highest level, the overall hazard was the danger of a fire started by a pressurised fuel release from a fuel filter onto the adjacent indicator tap or uncovered exhaust manifold. It can be argued that this was the result of an inadequate risk assessment. The operators and crewmember failed to recognise the potential threat to everyone on the vessel and to other ships in the area. As before, the MORT diagram can be used to guide the analysis of what might have caused this failure. The Risk Assessment LTA branch contains a number of detailed nodes that investigators can adopt as working hypotheses about the factors that led to an incident. For example, Table 11.5 argued that more regular inspections might have prevented the fire from developing if the crew had been able to raise the alarm sooner than they did. The failure to effectively implement such a barrier can be explained in terms of the node Inspection Plan LTA which is located under the path Why? Management LTA : Risk Assessment LTA : Safety Program Review LTA : Design and Development plan LTA in Figure 11.7. Similarly, the failure to provide a sufficient barrier to protect the control cables for the steering system can be explained in terms of the Design basis LTA node which appears at the same level as Inspection Plan LTA. Had the Nanticoke been built after the September 1984 regulations were introduced then the cables would have been distributed more widely throughout the vessel. An isolated fire would then have been less likely to damage all of the redundant steering systems.

Investigators can also use the MORT diagram to identify potential reasons why Controls failed to protect the system. For example, Table 11.6 suggested that inspections might have been less than adequate because crewmembers might have noticed the possibility of a leak well before the fire. In particular, engineers could have reported the problems in obtaining a fuel tight seal during the periodic maintenance that took place immediately before the incident. Both of these apparent inadequacies can be described in terms of less than adequate inspection plans and less than adequate maintenance plans. Similarly, the failure to inspect the engine room between 15:15 and 16:00 can be characterised as the result of less than adequate procedures. It is important to reiterate that these are subjective interpretations of the failures that were identified during the first stage of the analysis. For instance, it could be argued that the failure to inspect the engine room between 15:15 and 16:00 was not simply the result of inadequate operating procedures. Better protection might have been offered if operators had been expected to document their inspection activities. This would have led the same Inspection LTA failure to have been classified under the Monitoring points LTA node of the Why? branch. Similarly, it can be argued that the inspection failure was due to inadequate training about the importance of these activities. This, in turn, could be due to a managerial failure to identify such a training requirement; Why?:Risk Assessment LTA: Safety Program Review LTA: Design and Development Plan LTA: Operational Specification LTA: Training LTA. Alternatively, it might be argued that the lack of inspection was not due to any of these factors but to management's failure to motivate staff to perform necessary safety inspections: Why?:Risk Assessment LTA: Safety Program Review LTA: Design and Development Plan LTA: Operational Specification LTA: Motivation LTA. These observations illustrate a number of important points about causal analysis using the MORT approach. Firstly, a number of different causal factors can be associated with the items identified in the first stage of the analysis. Some of these factors are not mutually exclusive. So, for example, inadequate inspection procedures might be compounded by a lack of monitoring points. Even if inspection procedures had been well-defined, motivational problems can 'dissuade' individuals from effectively following monitoring requirements.

Secondly, the Nanticoke case study supports a number of important observations about the nature of any causal analysis. It is difficult to be certain about which causal hypotheses, the nodes of the Why branch in the MORT diagram, can actually be applied to this incident. The available reports and documentation provide very little information about the motivation of the crewmembers or about the written procedures that were available to key personnel. Further investigations would, therefore, be necessary before any conclusions could be reached about these potential causes. An important strength of the MORT approach is that it directs investigators towards these potential hypotheses that must then be supported by further investigations. This offers a strong contrast to many event-based approaches. There is often an implicit assumption that counterfactual reasoning over a temporal model of event sequences can provide sufficient information about the underlying causes of an incident. This is a strong assumption. Chapter 9.3 has shown how NASA and the US Department of Energy have partially addressed these concerns by recommending the use of Tier or Compliance analysis to supplement the counterfactual reasoning afforded by ECF modelling.

The other control failures identified in Table 11.6 can be analysed in a similar fashion. Inadequate maintenance was recognised by the manner in which modifications to the forward filter cover and bolt sealing surface left grooves that made it hard to achieve a fuel-tight joint. This can potentially be explained in terms of inadequate maintenance and inspection plans under the path Why?:Risk Assessment LTA: Safety Program Review LTA: Design and Development Plan LTA. Operability problems including the lack of any new gaskets and the problems associated with the reuse of deformed gaskets can be associated with a management failure to conduct an adequate hazard analysis. Supervisory problems such as the failure to identify and correct damage incurred during previous modifications to the filter can be interpreted as the result of inadequate procedures. For example, a fault reporting system might have altered the chief engineer to the watchkeeping engineer's problems in achieving a sufficient seal on the filter. The failure to monitor the engines adequately between 15:15 to 16:00 can be interpreted as a failure of supervision in the operational specification of the system. The lack of any emergency exit forced the chief engineer to follow hand rails out of the engine room. This can be seen as a failure in the design basis of the ship as it was being operated immediately before the incident. Additional emergency lighting might, arguably, have supported the chief engineer's exit from a hazardous situation. Finally, the failure to prevent a recurrence of four previous engine fires on Canadian ships within six months of the Nanticoke incident can be associated with a failure to review the overall safety programme over previous years. In particular, Transportation Safety Board of Canada argued that previous warnings, such as that contained in Ship Safety Bulletin 13/85, had not resulted in adequate barrier being placed between potential fuel sources and adjacent exposed, hot surfaces.

Sub-Tree: Management Less Than Adequate (LTA)		
Why?	What?	Description
Risk Assessment LTA	Hazard	Danger of a fire being started by a pressurised fuel release from a fuel filter onto the adjacent indicator tap or uncovered exhaust manifold.
	Target	People and systems in the engine room. Everyone on the vessel. Other ships in the same area as the Nanticoke.
Hazard Analysis LTA	Control: Operability problems	No new copper gaskets. Copper gaskets are deformed by use and pose more problems in obtaining a tight seal even if they have been annealed.
Inspection Plan LTA	Barrier: Did not use	More frequent inspections might raised the alarm sooner.
	Control: Inspection LTA	More regular inspections of filter assembly might reduced likelihood of fire. Crew members might have noticed the leak before ignition.
	Control: Inspection LTA	Engine room was not inspected between 15:15 and 16:00.

Table 11.7: MORT (Stage 2) Analysis Form

Tables 11.7 and 11.8 summarise the findings from the second stage of our MORT analysis. As can be seen, each of the nodes from the why branch in the MORT diagram can be represented as a row in the table. The what nodes that were identified during the first stage of the MORT analysis are then listed next to each of the why nodes if the corresponding (managerial) failures are perceived to have caused the more immediate failures that contributed to the incident. For example, the lack of adequate monitoring points to encourage compliance with inspection procedures is seem to have been a cause of the crews failure to adequately inspect the engine room between 15:15 and 16:00. It is important not to underestimate the significance of such tables. As mentioned, they provide a focus for continued discussion and analysis amongst the members of an investigation team.

The MORT analysis forms, illustrated in Tables 11.7 and 11.8, also act as a focus for other forms of analysis. For instance, the US Department of Energy have argued that investigators can sum the number of *what* factors associated with each why node to provide 'a measure of how widespread the element inadequacy is'. [205] In Tables 11.7 and 11.8 this can be done by counting the number of rows for each why? node. This would yield the following rankings for the Nanticoke case study. Inspection plan LTA is the only causal factor that is associated with three specific what failures. Risk Assessment LTA, Maintenance Plan LTA and Design Basis LTA are all associated with two specific failures. Hazard Analysis LTA, Monitoring Points LTA, Procedures LTA, Supervision LTA and Safety Program LTA are identified as the causes of a single failure in the accident/incident branches of the MORT diagram.

A number of objections can be raised to this form of analysis. The subjective nature of both stages in the MORT method can create considerable differences in the results that are obtained from this simple summation of accident factors. Similarly, it can be argued that different weights should be associated with each of the causal factors in the why branch of the MORT diagram. For instance, investigators may identify numerous instances in which operating procedures were inadequately specified. Changes in equipment design, in the operating environment and in regulatory requirements can prevent even the most assiduous operator from ensuring that all operating procedures are correctly documented. It might, therefore, be argued these problems are not as serious as less numerous maintenance failures. For instance, the Nanticoke incident might have had far worse consequences had the Halon system not been available to the Captain once his fire-fighting teams had been beaten back. Rather than develop more complex procedures for deriving aggregate weightings from MORT analysis form's, we adopt the more usual practice of assuming that investigators will use their skill and expertise to determine the overall significance of each row within Tables 11.7 and 11.8.

Previous paragraphs have described how the first stage of MORT analysis identifies what occurred during an incident. The second stage goes on to identify causal factors by asking why these failures arose. We have not, however, described the process by which root causes might be distinguished from the wider causal factors to the right of the MORT diagram. Several authors have argued that the concept of a 'root cause' originates with Johnson's early work on MORT [430, 444]. For example, Briscoe developed an analytical technique in which root causes are literally represented by the roots of the MORT diagram [96]. Investigators simply trace the more detailed why factors, identified in the Analysis Forms of Tables 11.7 and 11.8, up through the tree to identify the higher-level branches that represent the wider causes of managerial failure. The following list summarises the main categories that were identified by Briscoe's root cause analysis technique. Most of the items are relatively straightforward. Bridge elements represents the manner in which high-levels of management implement safety-related management policies throughout the various intermediate tiers of management within an organisation.

1. Policy
2. Policy Implementation
    - Line/staff responsibility
    - Accountability
    - Vigour and example
    - Methods and criteria analysis
3. Risk assessment
    - Safety-information systems
    - Hazard-analysis process
    - Safety-programme audit
4. Bridge elements
    - Management services
    - Directives
    - Budget
    - Information flow

Many of the causal factors that were identified for the Nanticoke case study can be broadly grouped under the 'hazard analysis process' root cause. Management failed to appreciate the dangers of the maintenance and inspection practices that were identified in Tables 11.7 and 11.8. Alternatively, if those dangers were recognised then it can be argued that there was an inadequate safety-programme audit because such practices were permitted to continue even after warnings such as that contained in Safety Bulletin 13/85.

Briscoe's approach is not the only checklist form of root cause analysis that might be applied after the second stage of a MORT analysis. For example, the International Loss Control Institute have developed a model of incident causation that extends the domino theory [85]. This approach proposes a number of further root causes in addition to those proposed by Briscoe [444]. These focus on common reasons behind failures at the workplace level:

Sub-Tree: Management Less Than Adequate (LTA)		
Why?	What	Description
Maintenance Plan LTA	Control: Inspection LTA	More regular inspections of filter assembly might reduced likelihood of fire. Crew members might have noticed the leak before ignition.
	Control: Maintenance LTA	Modifications to forward filter cover and bolt sealing surface left grooves that made it hard to achieve a fuel-tight joint.
Monitoring points LTA	Control: Inspection LTA	Engine room was not inspected between 15:15 and 16:00.
Design basis LTA	Barrier: Did not provide	Fire burnt through common cable tray containing all of the steering systems. Nanticoke was disabled until alternative power source was rigged for steering gear.
	Emergency actions LTA	No emergency exit from the control room. Chief engineer had to follow hand rails out of the engine room.
Procedures LTA	Control: 1st Line Supervision LTA	Failure to identify and correct damage incurred during previous modifications to the filter.
Supervision LTA	Control: 1st Line Supervision LTA	Failed to monitor engines during interval prior to the fire (15:15 to 16:00).
Safety Program Review LTA	Did not prevent 2nd accident	Four similar engine room fires occurred on Canadian ships within six months of the Nanticoke incident. Ship safety bulletin (13/85) had not resulted in adequate barrier being placed between potential fuel sources and adjacent exposed, hot surfaces.

Table 11.8: MORT (Stage 2) Analysis Form Continued

1. inadequate health and safety programme
2. inadequate health and safety programme standards
3. inadequate compliance with health and safety programme standards

Further additions might be made. Chapter 2.3 argued that the regulatory environment has a profound impact upon managerial behaviour. The decision only to apply the revised wiring requirement to vessels built after 1st September 1984 left the Nanticoke in a particularly situation when the fire burnt through the common cable tray that contained all of the steering systems. The decision to include such regulatory influences as a potential root cause within a MORT diagram depends upon the position of the investigator within an incident reporting system. In some schemes, typically those run by independent reporting agencies, it is possible for investigators to address these more general issues that might otherwise lie outside the scope of a conventional MORT analysis. If investigators decide to introduce regulatory and workplace factors, mentioned above, then these factors must appear as potential root causes in the upper levels of a revised MORT diagram. This increases the scope of the root cause analysis. Investigators must, however, navigate an increasingly complex diagram to identify those leaf nodes that best describe why particular failures occurred.

The MORT approach offers a number of significant benefits. In particular, it provides an early example of the way in which an engineering approach to safety, typified by barrier analysis, can be combined with broader managerial concerns. This blend of concerns has provides detailed insights into the way in which particular management activities contribute to many accidents and incidents [765]. The distribution and delegation of responsibility without adequate supervision often emerges as a common theme in MORT analyses. Similarly, the failure to implement well-specified safety plans can also be identified as a recurring pattern. There remains a considerable debate about whether or not these recurring themes are artifacts of the MORT analysis or whether they reflect common problems for different safety-critical systems [348]. A number of authors have, however, proposed automated tools that might automatically detect such recurring causal patterns amongst a 'database' of incident reports [457].

MORT offers a number of further benefits that relate more narrowly to the management of any investigation. The elements of the diagram direct investigators towards the potential causes of an incident. This helps to ensure that analysts consider a broad range of causal factors. The use of the tree can also provide necessary guidance for inexperienced investigators. It provides a common structure and format that encourages consistency in the investigatory process. The method associated with the tree is intended to ensure that investigators consider both what happened and why the incident occurred. The use of tabular check lists helps to communicate the products of a causal analysis to others within an investigatory team. Finally, the summary data that can be obtained from MORT tables, such as that illustrated in Table 11.5, can be used to monitor the changing causes of incidents across different geographical regions or organisational boundaries.

A number of limitations also restrict the utility of MORT as a tool for the causal analysis of safety-critical incidents. In contrast to STEP, this approach best be applied once investigators have already obtained a significant amount of information about an incident. Some proponents have argued that incident modelling, using ECF or accident Fault Trees, should be a prerequisite to any MORT analysis. In this view, counterfactual reasoning is used to identify causal factors that are then classified using the what branch of the tree. Instead of using Tier or Non-compliance analysis as in Chapter 9.3, investigators can then apply MORT to classify root causes against the why branch. Unfortunately, the perceived complexity of the MORT diagram and the potential overheads of such an integrated approach have dissuaded many analysts from exploiting these techniques [486].It is, therefore, seldom used in its full form without regulatory sanction. Munson argues that MORT is used more as a pro-active tool to support the analysis of a safety-critical design than it is as an accident investigation technique. This is due to the "nature of the nuclear industry, identifying possible loopholes in the safety system to eliminate hazards is more cost effective and publicly expedient than after the accident occurs" [553].

The leafs of MORT and mini-MORT diagrams may not capture the specific causes of an incident [292]. This should not be surprising. These diagram reflect the inevitable trade-off between large and unwieldy structures that embody many causal distinctions and more compact trees that provide a smaller number of more generic categories. As we have seen, investigators can extend MORT diagrams to address these limitations. This can, however, create inconsistencies within an incident reporting system. For instance, other investigators may not have used the new category in previous investigations. The extension of the MORT diagram can also create external inconsistencies between incident reporting systems if other organisations choose not to exploit the amended MORT diagram. Such problems can dissuade investigators from searching for causal factors that are not represented on the MORT diagram.

11.2.2 Prevention and Recovery Information System for Monitoring and Analysis (PRISMA)

As we have seen, ECF, MES and STEP help analysts to reconstruct the event sequences that contribute to incident and accidents. Different forms of counterfactual reasoning can then be used to distinguish between the causal factors and contextual details that are represented in these incident models. These techniques all focus on the specific events that occurred during a particular incident. Investigators must use a range of complementary approaches, such as Tier analysis, to identify the more generic root causes from the results of these more focussed techniques. In contrast, MORT relies upon investigators already having a relatively detailed understanding of the particular events leading to a mishap. The associated diagram and tabular form can be used to classify specific causal factors into a number of more general categories. It is important not to underestimate the significance of this distinction between MORT and the previous techniques. ECF, MES and STEP focus on 'singular causality' [678]. MORT focuses on the notion of 'general causality' that was introduced in Chapter 6.4.

A number of researchers have recognised the distinctions between particular and general causality that are embodied within ECF, MES, STEP and MORT. They have responded by developing more integrated approaches that are intended to support both the reconstruction of the specific events that lead to an incident and the identification of more general causal factors. The Prevention and Recovery Information System for Monitoring and Analysis (PRISMA) is one example of this dual technique [841, 842]. This approach is also different from those introduced in previous sections because it was specifically developed to enable organisations to monitor and respond to incident reports. It was not intended to support accident investigation.

Van Der Schaaf's motivation in developing PRISMA was to support the development of a quantitative database of incident data. This resource was to guide the detection and prevention of structural problems rather than the particular characteristics of individual incidents [845]. The PRISMA approach consists of three principle stages. The following paragraphs describe each of these stages and illustrates how they can be used during a causal analysis of the Nanticoke case study:

1. *Reconstruct the incident using a causal tree.*
2. *Use a classification model to identify generic factors.*
3. *Apply a classification/action matrix to identify potential counter-measures.*

Causal trees are similar to the Fault Trees that were introduced in Chapter 9.3. The overall structure of the tree reflects the chronology of an incident. The left-most branches indicate latent conditions or failures that occur relatively early in the course of events. The right-most branches are, typically, used to model recovery actions and interventions that mitigate the consequences of an incident. It is important to note, however, that causal trees are constructed using AND gates. Investigators must avoid the uncertainty that is implied by disjunction. Van Vuuren notes that "the main difference between a causal tree and a fault tree is that the top event in a causal tree is not a class of events but one particular incident, which actually occurred and for which the chain of causation can be discovered" [845]. In contrast to the MORT diagram, causal trees are intended to capture the 'who', 'what' and 'where', they do not explain 'why' an incident may have occurred. Figure 11.8 presents a causal tree for the Nanticoke case study.

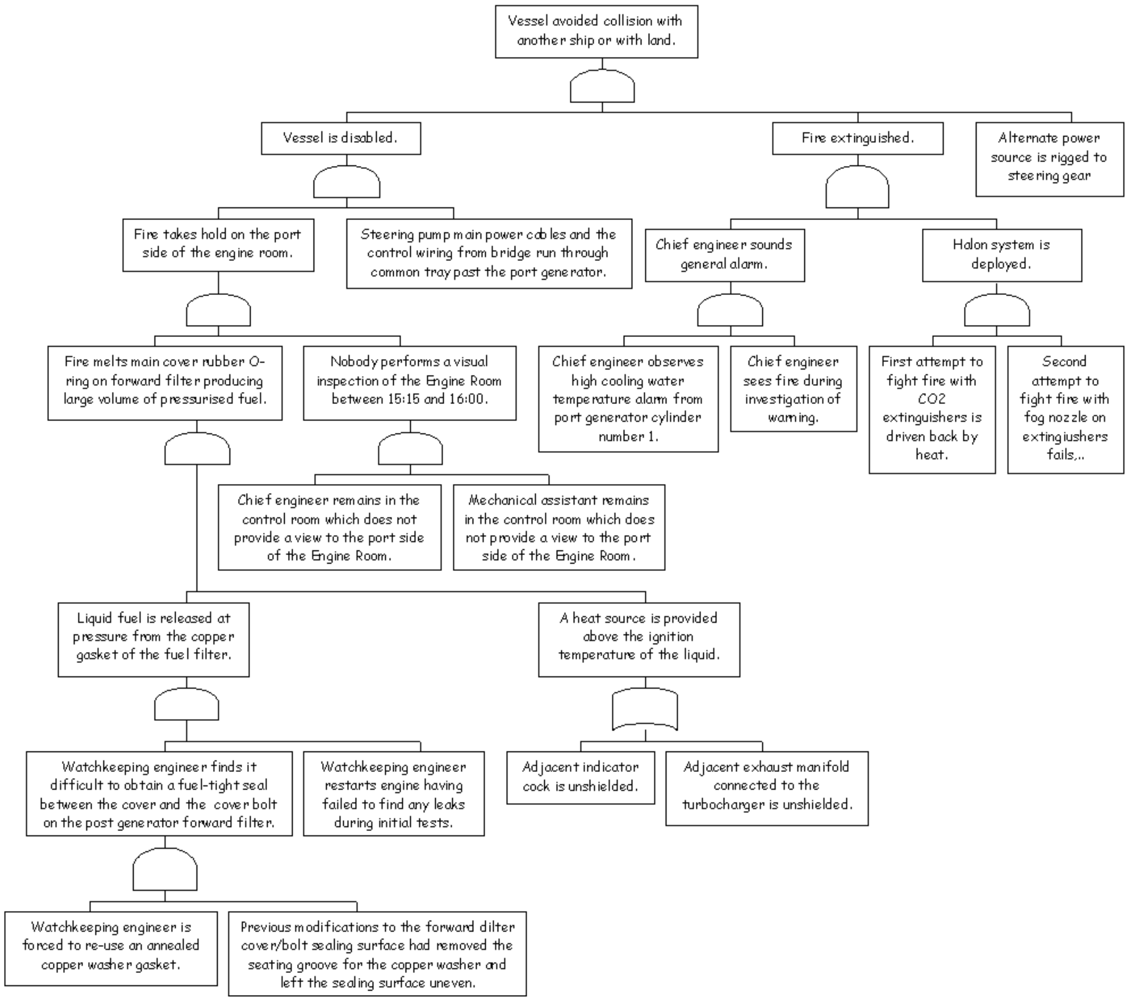

Figure 11.8: A Causal Tree of the Nanticoke Case Study

There are considerable differences between causal trees and the various modelling techniques in ECF, STEP and MES. For instance, analysts can annotate the nodes in a causal tree with natural language labels that do not distinguish between events and conditions. These annotations are intended to capture observations about the course of an incident in a flexible and informal manner. One consequence of this is that it can be difficult for investigators to distinguish the actions of particular individuals during an incident. Rather than grouping these along a single row, as in STEP, they can be distributed across the many different nodes of a causal tree. The lack of typing information also means that ambiguities and omissions can weaken the integrity of these diagrams. For instance, some of the proponents of this approach have published trees whose nodes are labelled He was standing next to the person or he saw the falling object. Such annotations work well for small examples but cannot easily be maintained for more complex incidents, such as the Nanticoke case study. Figure 11.8, therefore, explicitly identifies the key individuals who were identified during the primary investigation into this mishap.

There are further differences between the causal tree of Figure 11.8 and the checklist approach embodied in MORT. In particular, the nodes represent both positive or mitigating factors as well as the failures that contribute towards an incident. As can be seen, this diagram denotes the way in which the chief engineer eventually noticed the high cooling water alarm from the port generator, cylinder number 1. It also records the successful use of the Halon system to extinguish the fire after two attempts to use carbon-dioxide extinguishers were beaten back by the heat. These right-hand branches are a significant strength of the PRISMA approach to incident modelling. As we have reiterated, organisation learning depends not simply upon recognising the causes of failure but also on promoting those actions that help to combat previous failures.

Every leaf nodes represents a causal factor. At first sight, this might appear to lack the sophistication of the more elaborate counterfactual approaches from previous sections. It is important to remember, however, that causal trees are entirely constructed from AND gates. It, therefore, follows that if any of the leaf nodes are not true then the top level incident will not be true. In consequence, this approach mirrors the counterfactual decision procedure of ECF, MES and STEP. There are, however, some exceptions to these general comments. As can be seen from Figure 11.8 it may still be necessary to include an OR gate within the causal trees that represent particular incidents. As with the Allentown explosion in Chapter 8.3, it is likely that we shall never be able to determine the exact ignition source for the Nanticoke case study. Transportation Safety Board of Canada investigators identified the indicator tap and exhaust manifold as potential sources. They were, however, unwilling to commit themselves to which was the most likely cause of the ignition. This uncertainty is denoted by the OR gate in Figure 11.8. As we shall see, this introduces a number of theoretical problems for the application of the PRISMA technique.

The second stage in the application of the PRISMA approach uses a classification model to associate a more generic root cause with each of the causal factors that are denoted by leaf nodes. This focus on the leaf nodes is justified by the observation that internal nodes are often the result or consequence of these other events and conditions. For instance, in Figure 11.8 two leaf nodes represent the facts that Previous modifications to the forward filter cover/bolt sealing surface had removed the seating groove for the copper washer and left the sealing surface uneven and Watchkeeping engineer restarts engine having failed to find any leaks during the initial tests. These two factors helped to create a situation that is represented by the interior node Watchkeeping engineer finds it difficult to obtain a fuel-tight seal between the cover and the cover bolt on the port generator forward filter. The re-use of the annealed copper gasket and the damage caused by previous modifications are seen to be causes of the engineer's subsequent difficulties. They are the focus for the subsequent classification rather than the interior node that represents the consequence of those two factors.

The second stage of the PRISMA analysis also, typically, focuses on the left-hand side of the causal tree. Recovery or mitigating factors are typically located on the right-hand side of the tree because they, typically, occur after the initiating conditions. These factors are important because they provide insights into protection mechanisms that successfully mitigated the potential consequences of an incident. For instance, the right-hand nodes of Figure 11.8 represent the crews actions that ultimately extinguished the fire on the Nanticoke. They also describe how an alternative power supply was rigged to the steering gear so that the crew could regain control of their vessel. These mitigating factors are not considered during this second stage of analysis. They represent remedial actions rather than causal factors. It is important to provide a procedure that can be used to distinguish causal factors from other mitigating actions in a causal tree. This can be done using the counterfactual reasoning that was introduced in previous paragraphs. For each node in a causal tree then investigators must ask whether the incident would have occurred if that node had not occurred. If the answer is no then the node represents a true causal factor and it is used during the subsequent classification. If the answer is yes then the node is not carried forward into any subsequent analysis. For example, the omission of a mitigating factor is likely to have exacerbated an incident rather than prevented its occurrence.

Unfortunately, the presence of disjunctions in a causal tree can considerably complicate this use of counterfactual reasoning. For example, if ask 'would the Nanticoke incident have been avoided if the adjacent indicator tap been shielded' then the answer would be no. The ignition might have been caused by the exhaust manifold. Conversely, if we ask 'would the incident have been avoided if the adjacent exhaust manifold had been shielded' then the answer would also be no. The ignition might have been caused by the indicator tap! Such problems can be resolved by further empirical studies or mathematical modelling. As we have seen in the Climate Orbiter case study, it is important not to over-estimate our ability to reconstruct the events leading to many incidents. The Nanticoke mishap is not the only case study in which such problems arise. For example, there are a number of competing hypotheses about the event sequences that led to the loss of the Deep Space 2 probes. In Chapter 9.3 we focussed on the potential problems that may have arisen during impact with the Mars surface. However, the probes may also have been damaged during separation from the cruise stage of the Polar Lander. If we ask 'would the incident have been avoided if the probes successfully separated from the cruise stage' then the answer is no. The probes might have been destroyed on impact with the planet surface. Conversely, if we ask 'would the incident have been avoided if the probes were resilient enough to survive impact with the planet surface' then the answer would again be no. Even if they had been capable of surviving the impact, they may not have reached that stage of the mission if problems had occurred during separation. Previous chapters have argued that such problems can be avoided by applying counterfactual reasoning over several different competing failure scenarios. In this view, investigators assume that one of the competing sets of events occurred. For instance, that the Nanticoke ignition was started by the adjacent indicator tap and not be the exhaust manifold. Counterfactual reasoning can then be applied as before. The lack of shielding can, therefore, clearly be identified as a causal factor. This reasoning process can then be repeated for the alternative failure scenarios. We term this counterfactual reasoning by *proxy*. Any ambiguity, such as that represented by the OR gate in Figure 11.8 is replaced by an assumed version of events. This assumption can then, in turn, be substituted by alternative event sequences during subsequent analysis. For instance, the assumption that the exhaust manifold provided the ignition source can be replaced by an assumption that the indicator tap helped to cause the fire.

The leaf nodes that represent causal factors in the Nanticoke case study are summarised as follows:

- Steering pump main power cables and the control wiring from the bridge run through a common tray past the port generator.
- Chief engineer remains in the control room which does not provide a view to the port side of the engine room.
- Mechanical assistant remains in the control room which does not provide a view to the port side of the engine room.
- Watchkeeping engineer is forced to re-use annealed copper washer gasket.
- Previous modifications to the forward filter cover/bolt sealing surface had removed the seating groove for the copper washer and left the sealing surface uneven.
- Watchkeeping engineer restarts engine having failed to find any leaks during the initial tests.
- Adjacent indicator tap is unshielded.

- Adjacent exhaust manifold is unshielded.

As mentioned, these causal factors are then categorised using a classification model that guides the investigators analysis. These models are used to associate more general root causes with the specific causal factors that are obtained from the causal tree. They can therefore be thought of as a further variant of the checklist approach, introduced in Chapter 9.3. PRISMA was initially developed to exploit the Eindhoven Classification Model, illustrated in Figure 11.9. This model was derived from an investigation of the causes of safety-related failures in the chemical process industry [841]. Since that time, however, a number of more detailed models have been developed to support the analysis of incidents in the medical and steel production domains [845]. For example, Figure 11.10 illustrates a medical classification scheme. The Eindhoven Classification Model focuses on three main categories of failure: technical; organisational and human. These can then be sub-divided into a number of more detailed causal factors. For instance, causal factors that relate to human behaviour can be associated with rule, knowledge or skill-based performance. These distinctions reflect Rasmussen's model of cognition introduced in Chapter 2.3. Similarly, organisational root causes are divided into inadequate operating procedures or ill-advised management priorities.

The classification process follows a fixed order [845]. Investigators must first determine whether the causal factor relates to the technical work environment. If the answer is yes, then the investigator must use the model in Figure 11.9 to determine the nature of that technical failure. Was the root cause related to an engineering, construction or materials problem? If the causal factor cannot be associated with a technical root cause then investigators must consider the organisational context of the incident. If technical and organisational factors are ruled-out then human behaviour can be considered as a root cause. This ordering is entirely deliberate. As with MORT, the detailed architecture of the classification scheme reflects the perspective and priorities of its developers. In this case, the Eindhoven Classification Model places human behaviour last so that investigators are forced to consider other causal factors before 'blaming' individual operator error.

The Eindhoven Classification Model from Figure 11.9 can be used to identify root causes from the causal leaf nodes of Figure 11.8. Table 11.9 summarises the results of this analysis. As we have seen, the use of a common tray to route all of the steering power and control wiring was identified as a causal factor in the loss of control that followed the fire. The decision to employ this approach can be associated with a technical failure in the engineering of the vessel. In consequence, Table 11.9 associates the wiring layout with the TE root cause from the Eindhoven Classification Model. The same categorisation can be applied to the manner in which previous modifications had removed the seating groove for the copper washer and left the sealing surface uneven. Previous sections have argued that this damage reflects incorrect maintenance procedures. It can, however, be argued that the removal of the seating groove was a consequence of previous maintenance problems. This again illustrates how the application of causal analysis techniques, such as PRISMA, are not an end in themselves. They raise questions that can only be resolved through further investigation.

Table 11.9 associates the same root cause with both of the hypothesised ignition sources. The lack of shielding around the indicator tap and exhaust manifold is associated with a technical failure in the construction of the engine assembly. It could be argued that these problems relate more to the engineering or design of the engine and filter rather than to its construction. This example also illustrates how distinctions that are meaningful within one industry need not be important in other domains. The differences between engineer, construction and materials are clearly defined within Van Der Schaaf's initial studies of the chemical process industries [841]. They are, however, less clear cut for our maritime case study. Such observations illustrate the need to derive classification models that capture pertinent root causes within a particular application domain.

It is also possible to challenge our claim in Table 11.9 that the re-use of the annealed copper washer gasket stemmed from a failure in organisational operating procedures. The re-use of copper gaskets that had previously been deformed under high operating pressures should not have been permitted. Conversely, it can also be argued that this failure stems more from a technical failure to ensure that the engineers were supplied with adequate materials. This illustrates the importance of both documenting the outcome of any root cause analysis and the associated justifications that support a particular categorisation. These documents can be shown to other investigators and safety managers to validate the products of any causal analysis. Any conflicts might be resolved

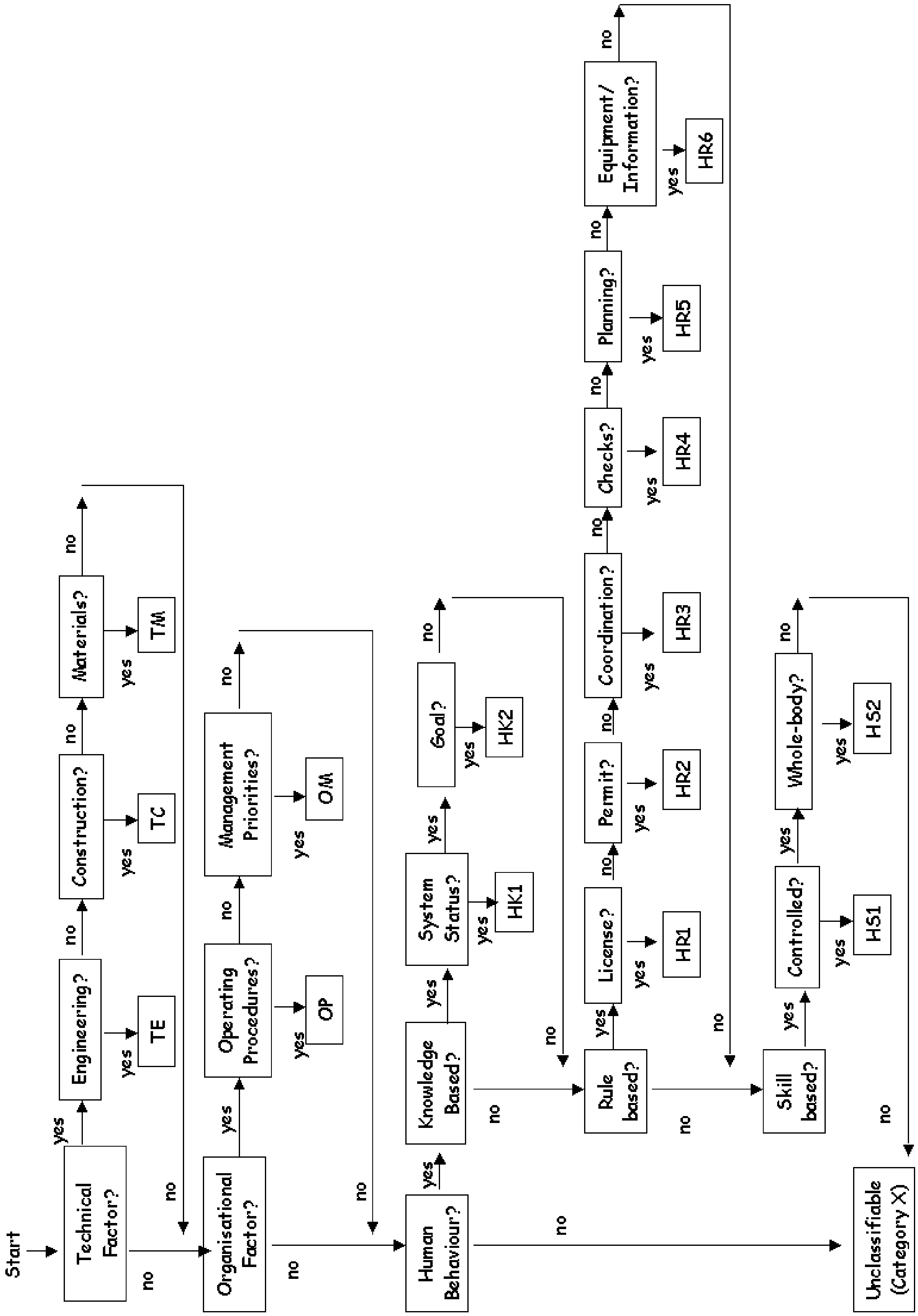

Figure 11.9: The Eindhoven Classification Model [841]

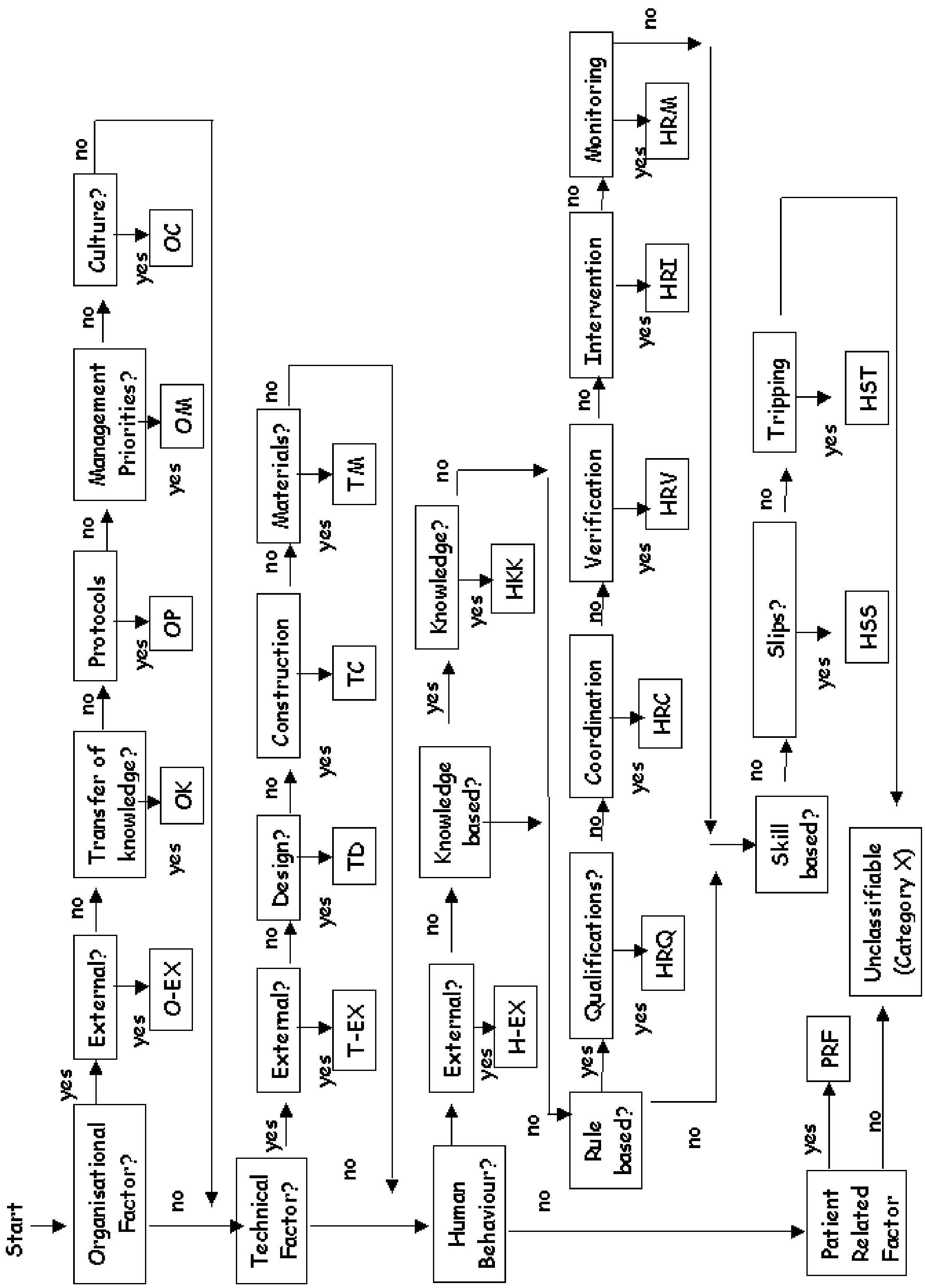

Figure 11.10: Classification Model for the Medical Domain [845]

Causal factor	ECM Classification
Steering pump main power cables and the control wiring from the bridge run through a common tray past the port generator.	TE - Technical Factor: Engineering.
Chief engineer remains in the control room which does not provide a view to the port side of the engine room.	HR4 - Human Behaviour : Rule Based : Checks.
Mechanical assistant remains in the control room which does not provide a view to the port side of the engine room.	HR4 - Human Behaviour : Rule Based : Checks.
Watchkeeping engineer is forced to re-use annealed copper washer gasket	OP - Organisational Factor : Operating Procedures.
Previous modifications to the forward filter cover/bolt sealing surface had removed the seating groove for the copper washer and left the sealing surface uneven.	TE - Technical Factor: Engineering.
Watchkeeping engineer restarts engine having failed to find any leaks during the initial tests.	HK1 - Human Behaviour : Knowledge Based : System Status.
Adjacent indicator tap is unshielded. Adjacent exhaust manifold is unshielded.	TC - Technical Factor: Construction.

Table 11.9: PRISMA (Stage 2) Summary Table

by encouraging analysts to associate multiple root causes with each of the causal factors that are identified during previous stages of analysis. This approach is not generally encouraged [845]. There is a danger that the unnecessary proliferation of root causes will hide important information about the factors that contributed to an incident.

Table 11.9 identifies a number of root causes that stem from human factors problems. The Chief engineer and the mechanical assistant remained in the control room from 15:15 to 16:00. They could not observe the port side of the engine room from this position and so failed to observe the fire as it began to take hold. This can be interpreted as a rule-based failure to perform necessary checks. It can be argued that the watchkeeping engineer's decision to restart the engine after failing to find any leaks was the result of a knowledge-based failure in their interpretation of the state of the system. Such findings must, however, be treated with caution. The Transportation Safety Board of Canada investigators provided very little information about the decision to restart the engine. It is, therefore, difficult to be certain of the root causes that may have influenced the engineer's behaviour. It was not anticipated that so many root causes would relate to human factors problems in our analysis of the Nanticoke case study. The ordering of the Eindhoven Classification Model considers there factors after other technical and organisational factors. The analysis may reflect the reliance upon human intervention in the Nanticoke case study. Our findings might also be unnecessarily biased by the evidence that was available in the aftermath of this incident.

The final stage in any PRISMA analysis is to identify 'recommended' actions that might address each root cause. PRISMA provides a classification/action matrix to support this task. These tables link each category of the classification model to a ranked list of interventions. These responses are ordered according to their perceived cost effectiveness. They may relate to improved acquisition or equipment design, to better procedures, information management or communication, to revised training practices or motivational activities [841]. The exact nature of the table will vary from industry to industry and from organisation to organisation. The effectiveness of particular recommendations can be affected by the wider safety culture in a company. It can also be influenced by the financial and other resources that are available to an investigator. In consequence, the entries in a classification/action matrix are likely to change over time. Safety reviews are liable to identify new rankings for the effectiveness of particular recommendations.

The classification/action matrices represent an important aspect of PRISMA that has not been addressed by the other causal analysis techniques in this Chapter. ECF, MES and STEP focus on the events leading to an incident or accident. MORT does provide means of analysing the response to an incident. Recommendations from any previous incidents should ensure that an oversight or omission becomes an assumed risk. These "are defined as only those risks that have been analysed and accepted by the proper level of management; unanalysed or unknown risks are not considered to be Assumed Risks" [204]. None of these techniques provides explicit means of ensuring a consistent response to similar incidents. Not does it provide means of monitoring the effectiveness of that response.

	Organisational Factors				
	External Factors (O-EX)	Knowledge Transfer (OK)	Operating procedures (OP)	Manag. priorities (OM)	Culture (OC)
Inter-departmental communication	X				
Training and coaching		X			
Procedures and protocols			X		
Bottom-up communication				X	
Maximise reflexivity					X

Table 11.10: Example PRISMA Classification/Action Matrix [845]

Table 11.10 illustrates the general format of the classification/action matrices that are advocated by the PRISMA approach. This particular example is derived from the medical classification model. The increased number of organisation categories in this model provides an interesting insight into the nature of medical incidents when compared with the abridged version in the original Eindhoven model, illustrated in Figure 11.9 [845]. As can be seen in Table 11.10, incidents that involve a failure in knowledge transfer within an organisation might result in revised training and coaching practices. Failures that stem from problems involving operating procedures will, as expected, result in revised procedures and protocols. The precise nature of such tables is determined by the context in which any recommendations will be applied. Individual organisations may also be forced to increase the level of detail that is represented within Classification/Action matrices such as that shown in Table 11.10. For example, a recommendation to improve training and coaching is not at a sufficient level of detail to encourage confidence that any recurrence will be avoided. The motivation behind this technique is summarised by Van Vuuren who argues that:

> "However, the incident data clearly shows decreased risk awareness and vigilance as main contributors to adverse group behaviours, leading to incidents. Therefore, an organisation should reflect on its safety experiences and try to learn as much as possible from them. The correct level of risk awareness and vigilance can be maintained by reporting and analysing the often abundantly available near misses. Based on these analyses, feedback to the organisation can be provided to show the dangers of day to day practice. This way, a continuous circle of learning from its own safety experiences and measuring the safety performance of the organisation results." [845]

It is, however, possible to apply elements of Table 11.10 to the Nanticoke case study. Previous stages of the analysis argued that the re-use of the annealed copper washer gasket stemmed from a failure in organisational operating procedures. The re-use of copper gaskets that had previously been deformed under high operating pressures should not have been permitted. As might be expected, Table 11.10 suggests that this root cause might be combatted by revising the procedures and protocols that govern current maintenance practices.

A number of limitations can be identified for the PRISMA technique. Some of these relate to particular features of this approach, others are more general criticisms of checklist techniques. PRISMA, like MORT, offers greatest support after primary and secondary investigations have been completed. It depends upon investigators being able to construct the causal trees that have been illustrated earlier in this section. The proponents of PRISMA do, however, urge that the application of this approach should be based around critical incident interviews based on a technique developed by Flanagan in the 1950's [252]. This interview technique encourages individuals to describe situations in which the success or failure of an operation was determined by specific causes. It is argued, by extension, that the same approaches can be used to elicit information about mishaps for which the causes are less certain. This utility of this elicitation technique has been validated by considerable fieldwork. It also integrates well with the generation of causal trees that are intended to capture both good and poor performance. Critical incident interviews can, however, only provide part of the evidence that is necessary for the causal analysis of complex, technological failures. For example, it is unclear how information from automated logging systems or from regulatory documents might be integrated into these 'anecdotal' accounts. Similarly, there is little guidance about how to address the increasing complexity of many near-miss incidents, which involve individuals and systems from many different organisations and working groups.

The practical application of PRISMA has been assessed in a number of studies. For example, this approach has been used to identify the root causes of incidents from NASA's Aviation Safety Reporting System [530]. Investigators were trained to use a variant of the Eindhoven Classification Model. They were then asked to independently classify the same group of incident summaries. The intention was to assess interrater reliability using the PRISMA method. The results indicated that subjectivity might be less of an issue than has been claimed for checklist approaches, however, the investigation raised more questions than it addressed. More interestingly, this study identified a number of fundamental misconceptions that arose when investigators were trained to apply the PRISMA technique. For example, one participant in the trial was unhappy that they were able to provide an unambiguous classification for all of the incidents that were studied. They then went back to the dataset until they could classify some incidents under the X - unclassifiable category. Such incidents are instructive for a number of reasons. Firstly, they point to the difficulty of training investigators to use even simplified forms of the existing analytical techniques. Secondly, they point to the way in which individual differences can influence the successful application of these techniques. None of the other participants expressed this concern that some incidents should not be classified by the existing model! It is important to emphasise that these concerns are not simply centred on the PRISMA approach but can potentially affect all of the analytical techniques described in this book.

It is also possible to identify a certain confusion about the distinction between causal factors and root causes in the PRISMA technique. Van Vuuren has argued that root causes can be identified as the leaf nodes in the left-hand branches (i.e., the non-mitigating branches) of a causal tree [845]. In his view, classification model simply provide a means of grouping these root causes into categories that are amenable to statistical analysis. Managers can use the results of the classification process to monitor, for instance, how many incidents are caused by problems with operating procedures in a given time period. This is an interesting approach because, in some ways, it is the antithesis of MORT. Root causes are represented by the upper nodes of the MORT diagram. In Van Vuuren's view of PRISMA, root causes are denoted by the lower leaf nodes of a causal tree. The difference becomes apparent if we compare the leaf node Steering pump main power cables and the control wiring from bridge run through common tray past the port generator from the PRISMA causal tree with the corresponding *why* branch from the MORT analysis Barrier: Did not provide. As can be seen, the MORT approach more closely resembles our requirement that root causes should be more general than the causal factors that characterise a particular incident. In consequence, the previous pages have adopted the convention of referring to the non-mitigating leaf-nodes of a causal tree as *causal* *factors* and the elements of a classification model as *root causes*.

A number of general criticisms can be made about checklist approaches such as PRISMA and MORT [444].Previous paragraphs have already argued that investigators may be dissuaded from searching for potential root causes that do not appear on a checklist. Further biases can affect the selection of items within a checklist. For instance, items at the top or the bottom of a list are more likely to be selected than those in the middle [457]. Similarly, if certain classes of causal factors occur more frequently in a checklist then there is an increased likelihood that those factors will be identified. MORT provides an extreme example of this in which all root causes can be linked to management failures, neglecting regulatory, environmental or other workplace factors. Kjellén has also argued that investigators and supervisors are more likely to choose those causal factors on a checklist that are difficult to verify or that involve limited management responsibility [444]. There is an increased tendency to select factors that relate to individual failures or to adverse factors that are 'beyond the control' of senior and middle management. This partly explains MORT's bias towards managerial factors.

Checklist approaches also suffer from the wide range of biases that have been noted in previous chapters. Attribution errors make it more likely that investigators will select transient or environmental causal factors if they are implicated in an incident [444]. This is less likely to occur when investigators belong to an independent investigation agency. We have also seen how the lack of event-based models can also create problems for checklist-based approaches. techniques such as ECF analysis, MES and STEP provide a map of events that can be used to trace the development of an incident over time. If additional evidence becomes available then this can be directly used to revise these temporal models. In contrast, it can be more difficult to trace the impact of new information on the causal analysis supported by checklist techniques. Information about particular events can be distributed throughout the stage 1 and stage 2 tabular forms that support any MORT analysis. Similarly, it can be difficult to reconcile the temporal and causal relationships that are embedded within the gates of a causal tree.

11.2.3 Tripod

Previous sections have reviews a number of techniques to support causal reasoning about adverse occurrences. None of these techniques has, however, explicitly recognised the distinctions between catalytic and latent failures that has been emphasised in previous chapters. In contrast, the Tripod techniques were deliberately developed to account for this important distinction. The Tripod research project started in 1988 from a collaboration between the Universities of Leiden and Manchester. This collaboration has produced a range of analytical techniques. Tripod-Delta supports the predictive analysis of potential causal factors without the need for accident and incident statistics. Tripod-Beta provides more focussed support for incident and accident investigation [702]. The underlying analytical techniques have been widely used within the petrochemical industry [374, 854]. It is important to emphasise, however, that Tripod is not simply an accident or incident analysis technique. It's proponents argue that it offers a coherent philosophy based on the precept that safety management is essentially an organisational control problem.

Figure 11.11 sketches the model of incident and accident causation that underpins the Tripod method. It also illustrates the three key concepts that motivate the use of the name Tripod. Incidents and accidents provide important information about underlying, or root causes, of systems failure. These underlying or latent conditions are referred to as General Failure Types. As we shall see, they stem from the organisational, managerial and regulatory practices that create the preconditions for failure. The final leg of the tripod is provided by the active failures or unsafe acts that trigger an incident. These unsafe acts initiate hazards that can be mitigated by the proper use of barriers or may ultimately develop to compromise the safety of the target [206].

Tripod also provides a framework for thinking and for measuring the disturbances that affect safe operations. This measurement is based upon the General Failure Types mentioned above. These have strong similarities to the branches in a classification hierarchy, such as a MORT diagram or the Eindhoven Classification Model. There are also important differences. For instance, Tripod-Delta's measurement of potential disturbances to safe practice does not rely upon incident or accident

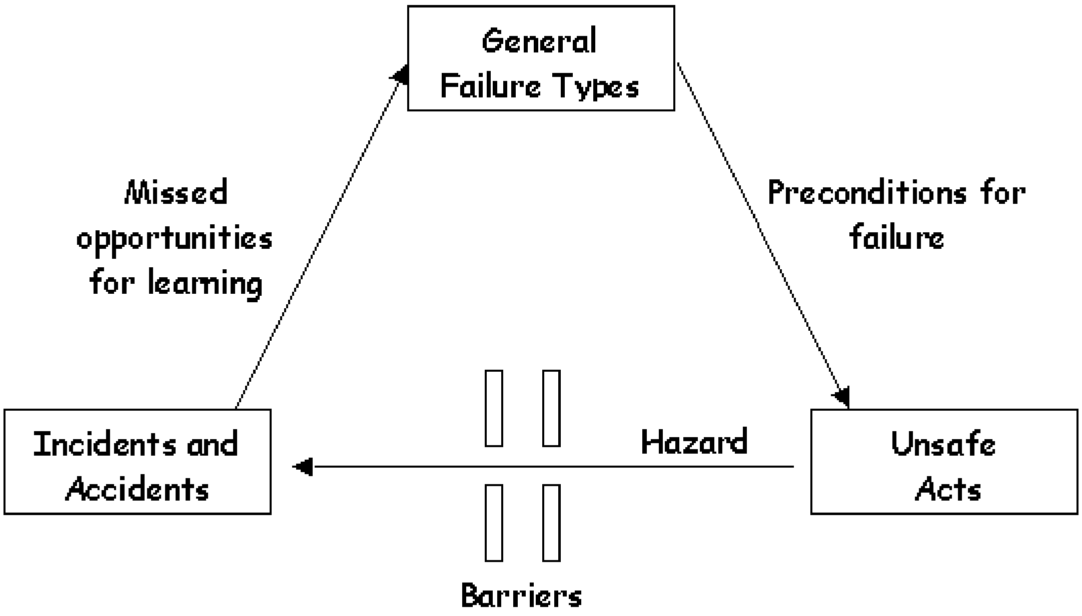

Figure 11.11: The Three Legs of Tripod

statistics. This contrasts strongly with the US Department of Energy's proposal to derive aggregate values for the root causes identified by MORT analysis.

Tripod relies upon an underlying model of causation. This assumes that incidents are caused by local triggering factors that combine with more latent General Failure Types. It is also assumed that organisations can do little to predict or address these local triggering factors. Reason uses the analogy that they are like mosquitos [702]. It does little good to swat at them individually; it is far better to drain the swamp in which they breed. In this case, the swamp represents the latent General Failure Types. These can be summarised by the following list. Each item emerged through close study of previous incidents rather than through any explicit empirical investigation. It should also be stressed that some failure types have consequences that promote other 'knock-on' failure types. For instance, inadequate maintenance management can lead to working conditions that increase the likelihood of operator error:

1. *Hardware.* Unsafe acts often result from the provision of inadequate equipment and materials. This can be the result of poor stock control, of problems in the supply chain, of component defects etc.

2. *Maintenance management.* Unsafe acts may also stem from the management rather than the execution of maintenance activities. For example, an incident may occur because necessary maintenance work was delayed or postponed.

3. *Design.* Unsafe acts can occur if designers fail to provide operators with sufficient information about the purpose and reliability of a device. Similarly, designers may provide inadequate information about the range of safe interventions that can be made with a device. They may also provide users with insufficient feedback about the state of a device.

4. *Operating procedures.* Unsafe acts may stem from procedures that either could not be applied in a given context or which contained dangerous advice or which contained advice that could not physically be complied with by an operator. Procedures may also be ambiguous in their application and in the guidance that they offer.

5. *Violation-inducing conditions.* Unsafe acts can stem from workplace or environmental pressures that encourage violations or discourage compliance. These factors may also promote erroneous behaviour, for instance, by exposing operators to hostile working conditions.

6. *Housekeeping.* Many incidents are caused by failures that are well known but which have not been adequately addressed over a long period of time. For example, problems in maintenance management can lead to hardware problems that become compounded over time.

7. *Incompatible goals.* Incidents can occur because individuals may be preoccupied or have goals that conflict with those that are intended to ensure the safety of the system in which they operate. The goals and working practices of groups can conflict with those of others within an organisation. Finally, there may be conflict between organisational objectives, such as profit or public approval, and safety.

8. *Communication.* Mishaps can be the result of system failures that impair communications channels. They can also stem from lost signals even when it is physically possible to transmit a message. For example, a safety warning might be delivered to the wrong person within an organisation. Even if a messages is successfully received, it can be misinterpreted or may arrive too late to ensure the safety of an application.

9. *Organisation.* Organisational structures can prevent individuals from responding to the lessons provided by safety-related incidents. For example, there may be divided responsibilities or conflicts of interest.

10. *Training.* Mishaps can occur if personnel lack the competence required to complete necessary tasks. This can occur if training is inadequately prepared, if it is curtailed, if it is not validated as providing the necessary instruction etc.

11. *Defence planning.* Mishaps can also occur if there are deficiencies in the detection, mitigation and remedial actions that are taken in the aftermath of an incident.

In common both with MORT and several other checklist approaches [387], General Failure Types stem from management decisions. Within each of these General Failure Types it is possible to distinguish two different levels of cause. Functional failures stem from decisions made by line managers, by designers, by planners etc. In contrast, source failures refer to more strategic decisions at senior management level. This has some similarities to the broad categories within Tier Analysis, described in Chapter 9.3.

As mentioned above, Tripod-Delta can be used in a pre hoc manner. It does not depend upon incident or accident statistics. This is important because, as we have seen, the insights that are provided by these information sources can be marred by under-reporting or by analytical biases. Reason argues that domain and task specialists can devise questions that will test for the presence of different General Failure Types before an incident occurs. For example, workers on an offshore platform might be asked 'was this platform originally designed to be unmanned?' or 'are shutoff valves fitted to a height of more than two meters?' [702]. These questions are intended to elicit highly focussed responses that are indicative of the more general General Failure Types, listed above. Software support has been developed to help administer these questionnaires. Approximately, twenty indicators are identified for each of the eleven General Failure Types. Once operators have completed these questions, the system compiles a bar chart that represents a Failure State Profile. This bar chart lists the General Failure Types according to the number of 'incorrect' questions that were answered by the operator. For example, the system asks twenty questions that relate to each General Failure Type. If eleven of the questions about hardware failures raised a potential cause for concern but only six of the questions about communication were answered in this way then hardware might be interpreted as a greater priority than communications issues. This represents a relatively crude interpretation of the analysis. It is recommended that the software be used three or four times a year and that any consequent decisions are based on trends rather than one-off values. For example, if we assume that an operator answered ten if ten out of twenty answers that the operator provided about hardware failure indicated that this was a significant cause for concern then this

General Failure Type would be ranked above any other failure types that The key point in all of this is that the questions, or indicators, help to trace the symptoms of a problem. The General Failure Types capture the underlying causes of future safety problems.

Tripod-Delta provides a general tool that can be used without incident and accident statistics. In contrast, Tripod-Beta was developed to provide incident analysis tools that can be used as an investigation progresses [217]. This explicit intention to support the investigatory process is similar to the motivation behind event cards in STEP. It contrasts sharply with the assumption in techniques, such as MORT or PRISMA, that the investigatory process has been largely completed. The Tripod-Beta software provides investigators with guidance about the elicitation process. As might be expected, investigators are prompted to go beyond the local triggers to identify latent General Failure Types. Hence, Tripod-Beta was deliberately intended to be compatible with Tripod-Delta.

Tripod-Beta analysis exploits many of the concepts that were introduced during the discussion of Barrier Analysis in Chapter 9.3. Investigators begin by identifying the targets that were affected by a potential hazard. They then have to trace the manner in which individual barriers were compromised during an incident. This is, typically, done by constructing a form of causal tree. At the root of the Tripod-Beta tree is an active failure that helped to compromise one of the barriers, mentioned above. The second level of the tree describes preconditions that had to be satisfied in order for the active failure to occur. For example, Figure 11.12 uses Tripod to analyse active and latent failures during the Nanticoke incident. This failure might have been prevented by barriers that were intended to avoid the release of fuel or by visual inspections once the initial fire had started. The first of these barriers was compromised by the engineer's active failure to ensure a fuel-tight seal for the filter gasket when he restarted the engine. The visual inspections were jeopardised by the restricted field of view that was afforded by the Chief Engineer's and Mechanical Assistant's decision to remain in the Engine Control Room.

In order for an active failures to occur it is necessary for a number of preconditions to be satisfied. These preconditions, typically, relate to the general failure types that were introduced in previous sections. Figure 11.12 provides several examples of this aspect of Tripod-Beta modelling. The watchkeeping engineer's difficulties in achieving a fuel-tight seal were exacerbated by the lack of new, spare copper washer gaskets. This precondition stemmed from a latent failure to identify the importance of these items within the spare parts inventory . This latent failure can, in turn, be associated with the *hardware* general failure type. These hardware failures stem 'from the provision of inadequate equipment and materials' and are the result 'of poor stock control, of problems in the supply chain, of component defects etc'. It can also be argued that the failure to ensure an adequate stock inventory helped to create and was created by *error enforcing conditions*. The fact that the engineer was forced to anneal an existing gasket introduced additional sub-tasks into the preventive maintenance programme. It can argued that this reduced the amount of time available for monitoring and inspection of the generator after it had been reassembled.

A number of further preconditions contributed to the engineer's decision to restart the generator, in spite of the problems that they subsequently reported for their maintenance activities. The modifications to the fuel cover removed the seating groove that helped to ensure an adequate seal. The Watchkeeping Engineer also failed to find any leaks during their initial observation of the generator after the preventive maintenance had been completed. These preconditions are, in turn, be associated with underlying general failure types. Unlike the problems with the stock inventory, mentioned above, it is possible to identify a number of common failure types that may have affected both of these preconditions . For example, *housekeeping* failures relate to problems that have been known for a long time and which have not been adequately addressed. It can be hypothesised that the Engineer did not express concern over the modifications to the forward fuel cover nor did they conduct prolonged inspections of the reassembled generator because the problems that they experienced in obtaining a seal were not unusual. A similar argument might also justify the use of the *communication* and *training* general failure types to characterise the reasons why key personnel failed to report the problems that they faced during maintenance procedures.

The previous paragraphs illustrate the way in which an informal argument must be constructed to explain and justify the decisions and judgements that are represented in Figure 11.12. This is important if other analysts are to understand and accept the reasons why, for instance, the failure to

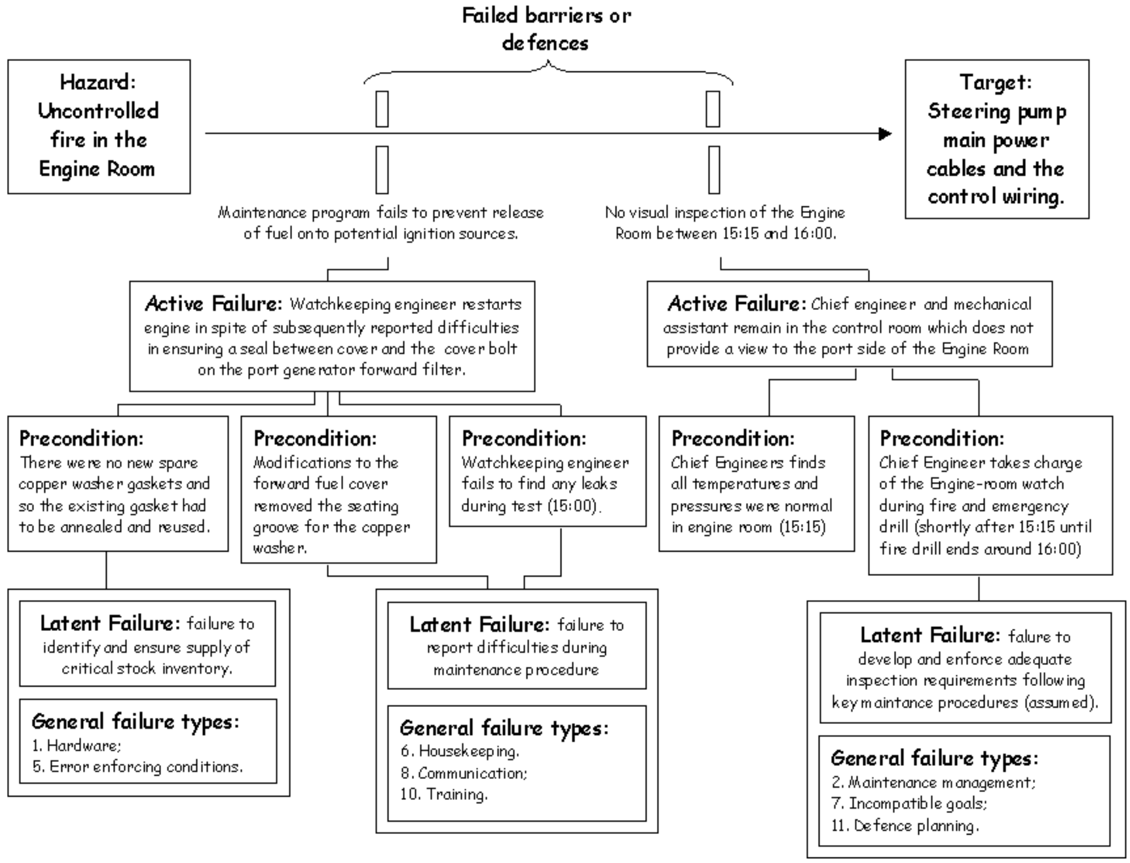

Figure 11.12: Tripod-Beta Event Analysis of the Nanticoke Incident (1)

report maintenance difficulties can be seen as an instance of a more general housekeeping problem. Chapter 8.3 has explained how the need to provide such rationale is a more general requirement for many analytical techniques. It is especially important when these techniques capture subjective judgements about the underlying causes of a mishap, such as the Nanticoke incident. Other investigators may disagree with the allocation of general failure types represented in Figure 11.12. The provision of a free-text rationale for that allocation can, therefore, be used during subsequent analysis. Additional evidence may also be sought to support assertions about the state of the seating groove during previous maintenance procedures and about the more general reporting of maintenance problems onboard the Nanticoke before this incident.

Preconditions can be thought of as causal factors . They are necessary for an active failure to occur. For instance, the lack of any spare, new gaskets was a necessary precondition for the Watchkeeping Engineer's failure to ensure a seal. Individual precondition need not, however, provide sufficient conditions for an active failure to occur. For example, if the Engineer had detected a leak during their subsequent tests then they might not have decided to restart the generator even if they had been forced to re-use an annealed gasket. The necessary and sufficient conditions for an active failure are represented by the conjunction of all of the preconditions associated with that failure. For example, it was necessary for there not to be any spare gaskets and for modifications to have removed the seating groove and for tests to indicate there were no leaks in order for the Watchkeeping engineer to start the generator. This analysis suggests further links between Tripod-Beta and other techniques, such as ECF analysis and MES, that exploit counterfactual reasoning. For each precondition, analysts must be sure that the associated active failure would not have occurred if that precondition had not been satisfied.

Figure 11.12 represents preconditions that can explain the Chief Engineer's and the Mechanical Assistant's failure to monitor the port side of the Engine Room. The Chief Engineer observed normal temperature and pressure readings in the Engine Room at 15:15. As can be seen, this precondition is not associated with a latent failure or with a general failure type . This is justified because it does not represent a failure. The Chief Engineer correctly monitored the available readings. This was as a precondition for the active failure because it may have reassured him that there were no problems after the preventive maintenance had been completed around 15:00.

The failure to monitor the port side of the engine room may also have been caused by the change in watch that occurred during emergency and fire drills. It is normal practice for Chief Engineers on merchant ships to assume control of the Engine Room during fire and emergency drills. This enables other members of the crew to participate in the exercise while ensuring that normal watchkeeping activities are not compromised. In the Nanticoke incident, the Chief Engineer relieved the watchkeeping engineer who had completed the generator maintenance. This enabled the watchkeeping engineer to proceed to his fire station. This hand-over may, however, have played an important role in the development of the fire. It can be argued that the fire and emergency drills created a context in which the crew were less likely to perform their normal inspection activities. Such interruptions to normal operating procedures can often result in reduced vigilance. Fire and emergency drills provide opportunities for social interaction that are less frequent under the demands of everyday operation. It can also be difficult to ensure that adequate information is handed over from one operator to another. In particular, the Watchkeeping Engineer did not report their difficulty in obtaining a fuel-tight seal. If these concerns had been expressed then the Chief Engineer might have maintained a direct visual observation of the Port-side generator. All of these concerns might have been addressed by the use of operating procedures to ensure that an adequate watch was maintained during the fire and emergency drill [623]. As before, this latent failure is associated with a number of more general failure types. It reflects a failure in *maintenance management*, a problem with *incompatible goals* and potentially with *defence planning*. The maintenance management concerns centre on the need to specify and follow adequate monitoring guidelines during the fire drill after the generator had been restarted. Mishaps are likely to occur if individuals are preoccupied or have goals that conflict with those that are intended to ensure the safety of the system in which they operate. It can be argued that the Chief Engineer's role in assuming the watch during the fire drill might have introduced social or technical demands that impaired their ability to continue monitoring the engine room. Finally, incident can also occur if there are 'deficiencies in the detection, mitigation and remedial actions that are taken in the aftermath of an incident'. This general failure type summarises the role that the active failure played in the incident as a whole, the crews' failure to monitor the port side of the engine room delayed the detection of the fire while it was still taking hold.

Tripod-Beta offers a number of benefits for the causal analysis of safety-critical incidents. In particular, the graphical representation of defences helps to ensure that analysts explicitly consider the way in which active and passive failures combine to jeopardise potential barriers. This is important because other techniques, such as ECF and MES, only consider barriers in an indirect manner. It is possible, however, to raise a number of minor caveats about the manner in which defences are represented in Tripod-Beta. Previous applications of this technique focus on the way in which defences have failed. For example, Figure 11.12 shows how the Nanticoke incident stemmed from a failure to prevent the release of fuel onto a potential ignition source and from a failure to inspect the engine while the resulting fire took hold. A continuing theme in this book has, however, been that near-miss incidents also provide important information about successful defences. This is important if engineers and designers are to accurately assess whether or not those defences can be relied upon to ensure the future safety of a potential target. Figure 11.13, therefore, shows how the conventional use of Tripod-Beta can be extended to consider the role of successful defences and barriers as well as those that are known to have failed. In spite of the Chief Engineers failure to perform a direct visual inspection of the port side of the engine room between 15:15 and 16:00, he did notice the high cooling temperature water alarm that eventually promoted the crews' response to this incident. Figure 11.13 also illustrates a number of additional defences that were not tested during this incident . This is important because, as we shall see, any recommendations must also consider what

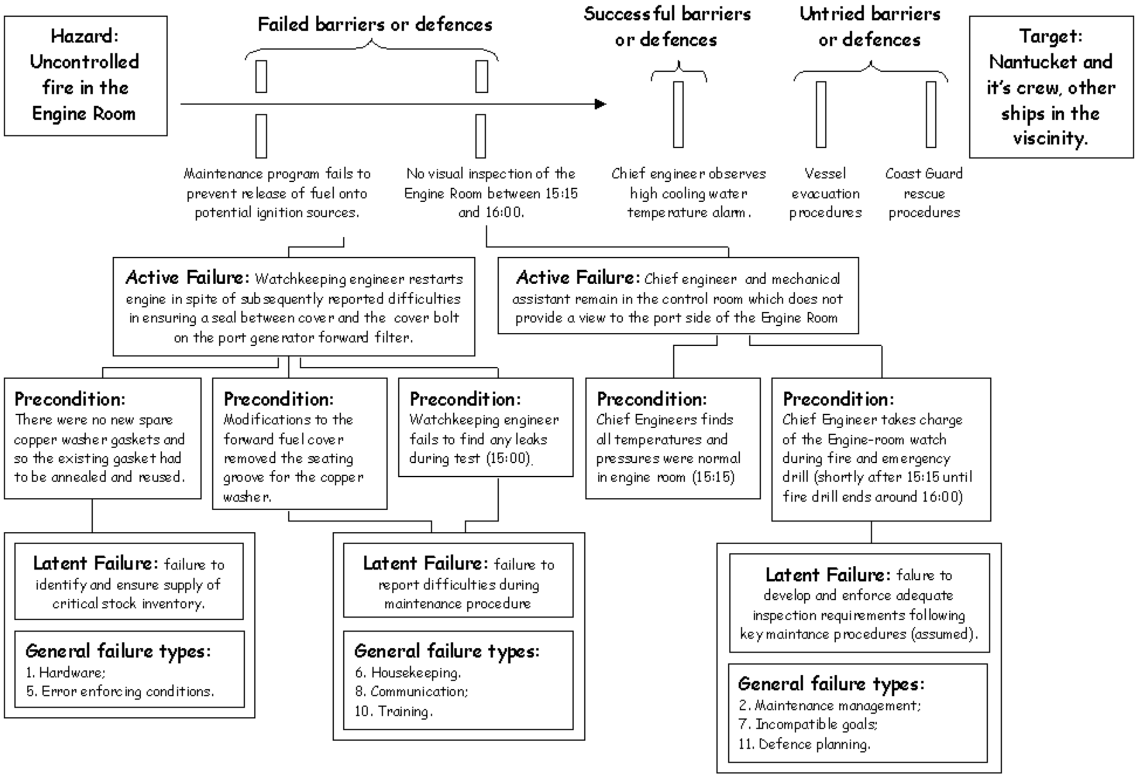

Figure 11.13: Tripod-Beta Event Analysis of the Nanticoke Incident (2)

might have happened if the successful defence had also been compromised. Tripod-Beta also offers a number of further benefits. These can be summarised as follows:

- *focussed use of a time-line.* The event analysis component of the Tripod-Beta technique includes a time-line that shows some similarity with those that are used in ECF analysis, in MES and STEP. This is represented by the horizontal arrow that potentially connects a hazard to a target in Figures 11.12 and 11.13. Unlike the alternative analysis techniques, Tripod-Beta focuses on those events that are associated with the failure of defences. This considerably simplifies the modelling of an incident or accident. The sparse approach advocated by Tripod-delta also omits information, such as the actors involved in an event, that forms an important component of the MES and STEP approaches;

- *explicit representation of active and latent failures.* The distinction between latent and active failures reflects much recent research into the nature of technological failure [702, 364]. It is intended to move the focus of an analysis away from the individual failures that characterise a particular incident to look for more general managerial and organisational causes. In other techniques, such as ECF analysis, this distinction is only recognised through the use of auxiliary techniques such as Tier analysis;

- *support for a checklist approach to root cause analysis from the eleven general failure types.* The general failure types in Tripod-Beta are similar to the leaf nodes within PRISMA classification models. They describe a number of recurring 'root causes'. They support analysts by directing their attention to recognised causes of previous incidents. This, in turn, can encourage greater consistency between investigators than might otherwise be possible with techniques that do not exploit a checklist approach.

- *balance between a high level of abstraction in the general failure types and more specific information from the use of preconditions.* It is possible to contrast the eleven general failure types supported by Tripod-Beta with the one thousand five hundred items in a full MORT diagram. It is far easier to perform an initial analysis using this limited number of general failure types than it is to perform an exhaustive search through a MORT diagram. Conversely, it can be more difficult to identify general failure types that accurately characterise the particular root causes of incidents within a particular industry or organisation. PRISMA avoids this problem by combining a relatively simple checklist, which is similar to aspects of Tripod-Beta, with a recommendation that analysts extend the classification scheme to reflect local conditions within particular industries. For instance, Van Vuuren's medical checklist includes an item for 'patient related factors' [845]. This item is not included within Van der Schaaf's PRISMA taxonomy for chemical incidents [842]. It can, however, be argued that such differences can introduce important inconsistencies between the results of causal analyses that were obtained using different classification schemes. Tripod-Delta avoids some of these problems by explicitly representing the relationship between specific details of an incident, in the annotations associated with active failures and with preconditions, and the more general root causes. These annotations can be used to stress particular aspects of an incident that cannot easily be captured using the restricted palette offered by the eleven general failure types.

- *tool support.* Finally, the application of Tripod-Beta is supported by a number of computer-based tools. This is significant because these systems can also be integrated with the constructive use of Tripod-Delta as part of a wider safety management programme. The Tripod-Beta tools provide a number of internal consistency checks that help analysts to construct the event analysis diagrams, illustrated in Figures 11.12 and 11.13. It is important to stress, however, that our analysis was conducted without the use of these tools. This provided greater flexibility, for example in the representation of successful barriers in Figure 11.13, that might not be so desirable if an organisation were keen to ensure greater consistency between the event diagrams that were produced by incident analysts.

The benefits of Tripod-Beta analysis must be balanced against a number of potential problems. In particular, this technique raises concerns that are similar to those that motivated Benner to omit conditions from the STEP approach. It can be difficult to distinguish between active failures and preconditions. For instance, Figure 11.12 argued that the Watchkeeping Engineer's failure to find any leaks was a precondition for their active failure in restarting the engine without reporting a potential maintenance problem. It might be argued that the failure to detect any leaks should be classified as an active failure in its own right. This would result in a graph in which an active failure is the result of both preconditions and of active failures. Each of these active failures might, in turn, be the result of further preconditions and further active failures and so on. The ECF analysis in Chapter 9.3 has illustrated the complexity of a similar approach. This technique might have even worse consequences for Tripod-Beta; ECF charts do not distinguish between active and passive failures.

To summarise, preconditions introduce a potential ambiguity into Tripod-Beta modelling. They capture information about the state of a system; modifications to the forward fuel cover removed the seating groove for the copper washer. They also capture event-based information; watchkeeping engineer fails to find any leaks during test (15:00). This creates ambiguity because these events may themselves represent active failures that can be associated with further pre-conditions. In practice, it is possible to develop a number of heuristics that reduce the consequences of such ambiguity. For instance, Figures 11.12 and 11.13 only consider the preconditions of those active failures that are directly associated with the failure of particular barriers. The analysis does not consider the preconditions of a precondition. If analysts wanted to consider the Watchkeeping Engineer's failure to find any leaks then that event would have to be associated with the failure of a particular barrier at the top level of the Tripod-Delta diagram.

The application of Tripod-Beta has also shown how analysts must provide considerable additional documentation to support the diagrammatic form illustrated in Figures 11.12 and 11.13. In particular, it is important to explain why particular latent failures can be associated with general failure types. Similarly, rationale must be provided so that other analysts can understand the relationship between a latent failure and a particular precondition. Our analysis of the Nanticoke case study illustrated this issue when we considered the possible impact that the fire and emergency drills might have had upon the monitoring of the port side of the engine room. In order to interpret the relationship between the precondition, latent failure and general failure types, analysts must understand the manner in which responsibilities and tasks are routinely handed-over so that other members of the crew can participate in the drill. It was also necessary to draw upon evidence from previous failures to explain the problems that can arise from the transfer of information during such hand-overs. This additional information illustrates the manner in which the Tripod-Beta event analysis diagram is not an end in itself. It provides a high-level framework for the causal analysis of incidents and accidents. It does not, however, replace the more general inferential and reasoning skills that are established by expertise and training in incident analysis.

11.3 Mathematical Models of Causation

Previous sections have introduced a number of semi-formal techniques that are intended to support the causal analysis of safety critical incidents. They can be classified as 'semi-formal' because it can be difficult to develop a coherent set of rules to describe the syntax and semantics of the associated notations. For instance, we have identified some of the problems that can arise when attempting to construct a precise definition of the preconditions that form an important component of Tripod's event analysis diagrams. Similarly, it can be difficult to derive a precise definition for what can and what cannot be represented in the leaf nodes of a causal tree. Investigators are free to use natural language annotations. This increases the flexibility of the approach. It can, however, also introduce potential ambiguity and inconsistency if a team of investigators must cooperate in the construction of a shared tree during a PRISMA analysis. A number of organisations have responded to these problems by developing more formal, mathematically based, causal analysis techniques.

11.3.1 Why-Because Analysis (WBA)

Why-Because Analysis stems from an initiative to increase the objectivity of accident investigations by encouraging "rigorous causal analysis" [469]. The technique is based around two complementary stages. These can be summarised as follows:

1. *Construct the Why-Because Graph.* The first stage in the WBA involves the construction of a graph that is intended to capture the significant causal relationships that led to an incident. The causal relationships are identified using the counterfactual reasoning that has been a feature of previous approaches. The method is, however, supported by a formal semantics for causation that is based on that provided by the philosopher and logician David Lewis, mentioned in previous chapters [490, 491].

2. *Prove that the Graph is Sufficient and Correct.* The previous techniques that have been presented in this chapter and in Chapter 9.3 would stop after stage 1 of the WBA. In contrast, however, this logic-based technique provides procedures for ensuring that the causal relations in a Why-Because graph actually satisfy the semantics for causation that is implied by Lewis' underlying model. In other words, there are rules for showing that the model of an incident reflects Lewis' view of causation. These techniques can also be used to ensure that there is a sufficient causal explanation for each identified fact that is not itself a root cause [469].

The following pages provide a brief introduction to these two stages of analysis. It is important to emphasise, however, that the benefits provided by the mathematical underpinning of WBA can also important impose considerable upon the analyst. The various stages of the technique can appear to be extremely complex even for investigators who have a background in mathematical logic. As we shall see, therefore, this approach may be most suitable for near-miss incidents that might under other circumstances have resulted in high-consequence accidents.

As mentioned, the first stage of WBA involves the construction of a graph that is intended to capture the causal relationships that lead to incidents and accidents. The nodes of these graphs represent four different factors: states; events; processes and non-events [499]. States are represented by collections of state predicates. These can be thought of as sentences that are true in that state. For example, the ignition of the Nanticoke fire might be represented by state in which it was true that 'fuel is being sprayed under pressure from the forward fuel filter of the port generator'. WBA uses angled brackets to denote individual states, $\langle State \rangle$. Events represent changes in state. For instance, the deployment of the Halon system is an event that transformed the state of the Nanticoke from one in which there was a fire to one in which there was no fire. WBA uses brackets to denote individual events, $[Event]$. Processes can be defined to describe mixtures of states and events that have some bounded duration. For example, the Nanticoke incident can be described in terms of a process in which the maintenance event transformed the state of the system into one in which a fire could occur. The ignition event changed the state of the system into one in which a fire was taking place and so on. WBA uses curling brackets to denote processes, $\{Process\}$. Finally, as we have seen, it is often necessary to consider the impact that errors of omission have upon the course of an incident. WBA, therefore, provides non-events using the following notation $(non - events)$

WBA proceeds by developing a history of the incident. Successive states of the system are liked using a temporal ordering relation that is denoted using the $\hookrightarrow$ symbol. For more information on the semantics of the $\hookrightarrow$ operator, see Lamport [473]. For now it is sufficient to observe that it forms part of a more complex Explanatory Logic that was developed by Ladkin and Loer to provide means of formally demonstrating the correctness of a causal argument [470, 499]. The initial stages of the Nanticoke case study can be represented by the following high-level history:

$$\langle Maintenance \rangle \hookrightarrow \langle Fire \rangle \hookrightarrow [Deploy\ Halon\ System] \tag{11.1}$$

It is important to emphasise that the temporal ordering, captured by the $\hookrightarrow$ symbol, does not represent causality. Loer illustrates the distinction between causation and temporal sequence [499]. A traffic-jam may occur immediately after I leave the highway, however, there need not be any causal relationship between these two events unless I have parked my car across the carriage-way. A number of axioms can be used to describe important properties that must exist between temporal and causal relations. For example, if a causal chain exists such that A causes B then the first element of this causal chain, A, must occur before the last element, B. This leads to the following inference rule:

$$\frac{A \Rightarrow\!\!\!\succ^* B}{A \hookrightarrow B} \tag{11.2}$$

If we know that A causes B, $A \Rightarrow\!\!\!\succ^* B$, then we can also conclude that A must precede B, $A \hookrightarrow B$. If this rule were not to hold then past events could be the result of situations that still lie in the future!

To summarise, we would like to be able to construct a causal model of an incident using the $\Rightarrow\!\!\!\succ^*$ operator. Most primary and secondary investigations result in temporal models, similar to those proposed in Chapter 8.3. These describe sequences that can be represent using the $\hookrightarrow$ operator. Unfortunately, there is no automatic means of translating temporal sequences into causal relations. Many different causal chains can produce the same high-level temporal sequence. For instance, the (11.1) sequence might have been caused by maintenance to the starboard generator, to the transmission system and so on. Investigators must apply their skill and expertise to identify the causes of the temporal sequences that can be reconstructed in the aftermath of an incident. Fortunately, WBA provides an informal procedure that helps in this task. This process starts by asking *Why did the final event in the sequence occur?*. For the Nanticoke example in (11.1) this would yield:

Why was the Halon system deployed?.
Because the second fire party withdrew from fighting the fire.

The analysis continues by asking, in turn, why did the second fire party withdraw? This was because they were ordered to abandon their attempt to extinguish the fire. As mentioned, the key

Why-Because questions are intended to guide the process by which the temporal $\hookrightarrow$ sequences are translated into more detailed causal relations, $\Rightarrow\!\!\!\succ^*$. However, this process may also help to identify factors that were not considered during the initial temporal sequence. For instance, the previous questions helped to identify that the failure of the second fire party was a reason why the Halon system was deployed. Our previous analysis did not include any information about either the first or the second fire party. Figure 11.14 illustrates how this recursive analysis can be used to identify the reasons why the Halon system was deployed. The first fire party's attempt to use carbon-dioxide extinguishers was beaten back by the heat of the fire. This led to a second fire party attempting to use charged hoses. This attempt was ordered out of the engine room which then led to the Chief Engineer discharging the Halon system.

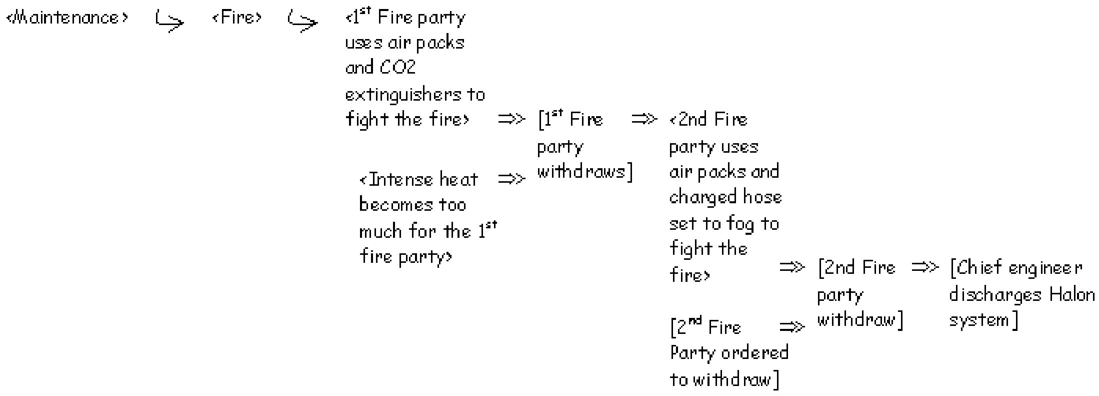

Figure 11.14: Why-Because Graph Showing Halon Discharge

A number of observations can be made about the Why-Because graph illustrated in Figure 11.14. As can be seen, the maintenance and fire states that were identified in (11.1) continue to be connected by the sequence relation, $\hookrightarrow$. However, the causal analysis has helped to identify states and events that are causally related, $\Rightarrow\!\!\!\succ$. Formally, $\Rightarrow\!\!\!\succ$ is the transitive closure of $\Rightarrow\!\!\!\succ^*$. Informally, $A \Rightarrow\!\!\!\succ B$ denotes that A is a direct causal factor of B. $A \Rightarrow\!\!\!\succ^* B$ represent situations in which there may be intermediate or 'knock-on' causal relations. For example, in Figure 11.14 we can say that the withdrawal of the first fire party is a direct causal factor behind the 2nd fire party's use of the hoses to fight the fire, denoted using $\Rightarrow\!\!\!\succ$. In contrast, the withdrawal of the first fire party is a knock-on cause of the Chief Engineer's action to discharge the Halon system, denoted using $\Rightarrow\!\!\!\succ^*$.

Why-Because graphs, typically, use a numerical indexing system rather than the free-text labels that are shown in Figure 11.14. $\langle Maintenance \rangle$ might be denoted by $\langle 1 \rangle$, $\langle Fire \rangle$ by $\langle 2 \rangle$ and so on. This has not been done because the graph is relatively simple and the labels are intended to help the reader trace the causes of the Halon deployment. However, this approach quickly becomes intractable as the scope of the graph increases.

It is possible to perform a number of consistency checks using the formal rules that underpin the graphical notation that is provided by Why-Because graphs. The simplest of these involves checking that the causal relationships are consistent with the previous temporal order described in (11.1) using the $\hookrightarrow$ operator. Or more formally, the analyst must ensure that the transitive closure of the causal relations in Figure 11.14 continue to preserve the temporal sequence of (11.1) [499].

It should also be noted that, as might be expected, it can be difficult to determine how best to represent an incident using the four factors that form the nodes of a Why-Because graph: $\langle State \rangle$; $[Event]$; $\{Process\}$; $(Non-events)$. As mentioned, analysts must decide whether a particular aspect of an incident is best represented as a state, en event, a process or as a non-event. It is relatively straightforward to distinguish an event from a non-event. It can, however, be more complex to determine what is an event and what is a process. For example, Figure 11.14 shows that the discharge of the Halon system was a discrete event. It can also be argued that the task of deploying this form of extinguisher is more likely to have been composed from a sequence of events and could, therefore, be better represented as a process. The Chief Engineer must form the intention to deploy the system. He must then ensure that everyone is accounted for and that none is left in the area in which the system will be deployed. There may have been a confirmation protocol to inform the Captain the system was to be deployed etc. This decision between an event or a process is typical of the choices that must be made when using many different causal analysis techniques. It reflects the level of detail that the analyst considers to be necessary when constructing a model of an incident or accident. The key point is that the model explicitly represents this information so that other analysts can review their colleague's view of an incident and, if necessary, request additional detail.

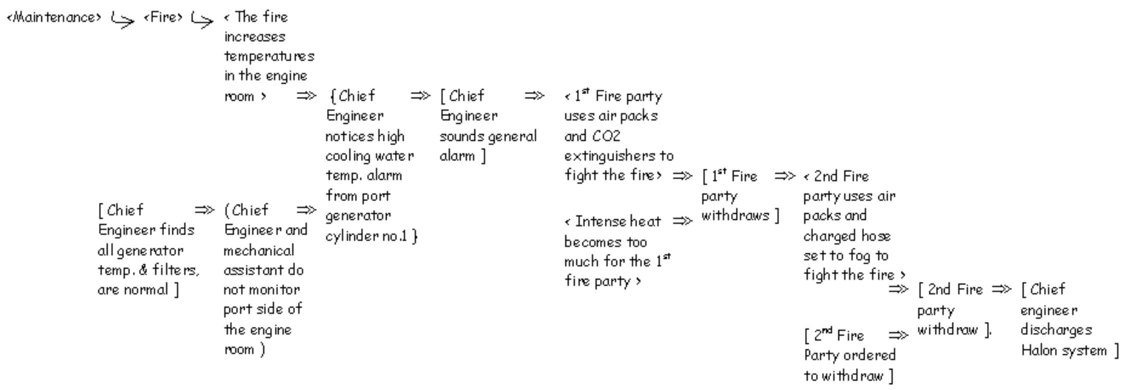

Figure 11.15: Why-Because Graph for the Nanticoke Alarm

Figure 11.15 extends the analysis to consider the reasons why the first fire party was called on to combat the fire in the first place. As can be seen, they were responding to a general alarm. Why had the general alarm been issued? Because the Chief Engineer had noticed the high cooling water temperature alarm. Why had the Chief Engineer noticed this alarm? Because the fire had increased temperatures in the engine room. The Chief Engineer had also noticed this alarm because he and the mechanical assistant had not monitored the port side of the engine room and so had not noticed the fire before it took hold. Why had they not monitored the port side of the engine room? Because an initial inspection had not shown anything unusual with the generators.

Figure 11.15 illustrates a number of further properties of Why-Because graphs. For instance, the reason that the Chief Engineer eventually observes a high cooling water temperature alarm is because they and the Mechanical Assistant fail to monitor the port side of the engine room. This is denoted as a $(non - event)$. In order to capture the semantics of these non-events, the Explanatory Logic of WBA draws upon deontic arguments of obligation and permission. The crewmembers violated the procedures and norms that obliged them to maintain a visual watch over the engine room. Ladkin and Loer provide a full justification for this application of deontics [470]. For now it is sufficient to observe that WBA provides a meta-rule that is intended to guide investigators in the identification of these non-events. Investigators must explicitly add a non-event, (E), to the history of states if $O\langle E\rangle$ is derivable and E does not occur, where O represents deontic obligation [470]. Figure 11.15, therefore, includes the non-event Chief Engineer and Mechanical Assistant do not monitor port side of the engine room.

Figure 11.15 also illustrates the way in which the $\{Process\}$ format provides powerful abstractions that can be used to describe complex causal sequences. For instance, there are likely to be a number of perceptual and cognitive mechanisms that led the Chief Engineer to notice the high cooling water temperature alarm. Subsequent analysis could recruit human factors experts to identify these factors. During any initial analysis, however, the details of this cognitive and perceptual process can be denoted as {Chief Engineer notices high cooling water temperature alarm from port generator cylinder number 1}. The process form is also used to represent the human factors mechanisms that led the Chief Engineer to conclude that there were no problems during their initial inspection of the engine room.

Figure 11.16 illustrates the results of applying WBA to the factors in the temporal sequence that was introduced in (11.1). As mentioned, the formal underpinnings of this analytical technique are intended to ensure that investigators can represent and reason about the products of their investigations. This helps to ensure that errors are avoided during the construction of relatively complex Why-Because graphs, such as that illustrated in Figure 11.16. These 'quality control' procedures take two principle forms. The first approach uses information about each node to ensure that a Why-Because graphs satisfy a number of high level properties. For example, investigators must ensure that each node has at least one causal factor that represents an event. They must also ensure that each node is classified exclusively as one of the four factors mentioned above.

Additional constraints can be imposed, for example, to ensure that investigators minimise the use of processes wherever possible. This injunction is justified because processes should not be used as a 'catch all' when investigators find it difficult to discriminate between events and states. In other words, they should not be used to mask or hide aspects of an incident that ought to be the subject of a more detailed investigation. For instance, Figure 11.16 might be refined to consider what exactly attracted the Chief Engineer's attention to the cooling water high-temperature alarm. It is important to stress that WBA was developed to support the investigation of accidents rather than near-miss incidents. More limited analytical and investigatory resources may, therefore, prevent individuals from obtaining the evidence that is necessary to resolve processes into their component states and events. There may be other processes, such as { 1st party decides to withdraw } in Figure 11.16, that may involve complex perceptual, cognitive and physiological 'states' or 'events'. Such processes are difficult to analyse. As we have seen, investigators may be forced to assume intention from observed behaviour. The proponents of WBA have developed the Perception, Attention, Reasoning, Decision, Intention and Action (PARDIA) model to help analyse such processes. Loer stresses that PARDIA should be used to classify rather than to understand error [499]. This is a fine distinction given that he constructs a normative model of intention. The details of this model are beyond the scope of the current work. It should be noted, however, that PARDIA focuses on cognitive, perceptual and physiological attributes of single operators. It, therefore, suffers from some limitations when applied to team-based incidents and accidents. As we have seen, group dynamics often lead to situations in which team-based behaviour cannot simply be described as the 'sum of its members'.

Automated tools have been developed to assist with the checks, described above. This is important because an error in writing an event node as a state, or an event node which only has states as causal factors, can result in a consistency review on the sub-graph leading to this node. Formal proof techniques provide an alternative means of ensuring the integrity of WBA. As we shall see, however, the costs of performing this analysis may dissuade investigators from going 'the full distance' on this form of analysis [469].

A number of benefits can be derived from the close relationship between WBA and philosophical work on the nature of causation [499]. For instance, investigators must often explain why one version of events is more plausible than another. Lewis has proposed the idea of *contrastive explanation* as a technique that can be used to support these arguments about plausibility [492]. If we have to decide between two versions of an incident we must assess the evidence that is derived through a primary and secondary investigations. In addition, we can also contrast the causal explanation of those histories as revealed using techniques such as the WBA. This approach resembles earlier arguments in this chapter; causal analysis often identifies the need to provide further evidence in support of hypothesised causal relations. An important application of this idea is that any causal analysis must not only explain why an incident occurred in a particular way, it must also explain why the system did not function in the manner intended. For example, Loer analyses an incident in which a DC10 landed at Brussels rather than its intended destination of Frankfurt Airport [499]. He uses WBA to contrast the actual incident, in which the aircraft landed in Brussels, with the "deontically-correct" world in which the aircraft was supposed to land at Frankfurt. His analysis proceeds by identifying the earliest contrast between what actually did happen and what was supposed to happen. In this case, the aircraft was transferred to Brussels Air Traffic Control rather than Maastricht . Figure 11.17 shows how this technique might be applied to the Nanticoke incident.

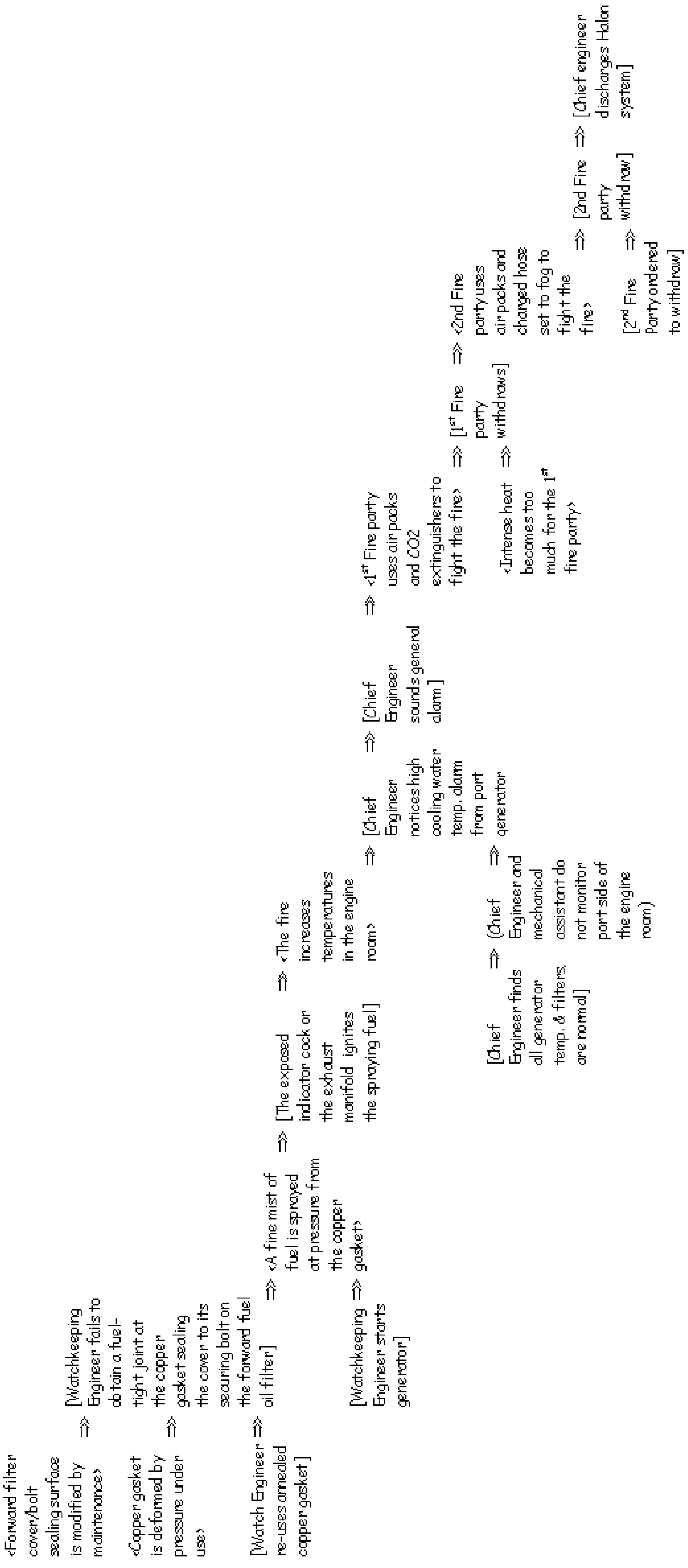

Figure 11.16: Overview of the Why-Because Graph for the Nanticoke Incident

The maintenance modifications that damaged the forward filter cover/bolt sealing surface occurred before the copper gasket was deformed or the maintenance engineer annealed the gasket. This event might, therefore, provide a good starting point for any contrast between what did happen and what ought to have happened. It can be argued that maintenance personnel should have noticed the damage and reported it through a management system. This should have resulted in the surface being made good before a fire could occur. The ? ⇒? symbols are used to distinguish causal links from this "possible" world in Figure 11.17. In practical terms, this analytical technique is useful because it can be used to identify non-events that might otherwise be omitted from an analysis. We could redraft the graph in Figure 11.17 by replacing the possible worlds with (Maintenance personnel do not notice the damage to the sealing surface). This particular application of contrastive explanations has much in common with barrier analysis. It can be used to explain the failure of a defensive mechanism that was intended to ensure that the system returned to a 'normative' state.

Figure 11.17: Possible 'Normative' Worlds for the Nanticoke Incident

A number of minor issues complicate our application of contrastive explanations. Loer's example of this technique is relatively straightforward [499]. The normative and non-normative paths diverge from the point at which the DC10 was handed to Brussels and not Maastricht Air Traffic Control. The Nanticoke incident is not quite so straightforward. It is also important to emphasise that our Why-Because graph ends with the deployment of the Halon system. This is justified because it is important to learn about the resolution of adverse incidents as well as the causes of any failure. One consequence of this is that we cannot simply look for the earliest contrast with a possible world in which the Halon was not deployed. We must also ensure that the alternative 'normative' world avoid a fire. As mentioned, there is a relatively simple divergence in Loer's DC10 case study. In contrast, the Nanticoke case study contains several points at which non-normative and normative behaviours can be distinguished. The Chief Engineer was supposed to monitor the engine room.

The two fire parties were supposed to be deployed before the fire required the use of Halon. By focusing on the earliest contrastive explanation, analysts might miss important lessons about other failures that contributed to the course of an incident.

There are further complications. We have identified the earliest contrast between what did and what should have happened as the maintenance on the filter cover and bolt sealing surfaces. Notice, however, that the previous Why-Because graphs did not specified any temporal sequence over this state and the other two initial causal factors that describe the deformation and re-use of the copper gasket. This sequence was inferred from the evidence that was obtained in the aftermath of the incident. It would, however, also be possible to contrast possible worlds in which the deformation of the gasket was monitored or in which the maintenance engineer did not have to re-use an annealed gasket. As before, this analysis can be used to help investigators explain why particular barriers failed to protect the system. For instance, additional nodes might be introduced into the Why-Because graph to indicate that the absence of a maintenance reporting system explains why these factors were not addressed prior to the incident. It is important, however, that investigators recruit evidence to justify their assertions about these potential defences. For instance, it is not clear that the incident would have been avoided even if the maintenance issues had been effectively reported. The absence of necessary parts or delays in maintenance scheduling might still have led to an adverse occurrence.

The previous paragraphs have described how an informal analysis of alternative possible worlds can be used to distinguish 'normative' from 'non-normative' behaviour. This is useful in identifying ways in which barriers, including regulations, working practices and automated systems, failed to prevent an incident from occurring. As we have seen, however, there are also situations in which investigators cannot distinguish between alternative causal explanations. For instance, we do not know whether the Nanticoke fire was ignited by the exposed indicator tap or by the exhaust manifold. We could use the $? \Rightarrow ?$ notation to describe two divergent causal paths. One might indicate that the indicator tap ignited the fire, the other might represent the exhaust manifold as the ignition source. This can create considerable additional complexity as almost half of the graph would be duplicated. Lewis suggests that these alternative explanations should be ranked by experts using some weighting mechanism [492]. If an alternative explanation was considered sufficiently unlikely then it can be omitted from subsequent analysis. There are a number of concerns about whether this is possible either in the general case or in the example of the Nanticoke fire [673, 469]. Loer advocates the retention of these different paths but acknowledges the consequent complexity [499]. We have, therefore, retained the node labelled The exposed indicator tap or the exhaust manifold ignites the spraying fuel. We have not duplicated the rest of the Why-Because graph, however, because the consequences of these two possible worlds are indistinguishable. Neither of these approaches provides a more general solution to this problem and it remains a subject for future research. It should also be noted that non-determinism complicates the application of all causal modelling techniques. This is most clearly seen in the closing sections of Chapter 9.3 where we recognised the difficulty of using ECF to model alternative causal hypotheses about the loss of the Mars Climate Orbiter and the Polar Lander.

The previous paragraph argued that investigators can construct different Why-Because graphs to represent alternative causal explanations. These alternative explanations can be thought of as 'possible worlds'. For instance, there is one possible world in which the Nanticoke fire was caused by the indicator tap and another in which it was caused by the exhaust manifold. This notion of alternative possible worlds provides WBA with a semantics for the counterfactual arguments that investigators use to identify causal factors. Chapter 6.4 distinguished causal factors using the argument:

> A is a necessary causal factor of B if and only if it is the case that if A had not occurred then B would not have occurred either.

Lewis [492] recasts this in the following manner:

> "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".

Ladkin and Loer formalise this definition as follows:

$$\frac{\begin{array}{l} A \wedge B \\ \neg\, A \mathrel{\Box\!\!\rightarrow} \neg\, B \end{array}}{A \Rrightarrow B} \tag{11.3}$$

Informally, $A \mathrel{\Box\!\!\rightarrow} B$ captures the notion that B is true in possible worlds that are close to those in which A is true. As can be seen, (11.3 uses this operator to express the counterfactual component of the Lewis definition. As mentioned, Loer and Ladkin provide a more detailed presentation of this application of Lewis' work [499, 470]. The key point, however, is that logic can be used to provide a clear semantics for informal concepts such as 'cause'. Investigators can also use associated proof rules to ensure both the consistency and sufficiency of informal reasoning about the causes of incidents and accidents.

The formal underpinnings of the Explanatory Logic in WBA help to determine whether those causes that are identified by an informal analysis provide a *sufficient* explanation for an incident or accident. Ladkin and Loer introduce the notion of a causal sufficiency criterion [470]. This is based on the argument that for causal relations $A_1 \Rrightarrow B, A_2 \Rrightarrow B, ..., A_n \Rrightarrow B$ the $A_1..A_n$ form a sufficient set of causal factors for B if it would be impossible for B not to happen if $A_1..A_n$ had happened. More formally $A_1..A_n$ form a sufficient set of causal factors for B if and only if:

$$\begin{array}{l} \wedge A_1 \Rrightarrow B \\ \wedge A_3 \Rrightarrow B \\ \wedge ... \\ \wedge A_n \Rrightarrow B \\ \wedge \neg\, B \mathrel{\Box\!\!\rightarrow} \neg\, (A_1 \wedge A_2 \wedge ... \wedge A_n) \end{array} \tag{11.4}$$

From this, Loer goes on to introduce the $\mathrel{\Box\!\!\Rightarrow}$ operator to denote both a necessary and sufficient causal relationship. He argues that the goal of the causal sufficiency criterion is to show that:

$$A_1 \wedge A_2 \wedge ... A_n \mathrel{\Box\!\!\Rightarrow} B \tag{11.5}$$

In order to establish such a relationship, analysts can exploit the following rules:

$$\frac{\begin{array}{l} C \\ \neg\, C \Rrightarrow \neg\, B \\ \neg\, B \Rrightarrow \neg\, C \end{array}}{C \mathrel{\Box\!\!\Rightarrow} B} \tag{11.6}$$

$$\frac{\begin{array}{l} A \Rrightarrow C \\ B \Rrightarrow C \end{array}}{(A \vee B) \mathrel{\Box\!\!\Rightarrow} C} \tag{11.7}$$

These rules provide a framework for reasoning about the sufficiency of the semi-formal Why-Because graphs. For example, the left most factors in Figure 11.16 describe a causal relationship between a number of maintenance failures and the initial release of fuel. This relationship can be represents as follows:

$$\begin{array}{l} \mathit{Step}\ 1\ \ (\mathit{Theorem}): \\ \langle \mathit{Forward\ filter\ cover/bolt\ sealing\ surface\ is\ modified\ by\ maintenance} \rangle \wedge \\ \quad \langle \mathit{Copper\ gasket\ is\ deformed\ by\ pressure\ under\ use} \rangle \wedge \\ \qquad [\mathit{Watch\ Engineer\ reuses\ annealed\ copper\ gasket}] \mathrel{\Box\!\!\Rightarrow} \\ \{\mathit{Watch\ engineer\ fails\ to\ obtain\ fuel-tight\ join\ at\ the\ copper\ gasket} \\ \qquad \mathit{sealing\ the\ cover\ to\ its\ securing\ bolt\ on\ the\ forward\ fuel\ oil\ filter}\} \end{array}$$

We can prove this relationship using Loer's meta-rule for deriving the causal sufficiency criterion [499]. The following paragraphs retain the labels that were introduced in the Why-Because graph. This is intended to make the steps of the proof more accessible. Later sections will, however, explain why these annotations might be replaced by predicates with a more precise interpretation. For now it is sufficient to observe that the use of these 'informal' labels makes it difficult to typeset the steps of the proof in a conventional format:

$$\begin{array}{l} \textit{Step } 2.1 \ \ (\textit{Using } 11.6): \\ \langle \textit{Forward filter cover/bolt sealing surface is modified by maintenance} \rangle \wedge \\ \quad \langle \textit{Copper gasket is deformed by pressure under use} \rangle \wedge \\ \quad\quad [\textit{Watch Engineer reuses annealed copper gasket}] \end{array}$$

Proof: We can assume that this conjunction is true providing that adequate evidence can be obtained in the aftermath of the incident. The Transportation Safety Board of Canada provided photographic evidence to support these assumptions [623].

$$\begin{array}{l} \textit{Step } 2.2 \ \ (\textit{Second obligation from } 11.6): \\ \neg\, (\langle \textit{Forward filter cover/bolt sealing surface is modified by maintenance} \rangle \wedge \\ \quad \langle \textit{Copper gasket is deformed by pressure under use} \rangle \wedge \\ \quad\quad [\textit{Watch Engineer reuses annealed copper gasket}]) \ \Box\!\!\rightarrow \\ \neg\, \{\textit{Watch engineer fails to obtain fuel} - \textit{tight join at the copper gasket} \\ \quad \textit{sealing the cover to its securing bolt on the forward fuel oil filter}\} \end{array}$$

$$\begin{array}{l} \textit{Step } 3.1 \ \ (\textit{Using De Morgan's Law}): \\ \neg\, \langle \textit{Forward filter cover/bolt sealing surface is modified by maintenance} \rangle \vee \\ \neg\, (\langle \textit{Copper gasket is deformed by pressure under use} \rangle \wedge \\ \quad\quad [\textit{Watch Engineer reuses annealed copper gasket}]) \ \Box\!\!\rightarrow \\ \neg\, \{\textit{Watch engineer fails to obtain fuel} - \textit{tight join at the copper gasket} \\ \quad \textit{sealing the cover to its securing bolt on the forward fuel oil filter}\} \end{array}$$

$$\begin{array}{l} \textit{Step } 4.1 \ \ (\textit{Using } 11.7): \\ \neg\, \langle \textit{Forward filter cover/bolt sealing surface is modified by maintenance} \rangle \ \Box\!\!\rightarrow \\ \quad \neg\, \{\textit{Watch engineer fails to obtain fuel} - \textit{tight join at the copper gasket} \\ \quad\quad \textit{sealing the cover to its securing bolt on the forward fuel oil filter}\} \end{array}$$

Proof: This is true if and only if the engineer obtains a seal in all nearest possible worlds to those in which the forward filter sealing surface is not modified. Given that there was a supply of new gaskets, in other words assuming that the second part of the disjunction in Step 3.1 is false, then the only other way in which the seal could be compromised was through modifications that were not authorised by the manufacturer [623].

$$\begin{array}{l} \textit{Step } 4.2 \ \ (\textit{Using } 11.7): \\ \neg\, (\langle \textit{Copper gasket is deformed by pressure under use} \rangle \wedge \\ \quad\quad [\textit{Watch Engineer reuses annealed copper gasket}]) \ \Box\!\!\rightarrow \\ \neg\, \{\textit{Watch engineer fails to obtain fuel} - \textit{tight join at the copper gasket} \\ \quad \textit{sealing the cover to its securing bolt on the forward fuel oil filter}\} \end{array}$$

$$\begin{array}{l} \textit{Step } 5.1 \ \ (\textit{Using De Morgan's Law}): \\ \neg\, \langle \textit{Copper gasket is deformed by pressure under use} \rangle \vee \end{array}$$

$$\neg\,[\textit{Watch Engineer reuses annealed copper gasket}] \;\Box\!\!\rightarrow$$
$$\neg\,\{\textit{Watch engineer fails to obtain fuel} - \textit{tight join at the copper gasket sealing the cover to its securing bolt on the forward fuel oil filter}\}$$

$$\textit{Step } 6.1 \;\; (\textit{Using } 11.7):$$
$$\neg\,\langle \textit{Copper gasket is deformed by pressure under use} \rangle \;\Box\!\!\rightarrow$$
$$\neg\,\{\textit{Watch engineer fails to obtain fuel} - \textit{tight join at the copper gasket sealing the cover to its securing bolt on the forward fuel oil filter}\}$$

Proof: This is true if and only if the engineer obtains a seal in all nearest possible worlds to those in which the copper gasket is not deformed under pressure. Additional expert validation is needed to support this argument.

$$\textit{Step } 6.2 \;\; (\textit{Using } 11.7):$$
$$\neg\,[\textit{Watch Engineer reuses annealed copper gasket}] \;\Box\!\!\rightarrow$$
$$\neg\,\{\textit{Watch engineer fails to obtain fuel} - \textit{tight join at the copper gasket sealing the cover to its securing bolt on the forward fuel oil filter}\}$$

Conjecture: This is true if and only if the engineer obtains a seal in all nearest possible worlds to those in which the engineer does not reuse an annealed gasket. This theorem is refuted in the following paragraphs.

$$6.3 \; \textit{Q.E.D. From } 11.7 \textit{ to } 6.1 \textit{ and } 6.2$$
$$5.2 \; \textit{Q.E.D. From De Morgan's law applied to antecedent of } 5.1$$
$$4.3 \; \textit{Q.E.D. From } 11.7 \textit{ to } 4.1 \textit{ and } 4.2$$
$$3.2 \; \textit{Q.E.D. From De Morgan's law applied to antecedent of } 3.1$$

$$\textit{Step } 2.3 \;\; (\textit{Third obligation from } 11.6):$$
$$\neg\,\{\textit{Watch engineer fails to obtain fuel} - \textit{tight join at the copper gasket sealing the cover to its securing bolt on the forward fuel oil filter}\} \;\Box\!\!\rightarrow$$
$$\neg\,(\langle \textit{Forward filter cover/bolt sealing surface is modified by maintenance} \rangle \wedge \langle \textit{Copper gasket is deformed by pressure under use} \rangle \wedge [\textit{Watch Engineer reuses annealed copper gasket}])$$

Proof: This is true if and only if the engineer obtains a seal in all nearest possible worlds to those in which the filter is not modified by maintenance and the copper gasket is not deformed under pressure and the engineer does not re-use an annealed copper gasket. If the filter cover had not been modified and the gasket had not been deformed by pressure and been reused then there is no evidence to suggest that the seal would have failed. As before, this argument must be carefully validated by domain experts [196].

$$2.4 \; \textit{Q.E.D. From } 11.6 \textit{ applied to } 2.1,\; 2.2 \textit{ and } 2.3 \qquad \Box$$

This proof illustrates how mathematically-based, specification techniques can be used to support the semi-formal structures in a Why-Because graph. As can be seen, the first stage in the proof was to derive a formal representation for the causal relationships that are represented in the left-hand nodes of Figure 11.16. This formalisation provided the theorem that we sought to establish through the use of Loer's meta-rules for the proof of a sufficient causal explanation. The key point here is that these meta-rules provide a template to guide further proofs of the remaining causal relationships in this diagram. Step 1 could be redrafted to formalise these relationships. Steps 2-6 can then be updated. Investigators simply provide the supporting arguments shown for steps 2.1, 4.1, 6.1, 6.2 and 2.3 [499].

This guidance is important because it can help investigators to identify potential weaknesses in their informal reasoning. For example, step 6.2 denoted a causal relationship that is true if and only if the engineer obtains a seal in all nearest possible worlds to those in which the engineer does not reuse an annealed gasket. On closer inspection, it is difficult to defend this argument. Modifications to the seating surface might have compromised the ability of the engineer to achieve a seal even if they had access to a supply of new copper gaskets. Even though we can question this proof step, the the overall proof need not fail. Step 3.1 shows how the argument depends on a disjunction. Step 4.1 has already established the first case and so we need not establish the remainder of the disjunction in order to demonstrate the remainder of the proof. This formal analysis yields several insights. In particular, it illustrates that the annealing of the gasket may not be a necessary cause of the leak. In contrast, the deformation of the gasket and the modifications to the sealing surface together provide the necessary and sufficient causes of the leak, denoted by $\Box\!\!\Rightarrow$.

The previous analysis identified a potential weakness in the previous arguments that have been presented throughout this chapter. The re-use and annealing of the copper gasket need not have been a causal factor in the leak. This argument could prompt investigators to pursue a number of different courses of action. Firstly, they might accept these criticisms and amend the Why-Because graph by omitting the node labelled [Watchkeeping engineer re-uses annealed copper gasket]. Alternatively, further validation might be needed before the results of this formal analysis can be accepted as part of the investigation. This is an important point because incident investigators who are skilled in a particular application domain are unlikely to be familiar with the reasoning techniques that were illustrated in previous pages. In consequence, expert validation is required to support the informal arguments that are made to support the 'Proof' stages for steps 2.1, 4.1, 6.1, 6.2 and 2.3.

The informal arguments that support the previous formal proof are important for a number of reasons. They help to ensure that non-formalists can validate the underlying assumptions that support the formal template or structure that supports the overall causal argument. They also indicate the depth to which investigators want to pursue the formal analysis. The key point here is that it is possible to pursue the formal reasoning beyond the level that was demonstrated in the previous example. For instance, the node (Chief Engineer and Mechanical Assistant do not monitor the port side of the engine room) could be represented by the following clause:

$$\begin{aligned}\neg\ (&attend(chief\_engineer, port\_engine\_room) \vee \\ &attend(mechanical\_assistant, port\_engine\_room))\end{aligned} \tag{11.8}$$

These clauses might then support the extension of formal reasoning techniques from the overall argument structure, shown as the meta-rule given above [499], into the informal arguments that are denoted by the 'Proof' stages for steps 2.1, 4.1, 6.1, 6.2 and 2.3. Ladkin and Loer note that this 'level' of formalisation depends:

> "... on how one wants to analyse the situation; how much one wants to say, what depth and detail of analysis one wants to pursue, the limitations of the language chosen to express the nodes. All of this is very much the choice of the investigator... A similar situation exists in pure axiomatic mathematics. One is provided with sufficient proof rules to get the job done, but what proofs are constructed and how are up to the individual wishes and skill of the user. Proofs may be more detailed or less detailed, easy to follow or cleverly slick, pro forma or creative. Yet the criteria for a valid proof remain constant throughout the enterprise. So with WBA. We have no wish to regulate whether an analysis is most subtle, or how it indicates what future steps to take to prevent recurrences, or whether it must use the latest theory of human-computer interaction. We wish to lay out criteria and reasoning rules for providing a formally-complete causal explanation, according to assumptions that an analyst makes in a particular case. We, thereby make the assumptions clear, explicit and precise, exhibit their role in the explanation, and make the reasoning clear." [470]

The source nodes in a Why-Because graph represent the reasons for an incident or accident. They represent necessary causal factors for an incident or accident. They can easily be identified because they do not have any incoming causal links. This has one very important consequence. Source nodes can be thought of as contingencies that might have be avoided precisely because they lack any necessary causal factors. Table 11.11 summarises the source nodes in Figure 11.16. The rows of this table describe the 'failures' and 'errors' that directly contributed to the incident. They also describe events that might have been appropriate in other contexts. For example, Watchkeeping engineer starts generator need not have caused any problems if everything else had been functioning correctly. However, this event in combination with the failure to obtain a fuel-tight seal led to the initial fire. The compilation of these tables can be used as a further validation for the analytical technique. For example, investigators may be required to justify any decision not to decompose processes into their component factors. The proponents of WBA also argue that source node lists can be used to develop procedures that might avoid particular combinations of adverse events. For example, engineers might be prevented from re-using annealed copper gaskets. Alternatively, maintenance modifications that jeopardise a fuel-tight seal might be closely monitored by supervisory staff.

Factors	Label
State	Forward filter cover and bolt sealing surface is modified by maintenance.
State	Copper gasket is deformed by pressure under use.
Event	Watch engineer re-uses annealed copper gasket.
Event	Watchkeeping engineer starts generator.
Process	Watchkeeping engineer finds all generators and filters are normal.
Process	1st fire party decides to withdraw.
Event	2nd fire party ordered to withdraw.

Table 11.11: Source Node Analysis of Nanticoke WBA Graph

As with the previous analytical techniques in this chapter, it is possible to identify a number of strengths and weaknesses that characterise WBA. For example, the entries in Table 11.11 can be compared to the Transportation Safety Board of Canada's findings about the cause of the Nanticoke incident:

> "The fire was caused by a leakage of fuel, which contacted an exposed exhaust manifold, from the forward fuel filter on the port generator. Contributing to the occurrence was the modification to the fuel filter cover, the re-use of the copper sealing gasket on the cover, the unshielded hot exhaust surfaces adjacent to the filter, and the less-than-adequate engine-room watchkeeping duty during the fire drill before the occurrence." [623]

As can be seen, there is a strong agreement between the informally derived observations of the investigation team and our application of the Why-Because technique. There are, however, a number of important differences. For example, the investigators stressed the significance of the proximity of an exposed ignition source which does not appear as a source node in Figure 11.16. This is a significant omission on our part. The ignition of the fire was represented on the graph as an internal node. We should have added a source state to denote the fact that the indicator tap and the exhaust manifold were exposed. This could have been avoided if the analysts had acquired greater expertise in WBA. It might also have detected during peer review or through a more sustained formal analysis of the causal model. Such omissions are, however, a powerful reminder that even sophisticated analytical techniques are ultimately dependent on the skill and expertise of the individuals who constrict and manipulate the abstractions that they provide.

Having acknowledged the strengths of a traditional 'informal' approach, it is also important to identify potential insights yielded by the more formal style of analysis. Table 11.11 does not simply focus on the causes of the incident itself. It also contains information about the failure of mitigating factors, such as the fire fighting teams. The discipline of listing source nodes can help to check whether the causes of these 'subsidiary' failures are considered in sufficient detail. Table 11.11 helps to reveal, for example, that we have not explained the process by which the first fire party decided to withdraw or the events that led to the order for the 2nd fire party to abandon their work. Additional analysis must be conducted to determine the precise reasons why these attempts were beaten back and, more importantly, whether they were an appropriate response given the state of the fire as it was observed by the crew. This aspect of the incident is, arguably, not considered in sufficient detail by the official report into the incident.

Strauch raises a number of caveats about the application of WBA to the Cali accident [166]. He argued that particular events on a Why-Because graph ought to be distinguished as being more important that others. For example, some decisions have a greater impact on the course of an incident than others. WBA would identify both as 'equal' causes:

> "...not decisions are equal at the time they are made ... each decision alters the subsequent environment, but that while most alterations are relatively benign, some are not. In this accident, this particular decision altered the environment to what became the accident scenario." [764]

These are interesting comments from an individual who has considerable first-hand experience of incident and accident investigations. They could, however, be applied to all of the causal analysis techniques that we have reviewed in this book. The possible exception to this criticism would be the analytical techniques devised to support the application of MORT. As we have seen, investigators can sum the frequency of *what* factors that are associated with *why* nodes to get a raw measure of their relative importance. Weights can also be used to discriminate between the importance of these different failures with common causes. Such techniques suffer from the difficulty of validating any weighting mechanisms that might be used. For instance, how would an investigate discriminate between the relative importance of the deformation of the gasket and the lack of monitoring during the early stages of the fire? Such distinctions are likely to introduce a degree of subjectivity that is intentionally avoided by other aspects of WBA.

There are also a number of deeper philosophical objections to Lewis' use of counterfactual reasoning as it is embodied within WBA. These objections have recently been summarised by Hausman's study of causal asymmetries [315]. Hausman's objections are beyond the scope of this book. Many of his caveats focus on the argument that causes are not counterfactually dependent on their effects. The exposed indicator tap was not counterfactually dependent on the ignition of the Nanticoke fire because the ignition might have been caused by an uncovered exhaust manifold cover. There are possible worlds in which no fire occurred because the exhaust manifold was covered that are at least as similar to the actual world as situations in which a fire did not occur because the indicator was guarded. As we have seen, these situations complicate the application of counterfactual reasoning. Hausman notes that we cannot assume a particular cause simply be observing a set of effects. Each set of effects may be produced by several different causes, even though investigators can identify a determined set of effects for each cause [507]. These observations explain Hausman's choice of 'causal-asymmetries' as the title for his work.

Further criticisms of Lewis' approach focus on the notion of multiple connections. Hausman argues that these occur if a cause d of a is, or in the absence of a, would be connected to b by a path that does not go through a. If there is a multiple connection between a and b, then b will not counterfactually depend on a. Such situations again provide an example of causation without a chain of counterfactual dependence. For instance, the Chief Engineer sounded the general alarm that led the first team of fire fighters to enter the engine room. Their exit caused a second team to be deployed. If we imagine a situation in which the alarm could have led the second team to be deployed whether or not the first had been beaten back then even if we could ensure the success of the first team then there is no guarantee that the second team would not have been deployed. In other words we cannot rely on the argument that if the first team had not been pulled out then the second team would not have been deployed. Both of these caveats affect the other analysis techniques this chapter and Chapter 9.3 that exploit counterfactual reasoning. It can be argued that these are minor caveats compared to the analytical benefits provided by Lewis' form of reasoning even if, as Hausman argues, 'one cannot defend a counterfactual theory of causation' [315].

The problems of demonstrating the cost-effectiveness of WBA is arguably more important than the theoretical objections proposed by the Hausman's philosophical critique. Semi-formal diagrams, such as Figure 11.16, are relatively cheap and easy to develop. There are some notable differences between this approach and the diagrams employed by MES and STEP. In particular, the ontology of Why-Because graphs including events, states, processes and non-events can be contrasted with the events and conditions of ECF charts. There are, however, considerable similarities. The spatial arrangement of causal relations and the process of informal analysis, including counterfactual reasoning, are comparable. Deeper differences stem from the role of formal reasoning to support the application of Why-Because graphs. These proofs are costly to develop both in terms of the time required and the level of expertise that is essential to guide this process. These formal proofs are important if investigators are to benefit from the strengths of the Why-Because approach. Ladkin and Loer introduce meta-templates that can be used to guide and simplify the formal validation of any causal analysis. Even so, WBA is a time-consuming process. Loer describes a case study during which the development of an 'intuitive' Why-Because graph with approximately 100 nodes required 300 hours. The associated formal proof required a further 1,200 hours [499]. These costs must be assessed against the potential benefits from identifying potential weaknesses in an accident or incident report:

> "We have already been able to identify reasoning mistakes in accident reports using this method. The three accident reports analysed all contained facts which were significantly causally related to the accident, which appear in the WB-graph analysis as causes, but which are not contained in the list of 'probable cause/contributing factors' of the report. We regard this as a logical oversight. (Formally, they appear in the WB-analysis as causal factors that are not themselves explained by any further causal factors; i.e., as source nodes with out-edges but no in-edges.) Some might speculate that there are administrative, political or other social grounds for excluding them from the list of original causal factors, but this is not our interest here. We regard the WB-graph analysis as demonstrating that logical mistakes were made, thereby justifying the use of the WB-analysis to put accident reporting on a rigorous foundation. " [291]

Ultimately, WBA provides many benefits in terms of the precision and rigour that it introduces to causal analysis. Unfortunately, the price that must be paid in order to obtain those benefits is likely to preclude the use of this technique in all but a handful of safety-critical incidents.

This chapter has exploited a deterministic view of the past. We have endeavoured to model a single chain of causal relations that together can help to explain the course of an incident. In our case studies, we have encountered situations where it has not been possible to determine which of a number of possible causal sequences actually led to a mishap. For example, it has not been possible to identify the ignition source in the Nanticoke incident. In general, however, we have attempted to avoid such ambiguity through further investigation. In contrast, the following sections examine ways in which probabilistic models of causation might be applied to support incident and accident analysis. These techniques stem from a scientific and philosophical tradition that questions the notion of deterministic cause [29]. Most of this work has focussed on the problems of using theories of causation as predictive tools. There are, however, important implications for the post hoc use of causal analysis to understand the events the lead to near miss incidents. For example, probabilistic views of causation affect our interpretation of the probability that an accident *might have* occurred. It should be emphasised that the following pages are more speculative than previous sections. We are unaware of any previous attempts to apply these techniques to support incident analysis.

11.3.2 Partition Models for Probabilistic Causation

The previous chapters in this book have assumed that 'causation' can be defined in terms of the necessary and sufficient conditions that must exist between objects in order to achieve particular effects. In particular, counterfactual arguments have been used to identify situations in which a set of effects would not have occurred if those necessary and sufficient conditions had not been fulfilled. It is important to note that a number of caveats can be raised to these general theories of causation.

For instance, previous sections have identified different forms of causal asymmetry. For instance, if necessary and sufficient conditions do not hold then an effect may still occur. This complicates the application of counterfactual argumentation when investigators use a form of 'backtracking' to identify causes from their effects. Similarly, many physicists maintain that occurrences are not determined [201]. In other words, we can never be absolutely certain that a set of effects will be produced even if necessary and sufficient conditions can be demonstrated to hold at a particular moment. In contrast, it is argued that a complete specification of the state of a system only determines a set of probabilities [315]. Some of the proponents of this view have argued that what happens in any given situation owes as much to chance as it does to cause. This analysis has profound implications. For instance, we might be persuaded to abandon the notion of 'sufficient' causes that do not account for this role of chance! In this view, causal analysis would owe more to probabilistic risk assessment and human reliability assessment than it does to the discrete mathematics of WBA or Causal Trees. This is an interesting conjecture. Such an approach might emphasise the role of performance shaping factors incident rather than discrete events [443]. Instead of focusing on the identification of a deterministic sequence of cause and effect relationships, which are difficult to validate given the problems of causal asymmetry mentioned above, investigators should focus on those conditions that made effects more likely within a given context. For instance, we might describe the Nanticoke incident in terms of the probabilities that either the indicator tap or the manifold ignited the fire.

It is important to emphasise, however, that probabilistic forms of analysis do not eliminate the need to consider causality. For example, supposing that a factory produced a faulty gasket and that this gasket eventually led to a fuel leak on board a ship. Investigators might argue that the gasket caused the leak even though the production of the gasket created a small probability that any particular vessel would be affected. Statistical mechanics has also identified mass populations for which particular relations are deterministic, however, the best means of describing mass effects is through the use of probabilistic techniques [315]. This is important within the field of incident analysis because, as we shall see, national reporting systems typify these mass phenomena. For instance, we might receive ninety-nine reports in which a fire is caused by the exposure of an ignition source to a fuel supply. In one report, however, the same circumstances might not have led to a fire. Although we have an apparently deterministic model of how a fire starts, there may be exceptions that persuade analysts to consider probabilistic aspects of causation. These exceptions characterise many different aspects of incident analysis and, more generally, of individual attitudes to causation. For instance, people often argue that fines cause reductions in health and safety violations even though they do not believe that the deterrence is perfect. Similarly, people will say that dropping a glass causes it to break even though they have seen similar situations in which the glass did not break. It is often argued that a more complete knowledge of the moment acting on the glass would enable causal explanations of why certain glasses break while others do not. However this indeterminism is equally apparent in the 'microscopic' causal relations that explain the physics of these different outcomes.

Probabilistics approaches to causal analysis raise many practical and theoretical questions. The frequentist approach derives the probability of an event from an analysis of comparative frequencies. We can use information about previous fires to derive numerical estimates for the number of times that ignition was caused by a manifold or by an exposed indicator tap. Previous sections have dismissed this approach because it can be difficult to validate the frequency of rare events. We shall return to this theme several times in the following pages. Alternatively, empirical analysis can be used to repeatedly recreate situations in which either of these sources might ignite leaking oil. Again, frequencies can be calculated to derive probability estimates. This approach raises questions about the validity of the experimental context in which the simulations are conducted.

Unfortunately, a number of factors complicate the use of probabilistic approaches to causal analysis. Raw event frequencies cannot, typically, be used to determine the probability of particular 'causes' in the aftermath of an incident. For example, an examination of previous fires might find that six were caused by indicator taps and ten were caused by exposed manifolds. Supposing, however, that nine of the ten manifold fires involved a different fuel leak than that on the Nanticoke. In this situation, any causal analysis must draw upon conditional probabilities. These represent the probability of an event given that some other factors hold. In this case we need to know the probability of ignition from each source given the fuel leak characteristics that held during the Nanticoke fire. This use of conditional probabilities has some significant benefits for incident analysis. Investigators are not dealing with prior probabilities describing future events where we know relatively little additional information about the potential state of a system. In the aftermath of an incident it is often possible to obtain the conditioning information that helps to support particular probability assessments. The following section, therefore, extends this analysis to consider Bayesian statistics. For now it is sufficient to observe that these techniques can be used to represent and reason about a hypotheses given particular evidence in the aftermath of an incident.

As mentioned, an important limitation of many probabilistic approaches to causation is that it can be difficult to validate numerical estimates of rare events. Fortunately, many probabilistic theories of causation avoid this problem by describing how particular causes make their effects 'more likely'. For instance, Hempel argues that a and b are causally connected in a context C if there is a very high probability that b is true given that a is true in C: $Pr(b \mid a \wedge C)$ [347]. For instance, we could say that the maintenance modification to the sealing surface of the Nanticoke's fuel filter, a, was a cause of the leak, b, because this modification made the leak very likely given everything else that was discovered about the incident including the failure to report such problems etc, C. Hempel's approach also avoids the need to assign precise numeric values to individual probabilities it also creates the problem that investigators must determine what is meant by 'very likely'. It is possible, however, that this theoretical objection can be addressed by experience in applying the technique within a particular domain. The following paragraphs explain how Hempel's ideas might contribute to a method for the causal analysis of adverse incidents:

1. *Record the context in which an incident occurs.* This step ensures that as much information as possible is derived from the primary and secondary investigation of an incident. Previous sections have mentioned the difficulty of predicting all of the information that might be relevant to a causal analysis and so investigators should collate as much data as possible. Chapter 14.5 will examine the practical problems that such a policy creates for information storage.

2. *Perform an initial deterministic causal analysis.* Having collated as much information about the context, C, in which an incident occurs, investigators can exploit one of the causal analysis techniques introduced in previous sections. For instance, STEP or WBA might be used to identify potential causal factors in the immediate aftermath of an incident. Chapter 11.5 will describe how these techniques can be used to derive initial recommendations that are intended to avoid any recurrence of an near-miss occurrence.

3. *Build up sufficient data to perform a statistical analysis of potential causes.* Over time an incident reporting system may gather information about a number of adverse occurrences that have similar outcomes, b. Investigators can then examine the contextual information that has been recorded for each incident, C, to identify those events, a, that have the highest relative frequency. These events need not, however, have any causal relationship to b. For instance, b might occur before a in the temporal ordering of events. Additional techniques, such as WBA, must therefore validate the causal relations that are induced by the statistical analysis of incident collections. This form of causal analysis does, however, avoid the bias that can arise from causal asymmetries. Analysts do not simply use deterministic models to search for a narrow range of causes that can be made to 'fit' the observed effects.

The approach, described above, has numerous potential benefits from its integration of deterministic and probabilistic models of causation. The initial use of deterministic approachs can help to direct resources to a number of clearly defined causal factors in the aftermath of an incident. Probabilistic techniques can be used to search for other causal factors through an analysis of the correlations that exist between common factors in similar incidents. As far as we are aware, this approach has not been explicitly described before. It is, however, increasingly being adopted by many commercial and regulatory organisations. Chapter 14.5 will describe how probabilistic information retrieval tools have been developed to exploit correlations between the terms that are used to describe both the consequences and the causes of incidents and accidents.

As mentioned, Hempel's initial formulation provided little guidance on the meaning of the term 'very likely'. Fortunately, a number of refinements have been made to these early ideas. One of these approaches holds that a is causally related to b in a context C if the probability of A and B in C is not the same as the probability of B in C and the probability of A in C:

$$Pr(B \wedge A \mid C) \neq Pr(B \mid C).Pr(A \mid C) \tag{11.9}$$

We assume that we cannot derive A or $\neg\, A$ from C. Upper case denotes types, lower case is used to denote tokens of a particular type; token a is of type A and so on. This inequality has some interesting properties that can be applied to guide the causal analysis of incidents and accidents. Recall from Chapter 8.3 that $Pr(a \wedge b) = Pr(a).Pr(b)$ depends upon the independence of both a and b. If there is a causal connection between A and B then we might expect that the occurrence of a would make b more likely. Conversely, if A is a barrier to B then an occurrence of a will make b less likely. Hausman argues that a is positively causally related to b when the probability of A and B given C is greater than the probability of B given C multiplied by the probability of A given C [315]. In other words, a causal relationship implies that the probability of there being a general fire alarm, a, and a Halon system being deployed, b, on board a vessel, C, is greater that the probability of a general fire alarm being issued multiplied by the probability of a Halon system being deployed in similar circumstances. The deployment of the Halon system might be a relatively rare event compared to the sounding of a general alarm. However, a causal relationship with the alarm might result in a much higher probability being associated with situations in which the alarm and the Halon deployment both occur than situations in which we only know that one of these events has occurred:

$$Pr(B \wedge A \mid C) > Pr(B \mid C).Pr(A \mid C) \tag{11.10}$$

The key point to understanding this formula is that causes do not make their effects probable. They simply make them more probable than they otherwise would have been. We can also say that a is negatively causally related to b when the probability of A and B given C is less than the probability of B given C multiplied by the probability of A given C [315]. For instance, the probability of an engineer failing to obtain a fuel-tight seal, b, and of that engineer reporting the problem associated with the sealing surface, a, are together less than the independent probabilities of the engineer reporting the problem multiplied by the probability of the engineer failing to obtain the seal. This follows because the fact that the engineer reported the maintenance problem makes it less likely that they will be satisfied by any subsequent attempt to form a seal on the damaged surface:

$$Pr(B \wedge A \mid C) < Pr(B \mid C).Pr(A \mid C) \tag{11.11}$$

From this line of argument, we can say that a's cause b's under circumstances C if a's precedes b's in the temporal sequence leading to an incident and it is the case that the probability of B and A in C is greater than the probability of B given that we know $\neg\, A$ and C. Or we can say that a's cause b's under C if a's precedes b's in the temporal sequence leading to an incident and it is the case that the probability of B and A in C is greater than the probability of B given only C:

$$\begin{aligned} &Pr(B \wedge A \mid C) > Pr(B \mid \neg\, A \wedge C) \vee \\ &\quad Pr(B \mid A \wedge C) > Pr(B \mid C) \end{aligned} \tag{11.12}$$

Unfortunately, this formalisation leads to further problems. For example, it may be that a precedes b and that $Pr(B \wedge A \mid C) > Pr(B \mid \neg\, A \wedge C)$ but that a and b and effects of the *same* cause. One way to avoid this is to examine the events prior to a to determine whether there is another event that might 'screen off' or account for both a and b. Further models have been developed to formalise this approach [766] and these, in turn, have been further criticised [315]. The key point here is to provide an impression of the complexity that must be address by any attempt to exploit probabilistic models of causation as a means of supporting incident analysis. The initial appeal of an alternative to deterministic models rapidly fades as one considers the complexity of an alternative formulation.

One important source of additional complexity is that causal factors may both promote and confound particular effects. In this refinement, some factor that causes as to occur can have an independent negative influence on the occurrence of b's. For instance, the probability that the Nanticoke fire would lead to the loss of the vessel was increased by the lack of effective monitoring when the initial fire developed on the port side of the engine room. This lack of monitoring might have been a result of having both the Chief Engineer and the Mechanical Assistant in the Control Room during the fire drill. However, the same circumstances that interfered with their monitoring responsibilities may also have reduced the probability that the fire would jeopardise the safety of the vessel because both crewmembers could initiate the eventual response to the incident. Similarly, the increasing probability of b from a by one causal path can be offset by negative influences from a along another causal path. For example, the fire drill procedures may have made it more likely that the vessel would be seriously damaged by distracting members of the crew from their normal activities. The same drills may have made it less likely that the vessel would be seriously damaged because members of the crew were already prepared to respond to the general alarm that was sounded by the Chief Engineer. The importance of these mitigating factors has been repeatedly emphasised in recent studies of incident reporting systems [843]. Unfortunately, these factors are not adequately represented within many deterministic causal analysis techniques.

The proponents of probabilistic theories of causation have responded to these observations by revising the previous formulations to include a partition S_j of all relevant factors apart from A and C. From this it follows that a's cause b's in circumstances C if and only if:

$$\forall j : Pr(B \mid A \wedge S_j \wedge C) > Pr(B \mid S \wedge C) \tag{11.13}$$

$\{S_j\}$ is a partition of all relevant factors excluding A and C. These factors represent the negative or positive causal factors, $c_1, ..., c_m$, that must be held fixed in order to observe the causal effect of a. We require that 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. For instance, a hot manifold is liable to have a negligible impact on an existing fire. We can, therefore, include a factor, c_i, in each subset to require that a fire must not have already started in order for a hot manifold, a, to ignite a fuel source, b. Each of the factors in $c_1, ..., c_m$ must be represented in each subset. Each factor must also either be present or absent; there may or may not be an existing fire. This results in 2^m possible combinations of present or absent factors. Some combinations of the factors $c_1, ..., c_m$ will be impossible. Hence some combinations of c_i can be excluded from S_j. For example, it is difficult to foresee a situation in which the engine room is flooded with Halon gas and the fire continues to burn. Yet both of these factors could prevent us from observing an ignition caused by a hot manifold. Other combinations may result in b being assigned a probability of 1 or 0 regardless of a. For instance, if the engine room were flooded with Halon then the fire should not ignite irrespective of the state of the exhaust manifold. As mentioned, these impossible combinations and combinations that determine b are omitted from S_j. All the remaining combinations of causal factors must be explicitly considered as potential test conditions and are elements of S_j. In other words, a's must cause b's in every situation described by S_j.

Some proponents of this partition theory dispense with any explicit representation of the context, C [153]. This approach relies entirely upon the partitioning represented by S_j. This is misleading. Causal relations may change from one context to another. For instance, the effects of a fuel leak may depend upon the pressure at which the fuel escapes. This, in turn, may depend upon the size and configuration of a generator. The meta-level point here is that we would like causal relations to hold over a variety of circumstances, these are characterised by S_j. We cannot, however, expect to identify causal relations that are not relativised to some background context [315].

A number of objections have inspired further elaborations to this partition model of causation [222]. In terms of this book, however, we are less interested in the details of these reformulations than we are in determining whether these models might support the causal analysis of incidents. The abstract model, outlined above, provides a structure for the analysis of incidents in the same way that Why-Because graphs and the associated proof templates provided by Ladkin and Loer also provide a structure for causal analysis. For example, we can apply the partition model to the Nanticoke example by identifying candidate causal relations. Investigators can use their domain expertise to determine those relations that are then subjected to a more formal analysis, this equates to stage 2 of the method proposed for Hempel's model given above. For instance, previous sections have argued that it is difficult to determine the ignition source for the Nanticoke fire. This causes problems for deterministic causal models. We might, therefore, exploit the partition model to represent a causal relationship between an ignition event, b, and the fuel oil coming into contact with an exhaust manifold. As mentioned, C represents all state descriptions for the system under consideration. We might, therefore, informally argue that C represents the state of any merchant vessel that relied upon diesel generators. This context might be narrowed if the formalisation of the incident is intended only to apply to a restricted subset of these ships. In contrast, it might be extended if the formalisation also captures important properties of other vessels, such as military ships that employ diesel generators. Irrespective of the precise interpretation, it is important that analysts explicitly identify this context that helps other investigators to understand the scope of the model. We can then go on to identify other causal factors that might be represented in subsets of the form $c_1, ..., c_m \in S_j$. Recall that d 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:

c_1 represents 'the room floods with Halon',
c_2 represents 'fuel is sprayed at pressure',
c_3 represents 'shielding protects the manifold'.

As mentioned, individual factors may either be present or absent during particular incidents. There are, therefore, 2^3 potential elements of S_j. In the following, the omission of an element from any set implies that the causal factors are omitted. The first sequence represents a situation in which all of the previous causal factors are present. The room floods with Halon and the fuel is sprayed at pressure and shielding protects the manifold. The second of the subsets indicates that all of the factors are true except for the last one; the shielding does not protect the manifold.

$$\{c_1, c_2, c_3\}, \{c_1, c_2\}, \{c_1, c_3\},$$
$$\{c_2, c_3\}, \{c_1\}, \{c_2\}, \{c_3\}, \{\},$$

We can, however, reduce the number of combinations that we need to consider in order to establish a causal relation between a and b. As mentioned, some combinations of these causal factors are impossible. Other combinations may entirely determine the effect irrespective of the putative cause. For example, we can ignore any subset that contains c_1. If the room floods with Halon then the fire will not ignite, b, whatever happens to the fuel and the manifold, a. Conversely, we can insist that all subsets must include c_2. If the fuel is not sprayed at pressure then the fire will not ignite even if the fuel oil comes into contact with an exhaust manifold; as the manifold may not reach the flash-point of the fuel. In order to establish causality, we must however consider whether a increases the probability of b taking all other combinations of the causal factors, c_i, into account:

$$\{c_2, c_3\}, \{c_2\}.$$

In other words, in order for a causal relation to hold between between an ignition event, b, and the fuel oil coming into contact with an exhaust manifold, a, we must show that the effect would still be more likely if fuel is sprayed at pressure whether or not shielding protected the manifold. The shielding might reduce the absolute probability of the ignition but may not necessarily reduce it to zero, as a Halon deployment might. We must, therefore, show that the cause still increases the probability of the effect in both of these conditions.

This application of the partition model has a number of practical advantages. For instance, investigators are not forced to quantify the probability that a cause will yield a particular effect. The partition model also offers some advantages when compared to more deterministic models. This approach provides an elegant means of dealing with uncertainty about the precise causes of an incident. In particular, previous analyses have experienced acute problems from the investigators difficulty in determining what caused the ignition of the fire on the Nanticoke. The partition model entirely avoids this problem. It is possible to characterise multiple potential causes using the relevant factors represented by c_i. For example, we could have extended C_i to include fuel oil comes into contact with an indicator tap. We can also use the same techniques to represent and reason about the impact of mitigating factors. This again was problematic in deterministic techniques. In the previous example, we had to demonstrate that an ignition was more likely to occur whether or not the manifold was protected by shielding. We also showed how the same approach can represent barriers, such as Halon deployment, that prevent an effect from occurring. It is important to stress that these arguments about the probability of an ignition must be validated [196]. The partition model helps here because analysts can explicitly represent the anticipated impact of contributory causes, of mitigating events and of potential barriers. In contrast, many deterministic techniques consider these issues as secondary to the process of enumerating those failures that led to an incident.

There are, however, a number of practical concerns that arise during the application of the partition model of non-deterministic causation. All of the relevant factors, c_i, in the previous example were carefully chosen to be events. This satisfies the requirement that 'd 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'. Previous informal examples in this section have argued that a hot manifold would not have ignited the fire if a fire had already been burning. Ideally, we would like to extend c_i to include an appropriate state predicate so that we can explicitly represent and reason about such a situation. Alternatively, we could refine the relatively abstract view of the context, C, that was introduced in this example. Further concerns stem from the problems of applying an abstract model of causation to support incident analysis. It is entirely possible that the previous example reflects mistakes in our interpretation of the theoretical work of Cartwright [153] and Hausman [315]. Further work is, therefore, needed to determine appropriate formulations and interpretations of these non-deterministic models. This brief example does, however, demonstrate that probabilistic approaches can avoid some of the problems that uncertainty creates for the deterministic techniques that have been presented in previous sections.

There are also a number of more theoretical concerns about the utility of partition models. The formula (11.13) ensures that a increases the probability of b irrespective of the values assigned to these other relevant factors. This provides a definition of causation in which the mitigating effects of these relevant factors must not offset the increased probability of an associated effect. This might seem a reasonable criterion for causality. It does, however, lead to a number of philosophical problems. For instance, it might be argued that the crew's failure to regularly inspect the engine room is a potential cause of major fires such as that on board the Nanticoke. It might equally be argued that, under certain circumstances, regular inspection of the engine room might lead to a major fire. For example, operators might miss an automated warning in the control room that indicated a potential problem elsewhere in the engine room [623]. Under the system described above, such circumstances would prevent investigators from arguing that lack of inspection is a cause of major fires! This is a general problem; there are contexts in which "smoking lowers one's probability of getting lung cancer, drinking makes driving safer and not wearing seat-belts makes one less likely to suffer injury or death" [315]. As before there are a number of refinements on the basic model outlines in (11.13). It remains to be seen whether any of these extensions might provide an adequate framework for the causal analysis of safety-critical incidents. It might seem to be far-fetched that probabilistic models of causation might yield pragmatic tools for incident analysis. Against this one might argue that Lewis' possible worlds semantics for counterfactual reasoning would have appeared equally arcane before the development of WBA.

11.3.3 Bayesian Approaches to Probabilistic Causation

There are a number of alternative semantic interpretations for the Pr function introduced in the previous section [150]. In particular, Pr may be viewed either as a measure of confirmation or as a measure of frequency. The former interpretation resembles the Bayesian view; probability is contingent upon the observation of certain evidence. The latter resembles the manner in which engineers derive reliability figures. Estimates of pump failures are derived from the maintenance records of plant components. This distinction has been a subject of some controversy. For instance, Carnap argued:

> "... for most, perhaps for practically all, of those authors on probability who do not

> accept a frequency conception the following holds. i. Their theories are objectivist (and) are usually only preliminary remarks not affecting their actual working method. ii. The objective concept which they mean is similar to (the frequency view of) probability." [150]

Brevity prevents a detailed explanation of the contrasting positions in this debate. It is, however, possible to illustrate the common origin for these two different approaches. Both the partition models and Bayesian views exploit conditional probabilities. These also formed the foundation for the treatment of probabilistic causality in the previous chapter. As before, we use the following form to denote that the probability of the event B given the event A in some context C is x.

$$Pr(B \mid A \wedge C) = x \tag{11.14}$$

From this we can derive the following formula, which states that the conditional probability of B given A in C multiplied by the probability of A in C is equivalent to the probability of A and B in C. In other words the probability of both A and B being true in a given context is equivalent to the probability of A being true multiplied by the probability that B is true given A:

$$Pr(B \mid A \wedge C).Pr(A \mid C) = Pr(A \wedge B \mid C) \tag{11.15}$$

We can use this axiom of probability calculus to derive Bayes theorem:

$$Pr(B \mid A \wedge C).Pr(A \mid C) = Pr(B \wedge A \mid C)$$
$$(Commutative\ Law\ applied\ to\ \wedge\ in\ (11.15)) \tag{11.16}$$

$$Pr(B \mid A \wedge C).Pr(A \mid C) = Pr(A \mid B \wedge C).Pr(B \mid C)$$
$$(Substitution\ of\ RHS\ using\ (11.15)) \tag{11.17}$$

$$Pr(B \mid A \wedge C) = \frac{Pr(A \mid B \wedge C).Pr(B \mid C)}{Pr(A \mid C)}$$
$$(Divide\ by\ Pr(A \mid C)) \tag{11.18}$$

The key point about Bayes' theorem is that it helps us to reason about the manner in which our belief in some evidence affects our belief in some hypothesis. In the previous formula, our belief in B is affected by the evidence that we gather for A. It should be emphasised that (11.18) combines three different types of probability. The term $Pr(A \mid C)$ represents the *prior* probability that A is true without any additional evidence. In the above, the term $Pr(B \mid A \wedge C)$ represents a *posterior* probability that B is true having observed A. We can also reformulate (11.18) to determine the *likelihood* of a potential 'cause' [201]. 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:

$$Pr(B \mid A \wedge C) = \frac{Pr(A \mid B \wedge C).Pr(B \mid C)}{Pr(A \mid B \wedge C).Pr(B \mid C) + \sum_i Pr(A \mid B_i \wedge C).Pr(B_i \mid C)} \tag{11.19}$$

The previous 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 accidents and incidents. In particular, (11.19) provides a means of using information about previous incidents to guide the causal analysis of future occurrences.

In the Nanticoke case study, investigators might be interested to determine the likelihood that reported damage to an engine room had been caused by the pressurised release of fuel from a filter. The first step would involve an analysis of previous incidents. This might reveal that fuel from a filter was identified as a cause in 2% of previous mishaps. Lubrication oil might account for 1%. Other fuel sources might together account for a further 3% of all incidents:

$$Pr(\textit{filter fire} \mid C) = 0.02 \tag{11.20}$$
$$Pr(\textit{lube fire} \mid C) = 0.01 \tag{11.21}$$
$$Pr(\textit{other fire} \mid C) = 0.03 \tag{11.22}$$

In order to gain more evidence, investigators might try to determine how likely it is that one of these fires would cause serious damage to an engine room. Further analysis might reveal that thirty per cent of previous incidents involving the ignition of filter fuel resulted caused significant damage to an engine room. Twenty per cent of lube fires and fifty per cent of fires involving other fuel sources might have similar consequences:

$$Pr(\textit{engine room damage} \mid \textit{filter fire} \wedge C) = 0.3 \tag{11.23}$$
$$Pr(\textit{engine room damage} \mid \textit{lube fire} \wedge C) = 0.2 \tag{11.24}$$
$$Pr(\textit{engine room damage} \mid \textit{other fire} \wedge C) = 0.5 \tag{11.25}$$

We can now integrate these observations into (11.19) to calculate the probability that a filter fuel fire was a cause given that a serious engine room fire has been reported. This following calculation suggests that there is a twenty-six per cent chance that such a filter fire had this effect:

$$\begin{aligned} &Pr(\textit{filter fire} \mid \textit{engine room damage} \wedge C) \\ &= \frac{Pr(\textit{engine room damage} \mid \textit{filter fire} \wedge C).Pr(\textit{filter fire} \mid C)}{\begin{array}{l}((Pr(\textit{engine room damage} \mid \textit{filter fire} \wedge C).Pr(\textit{filter fire} \mid C)) \\ + (Pr(\textit{engine room damage} \mid \textit{lube fire} \wedge C).Pr(\textit{lube fire} \mid C)) \\ + (Pr(\textit{engine room damage} \mid \textit{other fire} \wedge C).Pr(\textit{other sources} \mid C))\end{array}} \end{aligned} \tag{11.26}$$

$$= \frac{(0.3).(0.02)}{(0.3).(0.02) + (0.2).(0.01) + (0.5).(0.03)} \tag{11.27}$$

$$= 0.26 \tag{11.28}$$

A number of caveats can be raised against this application of Bayes' theorem. Many concerns centre on our use of evidence about previous mishaps to guide the causal analysis of new incidents. The previous calculations relied upon investigators correctly identifying when a fire had caused 'significant damage' to an engine room. These is a danger that different investigators will have a different interpretation of such terms. Chapter 14.5 describes how Bayesian techniques can account for the false positives and negatives that result from these different interpretations. For now it is sufficient to observe that our analysis of previous incident frequencies might bias the causal analysis of future incidents. For instance, we have made the assumption that these incidents occurred in comparable contexts, C. There may be innovative design features, such as new forms of barriers and protection devices, that would invalidate our use of previous frequencies to characterise future failures.

Dembski argues that it is seldom possible to have any confidence in prior probabilites [201]. 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 accidents and incidents. In spite of initiatives to share incident data across national boundaries, there are few data sources that validate the assumptions represented in (11.23), (11.24) and (11.25). This book has identified a number of different biases that affect the use of data from incident reporting systems. For example, Chapter 4.3 referred to the relatively low participation rates that affect many incident reporting schemes. This makes it difficult for us to estimate the true frequency of lube fires or filter fires. These incidents may also be extremely rare occurrences. It can, therefore, be very difficult for investigators to derive the information that is required in order to apply (11.19).

In the absence of data sources to validate prior probabilities, investigators typically rely upon a variant of the indifference principle. This states that given a set of mutually exclusive and exhaustive possibilities, the possibilities are considered to be equi-probable unless there is a reason to think otherwise. This would lead us to assign the same probabilities to fires being caused by filter fuel, to lube oil and to all other sources. Unfortunately, the pragmatic approach suggested by the indifference principle can lead to a number of paradoxes [371]. Objections can have also been raised against any method that enables investigators to move from conditional probabilities, such as $Pr(A \mid B_i \wedge C)$, to their 'inverse' likelihoods, $Pr(B_i \mid A \wedge C)$ [201].

The use of subjective probabilities provides a possible solution to the lack of frequential data that might otherwise support a Bayesian approach to the causal analysis of safety-critical incidents. Subjective probabilities are estimates that individual investigators or groups of investigators might make about the probability of an event. For example, a subjective probability might be an individuals assessment of the chances that lube oil could start a fire that might cause serious damage to an engine room. One standard means of estimating this probability is to ask people to make a choice between two or more lotteries. This technique is usually applied to situations in which it is possible to associate financial rewards with particular outcomes. Von Neumann and Morganstern provide a detailed justification for the general applicability of this approach [628]. Certain changes must, however, be made in order to explain how these lotteries might support the causal analysis of adverse incidents.

1. I might be offered a situation in which there is a certainty that if a lube oil fire occurs in the next twelve months then it will result in major damage to an engine room;

2. alternatively, I might be offered a form of 'gamble'. This requires that I select a token at random from a jar. This jar contains N tokens that are marked to denote that there is no serious damage to an engine room during the next twelve months. The remaining 100-N tokens are marked to denote that there is such an incident.

I will prefer option (2) if every token indicates that engine rooms remain undamaged, i.e. N=100. Conversely, I will prefer option (1) if every token indicates the opposite outcome, i.e. N=0. This requires additional explanation. Recall from option (1) that engine room damage will occur *if* there is a lube oil fire. In option (2), if N=0 then there is a certainty that engine room damage will occur. This explains the preference for (1), the individual makes a subjective assessment of the likelihood of the lube fire and then must trade this off against the potential for there not to be engine room damage in (2). There will, therefore, be a value of N for which the two situations are equally attractive. At such a position of indifference, $\frac{N}{100}$ is my estimate of the probability that a lube oil fire will cause serious damage to an engine room in the next year. Jensen notes that "for subjective probabilities defined through such ball drawing gambles the fundamental rule can also be proved" [400]. This fundamental rule is given by formula (11.15) that provided the foundation for Bayes' theorem (11.18).

A number of problems affect the use of subjective probabilities. An individuals' preference between the two previous options is not simply determined by their subjective estimate of the probability of a lube oil fire. It can also be affected by their attitudes towards risk and uncertainty. For example, one person might view that a 20% chance of avoiding engine room damage is an attractive gamble. They might, therefore, be willing to accept N=20. Another individual might be very unwilling to accept this same gamble and might, therefore, prefer option (1). These differences need say very little about the individual's view of the exact likelihood of a lube fire resulting in major engine room damage. In contrast, it may reveal more about their attitude to uncertainty. The first individual may choose the gamble because they have more information about the likelihood of engine room damage than they do about the lube fire in (1). Individual preferences are also affected by attitude to risk [690]. Experimental evidence has shown that different individuals associate very different levels of utility or value to particular probabilities. A risk adverse individual might view a 20% gamble as a very unattractive option whereas a risk preferring individual might be more inclined to accept the risk given the potential rewards.

In spite of the problems if deriving both frequentist and subjective probabilities, Bayesian inference has been used to reason about the dependability of hardware [86, 296] and software systems

[497]. In particular, a growing number of researchers have begun to apply Bayesian Networks as a means of representing probabilistic notions of causation. it is based around the concepts on contingent probability that, as we have seen, can also arguably be used to provide insights into the likelihood of certain causes. Figure 11.18 presents a Bayesian network model for one aspect of the Nanticoke incident. Investigators initially observed horizontal soot patterns on top of valve covers 4, 5 and 6 and a shadowing on the aft surfaces of these structures. These observations indicate that the fire originated on the port side of the engine, forward of cylinder head number 1. These effects might have been caused by a fire fed from one of two potential sources. This is indicated in Figure 11.18 by the two arrows pointing into the node labelled horizontal soot patterns.... The arrows point from a cause towards the effect. The + symbols indicate the cause makes the effect more likely. Conversely, a barrier might be labelled by a - symbol if it made an effect less likely. As can be seen, the two potential causes of the horizontal soot patterns include a lube oil leak from under that valve cover near cylinder head number 6. These effects might also have been caused by a fuel oil leak from one of the filters. Further investigations reveal that the valve covers were in tact and in place after the fire. This increases the certainty that the fire started from a filter leak rather than a lube oil leak under the valve covers. Another way of looking at Figure 11.18 is to argue that leaks from either the filter or from lube oil are consistent with the horizontal soot patterns. Only a fuel oil leak from the filter is consistent with the valve covers being in tact after the fire.

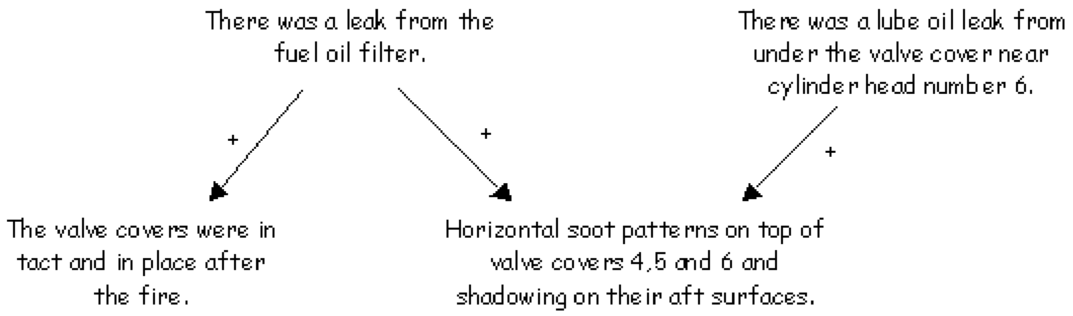

Figure 11.18: Bayesian Network Model for the Nanticoke Fuel Source

Before continuing to apply Bayesian networks to support our causal analysis of the Nanticoke incident, it is important to observe that some authors have argued that these diagrams must not be used as causal models. In contrast, they should only be used to model the manner in which information propagates between events. This caution stems from doubts over methods that enable investigators to move from conditional probabilities to their 'inverse' likelihoods, mentioned in previous paragraphs [201]. This point of view also implies further constraints on the use of Bayesian networks. For instance, it is important not to model interfering actions within a network of information propagation. Jensen provides a more complete introduction to these potential pitfalls [400].

	filter fire	$\neg$ *filter fire*
valve covers ok	1	0.98
$\neg$ *valve covers ok*	0	0.02

Table 11.12: Conditional Probabilities for the Bayesian Analysis of the Nanticoke Incident (1)

The quantitative analysis of Figure 11.18 begins with either a frequentist or subjective estimate of the likelihood of each cause. Recall that $Pr(\textit{filter fire} \mid C) = 0.02$ and that $Pr(\textit{lube fire} \mid C) = 0.01$. We can use these prior probabilities and the information contained in Figure 11.18 to derive the conditional probabilities for $P(\mathit{filter\ fire} \mid \mathit{valve\ covers\ ok})$. These are shown in Figure 11.12. If we know that there was a filter fire then it is certain that the valves would be in tact, this represents a simplifying assumption that can be revised in subsequent analysis. If there was not a filter fire then there is a 0.98 probability that the valves would be in tact but a 0.02 probability that they would not. The conditional probabilities shown in Figure 11.12 are represented in matrix form throughout the remainder of this analysis. We can calculate the prior probability that the valve covers are in tact using formula (11.15:

$$Pr(\mathit{valve\ covers\ ok} \mid \mathit{filter\ fire}).Pr(\mathit{filter\ fire}) = Pr(\mathit{valve\ covers\ ok} \wedge \mathit{filter\ fire}) \tag{11.29}$$

The following calculation introduces the conditional probabilities in Figure 11.12.

$$Pr(\mathit{valve\ covers\ ok} \wedge \mathit{filter\ fire})$$

$$= \begin{pmatrix} 1.0x0.98 & 0.98x0.02 \\ 0.0x0.98 & 0.02x0.02 \end{pmatrix} \tag{11.30}$$

$$= \begin{pmatrix} 0.98 & 0.0196 \\ 0.0 & 0.0004 \end{pmatrix} \tag{11.31}$$

In order to derive the prior probability $Pr(\mathit{valve\ covers\ ok})$ from $Pr(\mathit{valve\ covers\ ok} \wedge \mathit{filter\ fire})$ we have to use a procedure called marginalisation. This is characterised as follows:

$$Pr(A) = \sum_B Pr(A, B) \tag{11.32}$$

This can be applied to the matrix in (11.31) to derive $Pr(\mathit{valve\ covers\ ok}) = (0.9996, 0.0004)$. In other words the prior probability that the valve covers are in tact is just over 99%. Jensen provides more details on both the theoretical underpinning and the practical application of Bayesian networks [400]. The key point is that the underlying calculus provides investigators with a sophisticated analytical toolkit that can be used to supplement the less formal reasoning supported by the Bayesian network illustrated in Figure 11.18. The calculus can be used to derive prior and contingent probabilities depending on the nature of the information that is provided. Unfortunately, as can be seen from the preceding example, that application of these techniques can be complicated even for specialists who have considerable expertise in Bayesian analysis. For this reason, most practical applications of the approach rely upon the support of automated tools such as Hugin [401]. The previous calculations also relied upon the adaptation of models that were first developed to support medical diagnosis. This introduces the possibility that errors may have been introduced into the calculations as a result of attempting to reformulate the models to yield particular insights into the Nantcoke case study.

To summarise, the final two sections of this chapter have looked beyond the well-understood deterministic models of causation that have been embodied within incident and accident analysis techniques. The intention has been to determine whether investigators might benefit from recent developments in the theory and application of probabilistic models of causation. We have seen how this area promises many potential benefits. For example, the partition model and Baysian approaches can deal with the uncertainty that characterises the initial stages of an investigation. The importance of this should not be underestimated. *Given the increasing complexity and coupling of modern, safety-critical systems, it is inevitable that investigators will find it more and more difficult to determine a unique cause for many adverse incidents.* The Rand report into the National Transportation Safety Board (NTSB) repeatedly points to the increasing length of time that must be spent before analysts can identify the causes of many recent failures [482].

It is difficult to assess the true potential of these techniques because they have not been widely applied to support the causal analysis of adverse occurrences. In their current form there is little chance that they will be accessible to many investigators. Tool support must be provided. Methods and procedures must also be developed to help investigators learn how to apply these techniques without necessarily requiring a full understanding of the underlying theories that support the analysis. The use of Why-Because graphs as a central feature of WBA provides a useful prototype in this respect. The previous analysis has, however, identified several key issues that must be addressed before these more applied techniques will yield tangible benefits. In particular, there must be some means of assessing prior probabilities if investigators are to exploit Bayesian techniques for analysing causality through contingent probabilities. Dembski summarises this argument as follows:

> "Bayesian conceptions of probability invariably face the problem of how to assign prior probabilities. Only in special cases can prior probabilities be assigned with any degree of confidence (e.g., medical tests). so long as the priors remain suspect, so does any application of bayes' theorem. On the other hand, when the priors are well-defined, Bayes' theorem works just fine, as does the Bayesian conception of probability. To sum up then, there is no magic bullet for assigning probabilities" [201]

There may not be any general-purpose magic bullet but the previous pages have, at least, identified two potential solutions that might work as a means of assigning priors within the specialist domain of incident investigation. Firstly, we have shown how subjective probabilities can be derived using the lottery-based procedures of Von Neumann and Morgenstern [628] or of March and Simon [514]. In general these are difficult to apply because individual attitudes to risk make it difficult to interpret the expressed preferences that support inferences about subjective probabilities. We are not, however, dealing with a general population. Investigators are, typically, trained in the fundamentals of reliability and risk assessment. There is some prospect, therefore, that this method might yield better results than the more general studies of decision making under conditions of economic uncertainty.

The second, perhaps obvious, point is that we are not attempting to assign prior probabilities with complete ignorance about the nature of previous failures. In many ways, the entire purpose of an incident reporting system is to provide precise the sorts of quantitative information that is necessary in order to calculate the prior of Bayesian inference! It is paradoxical, therefore, to deny the usefulness of this data in a book that is devoted to the potential benefits of incident reporting. Unfortunately, as we have seen, we cannot trust the statistics that are extracted from national and international systems. Previous chapters have cited various estimates for the under-reporting of adverse occurrences. For instance, the Royal College of Anaesthetists estimates that only 30% of adverse medical incidents are voluntarily reported [716], Barach and Small estimate that this figure lies somewhere between 50 and 95% [66]. Chapters 4.3 and 14.5 describe techniques that can be used to assess the extent of this problem. For example, workplace monitoring can be used to identify the proportion of adverse incidents that occur within a given time period in a representative team. The results of this analysis can then be compared with incident submission rates by a similar workgroup. This is not a panacea. Even if we can assess the contribution rate within a reporting system, there is still no guarantee that we can trust the data that has been gathered about an incident. Consider the Nanticoke case study, if we wanted to gather data about the prior probability of fuel from a filter being involved in a fire, we would have to be sure that previous incidents were correctly analysed and indexed to indicate that this had indeed been a factor in previous incidents. The reliability of data about prior probabilities would be compromised if other investigators incorrectly diagnosed an incident as a filter fire when it was not. It data would also yield incorrect priors if investigators failed to diagnose this fuel source when it had contributed to an incident. Chapter 14.5 describes a statistical technique that can be used to identify and correct for these potential biases. For now it is sufficient to observe that this approach will only work if investigators have already performed a causal analysis of previous incidents. From this it follows that the application of Bayesian techniques may ultimately depend upon and support the use of more deterministic analysis.

11.4 Comparisons

Previous sections have reviewed a number of different techniques that can be used to support the causal analysis of safety-critical incidents. The diversity of these techniques makes it important that investigators and their managers have some means of assessing the support offered by these different approaches. Unfortunately, a range of practical, theoretical and also ethical issues complicate any attempt to perform comparative evaluations of causal analysis techniques:

- *the costs of learning new techniques.* Considerable training is required before investigators can apply some of the causal analysis techniques that we have considered. A significant level of investment would be needed to sponsor the evaluation of mathematical approaches unless investigators already had an appropriate training in the use of logic or of statistical reasoning. Similarly, it is difficult not to underestimate the problems associated with the independent application of Tier Analysis. Previous sections have emphasised the political and social pressures that affect the attribution of root causes to different levels within complex commercial organisations. Any investment in the evaluation of these techniques would carry the significant risk that they might not benefits the sponsoring organisation.

- *the costs of applying new techniques.* The investment that is required in order to train investigators to use particular analysis techniques must be supplemented to meet the costs associated with applying those techniques. This book has argued that computer-controlled automation supports increasingly complex application processes [677]. At the same time, incident investigations have been further complicated by the increasing appreciation that organisational, technical and human factors contribute to the causes of many 'failures'. These two influences have complicated the tasks associated with incident investigation. They are taking longer to complete and increasingly require the participation of multidisciplinary teams of investigators [482]. These increasing costs have not, to date, justified the allocation of resources to determine whether certain causal analysis techniques help to control the overall expenditure on incident investigations.

- *practice effects and the problems of fatigue.* Empirical test-retest procedures provide means of reducing the costs associated with the 'live' use of analysis techniques within multidisciplinary investigation teams. Investigators are presented with an example of a causal analysis technique being applied to a particular case study incident. The relative merits of that particular technique are assessed by asking investigators to answer comprehension questions, to complete attitude statements about the perceived merits of the approach and by timing investigators during these various tasks. The same procedure is then, typically, repeated for a number of further causal analysis techniques after a short break. This creates several experimental problems. For example, investigators can use the insights that were gathered from the first analysis technique to help answer questions about the second. One would, therefore, expect that the quality of the analysis might improve. On the other hand, investigators will also suffer from increasing fatigue as the evaluation proceeds. This, in turn, will impair their performance. These practice and fatigue effects can be addressed by counter-balancing. Different analysis techniques are applied in a different order by different investigators. One group might be presented with a STEP analysis and then a MORT analysis. This order would be reversed for another group. Such studies do not, however, provide any insights into the application of particular techniques over prolonger periods of time.

- *the problems of assessing learning effects.* The test-retest procedures, described in the previous paragraph, do not provide information about the long-term support that may be provided by a causal analysis technique. There studies also often yield subjective results that are strongly in favour of techniques that are similar to those which investigators are already familiar with. These potential biases create many problems. For instance, the results of a test-retest validation may simply indicate 'superficial' preferences based on a brief exposure to a relatively simple case study. These results may not be replicated if investigators actually had to apply a technique during a 'live' investigation. For example, we have described the results of an evaluation conducted using off-shore oil workers in which techniques that achieved the lowest subjective satisfaction ratings also yielded the highest comprehension and analysis scores [405]! Similarly, innovative techniques can often be undervalued if they provide significant long-term benefits that are not readily apparent during a cursory inspection.

- *the difficulty of finding 'realistic' examples.* Test-retest techniques reduce the costs associated with the validation of causal analysis techniques. The investment associated with training investigators is avoided because they, typically, are not required to apply the techniques themselves. The costs associated with applying the technique are, therefore, also avoided. Investigators are only committed to an initial assessment of existing case studies. This raises further concerns. In particular, the choice of case study may influence the investigators' responses. This is a significant issue because, as we have seen, techniques that focus on the managerial and organisational causes of failure may provide few insights into the failure of technical barriers. The test-retest procedure must, therefore, be replicated with several different case studies to provide a representative sample of the potential incidents that must be addressed. This, in turn, raises concerns that the individual preparing the case studies may also introduce potential biases that reflect their own experience in applying particular techniques. Some of these problems are addressed by accepting the costs associated with longitudinal studies of investigators applying causal analysis techniques. Given that high-consequence incidents will be rare events, even this approach provides no guarantee that investigators will meet a sufficient range of potential failures.

- *the difficulty of ensuring the participation of investigators.* Many of the previous problems relate to the difficulty of identifying an appropriate experimental procedure that can be used to support comparisons between causal analysis techniques. These issues often play a secondary role to the practical difficulties that are associated with ensuring the 'enthusiastic' participation of investigators in these studies. As we have seen, investigatory 'methodologies' are often intended to improve the *accuracy* of investigations by imposing *standard* techniques [73]. They constrain an individual's actions in response to a particular incident. It is, therefore, essential that to encourage the support and participation of investigators in the evaluation process. Any technique that under-values the existing skill and expertise of investigation teams is unlikely to be accepted. Similarly, the techniques that are being assessed must be adequately supported by necessary training material that is pitched at a level that can be understood by its potential users. Above all, the comparative evaluation of a causal analysis technique must not be misinterpreted as a comparative evaluation of incident investigators.

- *the ethical issues that stem from studying the causal analysis of incidents.* We have been involved in several studies that have performed empirical comparisons of different causal analysis techniques. These evaluations often raise a host of ethical questions for the organisations that are involved. If new techniques are introduced for a trial period then many industries require that these approaches should at least be as 'good' as existing approaches. This creates an impasse because such reassurances cannot be offered until after the evaluation has been conducted. This often forces investigators to continue to apply existing techniques at the same time as a more innovative technique is being trialed. At first sight, this replicated approach seems to offer many benefits. Investigators can compare the results that are obtained from each technique. It can, however, lead to more complex ethical issues. For instance, the application of novel causal analysis techniques can help to identify causal factors that had not previously been considered. In extreme cases, it may directly contradict the findings of the existing technique. Under such circumstances, it can be difficult to ignore the insights provided by the approach when the consequences might be to jeopardise the future safety of an application process.

The following pages build on this analysis. They provide a brief summary of several notable attempts that have been made to evaluate the utility of causal analysis techniques. As will be seen, the individuals and groups who have conducted these pioneering studies often describe them as 'first steps' or 'approximations' to more sustained validation exercises.

11.4.1 Bottom-Up Case Studies

Different causal analysis techniques offer different level of support for the analysis of different causal factors. For instance, MORT provides considerable support for the analysis of managerial and organisational failure. In contrast, early versions of this technique arguably lack necessary guidance for the technical analysis of hardware and software failures. In contrast, ECF analysis lacks any causal focus and, therefore, offers a broader scope. We have seen, however, that investigators must recruit supplementary tier analysis and non-compliance analysis to focus on particular human factors, managerial and organisational causes of an incident.

It is important to emphasise that the scope of causal analysis techniques is not static. Van Vuuren perceives a cycle in different industries [845]. A focus on individual blame and on isolated forms of equipment failure leads on to a focus on the organisational causes of incidents: This change in focus has altered the 'status quo' of safety related research and led to a number of innovative tools for causal analysis, including Tripod and PRISMA. Unfortunately, there has been a tendency for some organisations to accept that organisational failure is the end point in this process. In this Whig interpretation, accident and incident investigation has culminated in an acceptance of 'systemic' failure as the primary cause of incident investigation. Causal analysis techniques that identify the organisational precursors to systemic failures must, therefore, be chosen over those that focus more narrowly on the technical and human factors causes of incidents and accidents.

This argument raises a number of concerns. Firstly, it is unlikely that our view of incidents and accidents will remain unchanged over the next decade. The increasing development of incident reporting systems is likely to provide us with access to failure data on a scale that has never before been possible. In particular, the computer-based tools that are described in Chapter 14.5 already enable investigators to search through millions of reports to identify trends and causal factors that were not anticipated from the exhaustive, manual analysis of local data sources [412]. The current focus on organisational and managerial issues may, therefore, be superceded as we learn more about the causes of failure. Secondly, the focus on organisational issues is not an end in itself. We know remarkably little about the organisational and managerial causes of failure [444, 702, 840]. From this it follows that current techniques that specifically address these issues may actually fail to identify important causal factors. Indeed, many of this new generation of techniques have been attacked as premature. Researchers have pointed to particular theoretical weaknesses that are perceived to create practical problems for the investigators who must apply them:

> "The distinction between active and latent failure is the most important one in order to understand the difference in impact of different kinds of human failure. However, in his discussion Reason only focuses on the human contributions at different stages during accident causation, without providing insight into whether these human contributions result in technical, human or organisational problems. The eleven General Failure Types that are listed for Tripod are... a combination of technical, human and organisational factors, and are also a combination of causes/symptoms and root causes. For example, hardware problems are likely to be caused by incorrect design and the category organisation refers to failures that can cause problems in communication, goal setting, etc. This might be acceptable for an audit tool, however, it is not for incident analysis. Although claiming to focus on management decisions, no definition of management or organisational failure is provided. The lack of knowledge of how to model organisational failure in the area of safety related research states the importance of a bottom-up approach, using empirical incident data as a main input for new models and theories to be developed." [845]

These criticisms undervalue the pioneering role that Tripod played in re-focusing attention on the managerial and organisational factors that compromise barriers and create the context for latent failures. Van Vuuren does, however, make an important point when he urges that any evaluation of incident investigation techniques should be based on empirical data, derived from bottom-up investigations. He exploited this approach to assess the utility of the PRISMA technique. A series of case studies were conducted to demonstrate that this approach might support the causal analysis of incidents in a wide range of different domains. He also sought to validate PRISMA by applying it to different case studies within the same domain. For instance, he developed variants of the Eindhoven Classification Model to analyse incidents reported in the steel industry. He began by looking at coke production. Coke is a solid substance that remains after gases have been extracted from coal and is primarily used as fuel for blast furnaces. The company that he studied had an annual production of approximately five million tons of pig-iron. This required more than two million tons of coke from two different plants. His study focussed on one of these plants which employed 300 people in a 'traditional hierarchical organisation'. His study of fifty-two incidents revealed the distribution of causal factors illustrated in Table 11.13. The coke plant lies at the beginning of the steel making process. It provides fuel for the blast furnaces that produce pig-iron. He, therefore, conducted a second case study involving a plant that transformed pig-iron from the blast furnaces into steel. Table 11.14 presents the causal classification that was obtained for twenty-six incidents that were analysed using PRISMA in this second case study.

	Organisational	Technical	Human	Unclassifiable	Total
No. of root causes	111	67	126	13	317
Percentage	35%	21%	40%	4%	100%

Table 11.13: Distribution of root causes in Coke Production [845]

	Organisational	Technical	Human	Unclassifiable	Total
No. of root causes	73	46	57	5	181
Percentage	40%	25%	32%	3%	100%

Table 11.14: Distribution of root causes in Steel Production [845]

As mentioned, Van Vuuren was anxious to determine whether PRISMA could be successfully applied to a range of different domains. He, therefore, studied that application of the technique within both an Accident and Emergency and an Anaesthesia department. These different areas within the same healthcare organisation raised different issues in the application of a causal analysis technique. The work of the Accident and Emergency department fluctuated from hour to hour and was mainly staffed by junior doctors. In contrast, the Anaesthesia department provided well-planned and highly technical working conditions. It was mainly run by experienced anaesthetists. The insights gained from applying PRISMA within these institutions were also compared from its application in an institution for the case of the mentally ill. This institution had experienced nine incidents over a twelve month period that resulted in the death of eight of their residents and one near miss where the resident involved could barely be saved from drowning in the swimming pool at the institution. The direct causes of death varied between three cases of asphyxiation, three traffic accidents outside the main location of the institution and two drownings while taking a bath. The results of the causal analysis are summarised in Table 11.15.

	Organisational	Technical	Human	Patient related	Unclassifiable	Total
No. of root causes	29	3	24	11	4	71
Percentage	41%	4%	34%	15%	6%	100%

Table 11.15: Distribution of root causes in Mental Health Study [845]

Van Vuuren's work is important because it illustrates the use of a bottom-up approach to the validation of causal analysis techniques [845]. He provides direct, first-hand insights into the strengths and weaknesses of the PRISMA approach in a range of different application domains. This approach can be contrasted with the highly-theoretical comparisons that have been made by the proponents of other techniques. Unfortunately, the Van Vuuren's results cannot easily be applied to guide any decision between the different techniques that have been introduced in previous paragraphs. We simply lack the necessary data to make such a comparison. Techniques such as MORT have been widely applied in a range of different industries but there have been few recent attempts to systematically collate and publish the experience gained from the application of this approach. Other techniques, such as WBA and the statistical partition approaches, are relatively new and have only been validated against a small number of incidents and accidents.

Van Vuuren's approach is also limited as a basis for comparisons between causal analysis techniques. He was involved in the analysis of the case studies. It can, therefore, be difficult to assess how important his interventions were in the adoption and adaptation of the PRISMA technique. It must also be recognised that the case studies were not simply intended to provide insights into the relative utility of this approach compared to other causal analysis techniques. As can be seen, the results in Table 11.13, 11.14 and 11.15 provide no insights into how easy or difficult it was to apply PRISMA. Nor do they suggest that the findings of one investigation would be consistent with those of a previous study of the same incident. Van Vuuren was not primarily interested in the criteria that make one causal analysis technique 'better' than another. The primary motive was to learn more about the nature of organisation failure in several different industries. In contrast, Benner has applied a set of requirements to assess the utility of a wide range of investigatory methodologies.

11.4.2 Top-Down Criteria

The previous paragraphs have illustrated the diverse range of of causal analysis techniques that might be recruited to support incident investigation. This diversity is also reflected within investigatory organisations. Benner conducted a pioneering study into the practices of seventeen US Federal agencies: Consumer Product Safety Commission; Department of Agriculture; Department of the Air Force; Department of the Army; Department of Energy; Department of Labour; Mine Safety and Health Administration - Department of Labour; Occupational Safety and Health Administration (OSHA); Coast Guard; Department of Transportation; Federal Highways Administration - Department of Transportation; General Services Administration; Library of Congress; NASA; National Institute of Occupational Safety and Health; NTSB; Navy Department; Nuclear Regulatory Commission; National Materials Advisory Board - Panel on Grain Elevator Explosions [73]. He identified fourteen different accident models: the event process model, the energy flow process model, fault tree model; Haddon matrix model; all-cause model; mathematical models; abnormality models; personal models; epidemiological models; pentagon explosion model; stochastic variable model; violations model; single event and cause factors and a chain of events model. The term 'accident model' was used to refer to "the perceived nature of the accident phenomenon". Benner reports that these models were often implicit within the policies and publications of the organisations that he studied. He, therefore, had to exploit a broad range of analytical techniques to identify the investigators' and managers' views about the nature of accidents and incidents. Benner's study also revealed that these different models supported seventeen different investigation methodologies: event analysis; MORT; Fault Tree Analysis; NTSB board and inter-organisational study groups; Gannt charting; inter-organisational multidisciplinary groups; personal judgement; investigator board with intraorganisational groups; Baker police systems; epidemiological techniques; Kipling's what, when, who, where, why and how; statistical data gathering; compliance inspection; closed-end-flowcharts; find chain of events; fact-finding and legal approach; 'complete the forms'. The term 'accident methodology' refers to "the system of concepts, principles and procedures for investigating accidents" [73].

Benner's findings have a number of important consequences. He argues that the choice of accident methodology may determine an organisation's accident model. For instance, the application of the MORT technique would naturally lead to a focus on managerial issues. The use of Gannt charts would, similarly, suggest an accident model that centres on processes and events. Benner also observed the opposite effect; accident models can predispose organisations towards particular methodologies. An enthusiasm for epidemiological models leads to the development and application of an epidemiological methodology. He also identifies a third scenario in which an analysis method determines the accident model and investigation methodology but neither the model nor the investigatory methodology particularly influences each other. One interpretation of this might be situations in which organisations enthusiastically impose analytical techniques upon their investigators without considering whether those techniques are widely accepted as being consistent with the investigators' perception of an accident or incident.

A number of objections can be raised both the Benner's approach and to his analysis. For example, he used interviews to extract implicit views about models and methodologies. The findings of these meetings were supported by an analysis of documents and statutes. Previous sections in this book have enumerated the many different biases that can make it difficult to interpret these forms of evidence. Conversely, this distinction between model and methodology can become very blurred. The relatively broad definition of the term 'methodology' seems to imply that it contains elements of an accident model. The relationship between these two concepts is discussed but it is not the focus of Benner's work [73]. His investigation looks beyond the causal analysis that is the focus for this chapter, however, this work does identify a number of general problems:

> "Little guidance exists in the accident investigation field to help managers or investigators choose the best available accident models and accident investigation methodologies for their investigation... No comprehensive lists of choices, criteria for the evaluation or selection, or measures of performance (have) emerged to help accident investigators or managers choose the "best" accident model and investigative methodology." [73]

In order to address this problem, Benner proposed a series of criteria that might be used to guide a comparative evaluation of both accident models and their associated methodologies. A three point rating scheme was applied in which 0 was awarded if the model/methodology was not likely to satisfy the criterion because of some inherent shortcoming, 1 was awarded if the model/methodology could satisfy the criterion with some modification and 2 indicated that the model/methodology was likely to satisfy the criterion. Benner applied this scheme without any weightings to differentiate the relative importance of different criteria. He also acknowledges that the procedure was flawed "undoubtedly, ratings contained some author bias". The contribution of this work, arguably, rests on criteria that helped to guide his evaluation of accident models and methodologies.

The following list summarises Benner's conditions for models that reflect the perceived nature of accident phenomena. As will be seen, these criteria cannot be directly applied to assess the relative merits of causal analysis techniques. Most of the requirements support reconstructive modelling and simulation. Benner's methodological requirements have greater relevance for the content of this chapter. The model criteria are presented here, however, for the sake of completeness. This also provides an overview of Benner's more general comparison of investigatory techniques. It should be noted that we have redrafted some of the following criteria to reflect our focus on incident analysis rather than accident investigations:

1. *realistic.* This criteria focuses on the expressiveness of an incident model. Benner argues that it must capture the sequential and concurrent aspects of an adverse occurrence. It must also capture the 'risk-taking' nature of work processes.

2. *definitive.* Any model must describe the information sources that must be safe-guarded and examined in the aftermath of an incident. Ideally, the model must be composed from 'definitive descriptive building blocks' that enable investigators to set more focussed objectives during the investigatory process.

3. *satisfying.* The model must fit well with the investigatory agency's wider objectives, including any statutory obligations. It should not compromise the agencies 'credibility' or the technical quality of its work.

4. *comprehensive.* The model must capture both the initiating events and the consequences of an incident. It must capture all significant events. It must avoid ambiguity or gaps in understanding.

5. *disciplining.* The incident model must provide a rigorous framework that both directs and helpts to synchronise the activities of individual investigators. It should also provide a structure for the validation of their work.

6. *consistent.* This criterion urges that the incident model must be 'theoretically consistent' with the investigatory agencies safety program.

7. *direct.* The model must help investigators to identify corrective actions that can be applied in a prompt and effective manner. It should not be necessary to construct lengthy narrative histories before an immediate response can be coordinated.

8. *functional.* Accident models must be linked to the performance of worker tasks and to workflows. It should enable others to see how the performance of these tasks contributed to, mitigated or exacerbated the consequences of incident.

9. *noncausal.* "Models must be free of incident cause or causal factor concepts, addressing instead full descriptions of incident phenomenon, showing interactions among all parties and things rather than oversimplification; models must avoid technically unsupportable fault finding and placing of blame" [73].

10. *visible.* Models must help people to see relevant aspects of an incident. This should include interactions between individuals and systems. These representations must be accessible to investigators and to members of the general public who may themselves be 'victims' of an incident.

These criteria illustrate the way in which Benner's accident or incident models can be seen as models or templates for the incident reconstructions that have been described, for instance, in Chapters 7.3 and 9.3. The recommendation that models must capture the 'initiating events and the consequences of an incident' was a recurring theme of the earlier sections in this book. There are, however, some important differences between the approach developed in his paper and the perspective adopted in this book. For instance, Benner's criteria intend that accident models should be 'noncausal'. In contrast, we have argued that investigators cannot avoid forming initial hypotheses about the causes of an incident during the early stages of an investigation. Investigators must be encouraged to revise these working hypotheses as their work develops [851].

Benner's concept of an accident or incident model helps to determine what investigators consider to be relevant when analysing a particular 'failure'. In consequence although his model requirements focus on primary and secondary investigations, they indirectly determine the information that will be available to any causal analysis. In addition to the model criteria, list above, Benner proposes the following list of methodological requirements:

1. *encouragement.* This criteria argues that methodologies must encourage the participation of different parties affected by an investigation. Individual views must also be recognised and protected within such an approach.

2. *independence.* It is also important that methodologies should avoid 'blame'. The role of management, supervisors, employees must be recognised within the methodology.

3. *initiatives.* Personal initiatives must also be supported. It should produce evidence about previous failures that promotes intervention and shows what is needed to control future risks in the workplace.

4. *discovery.* Methodologies must support the timely discovery of information about an incident. It should also be clear when the discovery of such information may be delayed until a credible sample has been established or until "causality requirements are met" [73].

5. *competence.* This criteria argues that methodologies must leverage employees' competence. For example, it should be supported by training. This, in turn, should support the detection, diagnosis, control and mitigation of risk.

6. *standards.* Methodologies must provide credible and persuasive evidence for setting or reinforcing safety standards. It must also enable investigators to document and monitor the effectiveness of those standards over time.

7. *enforcement.* This criteria is intended to help ensure that a methodology can be used to identify potential violations. The methodology must explore deviations from expected norms. Compliance problems must be identified.

8. *regional responsibility.* In the US context, methodologies must help individual States to ensure that incident reports provide consistent and reliable accounts of accidents and incidents. More generally, methodologies must identify the role that regional organisations can play in identifying safety objectives from individual reports.

9. *accuracy.* Methodologies must validate the products of an investigatory process. It must assess the technical 'completeness, validity, logic and relevance' of these outputs.

10. *closed loop.* Methodologies must close the loop with design practices. Previous risk assessments often contain information about anticipated failure modes. These can be assessed against what actually did happen during an incident. In turn, future safety assessments can be informed by the results of previous investigations.

Benner identified the personal bias that affected his analysis. The detailed scores from his investigation are, therefore, less interesting than the criteria that drove the evaluation. In passing, it is worth noting that models which tended to represent accidents as processes were rated most highly according to the criteria listed above. These included the event process model and the energy flow process model. Elements of both of these techniques were incorporated into Benner's work on P-Theory and the STEP method mentioned previously. MORT was also ranked in the top three places according to these criteria. Similar findings were reported for the methodologies that were examined. Events analysis was rated most highly. MORT was ranked in second place assuming that it incorporated the ECF extensions described in Chapter 9.3 [430].

Many of the techniques that were assessed by Benner continue to have a profound impact upon existing investigatory practice. For instance, MORT is still advocated as a primary analysis technique by the US Department of Energy almost two decades after it was originally developed. Many aspects of accident and incident investigation have, however, changed since Benner first published his analysis. In particular, there has been a growing recognition of the organisational causes of adverse occurrences [702]. Software related failures also play a more significant role in many investigations [413]. The following paragraphs, therefore, use Benner's criteria to structure an evaluation of the causal analysis techniques that have been presented in previous pages.

Encouragement

This criteria was intended to ensure that methodologies encourage participation in an investigation. It is difficult, however, for many people to follow the detailed application of some causal techniques that have been presented in this chapter. This might act as a disincentive to participation during certain stages of a causal analysis. For instance, it can be hard to follow some of the statistical techniques if people are unfamiliar with their mathematical underpinnings. Similarly, the Explanatory Logic that supports WBA can be difficult to communicate to those without a training in formal logic. Fortunately, the proponents of mathematical techniques for causal analysis have recognised these objections. WBA is supported by a diagrammatic form that provides important benefits for the validation of any proof. Similarly, individuals can participate in the application of Bayesian techniques, for instance through the procedures of subjective risk assessment, without understanding all of the formulae that an investigator may employ during the more final aspects of the analysis.

These communications issues also reveal tensions within Benner's criteria. Mathematically-based techniques, typically, benefit from well-defined syntax and semantics. They provide proof rules that offer objective means of establishing the completeness and consistency of an analysis. These strengths support the accuracy criteria, assessed below, but are achieved at the expense of potentially discouraging some participation in the application of the analysis techniques. Conversely, the accessibility of Tripod, of ECF, MES and of STEP all encourage wider participation in the analysis. There is, however, a strong subjective component to the forms of analysis that are supported by these techniques. There is also a danger that participation without strong managerial control can compromise the findings of any analysis.

Independence

This criterion is intended to ensure that any methodology addresses the 'full scope' of an incident. It should consider the role of management, supervisors and employees without connotations of guilt or blame. Some techniques, including Tier analysis and MORT, provide explicit encouragement for investigators to consider these different aspects of an incident. As we have seen, however, there is a strong contrast between causal analysis techniques that offer checklist support and those that expect investigators to scope their analysis. It would be perfectly possible to apply tier analysis to an incident but for the analyst to overlook a particular sections of a management structure. In contrast, the MORT diagram provides explicit guidance on the roles that contribute to an incident or accident. The difficulty with this approach is that it can be difficult for analysts to match the details of a particular situation to the pre-defined scope in a checklist-based approach.

This criteria introduces further tensions between the Benner criteria. For example, techniques that encourage an independent assessment of the various contributions to an incident and accident can also lead to the tensions and conflict that discourage widespread participation. Many analysis techniques almost inevitably create conflict as a by-product of their application within the investigatory process. For instance, there are many reasons why non-compliance analysis creates resentment. It gains its analytical purchase from the observation that many incidents and accidents stem from individual and team-based 'failures' to follow recognised standards. Chapter 2.3 has also explained how observations in the health care and aviation domains have shown that operators routinely violate the myriad of minor regulations that govern their working lives. These violations do not have any apparent adverse consequences and often help individuals to optimise their behaviour to particular aspects of their working environments. In extreme examples, they may even be necessary to preserve the safety of an application process. Non-compliance analysis should reveal these violations. Investigators must, therefore, be careful to pitch their recommendations at a level that is intended to achieve the greatest safety improvements without *unnecessarily* alienating other participants in the investigatory process. Other causal analysis techniques raise similar issues. For example, Tier analysis is unlikely to promote harmony. Investigators successively associate root causes with higher and higher levels within an organisation. As mentioned, this encourages the independent analysis of the many different parties who can contribute to an adverse occurrence. However, it can lead to strong feelings of guilt, blame and anxiety as root causes are successively associated with particular levels in a management structure.

Initiatives

This criterion is intended to ensure that any accident or incident methodologies provide adequate evidence to encourage the focussed actions that address risks in a specific workplace. The previous sections in this book have not explicitly considered the means by which such recommendations for action can be derived from the products of any causal analysis. This is the central topic of Chapter 11.5. We have, however, considered how some analysis techniques can be used to derive particular recommendations. For instance, our analysis of PRISMA included a description of Classification/Action Matrices. These enable investigators to simply 'read-off' an associated action from a table once the causes of an incident have been determined by previous stages in the analysis. MORT offers similar support. Table 11.7 presented the 'Stage 2' analysis form proposed by the US Department of Energy. This is encourages analysis to enumerate the different ways in which an incident can be explained by the particular failures that are enumerated in the branches of a MORT diagram. The frequency with which specific items in the *why* branch are identified helps to establish priorities for action. These, in turn, help to justify the initiatives and interventions that are recommended by Benner's criteria. As we shall see in Chapter 14.5 similar summaries can be derived by collating the causal analysis of several incidents.

As before it is possible to highlight differences between the checklist and 'free form' approaches. Techniques, such as PRISMA, encourage a consistent approach to the recommendations and initiatives that are motivated by a causal analysis. Investigators have limited scope to alter the actions that are recommended by particular cells within a Classification/Action matrix. Conversely, less structured techniques enable investigators to tailor their response to the particular circumstances of an incident. The consequence of this is that without additional methodological support it is likely that different investigators will initiative different responses to very similar incidents.

Discovery

This criterion requires that an incident methodology should encourage a 'timely discovery process'. We have shown in previous paragraphs that there are tensions between Benner's criteria, for example encouragement can conflict with accuracy. This criteria illustrates a form of internal conflict. For example, checklist approaches are likely to provide relatively rapid insights because the items that they present can help to structure causal analysis. In contrast, techniques such as ECF analysis or Bayesian approaches to statistical causality are likely to take considerably longer given that investigators lack this guidance. In contrast, the application of 'raw' checklists is unlikely to yield entirely innovative insights. Investigators will be guided by the items that have already been identified by the author of the checklist. Free-form techniques arguably offer less constraints to the discovery process.

As mentioned, these methodology criteria were originally intended to support a comparison of accident investigation techniques. The causal analysis of safety-critical incidents creates new challenges. For instance, the statistical analysis of a body of incident reports can be used to yield new insights that might not emerge from the study of individual mishaps. The techniques that we have summarised in this chapter each pose different problems for the application of this form of discovery through data mining. For instance, the subjective nature of ECF analysis can make it very difficult to ensure that different investigatory teams will identify the same root causes for the same incident. This creates problems because any attempt to retrieve incidents within similar root causes will miss those records that have been 'miss-classified'. It will also yield reports that are potentially irrelevant from the perspective of the person issuing the query if they cannot follow the justification for a particular classification. In contrast, PRIMA's use of the Eindhoven Classification Model is intended to reduce inter-analyst variation by providing a high-level process to guide the allocation of particular categories of causal factors. Problems stem from the use of causal trees prior to the the use of the classification model. Subtle changes to the structure of the tree will have a significant impact on the number and nature of the root causes that are identified for each incident. This, in turn, can have a profound impact on the discovery of causal factors through the analysis of incident databases.

The argument in the previous paragraph assumes that causal analysis techniques have a measurable impact upon both the speed of an investigation and the likelihood that any investigation will yield new discoveries. As we shall see, some initial evaluations have shown that the investigators' background has a more profound impact upon these issues than their application of a particular technique. The discovery of particular causes can be determined by the investigator's familiarity with the nature of the corresponding more general causes. Individuals with human factors expertise are more likely to diagnose human factors causes [484].

Competence

The competence criterion requires that any methodology must help employees to increase the effectiveness of their work. This implies that they must be able to use a causal analysis technique in a cost-effective manner. Appropriate training must enable individuals to detect, diagnose, control and ameliorate potential risks. This criteria has clear implications for the more mathematical techniques that we have examined. The Explanatory Logic of WBA will only deliver the intended benefits of precision and rigour if those who apply it are correctly trained in its many different features. The partition approach to probabilistic causation provides a more pathological example of this. It is unclear precisely what aspects of the existing theories might actually be recruited to support incident investigation. The development of appropriate training courses, therefore, lies in the future. Conversely, these approaches offer means of objectively assessing the competence of individuals in the application of a causal analysis technique. Mathematical rules govern the use of statistical data and the steps of formal proofs. Tests can be devised to determine whether or not individuals understand and can apply these mathematical concepts. Such evaluations are more difficult to device for less formal approaches where the rules that govern 'correct' transformations are less clearly defined.

Techniques such as MORT and Tripod have already been widely adopted by commercial and industrial organisations [430, 702]. Training courses and commercial software can also be recruited to improve employee competence in the application of these techniques. PRISMA arguably rests halfway between these more commercial techniques and the more novel mathematical approaches to causal analysis. As we have seen, there is limited evidence that this approach can be used in a range of different contexts within several industries. These is, as yet, relatively little guidance on how investigators might be trained to exploit this approach. It is important to emphasise that the lack of training materials does not represent a fundamental objection to any of the techniques that have been considered in this book. Our experience in training investigators to conduct various forms of causal analysis has shown that most organisations tend to develop their own training materials. It is important that any causal analysis technique supports both their organisational priorities and also the reporting obligations that are imposed on them by other statutory and regulatory bodies.

Standards

The standards criterion requires that incident methodologies must enable investigators to identify potential flaws in their work. They must also provide a comprehensive, credible, persuasive basis for the advocacy of appropriate interventions. This criterion served as a prerequisite for the inclusion of causal analysis techniques in this book. It is possible, however, to identify a number of distinct approaches that are intended to improve the standard of incident investigation within the different approaches that we have analysed.

Arguably the weakest claim that is made for causal analysis techniques is that they provide intermediate representations, typically figures and graphs, that can be exposed to peer review during the investigatory process. ECF charting, non-compliance tables, MES flowchart all help to explicitly represent stages of analysis. This can help to identify potential contradictions or inconsistencies that might otherwise have been masked by the implicit assumptions that are often made by different members of an investigatory team.

Many of the techniques that we have studied also provide particular heuristics or rules of thumb that provide a basis for more complex forms of analysis. MES, STEP, ECF, MORT, WBA all exploit variants of counterfactual reasoning. This approach offers considerable benefits not simply because it encourages a consistent approach to the identification of causal factors. Counterfactual reasoning also provides investigators with a common means of identifying potential counter-measures. Recall that the counterfactual question can be expressed as 'A is a necessary causal factor of B if and only if it is the case that if A had not occurred then B would not have occurred either'. We might, therefore, consider preventing an incident, B, by devising means of avoiding B. As we have seen, however, causal asymmetry implies that investigators must be circumspect in exploiting techniques which advocate this style of analysis. Further questions arise both from the cognitive problems of reliably applying counterfactual reasoning [124] and from the practical problems of validating hypothetical reasoning about the failure of barriers, described in Chapter 9.3.

This criteria also urges that causal analysis techniques should be assessed to determine whether they provide a comprehensive, credible, persuasive basis for the advocacy of appropriate interventions. The interpretation of a 'comprehensive' approach depends upon the scope of the techniques. This was addressed in the previous section. It s more difficult to assess the credibility and persuasiveness of an approach. We are unaware of any studies that have directly addressed this issue as part of an evaluation of causal analysis techniques. Similar studies in other domains have, however, indicated that the application of a particular method may be less important than the identity of the individual or group *who* apply the technique [280].

Enforcement

Benner's criteria require that an incident methodology should reveal expectations about the norms of behaviour. This, in turn, helps investigators to identify non-compliance. It should also help to identify instances in which the enforcement of standards has been insufficient to prevent violations. As with the other criteria, each of the techniques that we have assessed can be argued to offer different levels of support for these aspects of 'enforcement'. For example, non-compliance analysis was integrated into our use of ECF in Chapter 9.3. This technique is deliberately intended to identify violations and to expose deviations from expected norms. The Explanatory Logic that supports the formal components of WBA also includes deontic operators that explicitly capture notions of obligation and permission. These have been used to represent and reason about the particular consequences of non-compliance with standards and procedures [118, 469].

A number of caveats can be made about this criterion and its application to causal analysis techniques. Firstly, it can be difficult to distinguish a willful violation from a slip or a lapse. The distinction often depends upon the analyst's ability to identify the intention of an individual or group. None of the causal analysis techniques that we have investigated support the inference of intention from observations of behaviour. The PARDIA components of WBA provide a possible exception to this criticism. Unfortunately, there are relatively few examples of this technique being used to analyse complex, operator behaviours. It remains to be seen whether this approach might be used to enable analysts to reason about the detailed cognitive causes of violation and non-compliance.

A second caveat to Benner's enforcement criteria is that numerous practical difficulties frustrate attempts to chart the differences that exist between everyday working practices and the recommendations of standards and regulations. Chapter 9.3 showed that it was extremely difficult for executives, managers and supervisors to keep track of the dozens of procedures and guidelines that had been drafted to guide the development of the Mars Polar Lander and Climate Orbiter projects. Previous paragraphs have also noted the high-frequency of apparently minor violations that have been noted as characteristic of expert performance within particular domains, especially aviation [674]. The 'enforcement' criterion, therefore, represents a class of requirements that are currently not adequately met by any causal analysis techniques. These criteria can be contrasted with other requirements, such as the need to provide 'standards' for causal analysis, which are arguably satisfied by all of the techniques that we have examined.

Regional responsibility

This criterion was initially drafted to ensure that individual States are encouraged to use a methodology and to take responsibility for its application within the context of U.S. Health and Safety "mandates" [73]. In contrast, we argue that causal analysis techniques must consider the different roles and objectives that distinguish regional organisations from the local groups that, typically, implement reporting systems. Some causal analysis techniques seem to be better suited to regional organisations. For instance, Chapter 9.3 shows how Tier Analysis associates root causes with higher levels in an organisational hierarchy. This process is likely to create conflicts between local investigators and senior members of a management organisation. Regional investigators are more likely to possess the competence and independence that are necessary to resist the pressures created by such conflicts. Other techniques, such as WBA, can be so costly in terms of the time and skills that are required to exploit them that regional and national groups must be involved in their application. In contrast, the methods that might be derived from probabilistic models of causality are likely to benefit from the information contained in large-scale datasets. Regional organisations may, therefore, be best placed to offer advice and support in the application of these techniques.

The aims and objectives of national and regional organisations are likely to be quite different from those of the local teams who are responsible for conducting an incident investigation. Regional organisations are more concerned to derive a coherent overview from a collection of incident reports than they are to understand the detailed causal factors that led to one out of several thousand or hundreds of thousands of incidents [444]. An immediate concern to mitigate the local effects of an incident are part of a wider concern to ensure that lessons are propagated throughout an industry. It is, therefore, important that regional organisations should understand and approve of the causal analysis techniques that are used to provide data for these aggregate data sources. They may also impose training requirements to ensure competence in the application of those techniques [423].

A number of further complications can, however, frustrate regional and national initiatives to exploit the products of causal analysis techniques. For instance, regional bodies are often anxious to ensure that investigators exploit similar techniques so that accurate comparisons can be made between individual reports from different areas of their jurisdiction. It is likely, however, that there will be pronounced local distortions even if different geographical regions all share the same causal analysis techniques. A succession of similar incidents or an accident with notably severe consequences can sensitise investigators to certain causes. This effect may be more pronounced for those who are most closely associated with previous incidents [412]. In consequence, very different results can be obtained when the same incidents are reclassified by investigators from different regions and even from different teams. These issues complicate attempts to share data across European boundaries in the aviation domain [423] and across State boundaries within US healthcare [453]. They are also largely orthogonal to the problems of identifying appropriate causal analysis techniques.

Accuracy

This criteria is similar to aspects of the 'standards' requirement that was introduced in the previous paragraphs. Accuracy is intended to ensure that incident methodologies can be tested for completeness, consistency and relevance. All three of these concepts are relevant to the causal analysis of safety-critical incidents. The first two directly motivated the application of formal proof techniques to support semi-formal argumentation in WBA. As mentioned, however, it can be more difficult to validate techniques that exploit precise mathematical concepts of consistency and completeness. A further caveat is that the use of formal techniques does not guarantee an error-free analysis [21]. It does, however, provide objective rules for identifying and rectifying these problems.

The other techniques that we have presented, such as STEP, MES and MORT, provide fewer rules that might be applied to assess the accuracy of a causal analysis. Instead, as mentioned, they rely upon diagrams and tables that are open for peer review. The less formal processes involved in achieving group consensus are intended to provide greater confidence than the formal manipulations of abstractions whose precise interpretation can defy even skilled mathematicians. A similar debate between informal, semi-formal and formal methods has dominated areas of computing science research for several decades [32]. The detailed comparison of the strengths and weaknesses of these different approaches lies outside the scope of this book. In particular, such comparisons have little value unless they can be substantiated by detailed empirical evidence. Later sections will briefly summarise the preliminary results that have been obtained in this area [530, 553]. Unfortunately, these findings provide limited insights into the application of particular approaches. They do not support more general conclusions about comparative benefits and weaknesses. In consequence, investigators must make their own subjective assessments of the claims that proponents make for the 'accuracy' of their causal analysis techniques.

Closed Loop

This final criterion requires that incident methodologies should be tightly integrated into other aspects of systems design and implementation. The data from incident reporting systems should inform future risk assessments. Information from past risk assessments, or more precisely the arguments that support those assessments, can also help to guide a causal analysis providing that it does not prejudice investigators' hypotheses. We have not suggested how any of the causal analysis techniques that we have examined might support satisfy such requirements. There are, however, many similarities between non-deterministic causal models and the techniques that support quantitative approaches to reliability and risk assessment. The prior probabilities of Baysian analysis can be derived from the estimates that are embodied in risk assessments, especially when reliable data cannot be derived directly from an incident reporting system.

Previous paragraphs have, however, provided an example of a risk assessment technique being used to guide the causal analysis of an adverse incident. Chapter 9.3 described how NASA's mishap investigation board identified a problem in the software that was designed to automatically re-establish communications links if the up-link was lost during the Polar Lander's Entry, Descent and Landing phase. This bug was not detected before launch or during the cruise phase of the flight. A Fault Tree analysis did, however, identify this possible failure mode after the Polar Lander had been lost.

Chapter 11.5 will return to this issue in greater detail. The relationship between risk assessment, design and incident reporting is often only considered as an after-thought by many organisations. In consequence, subjective assessments of component and system reliability are often exploited by one group within a company while others in the same organisation continue to collate data about the actual performance of those systems [417]. More generally, this same situation can arise when design groups are unwilling to approach other organisations within the same industry who have previous experience in the incidents that can arise from the operation of particular application processes. In consequence, development resources are often allocated to perceived hazards that cannot be justified by data about previous failures.

The previous paragraphs have shown how Benner's methodological criteria can be used to structure a comparison of causal analysis techniques. Some of the criteria, such as 'accuracy' and 'standards', are relevant objectives for all of the approaches that we have examined. Other requirements, such as 'encouragement', are less important for particular techniques. WBA and techniques derived from partition models of probabilistic causality focus more on 'accuracy' and 'competence'. The main weakness with this approach is that Benner fails to provide any objective procedures that might be used to determine whether or not a particular methodology satisfies a particular criterion. It is, therefore, possible for others to argue against our analysis. For example, it might be suggested that partition methods can encourage greater participation. Certainly, the diagrammatic forms of WBA do help to increase access to some aspects of this technique. There have, however, been no empirical studies to investigate the communications issues that might complicate the use of these formal and semi-formal techniques within the same approach. The following sections, therefore, briefly describe the limited number of studies that have been conducted in this area. These studies do not provide a firm basis for the comparative evaluation of causal analysis techniques. They focus on a limited subset of the available approaches. They also concentrate on incidents within particular industries. These studies do, however, illustrate the manner in which empirical evidence might be recruited to support assertions about the relative merits of these techniques.

11.4.3 Experiments into Domain Experts' Subjective Responses

Both Van Vuuren's bottom-up analysis of the PRISMA approach and Benner's application of model and methodology criteria were driven by the direct involvement of the individuals who were responsible for conducting the tests. Van Vuuren participated in the analysis that is summarised in Tables 11.13, 11.14 and 11.15. Benner performed the ratings that were derived from the lists of criteria presented in the previous section. This level of personal involvement in the validation of causal analysis techniques should not be surprising. Previous sections have summarised the practical, theoretical and ethical issues that complicate the evaluation of different causal analysis techniques. Many researchers, therefore, side-step the problems of investigator training and recruitment by conducting subjective studies based on their own application of alternative techniques. In contrast, Munson builds on the work on Benner [73] and Van Vuuren [845] by recruiting a panel of experts to validate his application of causal analysis techniques [553]. Munson began by applying a number of analysis techniques to examine a canyon fire that had previously been investigated by the U.S. Fire Service. In particular, he applied Fault Tree analysis, STEP and Barrier analysis to assess the causal factors that contributed to this incident. He then recruited his panel by choosing wildland firefighting experts rather than 'professional' investigators; this "most accurately emulates real world situations where investigators may have some investigative experience but their primary occupation and training is not in these techniques" [553]. None of the evaluators had any prior experience with accident analysis techniques. This helped to avoid any preference for, or experience of, existing approaches. Each member of the panel had a minimum of fifteen years experience in wildland fire suppression and were qualified to 'Strike Team Leader' level. Individuals were selected on a 'first come' basis.

Munson acknowledges that this may have introduced certain biases, however, he endeavoured to ensure that none of the panel consulted each other about their ratings. He was also aware that the panel members may have held preconceived ideas about the causes of the case study; "since they were evaluating the investigation methods and not the reinvestigation of the fire, bias should have been reduced" [553].

The members of the panel were asked to compare Munson's Fault Tree analysis, STEP analysis and Barrier analysis of the canyon fire by rating each technique against a number of criteria. These requirements were based on a subset of the Benner criteria [73]. As can be seen, some of these requirements apply more to reconstruction and modelling than they do to causal analysis. This can be justified by the mutual dependencies that we have stressed in previous chapters. Munson's criteria can be summarised as follows:

1. *Realistic.* Any analysis must capture the sequential, concurrent, and interactive nature of the flow of events over time.

2. *Comprehensive.* Any analysis must identify the beginning and the end of an accident sequence and there must not be any gaps in the investigator's understanding of an incident.

3. *Systematic.* Any analysis must be supported by a logical and disciplined method that allows for peer review and mutual support by all of the members of an investigation team.

4. *Consistent.* The method must be consistent and it should be possible for investigators to verify that any conclusions are correct from the information that is available.

5. *Visible.* Any analysis must discover and present the events and interactions throughout an accident sequence so that colleagues can understand the manner in which they contribute to an incident.

6. *Easy to learn.* It should be possible for investigators to learn how to apply a technique by attending a one week course. This criterion reflects Munson's focus on the fire fighting community and he acknowledges that it should not be considered when attempting to assess the 'best' analysis technique.

The experts were asked to use a ranking system that was similar to that described in the previous section; "The rating scale follows Benner's [73] approach in that until a more comprehensive scale is developed to better differentiate levels of compliance to the criterion, a more simple direct measurement scale is appropriate" [553]. For each model, they were asked to rate each criterion. A score of zero was used to denote that they did not believe that the approach met this criterion. A score of one was used to denote indicate that they believed that the approach addressed the criteria but not completely and improvement would be required. A score of two was to be awarded if the analysis technique satisfied the criterion. No weighings were applied to the results of this process because no criterion was perceived to have more significance than any other. The results from summing the individual scores showed that STEP received the highest rating; 52 from a possible 60 (87%). Fault Tree Analysis received 51 out of 60 (85%). Barrier Analysis obtained 42 out of 60 (70%). STEP was rated as the most 'comprehensive' (100%) and most 'consistent' (100%). Both Fault Tree Analysis and STEP were rated as the 'easiest to use' (90%). Barrier analysis was rated the most 'realistic' technique (90%). Fault Tree Analysis was rated as the most 'systematic' method (100It was also the most visible (90%). Two evaluators rated it as the best overall approach. Two rated STEP the highest. One assigned equal value to both STEP and Fault Tree Analysis. Barrier Analysis was not assigned the highest rating by any of the evaluators.

Munson also analysed his data to assess the level of agreement between his panel of assessors. Multivariate techniques were not used; "the number of evaluators and criteria were considered too small and would not constitute any meaningful insight" [553]. Instead, Perreault and Leigh's [676] index was used to assess inter-rater reliability. Indexes above 0.85 are considered to indicate a high degree of consensus. Levels below 0.80 require further analysis. Munson provides the following equation for the reliability index. F is the frequency of agreements between the evaluators, N is the total number of judgements and k is the number of categories:

$$I_r = [(F/N) - 1/k)][k/k - 1)]^{0.5} \tag{11.33}$$

The panel's evaluation of the six criteria for the STEP method yielded an index of 0.84 [676]. Fault Tree Analysis received 0.86 over all of the criteria. Barrier Analysis achieved a reliability index of 0.79. The inter-rater reliability for all methods was 0.84. As can be seen, only the Fault Tree assessment indicated a high degree of consensus but all other measures fell into the acceptable region identified by Perreault and Leigh [676]. If we look at levels of agreement about individual criteria it is possible to see some interesting patterns. For example, there was little agreement about whether or not Fault Tree analysis was a 'realistic' technique (0.63). STEP received the highest rating for 'comprehensiveness' and achieved an index of 1.0 for inter-rater reliability. Fault Tree analysis and Barrier analysis achieved a similar level of consensus but at a lower over rating about the 'comprehensiveness' of the techniques. Munson provides a more sustained analysis of these results [553].

The experts were each asked to provide additional comments about the applicability of each method. Munson cites a number of the responses that were provided. Ironically some of these comments reveal the experts' lack of understanding about the technical underpinning of the methods that they were asked to evaluate. The attitudes to Fault Tree analysis are particularly revealing in this context, given the key role that they play within many areas of reliability and risk assessment:

> "One evaluator liked the way Fault Tree Analysis visually presented complex events and the way it showed accidents as a chain-of-events as opposed to a single random occurrence... They thought that this method might be better at uncovering managerial/administrative latent factors contributing to the incident than the other two methods. In contrast, one evaluator responded that the STEP method appeared more stringent in revealing underlying human causal factors. They commented that STEP (and Control/Barriers Analysis) provided an approach that was more likely to distinguish more abstract human factors from hard factual data considerations and therefore be better at raising questions into human error causes... All evaluators expressed concern that Control/Barriers Analysis was inadequate in determining causal factors when applied to wildland firefighting. It had strengths in identifying needed and/or compromised barriers at an administrative level but the dynamic and highly variable aspect of the firefighting environment made its application to investigations inadequate" [553].

Munson concludes that STEP is the most 'desirable' method for the investigation of wildland firefighter entrapments. The small differences between the scores for this technique and for Fault Tree analysis suggest, however, that there are unlikely to be strong differences between these two techniques. Both were rated more highly than Barrier Analysis.

A number of questions can be raised both about the methods that Munson used and about the results that he obtained from them. Firstly, Munson was not qualified in accident or incident investigation when he undertook this study. The manner in which he applied the three techniques need not, therefore, have reflected the manner in which they might have supported an active investigation by trained personnel. Secondly, a number of caveats and criticisms have been made about his application of particular techniques. For example, Fault Tree analysis of the canyon fire breaks some of the syntactic conventioned that are normally associated with this approach, see Chapter 9.3. Paradoxically, these differences aid the presentation of Munson's analysis. They also make it difficult to be sure that the results from this study could be extended to the more general class of Fault Trees that obey these syntactic conventions. Thirdly, this study focuses on experts who only represent a very small cross-section of the community who are involved in accident and incident investigations. This is a universal weakness shared by all previous validation studies that we have encountered. Chapter 3.7 has shown that incident and accident reporting systems involve individual workers, supervisors, investigators, safety managers, regulators and so on. Benner's original 'encouragement' criteria captures some aspects of this diversity. However, experimental validations focus on the utility of causal analysis techniques for investigators or, as in this case, domain experts. Regulators might take a very different view. Fourthly, a number of minor caveats can be raised about the choice of statistical techniques that were used to analyse the data from this study. Multivariate analysis might have been applied more systematically. This could have yielded results that are easier to interpret than the piecemeal application of Perreault and Leigh's index. Finally, Munson's study specifically addresses the fire fighting domain. Several of the criteria were specifically tailored to reflect the working and training practices of this application area. Further studies are required to replicate this work in other domains.

It is important to balance these criticisms and caveats against the major contribution that has been made by Munson's work. The opening paragraphs of this section reviewed the many pragmatic, theoretical and ethical barriers that complicate research in this area. His approach successfully addresses many of these potential problems. Muson shows that it is possible to provide further evidence to support Benner's subjective analysis.

11.4.4 Experimental Applications of Causal Analysis Techniques

Previous sections have described a number of different approaches to the validation and comparative evaluation of causal analysis techniques. Van Vuuren adopted a bottom-up approach by applying PRISMA to support a number of incident reporting systems within particular industries. Benner adopted a much more top-down approach when he developed and applied a set of criteria in a subjective evaluation of accident models and methodologies. Munson used this approach as the foundation for an expert panel's evaluation of causal analysis techniques for fire fighting incidents. A limitation of Benner's approach is that it was based upon the subjective analysis of the researcher. Munson avoided this by recruiting a panel of experts. They did not, however, apply any of the methods and only provided subjective ratings based on a case study that was developed by Munson himself. Van Vuuren's study involved investigators in the application of the PRISMA technique. He, however, played a prominent role in coaching the use of this approach; "guidance was necessary to pinpoint these mistakes or lapses and by doing this to improve the quality of the causal trees and stimulate the learning process regarding how to build causal trees" [845]. This intervention was entirely necessary given the ethical issues that surround the validation of incident investigation techniques using 'live' data. The closing sections of this chapter describe an alternative approach. McElroy attempted to integrate elements of Munson's more controlled experimental technique and Van Vuuren's concern to involve potential end-users in the application of particular approaches [530].

McElroy's evaluation began with a sustained application of the PRISMA technique. He used a variant of the Eindhoven Classification Model to identify the root causes of more than one hundred aviation incidents from the ASRS dataset. This yielded approximately 320 root causes; the majority of which related to human factors issues. In order to validate his results, he recruited a number of experts to repeat his analysis of selected incidents from the study. The intention was then to compare the causal trees that they produced and the resulting root cause classification with McElroy's findings from the initial analysis. He rejected Munson's approach of recruiting domain experts, such as pilots or air traffic controllers. This was partly motivate by pragmatic reasons, such as the difficulty of securing access to participants for the study. It was also motivated by the difficulties that Munson and Benner had foreseen in training domain experts to apply novel analysis techniques, rather than simply requiring them to comment on the use of the approach be someone else. In contrast, McElroy recruited participants who had specific expertise or training in incident and accident analysis. This approach also raised problems; he found it difficult to secure the involvement of participants with similar expertise and training. Both of these factors are significant given Lekberg's results, which show that the investigator's training will influence their causal analysis of safety-critical incidents [484]. In the end he was only able to assess the application of the technique by two participants. In consequence, his findings cannot be used to support general conclusions about the PRISMA technique. They do, however, provide a glimpse of some of the individual differences that might affect the application of causal analysis techniques by incident investigators.

As mentioned, McElroy provided his participants with short synopses of incidents that had previously been submitted to the ASRS. The following paragraph provides a brief extract from the summary that McElroy used in his study:

> ACCESSION NUMBER : 412640
> DATE OF OCCURRENCE : 9808
> NARRATIVE : DEPARTING NEWPORT ARPT, AT THE TIME OF DEP, THE W HALF OF THE ARPT WAS STARTING TO FOG IN. I HOVER-TAXIED TO THE FAR E END OF THE ARPT AND WAS ABLE TO TAKE OFF IN BLUE SKIES AND UNLIMITED VISIBILITY. THIS ARPT IS SET UP FOR A CTL ZONE WHEN THE VISIBILITY IS LESS THAN 3 MI AND A 1000 FT CEILING. THERE WAS ANOTHER HELI IN THE PATTERN WHOM I WAS IN RADIO CONTACT WITH. HE GAVE ME PERMISSION TO TAKE OFF FIRST AND THEN HE WENT IN AND LANDED. ALL OF THIS WAS DONE VFR ON THE E END OF THE FIELD WHILE THE W END WAS FOGGED IN. THE STANDARD FOR THE OTHER ARPTS WITH CTL TWRS HAS BEEN IF I WAS INSIDE OF THEIR CTL ZONE AND IT WAS IN EFFECT, THEY HAVE ALLOWED ME TO WORK INSIDE THE CTL ZONE WITHOUT A SPECIAL VFR IF I WAS IN THE STANDARD VFR CONDITIONS. ALL I NEEDED TO DO WAS MAKE RPTS OF MY LOCATIONS WHILE WORKING IN THEIR AIRSPACE. AS LONG AS I WAS VFR, I DID NOT NEED A SPECIAL VFR TO BE INSIDE THE AIRSPACE. MY POINT TO ALL OF THIS IS THAT IT IS NOT TAUGHT TO NEW STUDENTS THIS WAY SO IT BECOMES MORE LIKE JUST A STORY WHEN AN OLDER PLT DOES SOMETHING LIKE THIS. IT IS LEGAL TO DO BUT NOT GOOD FOR STUDENTS TO SEE. NOT SURE OF HOW OR WHERE TO MAKE A POINT OF THIS, OR IF MAYBE IT IS NOT A RELATIVE POINT TO MAKE AT ALL. HOPE THIS IS NOT TOO CONFUSING, AND THANK YOU FOR YOUR TIME.

The first participant produced the Causal Tree shown in Figure 11.19. McElroy's study focussed more on the application of this diagram to support PRISMA's root cause analysis. A number of insights can, however, be derived from this initial stage of his evaluation. The tree took several hours to construct but, as can be seen, it is essentially a sketch of the incident. It includes inferences and judgements that cannot directly be supported from the synopsis. For instance, one note is annotated to denote that the helicopter pilot took off illegally, happy he was on a visual flight rule. Nowhere does the report state that the pilot was 'happy' with the state of affairs. Similarly, the causal tree refers to the maneuver as 'illegal' although the pilot believes that there actions were 'legal' within the control zone of the airport tower. This ambiguity reflects a lack of contextual information and the participants' limited domain knowledge. It was not, however, addressed in McElroy's analysis. A key point here is that although this evaluation ran for several hours, the participants never had the opportunity to move beyond this high-level sketching to the more meticulous analysis that would be need to demonstrate the sufficient and correctness of a causal 'explanation'. One might, therefore, infer that such an experimental procedure would have to be significantly revised if it were to be used to assess the utility of one of the mathematical techniques that we have described.

As mentioned, McElroy's aim was to determine whether participants who were training in incident analysis would confirm his own application of PRISMA. The first participant was, therefore, asked to use their diagram in Figure 11.19 to drive the categorisations of root causes using a variant of the Eindhoven Classification Model. They identified the following list of potential causes:

- Operating procedures. This is represented by the node labelled OP in the Eindhoven Classification Model of Figure 11.9. The participant identified that the incident was the result of inadequate procedures.

- Management priorities. This is represented by MP in the Eindhoven Classification Model. The participant identified that top or middle management placed pressure on the operator to deviate from recommended or accepted practice.

- Permit. This is represented by HR2 in the Eindhoven Classification Model. The participant identified that the operator failed to obtain a permit or licence for activities where extra risk was involved.

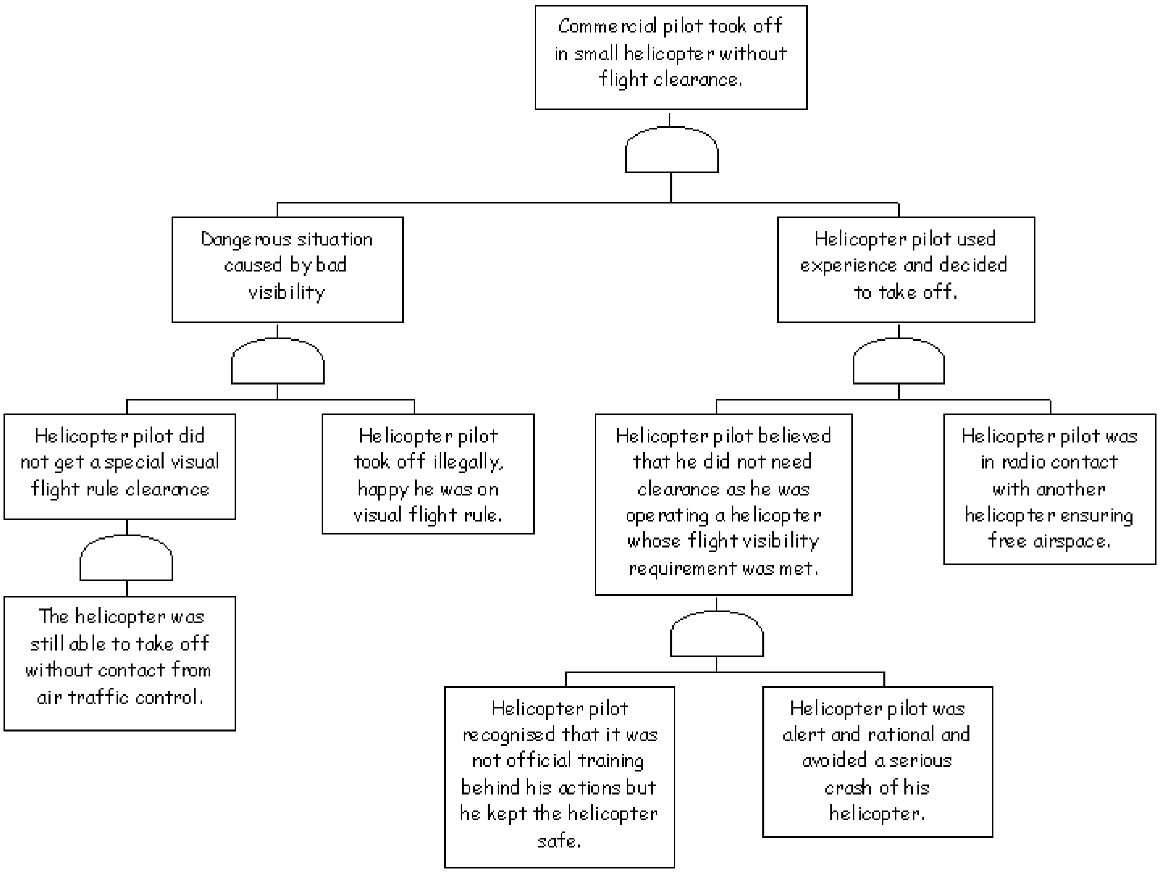

Figure 11.19: Causal Tree from McElroy's Evaluation of PRIMA (1)

- Planning. This is shown as HR5 in the Eindhoven Classification Model. The participant identified that the activity was not fully planned. Appropriate methods were not identified nor were they carried out in a well-defined manner.

- Unclassifiable behaviour. This is shown as X in the Eindhoven Classification Model. The participant also identified that some of the causal factors denoted in Figure 11.19 could not be classified using the model.

In contrast, McElroy's analysis only identified management priorities and planning as causal factors in this incident. The other three causes identified by the participant were not identified in the initial analysis. In addition, McElroy's analysis identified Goal? (HK2) as a potential cause that was not recognised by the first participant. This root cause categorisation denotes that the operator failed to identify appropriate goals or priorities for their actions. This comparison raises several issues. Firstly, the study tells us what categories the participant felt were important to the causes of the case study. It does not tell us *why* they believed them to be important. This is important because both McElroy and the first participant identified planning as a causal factor, it is entirely possible however that they had entirely different reasons for making this categorisation. Conversely, we do not know the reasons why they differed over specific elements in their causal analysis. Secondly, it is difficult to determine the justification for some of the reported conclusions made by both McElroy and the participant. Although the previous quotation is an abbreviated from of the incident report that was supplied during the study, there is no explicit indication that management priorities had caused the pilot to behave in the manner that they reported. This illustrates the more general point that we have made repeatedly in this book; it is not sufficient simply to present a causal analysis without providing a detailed justification of the reasons supporting that analysis.

As mentioned, the second evaluation focussed on an incident from the ASRS' air traffic dataset:

> ACCESSION NUMBER : 425120
> DATE OF OCCURRENCE : 9901
> NARRATIVE : WX WAS SUNNY BUT COLD, A DAY OR 2 AFTER A SNOW/ICE STORM. SABRELINER WAS TAXIING OUT FOR IFR DEP. ATC OBSERVED THE FUSELAGE WAS COVERED WITH SNOW AND ICE. ATC ADVISED THE PLT 'IT APPEARS THERE'S A LARGE AMOUNT OF SNOW AND ICE ON THE TOP OF YOUR ACFT.' THE PLT STATED 'IT'S NOT A LOT, IT'S A LITTLE, AND IT WILL BLOW OFF WHEN WE DEPART.' ON TKOF ROLL, ICE WAS OBSERVED PEELING OFF THE FUSELAGE. THIS CONTINUED AS THE ACFT CLBED OUT. ICE WAS OBSERVED FALLING ON OR NEAR A HWY JUST OFF THE DEP END OF THE RWY. THE ACFT WAS SWITCHED TO DEP, BUT A FEW MINS LATER RETURNED FOR LNDG. AS THE ACFT TAXIED IN, SIGNIFICANT ICE FORMATION WAS OBSERVED ON THE ELEVATORS. THE ACFT TAXIED TO AN FBO AND WAS DEICED BEFORE TAXIING BACK OUT FOR DEP. I SPOKE WITH THE FBO LATER. THEY SAID THEY HAD SEEN THE PLT CLRING SNOW AND ICE OFF THE ACFT BEFORE HE FIRST DEPARTED. HOWEVER, THE UPPER SURFACE OF THE ELEVATORS WAS TOO HIGH FOR THE PLT TO SEE FROM THE GND.

The second participant produced the Causal Tree shown in Figure 11.20 for this incident report. McElroy's again analysis focussed on the causal factors that were identified using a variant of PRISMA's Eindhoven Classification Model. As before, however, this diagram yields several insights into the assessment of causal analysis techniques. There is a far greater level of detail in this tree than in Figure 11.19. There is insufficient evidence to determine whether this is an artifact of individual differences between the participants or whether it stems from differences in the two incidents that they studies. As with many aspects of McElroy's work, it provides tantalising hints of deeper issues. He did not counter-balance the study so that each participant was asked to analyse each incident. This had been an intention behind the study but he ran out of time. Rather than rush the participants to perform a partial study of two incidents, he chose to allow them more time with a single incident.

Both Figures 11.19 and 11.20 are sketches. They record the participants' initial thoughts about the incidents. They follow the high-level structure proposed by the causal tree approach; left branches represent the 'failure' side while the right branch denotes 'recovery' events. There are also examples in both trees where that participants either deliberately neglect the syntax of there trees or else they did not follow the syntactic rules that were presented. In Figure 11.19, there is a minor violation with an AND gate that includes a single event. It can be argued that this represents a stylistic issue rather than a violation f any explicit syntactic rule. In this case, however, it is uncertain how to interpret the relationship between Helicopter pilot did not get a special visual flight rule clearance and The helicopter was still able to take off without contact from air traffic control. Figure 11.20 raises more questions. No checklist or protocol is linked to two events without any intervening gate. The event labelled Pilot dismiss ATC concerns is provided as an input to two different AND gates. Such techniques break the independence assumptions that are important for the analysis of more 'conventional' fault trees. These rules were, almost certainly, not presented to the participants in McElroy's study. Such annotations are, therefore, of considerable interest because they illustrate ways in which users are shaping the notation to represent the course of an incident. In future studies, it would be important to know what was intended by the event labelled No checklist or protocol. This would enable us to determine whether the notation fails to support a necessary feature or whether the training failed to convey significant syntactic constructs to the participants. Given that participants were unlikely to derive a reliability assessment from Figure 11.20 it can be argued that the independence assumption has not value for the practical application of causal trees?

As with the first participant for the helicopter case study, the second participant was also asked to use their causal tree to drive the categorisation process that is supported by the Eindhoven Classification Model. The following list summarises the categories of root causes that were identified

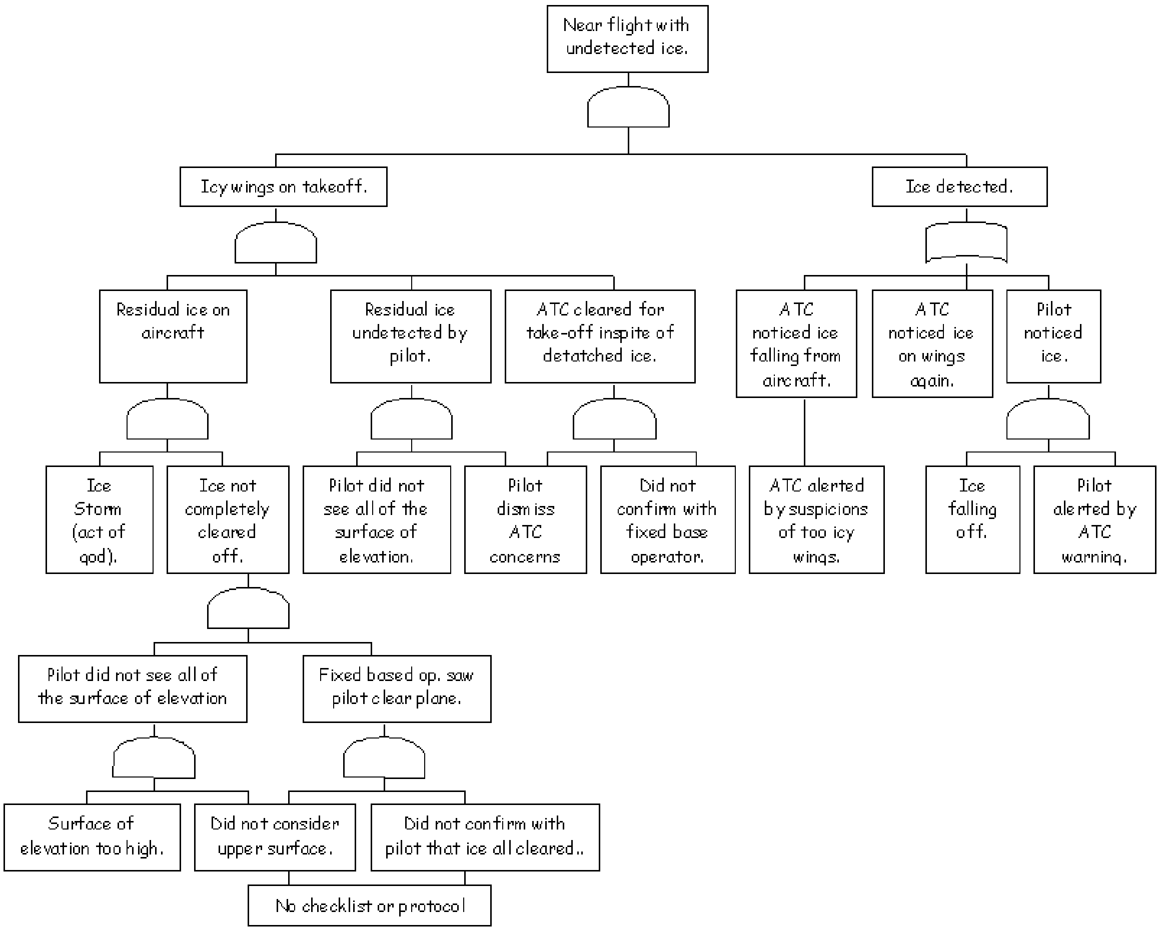

Figure 11.20: Causal Tree from McElroy's Evaluation of PRIMA (2)

during by the second participant:

- Operating procedures. This is represented by the node labelled OP in the Eindhoven Classification Model of Figure 11.9. The participant identified that the incident was the result of inadequate procedures. This category was also identified by the first participant for the helicopter case study.

- System Status. This is shown as HK1 in the Eindhoven Classification Model. The participant identified that the operator did not have an accurate knowledge of the "state and dynamics" of the system at key points during the incident [530].

- Permit. This is represented by HR2 in the Eindhoven Classification Model. The participant identified that the operator failed to obtain a permit or licence for activities where extra risk was involved. This category was also identified by the first participant for the helicopter case study.

- Checks. This is represented by HR4 in the Eindhoven Classification Model. The participant indicated that the operator had failed to conduct sufficient checks on the local system state to ensure that it complies with the expected conditions.

- Planning. This is shown as HR5 in the Eindhoven Classification Model. The participant identified that the activity was not fully planned. Appropriate methods were not identified nor were they carried out in a well-defined manner. This category was also identified by the first participant for the helicopter case study.

- Unclassifiable behaviour. This is shown as X in the Eindhoven Classification Model. The participant also identified that some of the causal factors denoted in Figure 11.19 could not be classified using the model. This category was also identified by the first participant for the helicopter case study.

McElroy's initial analysis had also identified Checks (HR4), Planning (HR5) and Unclassified behaviour (X) as root causes for this incident. The other categories were omitted. In addition, McElroy also identified License (HR1) as a causal factor. He argued that the operator in question must be qualified to do the job. He also identified Management Priorities (MP) as an issue in this incident. Top or middle management placed pressure on the operator to deviate from recommended or accepted practice. As noted in previous sections, it is difficult to reconstruct the thought processes that either of the participants used to justify their categorisation. McElroy notes in several places that the participants lacked the additional information that would have supported hypotheses about, for instance, the organisational causes of an incident. These are intriguing results. McElroy's results perhaps reflect the participants' suspicions that there must have been organisational causes to explain the operators' behaviour. If this is true then perhaps we are experiencing the consequences of the recent emphasis on the managerial and organisational causes of failure. These will be diagnosed as potential causes even when investigators are not provided with sufficient evidence to confirm these potential causes!

A number of methodological criticisms can be raised about McElroy's study. As mentioned, the lack of alternative data sources often forced the participants to make inferences and assumptions about potential causal factors. This led to causal trees and root cause classification that resembled rough 'sketches' of an incident. There criticisms can be addressed by acknowledging the severe time constraints that affected McElroy's work. They can also be countered by arguing that these high-level interpretations may resemble the level of detail that can be derived from an initial analysis of an incident report prior to a secondary investigation. It also provides an accurate impression of the 'rough' forms of causal analysis that can be conducted for contributions to anonymous incident reporting systems. In such circumstances, investigators are also constrained by the information sources that are available to them without compromising the identity of the contributor.

Further objections can be raised about the lack of empirical support for McElroy's work. He does not attempt to quantify agreement between his own causal analysis or that of the other participants. Given the limited data that he was able to obtain, this should not be surprising. He does not, however, speculate on measures that might be used. These is a vast range of techniques that can be used to represent and compare the structure of arbitrary tree structures [450, 451]. These algorithms might be used to detect patterns within the structure of causal trees. For example, Lekberg argues that an investigator's education background can bias their causal analysis [484]. Similarity metrics, for example based on vector representations, might be used to determine whether investigators from similar educational backgrounds produce measurably similar tree structures for the same incidents.

There are certain ironies about McElroy's approach. He framed his study as an experimental comparison between his own analysis and that of participants who were trained in incident analysis. he controlled the materials that were available to the participants and gave them the same training in the PRISMA technique. Having established these conditions, he lacked the resources to perform the additional tests that would have thrown light on many important issues. For instance, he did not counter-balance the incidents that were presented to the participants. This makes it difficult to determine whether any observed effects stem from the participant or the incident being studied. Similarly, McElroy only obtained access to two trained participants. Such a sample is inadequate to support any general conclusions. It should be stressed, however, that McElroy views his study as an initial experiment. It was intended to act as a marker for future studies that might attempt to assess whether investigators can *use* a causal analysis technique rather than just assessing their subjective attitudes towards someone else's application of an approach, as Munson had done [553].

This section has summarised recent attempts to assess the strengths and weaknesses of different causal analysis techniques. We have seen that these studies have only provided preliminary results.

The main conclusion from all of the work that we have cited in this section is that further research is needed to validate the many benefits that are claimed for the approaches that have been summarised in this chapter. It is, however, also possible to identify a number of more specific conclusions that might guide the future validation of causal analysis techniques:

- *Consider a broad range of stakeholders.* Previous studies have almost exclusively focussed on the investigators' assessment of causal analysis techniques. This is natural given that they are likely to determine whether or not a particular approach can be applied in the field. It should not be forgotten, however, that there are many other groups and organisations that must participate in, or validate, incident investigations. For instance, Chapter 2.3 discussed the role that regulators play in many reporting systems. A technique that satisfies investigators but does not convince regulatory bodies is unlikely to be acceptable. Similarly, it is important that any potential causal analysis technique should also be acceptable to those who must pay for its application. If this is not the case then there will be continued and increasing pressure either to reject the approach or to 'cut corners' in order to save resources.

- *Consider longitudinal factors as well as short-term effects.* All of the studies that we have presented are based around relatively short-term assessments of particular techniques. In particular, Munson and McElroy's evaluations took place over several hours. They do not, therefore, provide much evidence about the long-term benefits that might be provided by the consistent application of causal analysis techniques. There are also a range of detailed issues that are difficult to examine without more sustained studies. For instance, it is often necessary for investigators to revisit an analysis at some time after it was originally conducted. They may want to determine whether or not a subsequent incident has the same causal factors. In such circumstances, it is important not simply to identify the results of a causal analysis. It is equally important to understand the reasons *why* a particular decision was reached.

- *Consider the range of incidents in an evaluation.* It can be difficult to ensure that any assessment presents its participants with an adequate range of incidents. If this is not done then the utility of a causal analysis technique may be demonstrated for a sample of incidents that do not reflect the problems that are reported to the sponsoring organisation. There are further aspects to this issue. It may not be sufficient to base an evaluation on an accurate sample of current incidents. Given the overheads associated with training staff and financing the implementation of a new causal analysis technique, it is likely that any approach will be used for a significant period of time. If this is the case then any validation must also consider whether incidents will change during the 'lifetime' of an analysis technique. For example, Chapter 2.3 has argued that the increasing integration of computer-controlled production systems is posing new challenges in forensic software engineering. None of the techniques presented here, with the possible exception of WBA, explicitly addresses these challenges [413].

- *Consider the impact of individual or team-based investigation.* The studies of Munson, Benner and McElroy focussed on the performance and subjective assessments of individual investigators. Munson even went out of his way to prevent 'collusion' between the participants in his study. In contrast, Van Vuuren's evaluation involved teams of engineers, domain specialists, managers and safety experts. This reflects his intention to assess the application of this technique without the usual experimental controls that were accepted by Munson and McElroy. It is difficult, however, to determine whether team composition had any effect on the causal analyses reported by Van Vuuren. His published work says remarkably little about these issues. Work in other areas of groupwork have indicated that such factors can have a profound impact upon the successful application of design and analysis techniques [489, 557]. For example, the ability to use drawings and sketches as a medium of negotiation and arbitration can have a powerful effect during group confrontations. Attention may be focussed more on the shared artifact and less of the individuals involved in the discussion. We do not know whether these effects are also apparent in the application of causal analysis techniques.

- *Consider causal analysis in relation to other phases of investigation.* Benner reiterates the importance of evaluating any analytical technique within the context of a wider investigation

[73]. Analysis techniques are unlikely to yield sufficient explanations if investigators have not been able to elicit necessary information about the course of an incident. This argument was first made in the context of accident investigations. Unfortunately, cost limitations and the constraints of confidentiality/anonymity can prevent investigators from obtaining all of the data that they may need to complete a causal analysis. All of the techniques introduced in this chapter, with the exception of PRISMA, were developed to support accident investigations. These are, in one sense, information rich environments. In contrast, the particular characteristics of incident reporting systems may make relevant information very difficult to obtain. Any assessment must not, therefore, provide participants with information that they might not otherwise have available during the application of a particular technique.

- *consider which stage of an investigation is being assessed.* As we have seen, McElroy's initial evaluation of the application of an analysis technique produced results that were compatible with the early stages of an investigation. The participants produced trees that 'sketched' the outline of an incident. They did not produce polished artifacts that might provide consistent and sufficient causal explanations. Techniques that are intended to provide such quality control must, therefore, be validated in a way that enables investigators to achieve some degree of competence in the more 'advanced' use of the approach.

This is not an exhaustive list. Previous attempts to validate particular approaches have done little more than to sign-post areas for further work. It is equally important not to underestimate the importance of the small number of pioneering studies that have begun to validate the claimed benefits of causal analysis techniques.

11.5 Summary

This section has reviewed a broad range of techniques that can be used to support the causal analysis of safety-critical incidents. The opening sections build on our application of ECF analysis in Chapter 9.3 by introducing alternative event-based techniques. The related approaches of Multilinear Event Sequencing (MES) and Sequentially Timed and Events Plotting (STEP) were presented. These techniques all encourage analysts to use semi-formal, graphical or tabular notations to construct causal models of the events that lead to particular incidents. These notations provides great flexibility; investigators have considerable freedom in the manner in which they construct a causal model. Counterfactual reasoning is then, typically, applied to identify root causes from the candidate causal factors that are represented in a semi-formal model. Unfortunately, the flexibility offered by these approaches can also be a weakness. There are few guarantees that different investigators will derive the same results using this approach. Similarly, it is also unlikely that the same investigator will be able to replicate the details of their analysis at a later date.

Event-based techniques were, therefore, contrasted with approaches that exploit check-lists. These techniques provide investigators with a restricted choice of causal factors. Management Oversight and Risk Tree (MORT), Prevention and Recovery Information System for Monitoring Analysis (PRISMA) and Tripod all exploit variants of this underlying idea. The enumeration of causal factors guides and prompts investigators. It can also help to encourage consistency in an analysis. This is particularly important if national or regional comparisons are to be made between the causal factors of incidents that occur at a local level. Aggregated statistics would be unreliable if different investigators identified different causal factors behind the same incident. Of course, the price of consistency is that it may be difficult to identify an appropriate causal factor from the list of choices that are offered by these techniques. MORT and PRISMA address this potential caveat by encouraging investigators to extend the basic enumerations to reflect regional or domain-dependent variations in the incidents that are reported.

A further limitation of checklist approaches is that it can be difficult to check whether a particular analysis provides a consistent or sufficient explanation of a safety-critical incident. This chapter, therefore, introduced a range of formal causal analysis techniques. These approaches exploit mathematical systems of reasoning and argument to provide clear and concise rules about what can and what cannot be concluded about the causes of an incident. In particular, we have introduced WBA, partition techniques for non-deterministic causation and Bayesian approaches to subjective, probabilistic causation. Although these techniques are not widely used, they offer a number of potential benefits. They avoid many of the limitations that others have identified for the existing techniques that we have introduced in previous paragraphs [453, 482]. The rules that govern the application of these techniques provide objective criteria for verifying that an analysis is correct. The importance of ensuring the consistency and completeness of any analysis is also increasing significant given the rising cost of litigation in the aftermath of adverse occurrences. The modular approach supported by WBA and partition methods provides one means of addressing the increasing complexity of many safety-critical incidents. These benefits will only be realised if we can develop techniques that will enable non-formalists to participate in their application. At present, it can be difficult for those without an extensive background in mathematics to understand the strengths and the limitations of a particular formal analysis. Fortunately, many of the underlying mathematical models that support these causal analysis techniques can also be incorporated into software tools. There is also considerable potential for developing graphical and tabular representations that can be used to communicate more formal aspects of a causal analysis.

This chapter went on to describe attempts to validate some of the causal analysis techniques that we have described. Van Vuuren conducted bottom-up studies that involved the implementation of the PRISMA approach within several different industries [845]. He was then able to perform a comparative analysis of the different role that organisational factors played in a variety of different contexts. He did not, however, perform a detailed analysis of investigators' experiences in applying the causal analysis technique. In contrast, Benner provided a generic set of criteria that can be applied in a top-down manner to assess different accident models and methodologies [73]. By extension these same criteria might also be applied to assess different approaches to causal analysis. He relied largely upon his own subjective assessments. Munson, therefore, recruited an expert panel of fire fighters to apply similar criteria to a case study that had been analysed using Fault Trees, STEP and Barrier Analysis [553]. He was able to replicate results that suggested there were strong subjective preferences for STEP and Fault Trees over Barrier Analysis. Unfortunately, this study did not demonstrate that potential investigators might be able to apply any of these techniques themselves. McElroy, therefore, combined elements of Van Vuuren and Munson's approach when he asked a panel to apply the causal trees and Eindhoven Classification Model of the PRISMA technique [530]. This study revealed striking differences between the manner in which some people have proposed that causal analysis techniques should be used and the way in which investigators might actually use them in the early stages of an investigation. Rather than a detailed and careful analysis of the causal factors leading to an incident, the participants used them to sketch high level causes. They were less concerned with the consistency and sufficiency of an argument than they were with providing a clear overview of the incident itself. This, in part, reflects the important point that causal analysis techniques may have to play a variety of different roles during different stages of an investigation.

Our analysis of previous attempts to validate causal analysis techniques has revealed how little we know about the comparative strengths and weaknesses of these different approaches. We know from recent reports that current techniques are failing to support investigators tasks in many industries [482, 453]. This is another area that requires considerable further applied research so that practitioners can have greater confidence in the methods that are being proposed. The importance of this point cannot be underestimated. Several research teams are currently developing 'systemic' approaches to the causal analysis of incidents and accidents. These techniques are intended to address the challenges that are being posed by the failure of increasingly complex, tightly coupled systems. Unfortunately, less attention has been paid to the problem of demonstrating the practical benefits of these techniques than is currently being invested in their development.

It is worth emphasising that increasing complexity is one of several challenges that must be addressed by novel analysis techniques. They must also be proof against the sources of bias that influence the findings of many reports. Ultimately, it is not enough to show that any analysis technique can 'correctly' identify the causes of an incident. It must also demonstrate that it cannot easily be used to identify 'incorrect' causes. This is a significant burden given the many different forms of bias that might affect a causal analysis:

1. *Author bias.* This arises when individuals are reluctant to accept the findings of any causal analysis that they have not themselves been involved in.

2. *Confidence bias.* This arises when individuals unwittingly place the greatest store in causal analyses that are performed by individuals who express the greatest confidence in the results of their techniques. Previous work into eye-witness testimonies and expert judgements has shown that it may be better to place greatest trust in those who do not exhibit this form of over-confidence [224, 760].

3. *Hindesight bias.* This form of bias arises when investigators criticise individuals and groups on the basis of information that may not have been available to those these participants at the time of an incident. More generally it can be seen as the tendecy to search for human error rather than deeper, organisational causes in the aftermath of a failure.

4. *Judgement bias.* This form of bias arises when investigators perceive the need to reach a decision within a constrained time period. The quality of the causal analysis is less important that the need to make a decision and act upon it.

5. *Political bias.* This arises when a judgement or hypothesis from a high status member commands influence because other respect that status rather than the value of the judgement itself. This can be paraphrased as 'pressure from above'.

6. *Sponsor bias* . This form of bias arises when a causal analysis indirectly affects the prosperity or reputation of the organisation that an investigator manages or is responsible for. This can be paraphrased as 'pressure from below'.

7. *Professional bias* . This arises when an investigators' colleagues favour particular outcomes from a causal analysis. The investigator may find themselves excluded from professional society if the causal analysis does not sustain particular professional practices. This can be paraphrased as 'pressure from beside'.

8. *Recognition bias.* This form of bias arises when investigators have a limited vocabulary of causal factors. They actively attempt to make any incident 'fit' with one of those factors irrespective of the complexity of the circumstances that characterise the incident.

9. *Confirmation bias.* This arises when investigators attempt to interpret any causal analysis as supporting particular hypotheses that exist before the analysis is completed. in other words, the analysis is simply conducted to confirm their initial ideas.

10. *Frequency bias.* This form of bias occurs when investigators become familiar with certain causal factors because they are observed most often. Any subsequent incident is, therefore, likely to be classified according to one of these common categories irrespective of whether an incident is actually caused by those factors [396].

11. *Recency bias.* This form of bias occurs when the causal analysis of an incident is heavily influenced by the analysis of previous incidents.

12. *Weapon bias.* This form of bias occurs when causal analyses focus on issues that have a particular 'sensational' appeal. For example, investigators may be biased to either include or exclude factors that have previously been the focus of press speculation. Alternatively, they may become fixated on the primary causes of an incident rather than secondary causes that may determine the severity of an incident. For example, an investigation may focus on the driver behaviour that led to a collision rather than the failure of a safety-belt to prevent injury to the driver. This is a variant on the weapon focus that is described by studied into eye-witness observations of crime scenes [759].

The elements of this list illustrate the point that the success or failure of a causal analysis technique is, typically, determined by the context in which it is applied. For example, investigators can (ab)use causal analysis techniques by constructing causal chains that support particular, pre-determined conclusions. Such practices can only be discouraged by peer review during the various stages of a particular technique and by offering investigators a degree of protection against the sources of bias listed above. It should also be emphasised that causal analysis techniques are only one component in an incident reporting system. We cannot, therefore, assess the success or failure of such a system simply in terms of the sufficiency and completeness of the causal analyses that it produces. Such a validation must consider the success or failure of the recommendations that are justified by any causal analysis. These issues are addressed in the next chapter.