CHAPTER 1

What is System Safety?

Engineering systems, such as cars, aircraft, bridges, road systems, chemic-process plants, fuel storage and delivery depots, electric-power plants, and so on, can sometimes cause damage when they go wrong: people may be hurt or killed; the environment might be degraded or poisoned; property and equipment might be destroyed.

1.1 Safety is About Avoiding Damage

System safety is concerned with analysing the operation or misoperation of a system which might cause damage, and then assuring as far as possible that such damage won't happen. But sometimes it does. Events in which the operation or misoperation of a system resulted in damage are called accidents, and analysing accidents is also part of system safety.

There are two reasons to analyse accidents. One is to avoid similar accidents in the future by identifying causes, and inhibiting or mitigating one or more of those causes in future operations. The other is to assure a fair distribution of responsibility for the accident event, along with appropriate compensation for those affected. Although engineering accident investigators often profess an interest solely in the first of these, the second is a significant activity, as old as law itself, widely practiced, and socially valuable.

Accidents can occur through anomalous operation of a system, for example when the temperature in a chemical reactor becomes abnormally high and sets off a sequence of unwanted events resulting in an explosion. Accidents can also occur during normal operation, for example when a car is travelling along a road at a normal speed and suddenly another moving entity presents itself on a collision course and a collision ensues.

A system with operation which may result in an accident through normal or anomalous operation is often called a safety-critical system. A system which could contribute to accidents with other systems is often called a safety-relevant system. I'll use the term "safety-critical" to mean both here, and call them both SCS.

In order to try to avoid accidents with SCSs, they are analysed beforehand. Such a safety analysis (one geared towards avoiding or mitigating accidents) may be performed hand-in-hand with an analysis intended to enhance the reliability of the system (the proportion of time when a system successfully performs its required function). A separate analysis is needed for safety, for the circumstances in which an accident may happen during normal operations is of interest to the safety analyst but not to the reliability analyst.

1.2 Hazard Analysis (HazAn)

A safety analysis of a SCS usually starts with a hazard analysis. A hazard is a set of circumstances which precurses an accident. For example, the temperature rising in the reactor, or an entity presenting on a collision course with the vehicle. The point of identifying hazards is that there might be opportunities to take action: emergency cooling systems might be activated; a manoeuvre to avoid the incursing entity might be initiated. This might not always avoid the accident, but mitigating the consequences is also regarded as important: the reactor might still catch fire, but at least it doesn't explode; the vehicle collides with the entity, but at a lower speed and on a trajectory which results in less damage than otherwise.

A hazard analysis, or HazAn, consists of hazard identification, followed by an assessment of the severity, the damage caused in an accident following from the hazard.

Hazard Identification

As an example of hazard identification, we consider a generic communications bus for an automobile, in Figure 1.1. This example is considered in more detail in Chapter

9 of [12].

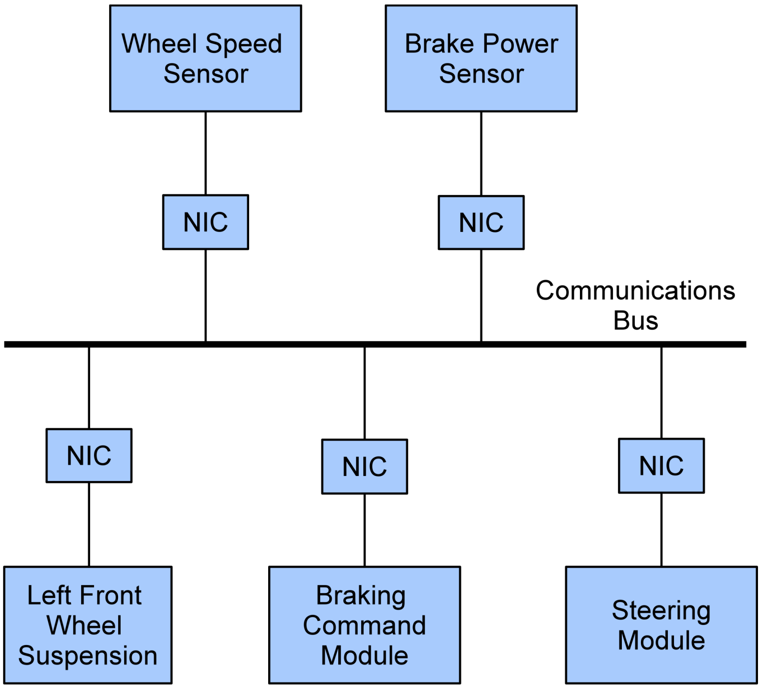

Figure 1.1: A Generic Communications Bus for an Automotive Road Vehicle

Various automobile components which need to communicate with each other over the bus are attached through Network Interface Controllers (NIC). Say the Braking Command Module (BCM) wants to send a message to the Left Front Wheel Suspension (LFWS). Then it will generate the command and specify the receiver, and give this information to the NIC. The NIC then interacts with the bus to deliver the message, say in the following way. Buses often have clocks attached to them. Individual components communicate on the bus in fixed-length time periods, specified by the clock signal. In the first time period, LFWS may send messages, say. In the next period, the BCM, in the next period, the Steering Module (SM), and so on until all have had one period; and then it begins all over again. Such a scheme is called time-triggered communication, or round-robin scheduling. In the time periods which are not "theirs", the NICs are all listening, and processing any messages intended for "their" component.

Hazard	Safety Requirement
NOT Integrity(Bus)	Not yet identified
NOT Integrity(NIC)	Not yet identified
Lost(msg)	Not yet identified
Corrupted(msg)	Not yet identified
InappropriateReceiver(msg,NIC)	Not yet identified
OutsideInterval(msg)	Not yet identified
PartlyOutsideInterval(msg)	Not yet identified
IntermittentlyAttached(NIC)	Not yet identified
CorruptedSending(msg,NIC)	Not yet identified
CorruptedReceiving(msg,NIC)	Not yet identified

Figure 1.2: A Table of Hazards Associated with the Communications Bus

In Figure 1, we can imagine that miscommunication could be hazardous. Say the SM is "steering right", and BCM is activating the brakes on driver command. The BCM might need the LFWS to stiffen, to resist the depression caused by braking while cornering. BCM puts this message on the bus. If LFWS hears no message (it is lost or destination is corrupted), or it reads "weaken" rather than "stiffen", we can imagine the car could well have a problem. Hence the list of hazards could look like Figure 2.

Here "Integrity" means the bus, or the NIC, is fulfilling its function as intended. If the bus is partly severed, say, or a NIC has a fused component on a chip, it may well not do so, and we say the component has lost integrity. The term "msg" refers to some message on the bus or being processed by a NIC. "Corrupted" means the contents of a message have been changed. "InappropriateReceiver" means that some NIC is receiving and acting on a message not intended for it. "OutsideInterval", resp. "PartlyOutsideInterval" means a message is not transmitted in the time slot, or partly across the boundary of a time slot, assigned to its NIC. Safety requirements which are assigned to a hazard have not yet been identified, because the severity of each hazard and mitigation or avoidance of the most severe consequences have not yet been considered.

The analysis performed so far could be part of Failure Mode and Effect Analysis (FMEA), a common form of analysis and quality control in the automotive industry. In an FMEA, it is not hazards per se which are listed, but failures of equipment and systems. Obviously, a failure can be a hazard, for example, if the brake hydraulic lines rupture. However, not all failures are hazards: if the engine fuel-injection develops a problem, the engine may produce less power than expected, but this is generally not dangerous. It is important to notice also that some hazards are not associated with failures, for example if a car is travelling too fast for environmental conditions. A failure results in a lack of a required function, and quality control is concerned with providing all required functions for as much of the operating period as possible. So quality-control engineering is primarily concerned with failures, whereas safety engineering primarily with hazardous failures, as well as control of hazards which do not represent failures but which are inherent in the operation. However, techniques such as FMEA are helpful for both failure analysis and hazard analysis.

Another common method for identifying hazards is Hazard and Operations Analysis (HAZOP), used widely in the process industries: chemical plants, power generation plants and so on. A HAZOP consists in describing operational characteristics, say a fluid flow, or a temperature, and then considering what happens when the value of this characteristic deviates: say it becomes too high, or too low, or flows in the opposite direction. To this end, HAZOP uses a series of guidewords, as in Figure 1.3, which are adverbs.

The first step of a HAZOP is to pick a characteristic, apply the guidewords to the characteristic, and figure out what this can mean! Consider, for example, fluid flow in a pipe. “Less” and “More” have obvious interpretations, as do “Slower” and “Faster”, namely that the flow can be faster or more than required, and also slower or less than required. But what about “Early” or “Late”? What might it mean to say that a flow is “early”? There seems to be no reasonable interpretation. So “slower/faster flow” could be something we need to consider, but “early/late flow” not. HAZOP when well practiced has proven itself invaluable in hazard analysis of process plants, and indeed other SCSs.

The goal of a hazard identification is to find all precursors of accidents. There is a certain amount of choice as to which conditions are chosen as precursors: at what temperature the reactor is considered to be in a hazardous condition; which near-field entities are regarded as possibly providing a collision threat. But whatever choice is made, the purpose of a hazard identification is to find them all, in the sense of identifying at least one precursor of each possible accident.

It is widely regarded by hazard analysts that there is no way really to tell whether

No
More
Less
As well as
Part of
Reverse
Other than
Early
Late
Before
After
Faster
Slower
Where else

Figure 1.3: HAZOP Guidewords

you have got them all – you just have to do the best you can. That is not quite true – there are indeed ways of enumerating hazards that allow you sometimes to tell whether you have them all [11], but for practical purposes an analyst must remain always open to the possibility that heshe has missed something important.

1.3 Severity Assessment – Completing the HazAn

Once one believes that all hazards have been identified, the next step is to identify the severity, the consequences that might follow from each one. So you consider the system in a given hazard state. Ask: what can now happen? This, and then that or that, and then ...... this way it leads to a severe accident, and that way only moderate damage. There is a divergent "tree" of possible behaviors from the hazard to an accident (or indeed a non-accident outcome).

For example, suppose a system contains some electronic equipment controlling electric current of a strength which could lead to injury through electric shock, or fire. Electric shock, or fire, is an immediate cause of damage. Since in a country such as Germany an order of magnitude more people die through fires started by faulty electrics and electronics than die through electric shock, let us order the severity of these events using three levels of increasing severity: “no dangerous effect” has Severity 0; electric shock has Severity 1; fire has Severity 2.

Suppose the electronics is damaged. There are a number of further events which could lead to damage, or not. If the electronics is damaged, then the protection mechanisms (we consider here just protection against residual current, and against overcurrent) may be intact, or may be rendered dysfunctional. If they remain intact, nothing dangerous happens (Severity 0). If overcurrent protection is damaged, then overcurrent is possible, which could lead to a fire (Severity 2). If residual current protection is damaged, then normally non-conducting parts of the system could become “live” and people in contact suffer an electric shock (Severity). This is straightforward and well understood. These events can be arranged in graphical-tree form, as in Figure 1.4.

This kind of representation of the possible consequences of a hazard has been formalised and is in wide use. Qualitative Event Tree Analysis (ETA) is an attempt to write down the possible future behaviors in this way that issue from a hazard state. The syntax of Event Trees is standardised, and looks visually somewhat different from Figure 4, but the intellectual content is similar. ETA has many more features, though, than appear in Figure 4. A more detailed explanation may be found in [1, 5].

When all the hazards have been identified, and all the behaviors that ensue from a hazard, including the behaviors that result in an accident have been enumerated, then one has completed the HazAn. If the HazAn has been performed correctly and exhaustively, then one knows all the ways things can go pear-shaped with the system. But how likely are events to turn out badly, as described?

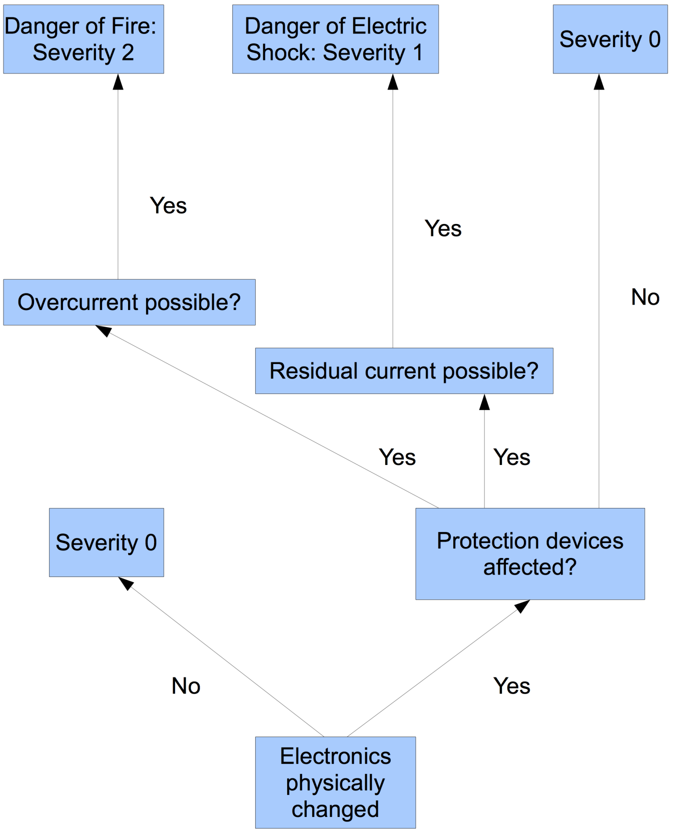
Danger of Fire: Severity 2
Danger of Electric Shock: Severity 1
Severity 0
Yes
Yes
No
Overcurrent possible?
Residual current possible?
Yes
Yes
Severity 0
Protection devices affected?
No
Yes
Electronics physically changed

Figure 1.4: An Event Tree Showing Possible Consequences of a Hazard

1.4 Risk Analysis – Assessing Likelihood

When a HazAn has been completed, assessing likelihoods of hazards and their outcomes is called risk analysis.

There are two components to an assessment of likelihood. One is the likelihood that a given hazard will occur. The other is the likelihood that a given type of accident will occur as a consequence of this hazard. By “type of accident” I mean the following. One often classifies accidents into severity classes. For example, here is a four-class classification: minor injuries and/or minor damage; some severe injuries and/or significant damage; one or a few deaths and/or major damage; many deaths and/or severe damage. Severity of road accidents amongst the OECD countries, for example are classified purely by human injury: no injury; minor injury; severe injury; one or more deaths. However, amongst these countries the criteria for belonging to one or another category differ considerably! Some count “severe” injury as anything requiring treatment for 30 days or more; others count “severe” injury as any treatment necessitating hospitalisation, no matter how brief. Some count a “death” as any fatality consequence to an accident within 30 days; others only count a “death” if it ensues within 24 hours. A “type of accident” is an accident with a specific severity, according to some severity classification scheme.

So how may we judge the likelihood that a given hazard will occur? One traditional method used for systems with mechanical and other components is Fault Tree Analysis (FTA). FTA is useful in cases in which the hazard has arisen because something has gone wrong or has broken – that is, a failure has occurred (but note that not all hazards are necessarily the result of failures!) A fault is supposed to be whatever abnormality caused the failure, and the idea of qualitative FTA is to proceed systematically to enumerate the things that can have gone wrong with subsystems and components that would have caused the failure, and then enumerate the faults that can have caused those, and so on until one arrives at the simplest kinds of component malfunctions. The most compelling feature of FTA is that the development and results are displayed as a graphic, a (combinatorial) “tree”.

For example, Figure 1.5 shows a schematic drawing of a pressure tank. The pressure tank has three inputs: steam, unspecified hydrocarbon, and a catalyst. The inputs react inside the tank, to produce an unspecified product, which a valve (on the right) lets out. The reaction involves (we assume) a possible increase in pressure in the

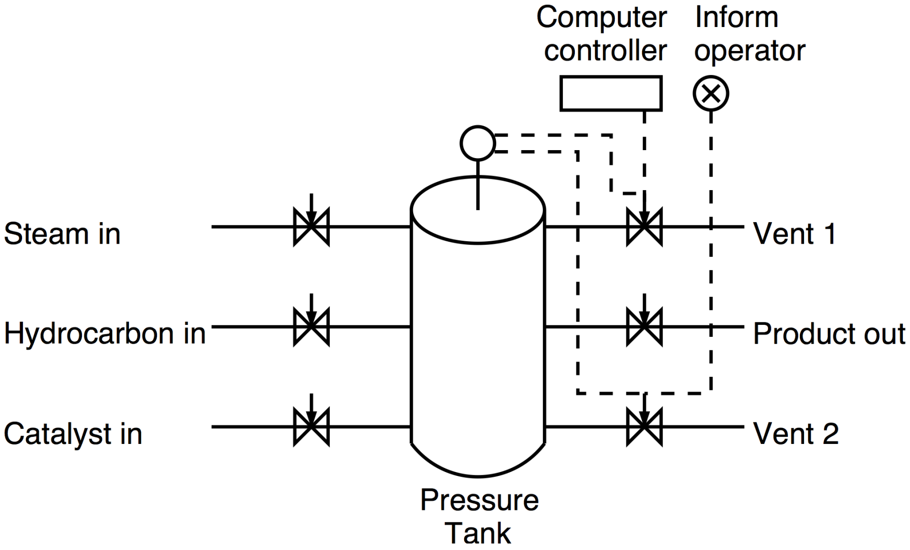

Figure 1.5: A Schematic Diagram of a Pressure Tank

tank, and in case this pressure should get too high, there are two relief vents, Vent 1 and Vent 2. Vent 1 is computer controlled, on the reading of a pressure sensor installed in the top of the tank. Vent 2 is human-operator controlled. The reading of the pressure sensor is displayed to an operator, who can then open Vent 2 manually. There is a potential weakness in that both emergency-relief systems depend on the same pressure sensor: if it reads too low when the pressure is too high, then both relief systems fail to operate and there is a danger of tank rupture ("explosion"). A fault tree showing the ways in which things can go wrong is shown in Figure 1.6.

The events in Figure 1.6 are shown in a semi-formal style of language, using pseudo-mathematical expressions such as "Ruptured(Tank)", which means that the property "Ruptured" applies to the object "Tank". Or, in plain language, the tank is ruptured. The construal of the engineering situation as objects, and properties which those objects have, as well as construing events as a change in properties of objects, has advantages for formal and semi-formal analysis and is part of our HazAn technique OHA. We don't consider it further here. An introduction to OHA may be found in the text [12].

The fault tree is what is called in branches of informatics an AND-OR tree, a

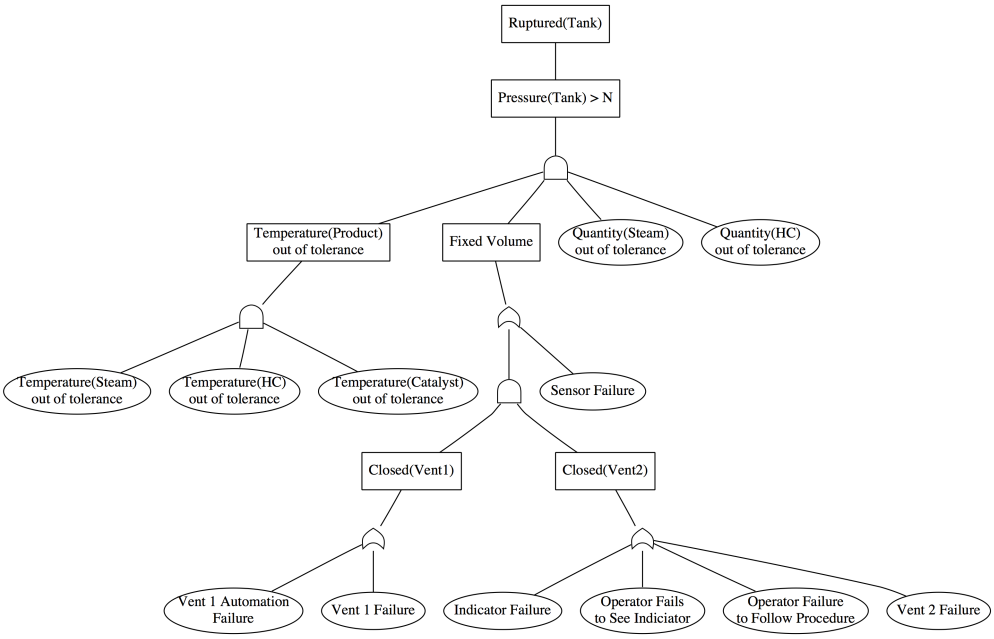
Ruptured(Tank)
Pressure(Tank) > N
Temperature(Product) out of tolerance
Fixed Volume
Quantity(Steam) out of tolerance
Quantity(HC) out of tolerance
Temperature(Steam) out of tolerance
Temperature(HC) out of tolerance
Temperature(Catalyst) out of tolerance
Sensor Failure
Closed(Vent1)
Closed(Vent2)
Vent 1 Automation Failure
Vent 1 Failure
Indicator Failure
Operator Fails to See Indiciator
Operator Failure to Follow Procedure
Vent 2 Failure

Figure 1.6: A (Qualitative) Fault Tree of the Pressure Tank

graphical representation of a series of assertions connected with the conjunctions AND or OR. Such an object is called a Boolean expression, after the logician and algebraist George Boole and his Boolean algebra. The fault tree is to be read as follows. The tank ruptures (the “Top Event” of the tree; the damage) if the Pressure of the Tank exceeds some unspecified value N (the first two levels). The Pressure in the Tank exceeds this value N if the Temperature of the Product is out of tolerance AND the Quantities of Steam and Hydrocarbon are out of tolerance AND there is no venting (“Fixed Volume”). The symbol joining these expressions means “AND”. It is the old symbol for an “AND” gate in digital electronics (in so-called combinatorial logic). The volume is fixed if either there is a sensor failure (so that the pressure sensor reads in order, but the pressure in the tank is not in order) OR both Vents remain closed. The symbol joining these two expressions is the old combinatorial-logic “OR”-gate symbol.

Fault Tree graphical representations are standardised, and differ somewhat from Figure 1.6 (for example, the layers are connected using joined horizontal and vertical line segments, rather than the direct angled lines as used in Figure 1.6).

Fault tree analysis (FTA) works best when there are just a few lower-level faults or failures which cause a failure at a higher level – the “tree” at such a point has few “branches”. A colleague who works as a safety engineer says that most of the fault trees he and his clients deal with have thirty to forty “nodes” (bubbles with text). However, some FTs for complex systems have thousands of nodes. Ways have been devised of splitting up FTs into segments with joining conventions, so that such big trees can be displayed accurately in normal documents. We don’t deal with any of those complications here.

FTA doesn’t work so well with digital systems because of the often practically innumerable number of possibilities there are for how program logic can go wrong.

This is still all qualitative – no likelihoods or other numbers have appeared. The FT contains events which indicate faults. To arrive at some estimate of how likely it is that the given hazard/failure will occur, faults in the qualitative FTA are assigned likelihoods, according to the following scheme. Faults in the “leaves” of the tree (at the bottom, with no other nodes below them) are assigned a likelihood, derived from engineering experience. Then the AND or OR node above the leaves is assigned a derivative likelihood, in the following fashion. If the node is an AND node, the probabilities of its leaves are all multiplied together. If the node is an OR node, the probabilities of its leaves are added together. The justification for these operations is as follows. The “leaf” events are assumed to have occurred independently of one another. The laws of probability say that if event A occurs independently of event B, the probability of event (A AND B) is the product of those probabilities, and the probability of event (A OR B) is the sum.

Given that the nodes above the leaves have now been assigned a probability, one can assign probabilities using the same procedure to the next layer of nodes, and so on up the tree until the Top Event is assigned a probability. The Top Event is the chosen hazard or accident, so one has assigned a likelihood to it as needed for the risk analysis. This is called quantitative FTA. It is one of the most widely-used methods in risk analysis. A more detailed explanation of FTA may be found in (BeCo01,KuHe96].

The major problem with quantitative FTA, though, is all those independence assumptions that have to be made to allow the calculations to go through to the Top Event. too often the independence assumption is wrong – either something occurs which triggers multiple faults that one has assumed in the FTA to be independent, so-called common-cause failure, or there are causal influences between components that have gone unremarked by the analyst and the triggering of the faults follows this causal influence. Indeed, there are well-known cases in which FTA analysis has been wildly mistaken – where the analysts concluded that specific failures were unlikely to happen once within the lifetime of the universe, but in fact these very failures started occurring with dangerous regularity. Quantitative FTA is not that reliable a guide to rates of failure.

A further difficulty with FTA is that it uses solely an analyst's intuition and experience for determining how failures arise from faults in components. There is no technical criterion which may be applied to check whether a proposed causal connection between component failure and (sub) system failure is indeed causal, or, if not, why not. When systems are complex, it is very easy to overlook things – but it is precisely when systems are complex that techniques such as FTA, which reduces failures of subsystems to straightforward failures of their components, is so useful. Not having some check whether you have got it "right" can be problematic.

FTA is often used to guide accident analysis also. If you know that in the course of an accident a particular failure has occurred, but you don't necessarily know how, then a fault tree for that failure will enumerate the possibilities (if the fault tree is correct!) and accident analysts can then search to see which of those causative failures actually occurred. Getting a correct fault tree is a major issue in these circumstances, and as noted above there is no way to check for correctness.

1.5 Accident and Incident Analysis - WBA

RVS-BI has a method for accident and incident analysis, Why-Because Analysis (WBA). We use a specific test for causality proposed by the philosophical logician David Lewis in 1973, which he suggested is ultimately derived from David Hume. We call it the Counterfactual Test. It is applied to two events, or an event and a state, and tells whether the one is a necessary causal factor of the other. In this way, the problem of correctness is solved. The Counterfactual Test is shown in Figure 1.7.

When applied to all phenomena of possible interest, the result is a causal picture

Event A occurred
Event B occurred
Suppose A had not occurred, but “everything else” had stayed“the same”.
Would B still have occurred?
If NO, then A is a necessary causal factor (NCF) in B
If YES, or COULD HAVE, then A is not an NCF of B

Figure 1.7: The Counterfactual Test

(actually a combinatorial graph), called a Why-Because Graph (WBG) of the factors which have been noted which are causal to the accident, and their interrelations. A WBG of an runway-overrun accident to a commercial aircraft, an Airbus A320, which occurred at Warsaw airport in 1993, is shown in Figure 8.

There is a further test in WBA which can indicate whether one is missing factors, although it cannot tell you exactly what is missing, only that there is something missing. The systematic extraction of factors from witness and other evidential reports, and then the testing of factors in pairs according to the Counterfactual Test, and the drawing of the resulting causal graph, the WBG, constitute WBA. WBA seems to be helpful in practice because the objective test, the Counterfactual Test, seems to be relatively easy to apply even for people new to WBA, and the resulting graphic, the WBG, is easy to read when a good layout engine is used to draw it.

More detail of WBA may be found at [15].

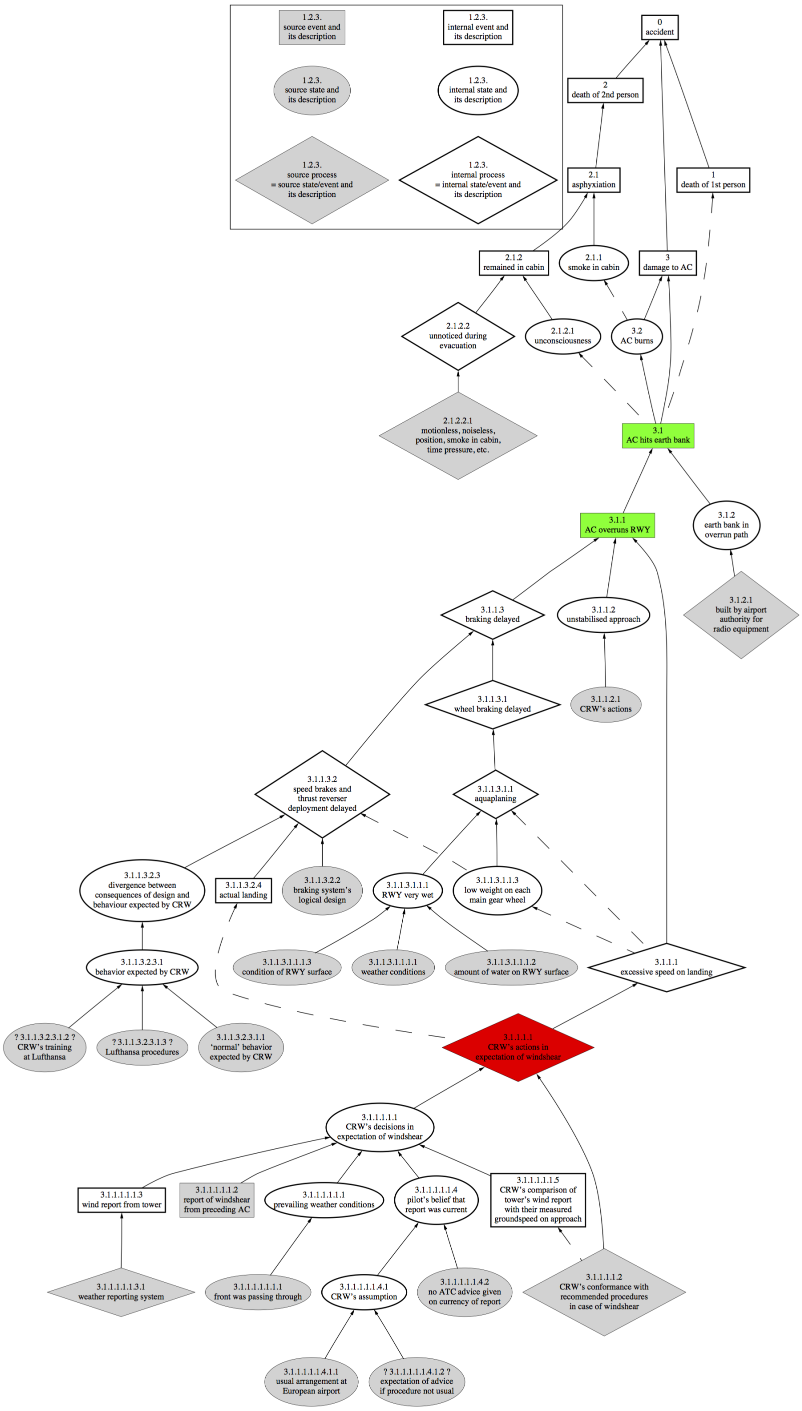
1.2.3. source event and its description
1.2.3. internal event and its description
1.2.3. source state and its description
1.2.3. internal state and its description
1.2.3. source process = source state/event and its description
1.2.3. internal process = internal state/event and its description
0 accident
2 death of 2nd person
2.1 asphyxiation
1 death of 1st person
2.1.2 remained in cabin
2.1.1 smoke in cabin
3 damage to AC
2.1.2.2 unnoticed during evacuation
2.1.2.1 unconsciousness
3.2 AC burns
2.1.2.2.1 motionless, noiseless, position, smoke in cabin, time pressure, etc.
3.1 AC hits earth bank
3.1.1 AC overruns RWY
3.1.2 earth bank in overrun path
3.1.2.1 built by airport authority for radio equipment
3.1.1.3 braking delayed
3.1.1.2 unstabilised approach
3.1.1.3.1 wheel braking delayed
3.1.1.2.1 CRW's actions
3.1.1.3.2 speed brakes and thrust reverser deployment delayed
3.1.1.3.1.1 aquaplaning
3.1.1.3.2.3 divergence between consequences of design and behaviour expected by CRW
3.1.1.3.2.4 actual landing
3.1.1.3.2.2 braking system's logical design
3.1.1.3.1.1.1 RWY very wet
3.1.1.3.1.1.3 low weight on each main gear wheel
3.1.1.3.2.3.1 behavior expected by CRW
3.1.1.3.1.1.1.3 condition of RWY surface
3.1.1.3.1.1.1.1 weather conditions
3.1.1.3.1.1.1.2 amount of water on RWY surface
3.1.1.1 excessive speed on landing
? 3.1.1.3.2.3.1.2 ? CRW's training at Lufthansa
? 3.1.1.3.2.3.1.3 ? Lufthansa procedures
3.1.1.3.2.3.1.1 'normal' behavior expected by CRW
3.1.1.1.1 CRW's actions in expectation of windshear
3.1.1.1.1.1 CRW's decisions in expectation of windshear
3.1.1.1.1.1.3 wind report from tower
3.1.1.1.1.1.2 report of windshear from preceding AC
3.1.1.1.1.1.1 prevailing weather conditions
3.1.1.1.1.1.4 pilot's belief that report was current
3.1.1.1.1.1.5 CRW's comparison of tower's wind report with their measured groundspeed on approach
3.1.1.1.1.1.3.1 weather reporting system
3.1.1.1.1.1.1.1 front was passing through
3.1.1.1.1.1.4.1 CRW's assumption
3.1.1.1.1.1.4.2 no ATC advice given on currency of report
3.1.1.1.1.2 CRW's conformance with recommended procedures in case of windshear
3.1.1.1.1.1.4.1.1 usual arrangement at European airport
? 3.1.1.1.1.1.4.1.2 ? expectation of advice if procedure not usual

Figure 1.8: The Warsaw Runway-Overrun Accident WBA

1.6 Risk Evaluation

The risk associated with a hazard, or an outcome, is normally considered in system safety engineering as a combination of likelihood of the circumstance occurring and the severity associated with that occurrence (see Chapter 9. The severity associated with a hazard is usually the worst outcome associated with the hazard). The purpose of system safety engineering is generally to reduce the risk as far as reasonably possible.

For complex engineered systems in which there are many differently-engineered components working together to achieve the overall purpose, such as a commercial transport aircraft, a need has been seen for some abstract guidance as to what is an acceptable risk (combination of likelihood and severity) and what is inacceptable.

The airworthiness regulations for commercial transport aircraft in Europe and associated lands are set by the European Aviation Safety Agency in Cologne, and are known as CS-25. CS-25.1309 is the section on equipment, systems and installations. CS-25.1309 (b) says the following (the likelihood and severity terms are defined in the next subsection):

> *(b) The aeroplane systems and associated components, considered separately*
> *and in relation to other systems, must be designed so that –*
> *(1) Any catastrophic failure condition*
> *(i) is extremely improbable; and*
> *(ii) does not result from a single failure; and*
> *(2) Any hazardous failure condition is extremely remote; and*
> *(3) Any major failure condition is remote.*

This regulation relates abstractly the severity of an outcome with its likelihood. It may be represented in a *risk matrix* as in Figure 1.9. The matrix is constructed from the text by noting that the likelihood categories are ordered:

$$\textit{extremely improbable} < \textit{extremely remote} < \textit{remote}$$

and reading the text "*Any <severity> condition is <likelihood>*" as "*Any <severity> condition is* at least *<likelihood>*". With this interpretation, the risk matrix represents exactly the likelihood-severity combinations which are regarded as acceptable or inacceptable in CS-25.1309. Note that there is an extra condition in CS-25.1309 besides the likelihood-severity combinations, namely that "*any catastrophic failure*

	Major	Hazardous	Catastrophic
Remote	Acc	Inacc	Inacc
Extremely remote	Acc	Acc	Inacc
Extremely improbable	Acc	Acc	Acc

Figure 1.9: Risk Matrix Derived from CS-25.1309

condition . . . does not result from a single failure. This condition is not represented in the risk matrix.

1.6.1 Definitions of Likelihood and Severity Categories

Risk matrices look simple, but the complexity of evaluation often lies in the definitions of the terms used in them, as we shall see here. What follows is somewhat involved, and is intended for illustration rather than to be understood in detail. Nevertheless, exact quotes from the regulations are appropriate, even if they are lengthy. For example, consider the simply-expressed requirement in the risk matrix that it is inacceptable for a condition classified as *catastrophic* to have as "high" a likelihood as *extremely improbable*. This translates using the definitions below into the compliance stipulation that software-based equipment (such as fly-by-wire control) whose failure might lead to loss of control (a catastrophic condition) should have an "average probability" of dangerous failure per flight hour of the order of 1×10^{-9} or less (a translation exercise left for the reader). This level of assurance is regarded by many authorities as inattainable in practice purely through software testing [10], which is still a method upon which much safety-critical software assessment is based, for example [7]. The risk matrix displays this requirement simply, without any hint of these difficulties in its application.

For a more detailed risk matrix that has been used in commercial air transport airworthiness certification, as well as detailed discussion of the categories defined below, as well as other useful techniques (such as a short introduction to fault trees and some probabilistic methods) see [14].

The definitions of the terms used in the risk matrix of Figure 1.9 are defined in

Book 2 of [4], the Acceptable Means of Compliance (known as AMC-25). Severity of failure conditions are classified in [1, AMC-25.1309 Section 6 Background, paragraph c (2) a] into

Minor Failure Conditions which would not significantly reduce aeroplane safety, and which involve crew actions that are well within their capabilities.

Major Failure Conditions which would reduce the capability of the aeroplane or the ability of the crew to cope with adverse operating conditions to the extent that there would be, for example, a significant reduction in safety margins or functional capabilities, a significant increase in crew workload or in conditions impairing crew efficiency, or discomfort to the flight crew, or physical distress to passengers or cabin crew, possibly including injuries.

Hazardous Failure Conditions, which would reduce the capability of the aeroplane or the ability of the crew to cope with adverse operating, conditions to the extent that there would be:

- (i) A large reduction in safety margins or functional capabilities;
- (ii) Physical distress or excessive workload such that the flight crew cannot be relied upon to perform their tasks accurately or completely; or
- (iii) Serious or fatal injury to a relatively small number of the occupants other than the flight crew.

Catastrophic Failure Conditions, which would result in multiple fatalities, usually with the loss of the aeroplane.

The associated likelihoods are classified in [1, AMC 25.1309 Section 6 Background, paragraph c (2) b] into

Probable Failure Conditions are those anticipated to occur one or more times during the entire operational life of each aeroplane.

Remote Failure Conditions are those unlikely to occur to each aeroplane during its total life, but which may occur several times when considering the total operational life of a number of aeroplanes of the type.

Extremely Remote Failure Conditions are those not anticipated to occur to each aeroplane during its total life but which may occur a few times when considering the total operational life of all aeroplanes of the type.

Extremely Improbable Failure Conditions are those so unlikely that they are not anticipated to occur during the entire operational life of all aeroplanes of one type.

Since it has been shown difficult in advance to anticipate the expected operational lifetime of transport aircraft, some of which such as the Boeing 737 and Airbus A320 proving unprecedentedly popular and going through many versions over many decades, quantitative probabilities are also associated with these qualitative likelihoods as follows, from [1, AMC 25.1309 Section 6 Background, paragraph c (2) c]

- Probable Failure Conditions are those having an Average Probability Per Flight Hour greater than of the order of 1×10^{-5}.
- Remote Failure Conditions are those having an Average Probability Per Flight Hour of the order of 1×10^{-5} or less, but greater than of the order of 1×10^{-7}.
- Extremely Remote Failure Conditions are those having an Average Probability Per Flight Hour of the order of 1×10^{-7} or less, but greater than of the order of 1×10^{-9}.
- Extremely Improbable Failure Conditions are those having an Average Probability Per Flight Hour of the order of 1×10^{-9} or less.

1.7 Risk Matrices

Some general principles in constructing risk matrices may be observed:

- If outcome A is acceptable for likelihood L, then it is acceptable for any likelihood lower than L.
- If outcome A is inacceptable for likelihood L, then it is inacceptable for any likelihood higher than L.

Risk matrices are a widely-used tool in qualitatively or semi-quantitatively evaluating risk with engineered systems. They are used, for example, for the evaluation of electronic railway systems in Europe under the Common Safety Methods [6], pursuant to European Directive 352/2009 [5]. For those who read German, a useful introduction to the details of risk analysis and assessment in rail automation is given in [3]. Another qualitative or semi-quantitative method of evaluating risk is risk graphs, commonly used with machinery as indicated in the intentional standard for safety in control systems for machinery, [8, Annex A].

1.8 Theory and Practice – Experience is Essential in Safety Analysis

Unlike other engineering disciplines in which academics are prominent, system safety depends a lot on experience. Most of the problems are quite practical, and do not have an abstract statement and a theoretical technical solution, as do many problems in, say, computer science, in which much of the engineering consists in getting theoretical solutions to work in an environment of actually-running hardware and software.

System safety is not a discipline particularly amenable to abstract statement – safety is not an abstract concept at all, but humanly based in everyday life and work. Good ideas are particularly prone to stumble on actual examples from everyday engineering. However, techniques from more theoretical work, such as David Lewis's and Hume's conception of causality, do have practical application, as we have shown. Techniques of formal refinement apply to hierarchical analysis of system hazards and safety requirements. There is a theoretical criterion for relative completenes of a hazard enumeration which can sometimes, or often, applied to practical hazard analyses. Our research consists not only in identifying such techniques, and working out how to apply them in real engineering contexts, but also working a lot of specific cases, to illustrate how these techniques really do apply, to identify the stumbling blocks and to overcome them, and suggest to others how they may be overcome. Some of our case studies we distribute openly as examples, and some of them are performed for private clients of our tech-transfer organisations.

Bibliography

[1] Book 2, *Acceptable Means of Compliance*, of [4].

[2] Tim Bedford and Roger Cooke, *Probabilistic Risk Analysis: Foundations and Methods*. Cambridge University Press, 2001.

[3] Jens Braband, *Risiskoanalysen in der Eisenbahn-Automatisierung*, Eurail Press, Hamburg 2005.

[4] European Aviation Safety Agency, *Certification Specifications for Large Aeroplanes CS-25*, (version of 2003-10-17. Latest version as of writing is 2012-01-27). Current version available from https://www.easa.europa.eu/certification-specifications/cs-25-large-aeroplanes

[5] European Commission, *Commission Regulation (EC) No 352/2009 of 24 April 2009 on the adoption of a common safety method on risk evaluation and assessment as referred to in Article 6(3)(a) of Directive 2004/49/EC of the European Parliament and of the Council*. Available from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:108:0004:0019:EN:PDF , accessed 2017-08-21.

[6] European Rail Agency, *Common safety methods for risk assessment*, available from http://www.era.europa.eu/Core-Activities/Safety/Safety-Management-System/Pages/Risk-Assessment.aspx , no date, accessed 2017-08-21.

[7] International Electrotechnical Commission, *IEC 61508, Functional safety of electrical/electronic/programmable electronic safety-related systems - Part 3: Software requirements*, 2nd Edition, 2010.

[8] International Organization for Standardization, *ISO 13849-1:2015, Safety of machinery – Safety-related parts of control systems – Part 1: General principles for design*, 2015.

[9] Hiromitsu Kumamoto and Ernest J. Henley, *Probabilistic Risk Assessment and Management for Scientists and Engineers*, Second Edition. IEEE Press, 1996.

[10] Peter Bernard Ladkin, *Causal System Analysis*, e-textbook. RVS-BI, 2001. Available from https://rvs-bi.de/publications/books/CausalSystemAnalysis/index.html

[11] Peter Bernard Ladkin, *Formal Definition of the Notion of Safety Requirement*, on-line article on the abnormaldistribution blog. RVS-BI, 2010. Available from https://abnormaldistribution.org/index.php/ 2010/11/09/formal-definition-of-the-notion-of-safety-requirement/

[12] Peter Bernard Ladkin, Bernd Sieker and Jan Sanders, *Safety of Computer-Based Systems*, e-textbook. RVS-BI, 2011. Available from https://rvs-bi.de/publications/books/ComputerSafetyBook/index.html

[13] Bev Littlewood and Lorenzo Strigini, *Validation of ultrahigh dependability for software-based systems*, Comm. ACM 36(11):69-80, November 1993.

[14] E. Lloyd and W. Tye, *Systematic Safety: Safety assessment of aircraft systems*, Civil Aviation Authority, London, 1982.

[15] Many authors, *The Why-Because Analysis Home Page*. Available at https://rvs-bi.de/research/WBA/ . RVS-BI, no date.