RAMS
Reliability, Availability,
Maintainability, and Safety

# PK8200 - Risk influence modelling

Course compendium

Jørn Vatn

December 18, 2024

Department of Mechanical and Industrial Engineering
Norwegian University of Science and Technology

# Preface

This course compendium is developed for the course PK8200 - Risk influence modelling.

The reader is assumed to have background in basic probability theory and some skills in using ICT-tools. Many of the problems can be solved by using a dedicated Excel file with built-in functions which implement many of the formulas given in the compendium. It is also indicated how Python programming can be used.

Some of the problems can be solved without using ICT-tools, that is with an ordinary calculator. However, this will be tedious, and it is recommanded to practice with either Excel or Python.

Trondheim, December 18, 2024
Jørn Vatn

# Contents

# Chapter 1

# Introduction

## 1.1 Background

blbl

# Chapter 2

# Introduction to Risk Influence Modelling

## 2.1 Introduction

This memo gives an introduction to risk influence modelling and serves as an introduction to what we will cover in the course PK8200 – Risk influence modelling.

## 2.2 What is risk, and principle elements of a risk picture

A variety of risk definitions exist in the literature. In this presentation we only discuss risk definitions in relation to quantitative risk analysis. The baseline when risk is to be defined is to address threats to values of concern by the relevant stakeholders. This again forces us to be explicit regarding the *magnitude of the threat*, and the actual *consequences*, i.e., what values are at stake. It is common to link the definition of risk to *events* in one way or another. Often the term hazardous event is used to be explicit regarding the events that may result from the hazards or threats of concern. This then leads us to a conceptual risk definition saying that *risk is the uncertainty regarding the occurrence and severity of hazardous events.* To make the risk definition operational, we introduce three elements, events $(e_i)$, probabilities $(p_i)$ expressing the uncertainty regarding occurrence of the events, and finally the severity with respect to the values at stake $(\mathbf{S_i})$. This yields the quantitative expression of risk:

$$R = \{< e_i, p_i, \mathbf{S_i} >\} \tag{2.1}$$

When risk is expressed in terms of Equation (2.1) this is always done conditionally on a set of aspects which here is denoted the $\mathscr{D}$, $\mathscr{U}$ and $\mathscr{V}$. $\mathscr{D}$ represents the result of dialog processes and risk communication among stakeholders that elaborate on the values and preferences domain, such as who are exposed to threats, and whose needs in the society should be focused on. Further $\mathscr{U}$ represents the relevant information, the theories, the understanding, and the assumptions which are the basis for the risk assessor, and finally $\mathscr{V}$ represents the result of any verification processes, e.g., third party verification. See Vatn (2012) for further discussions.

In the traditional or classical definition of risk, the probabilities in Equation (2.1) are interpreted as true properties of the system being analysed. Since we have limited data and information regarding the system it is impossible to reveal the exact values of these probabilities. It is then common to present uncertainty intervals for the risk measure. In the epistemic interpretation it is the other way around. Then the basis is that there is uncertainty regarding whether the undesired events will occur (lack of knowledge), and the corresponding severity. Probabilities are used to express this uncertainty, and there is no additional uncertainty in the probability statements. However, as part of the documentation of the risk analysis uncertainty is qualitatively stated in terms of discussion of assumptions and simplifications. In relation to Equation (2.1) such arguments are stated as part of $\mathscr{U}$. Methods and models used in risk analysis are often not affected by the interpretation of risk in Equation (2.1). However, the way uncertainty is interpreted and presented will vary between the classical and the epistemic interpretations of risk.

To be explicit on how a risk picture would look like Figure 2.1 shows an example of such a risk picture related to one particular hazardous event (gas leakage). The uncertainty regarding the occurrence of the hazardous event is expressed by the probability figure $p = 0.1$, the severity is described by possible threats to the values at stake, i.e., possible fatalities. Here the number of fatalities is split into categories to simplify the presentation of uncertainty regarding the number of fatalities given that the hazardous event occurs. The frequencies, $F_i$, are the unconditional probability of each of the end consequences (fatality category).

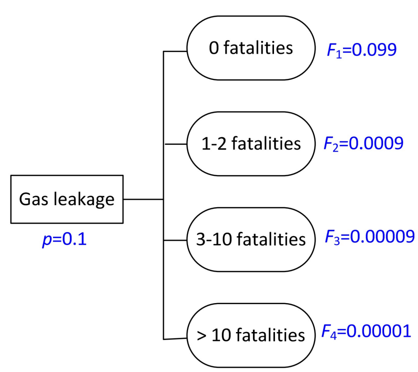


Figure 2.1: Example of risk picture related to one particular hazardous event

The following remarks could be made in relation to Figure 2.1:

- Only one hazardous event is elaborated, meaning that we are filtering the risk picture
- Only number of fatalities is emphasized, leaving out e.g., injuries, environmental damage and loss of production (filtering)

- Conditions leading up to the hazardous event are not visualized, leaving out the entire causal analysis (hiding background information)
- Conditions affecting the probability distribution over the end consequences are not visualized (hiding background information)
- Location (x-y-z) and time of the event is not visualized (no zooming)
- No distinction is made regarding which gas leakage type is considered (aggregating over all types of gas leakage events)
- No distinction is made regarding which personnel groups are exposed (aggregating)
- Some consequences are merged into one category, i.e., more than 10 fatalities corresponds to the event that 11 persons are killed, the event that 12 persons are killed etc. All these events are merged into one event.

The above comments points towards various types of operators applied to the risk picture. Before we these operators the risk picture is now defined as:

***Risk picture:*** A set of undesired events, the causes and factors that may contribute to the event, the possible consequences of the event with corresponding influencing factors, and uncertainties related to all these issues.

The operators applied on the risk picture are:

***Filtering***. With filtering we mean to filter out several aspects of the risk picture. Primarily filtering means to focus on only one hazardous event and/or a limited set of end consequences, e.g., only number of fatalities.

***Aggregation***. With aggregation we mean the process of summing more than one event, more than one cause etc. to give a sum of various events, causes and so on.

***Merging***. With merging we mean the process of grouping several similar outcomes into one category representing several outcomes.

***Zooming***. With zooming we mean to view part of the risk picture for a specific location (in space and/or time).

***Hiding/unhiding***. With hiding we mean to hide important information when presenting the complete risk picture. Typically we hide causes behind the hazardous event, factors that influence whether causes could lead to the hazardous event or not, and factors that influence the severity of the hazardous event, i.e., the probability distribution over the possible end consequences.

Due to limited resources when preparing the risk picture simplifications are made, merging is performed etc, hence it is not always possible to unmerge events, split up into a detailed set of failure causes etc.

## 2.3 Risk assessment

Risk assessment is the process of assessing the risk picture. Essentially this means to identify relevant hazardous events of concern, represent the corresponding values threatened in terms of a set of end consequences, and the uncertainty involved usually expressed by probabilities. Figure 1 shows one such hazardous event with related information.

In the following we discuss key elements in this process. The elements typically follow in chronological order, but going backwards is also often required in the analysis process.

1. *Identification*. Identification is two folded, first is the identification of hazardous events (or other undesired events) being the starting point of the analysis. The issue of identifying the *hazardous events* is not considered to be any principal problem. The second part of the identification process is much more demanding, i.e., to *identify relevant causes* to events, and factors and conditions affecting both the hazardous events, but also the severity given the event. In risk analysis it is equally important to identify conditions and factors that affect the situation in a positive way, as those having a negative impact.

2. *Structuring*. Structuring is an important step required before the modelling may start. Structuring means to present tacit knowledge, system understanding etc in such a way that the risk analyst is able to start modelling. In order to cope with the problem of identifying all causes behind events, so-called complexity attributes (Vatn, 2012) should be identified and structured.

3. *Modelling*. In risk analysis there are two types of models, probabilistic and deterministic. A deterministic model is primarily used to describe relations between physical quantities and other real world observables. Examples of such models are fire and explosions models for calculating pressures given an ignition of a specific gas cloud. A probabilistic model is a model that *enables the risk analyst to apply the law of total probability* in an efficient way when expressing uncertainty, i.e., performing probability calculus. It is important to

emphasize that a probabilistic model is not a model of the world, but it is a model used to express uncertainty regarding observables in the real world. Examples of modelling tools are Markov analysis, fault tree analysis (FTA) and event tree analysis (ETA) and Bayesian belief networks (BBN).

4. *Identification of the need of data.* All models will require input data such as failure rates, human error probabilities etc. Depending on the format and level of detail in the modelling, the main objective of this step is to be specific on the need for data and model parameters.

5. *Data collection and assessment of model parameters.* When the need for data is specified the next step is to collect data and in some cases also estimate/assign model parameter based on raw data and use of expert judgements.

6. *Run the model to establish the risk picture.* When the relevant models have been identified and built, they are feed with data. Then it is possible to run the models, and achieve the risk picture.

## 2.4 Building blocks for risk modelling

Mohaghegh et al. (2009) present a socio-technical risk analysis (SoTeRiA) framework which is a hybrid technique formalization, merging various classes of modelling techniques. In the following we briefly list the "building blocks" of SoTeRiA. For further elaboration we refer to Mohaghegh et al. (2009) and the references therein. The two last classes of techniques are new compared to Mohaghegh et al. (2009).

***Formal probabilistic risk analysis techniques*** refer to methods that apply a logical construct to describe the system. They include classical probabilistic risk assessment techniques, such as Event Sequence Diagram (ESD), Event Tree (ET), Barrier Block Diagram (BBD) and Fault Tree (FT).

***Regression-based techniques*** are common in economics and the social sciences. These techniques are used to distinguish true statistical causality from "spurious correlation". The process involves defining a set of variables and their relations, then "testing" all of the relations simultaneously. This is done by applying various techniques, such as Path Analysis or Structural Equation Modelling.

***Bayesian belief nets*** (BBN) defines a methodology for representing causal connections that are "soft," "partial," or "uncertain" in nature. The applications of BBNs have grown enormously over the past 30 years, with theoretical and computational development in many areas. A less strict formal realisation of such “soft” causal connections is modelling by use of risk influencing factors (RIF). A RIF is then a condition or factor that influences one or more parameters in e.g., a formal probabilistic risk analysis technique. BBN is often used to model interactions between RIFs. RIFs are discussed later in this document, and will be among the most central concepts in this course.

***Process modelling techniques*** aims at modelling the primary production processes of the organization. At a first step semi-formal process technique are adapted and applied to represent the various processes (e.g., work process) in an organization. Then, for quantification purposes, it needs to be converted to a formal technique that is consistent with other techniques in the quantification framework.

***Deterministic dynamic techniques*** are applied when there is enough information to establish “deterministic” relations among factors of the model or some parts of it. The deterministic modelling technique can be either analytical or simulation based. Examples of simulation-based techniques are Agent-Based Modelling (ABM) and System Dynamics (SD). SD shows significant capabilities for modelling certain human behaviour and decision-making processes, making it a good technique for modelling aspects of organizational behaviour. The strength of SD also lies in its ability to account for non-linearity in dynamics, feedback, and time delays.

***Energy related impact techniques*** are used to model the situation after loss off control of energy sources representing a potential harm. These models include gas dispersion modelling, ignition modelling, fire and explosion modelling and structural integrity modelling. The models basically utilize deterministic physical and chemical laws, biological knowledge related to critical heat and pressure levels etc. Since initial and boundary conditions are not known in advanced, and many model parameters are uncertain, the deterministic models are supported by probabilistic approaches.

***Recovery analysis techniques*** aim at modelling the recovering process after an emergency situation has occurred. It is reasonable to claim that the system now has become rather intractable (Hollnagel et al., 2006) which means that one single event tree model cannot capture the situation. Several scenarios need to be developed and there is a challenge to select a limited number of scenarios described at an appropriate level of detail to represent the recovery process. Also here the above listed techniques are relevant.

Note that in the literature a distinction is often made between systems which may be described in terms of linear cause and effect relations and emergent systems with complex cause and effect relations. The latter is often denoted complex systems. The above classification does not explicit reflect differences between linear and complex cause and effects. A rough classification would be to say that the formal probabilistic risk analysis techniques apply for linear systems, whereas the remaining classes of techniques apply for emergent and complex systems. But rather than classifying the techniques, we believe that the challenge of approaching emergent systems is not to choose the appropriate technique, but rather to combine the class of techniques in a way that enable us to express our system knowledge in the most appropriate way.

Although the list of techniques provided above are quite impressing we do not claim to have any final approach to determine the true risk. The techniques are only tools we apply to express our 6 uncertainty where the aim is to provide valuable decision support. This is also highlighted by e.g., Apostolakis (2005) where he discusses the value of risk analysis. He also points out limitations with current risk analysis practices, and points towards areas for improvement and further research. One such area is human performance during accident conditions. This points towards what is denoted *recovery analysis techniques* above. Challenges obvious occur when expressing human performance, but even more challenging would be to qualitatively express relevant accidental scenarios.

## 2.5 Risk modelling and risk influence modelling in a historical perspective

This section reviews important aspects of risk modelling required in order to present a risk picture. Vatn and Haugen (2012, RIO book) distinguish between three types of risk analyses for use in the offshore oil and gas industry:

1. Strategic risk analyses are primarily aimed at developing a safe design and safe operating procedures. The objective is typically to assess a proposed design or an operation, to evaluate whether the risk level is acceptable and to identify potential risk reducing measures. They are characterized by being performed with a global perspective, in the sense that they are considering the effect on the risk level for a whole installation. These studies are primarily quantitative. They focus very much on technical aspects, with operational input primarily being limited to activity levels, such as the number of offshore supply vessels visiting the installation, the number of lifts being undertaken by cranes, the number of wells drilled etc. For the hydro carbon events the analyses usually starts with a blowout/gas leakage, and then proceed with event trees (ETA = Event Tree Analysis), fire and explosion models etc. to derive the risk picture. This means that they are not focusing on causes for the hazardous event as indicated in Figure 2.1. To some extent the barriers are mod-

elled by fault tree analysis (FTA). These type of analyses are often referred to as total risk analyses (TRA) or quantitative risk analyses (QRA).

2. Qualitative design analyses are more detailed and more specific than the strategic analyses, and they will typically have a system focus. The most prominent examples of such studies are HAZOP and FMEA. This could be performed e.g., on the mud system or on the Blowout Preventer (BOP). These are performed to verify the design in detail, to ensure that safe (and reliable) operation is possible. They usually have a strong technical focus and are typically the responsibility of the design team or onshore staff working with technical safety, similar to the strategic studies.

3. Operative risk analyses are different from the strategic studies in almost every respect:

   a) They are typically performed as qualitative studies, sometimes using a risk matrix to classify the identified hazards/events and to determine acceptability of the identified hazards.

   b) They are performed on a much more limited problem area, typically an operation that is being planned or is about to be performed or as support for a specific, limited decision. The analyses may often address major accident risk (although not necessarily), but the link to the global risk picture for the installation, as expressed through the strategic analyses, is usually weak in these studies.

   c) The responsibility for these studies may be different personnel groups, including onshore planning/operations groups or offshore personnel, responsible for performing the work.

As discussed by Vatn and Haugen (2012) the three types of analyses are usually separate analyses where they are seen as independent analyses not being able to support each other. Several attempts have been made in order to improve the quality of such analyses. An early attempt to make the strategic risk analyses more dynamic was the ORIM methodology (Øien, 2001). The idea behind this approach was to take existing QRA models and investigate these with respect to the most important parameters, e.g., a Birnbaum like measure, say $I^{B}(i)$. Such a parameter could be the gas leakage frequency, or the safety integrity of a barrier. The next step was to investigate which of these parameters were most likely to change their value during the period between updates of the QRAs (typically every five year). The combination of *important parameters* and parameters that were *expected to vary* gave a list of parameters to include in a follow up regime.

Combining these gave raise to a methodology where the “critical” parameters were assumed to depend on a set of so-called risk influencing factors (RIFs). To monitor the RIFs, a set of risk indicators (RIs) were identified. For example for the RIF compentence, the number of certificates could be one RI, and the number of years in the department could be another RI. To calculate the RIF a weighted sum of the RIs were used. Finally to get an updated risk profile, the critical parameter, say $\lambda$ was updated according to

$$\lambda = \lambda_0 e^{\beta_0 + \beta_1 \mathrm{RIF}_1 + \beta_2 \mathrm{RIF}_2 + \ldots} \tag{2.2}$$

where $\lambda_0$ is the industry average value of the critical parameter, $\beta_0, \beta_1, \ldots$ are regression parameters and the RIFs are the numerical values of the risk influencing factors assessed by a weighted sum of risk indicators. The updated parameter, $\lambda$ is then inserted in existing the QRA to get an updated risk picture reflecting the current state of the risk influencing factors.

To investigate the RIF-model in eq. (2.2), check the literature for Cox-proportional hazard models.

To complete the approach a set of risk indicators (RI) were identified that were measuring the status of a RIF. The follow-up regime now was to obtain the values of the RIs every 3 months, calculate new RIFs based on the RIs, propagate this through the RIF model to get a change in the parameter. Now using the importance measure of the parameter, a new total risk picture could be established as well as the change in the risk level.

For various reasons the ORIM method was not implemented in the industry. One of the weaknesses in the ORIM was that it did not extend the existing QRA with respect to shed light on the failure causes. In the BORA project (Haugen et.al 2007) the focus was therefore changed to barriers and operative issues (BORA = Barrier and operative risk analysis). For critical conditions such as a gas leakage which in the QRA only is modelled by a single number found in generic databases, a detailed task analysis was conducted to reveal critical tasks that could lead to a gas leakage. For example review of real gas leakages revealed that many leakages were caused by maintenance of components and systems. This gave a much better understanding of what could go wrong, what where the critical tasks, and what RIFs influenced the probabilities in the model. This kind of modelling was very similar to HRA (Human Reliability Analysis) used in the nuclear industry. The BORA method is therefore seen as a significant improvement of the ORIM method since it have a much stronger link to the causes behind the initiating event (or failure of a critical barrier). A weakness of the BORA method was that it did not aim to update the QRA to see the total impact of the revised parameter (e.g., leakage frequency). The OMT project (Vinnem et.al 2011) was a follow-up of the BORA project where the methodology was refined. A weakness in the BORA model was that all the RIFs were threated in a flat hierarchy with only one level. It was recognised that there is a structure in the RIFs, where typically sharp-end RIFs are assumed to be influenced by blunt end RIFs (management issues). Further it was recognised that measuring the RIFs is demanding. Some companies are conducting internal audits to assess the current status of the RIFs on a regular basis (typically every second year). It is believed that this represents uncertainty, hence the RIFs were treated as random quantities and modelled by Bayesian Belief Networks (BBN) taking a two level hierarchical structure into account. The OMT project also improved the interaction modelling between the RIFs, and proposed ways to treat common cause errors similar to methods used in HRA.

## 2.6 Risk influencing modelling – Principal content

A primary risk model such as a combination of fault- and event trees aims to capture the core element of the course of events in the various accident scenarios. Such models represent formal probabilistic risk analysis techniques. Roughly speaking this part of the model describes the linear cause and effects identified, structured by means of formal logical statements. The more "soft" relations cannot be expressed by formal logical structures such as AND and OR gates. For example the level of competence is assumed to influence the error probability of a critical task, but we cannot model this by e.g., fault tree. Risk Influencing Factors (RIFs) and Performance Shaping Factors (PSFs) are often used to structure this part of the modelling. The various RIFs or PSFs are often structured by means of BBN techniques in order to model the influence on the basic events and barriers in the primary accident scenario model. In the following, we introduce important definitions used in risk influence modelling:

### QRA parameter

*Conditions that affect risk and included as single parameter in the quantitative risk model.* Examples of QRA parameters are failure rates, on demand probabilities for safety systems and ignition probabilities. Primarily we assume that QRA models comprise formal probabilistic techniques like fault- and event trees, and energy related impact techniques, e.g., an ignition model.

### 2.6.1 Risk influencing factor (RIF)

*A RIF is a factor or condition that influences the risk.* There are three principal types of RIFs, (i) RIFs that are equal to a QRA parameter, (ii) RIFs where the value of a RIF is assumed to influence a QRA parameter, and where the influence is described in probabilistic terms, and (iii) RIFs that influence a parameter that is not explicitly modelled in the QRA. The situation in (ii) is the most important covered in this course. In most cases, the value of a RIF is not known, and the RIF is therefore treated as a random quantity.

### 2.6.2 Risk indicator (RI)

*A quantity or condition that may be assessed or measured, where the value of the risk indicator is an indirect measure of a corresponding risk influencing factor.* There are two principal types of risk indicators, (i) resultant risk indicators which are reflecting the corresponding risk influencing factors, and (ii) controllable risk indicators, where it is possible by various efforts to set the value of the risk indicator, and hence change the value of the corresponding RIF. If the RIFs are treated as random quantities in the modelling, it is possible to derive the conditional probability distribution of a RIF given the value of one or more RIs. Among the models we consider, only the Risk_OMT treats RIFs as random quantities. In the Risk_OMT risk indicators are denoted scores.

### 2.6.3 Score (S)

*A realisation of the true underlying value of a RIF.* The term score were introduced in the Risk_OMT approach where the term score denote the summarized information regarding one RIF from interviews, surveys etc. A score is thus treated as a realisation (observation) of the true underlying RIF.

### 2.6.4 RIM parameter

*A RIM parameter is a parameter in the risk influence model (RIM).* There are several levels of RIM parameters, for example RIM parameters used in models to assign a QRA parameter given the value of a RIF, and RIM parameters used to link risk indicators (RI) or scores (S) to the RIFs.

## 2.7 Challenges in RIF modelling

In this section we discuss some of the main challenges with developing RIF models. A starting point for the discussion is that some quantitative risk analysis already exists. For example in the offshore oil and gas industry so-called total risk analyses (TRA) often also referred to as QRA exist for all installations. There are several reasons why we want to go into RIF modelling:

1. To get a more realistic risk picture taking “soft” factors into account
2. Link soft factors explicitly to the risk picture such that we can quantify the effect of risk reducing measures related to these soft factors
3. Establish a framework for updating the risk picture based on change of the value of risk indicators rather than updating the entire QRA to achieve a living risk picture.

4. By conducting the analysis gain more insight into risk influencing conditions, and hence be able to eliminate risk factors directly.

## 2.8 Identification of RIFs

Two identify RIFs we may choose between a top down, or a bottom up approach. In a top down approach we start with an existing QRA and search for the most important RIF parameters. To quantify the importance of a RIF parameter in this context we take into account both (i) a technical importance measure like Birnbaums measure of importance (i.e., the change in total risk by a small change in the parameter value), and (ii) the likely change in the parameter. For example in the offshore oil and gas industry gross accidents are mainly linked to well-control events, process events caused initiated by gas leakages, and structural damages caused by ship collisions. Thus the gas leakage frequency is a parameter with a high value according to the Birnbaums measure. If further, the gas leakage frequency is likely to change, or strongly affected by RIFs for which we have not really looked into, we have a good starting point to look for the most critical RIFs. In a bottom up approach we start with all possible RIFs we consider to have an impact on the total risk, and perform a screening without explicitly considering existing risk models.

Independent of a top down, or bottom up approach, we need to define the RIFs such that they relate to the existing risk models and QRA parameters. In some situations the RIF would be identical to the QRA parameter which is the easiest situation. In other situations, there will be a more or less direct link between the RIF and the QRA parameter. For example in human reliability analysis (HRA) the RIFs are denoted PSFs (performance shaping factors) and are linked directly to the so-called HEP (human error probability) which is a QRA parameter. The most challenging situation is when the level of details in the QRA is insufficient to really match the RIFs. For example in offshore oil and gas QRAs, the gas leakage frequency is modelled as one single number not taking the various failure causes into account. Failure causes for gas leakages spans over a range of technical, procedural and human error related issues, and if no model exist to map the failure cause level, it is also hard to link RIFs to an existing QRA parameter.

## 2.9 Extension of existing QRA models

As discussed above, existing QRA models often lack the level of details making them appropriate for linking RIFs to the QRA parameters. It is therefore often required to extend the existing QRA. Referring to the gas leakage example discussed above, a very simple approach is to split the gas leakage frequency into a set of failure causes followed by an assessment of the relative importance of each failure cause. The next step is then to link the various RIFs to one of the failure causes giving a good starting point for the modelling. Note that a cause here may be the failure of a safety barrier not directly linked or modelled in the QRA. In other situations it is required to develop new risk models to get confidence in the mapping of RIFs to the risk model. For example in the Risk_OMT methodology several task analyses were carried out to really catch critical activities during maintenance that influences the gas leakage frequency. An advantages developing detailed models is that we also get a better qualitative understanding of those issues that may cause e.g., a gas leakage. Such understanding is of great importance when searching for explicit risk reducing measures.

## 2.10 Defining the scale of the RIFs

Every RIF in a risk influence model has a value (known or unknown) which depends on the scale being used. Now, let $r$ be the value of a RIF. A neutral scaling regime would be to define $r = 0$ to be the industry average of a RIF, $r = -1$ be the worst case we can imagine, and $r = +1$ be the best we can imagine within a reasonable time horizon. Another approach is to use an arbitrary scale, for example in the ORIM model $r = 1$ corresponds to the worst case, and $r = 5$ to the best case. In the BORA and Risk_OMT models character values were used, where $r = A$ corresponds to best practice, and $r = F$ corresponds to the worst case, or unacceptable state. It is recommended to use the same scale for all RIFs. Later on we will also discuss risk indicators (or scores) as a means to assess the value of a RIF. It is recommended to use a matching scale for risk indicators/scores. However, this is more demanding, since for example a risk indicator may be measured in terms of e.g., $n$ = number of personnel having a certain formal certificate. In such cases, a mapping is required, for example $n = 0$ corresponds to the character $F$, $n \in [1,2]$ corresponds to $E$ and so on. Such a mapping is not straight forward and requires careful considerations.

## 2.11 Structuring RIFs

Usually more than one RIF are influencing a QRA parameter. In simple models a weighted sum is calculated to represent all the RIFs, and this sum is then used to adjust the QRA parameter. In more advanced modelling, two aspects are considered. The first aspect relates to the fact that a weighted sum will not take into account interaction effects between the RIFs. In some situations a bad value of two or more RIFs is considered more critical than the individual contribution from these two bad values, i.e., there are some interaction effects we would take into account. The second aspect relates to dependencies between RIFs. For example if we split into "sharp end" RIFs (e.g., time pressure) and "blunt end" RIFs (e.g., management of work organization) the latter RIF is assumed to influence the first RIF. If we collect data and combine into scores to reflect the value of the RIFs on different levels, we need to develop an influence model to connect the RIFs, and BBN methods will be a good starting point.

## 2.12 Linking risk indicators and scores to the RIFs

In simple models we link risk indicators to the RIFs by simple weighting formulas. If RIFs are treated as random quantities, the risk indicators (or scores) are only considered to be *indicators* for the true underlying value of the RIF. We then need to express how strong evidence a value of a risk indicator really is. A simple way to express this is to say that given a value $r$ of a RIF, the risk indicator or score, will be a random variable, say $S$, where $\mathrm{E}[S] = r$, and in addition a precision parameter is required to express $\mathrm{Var}(S)$. This will enable inference, i.e., assessing a probability distribution over the RIF, or simultaneous distribution over a set of RIFs by e.g., BBN methods.

## 2.13 Linking RIFs to the QRA parameters

Independent of whether RIFs are treated as random quantities, or fixed known values, it is necessary to link the various RIF values to corresponding QRA parameters. If r is the value of a RIF, and p is a corresponding QRA parameter, we need to establish a functional relation:

$$p = f(r) \tag{2.3}$$

In order to establish such a relationship we often ask what will be the value of $p$ when the RIF take the best and the worst value respectively. For values of the RIF in-between we often choose between linear or geometric relations.

Let $p_{\mathrm{L}}$ and $p_{\mathrm{H}}$ be the lowest and highest value the QRA parameter of interest can take respectively. Further let $r_{\mathrm{L}}$ and $r_{\mathrm{H}}$ be the lowest and highest value a RIF value can take respectively[1] (considering only one RIF). A linear interpolation is now given by:

$$p(r) = p_{\mathrm{L}} + \frac{(r - r_{\mathrm{L}})(p_{\mathrm{H}} - p_{\mathrm{L}})}{r_{\mathrm{H}} - r_{\mathrm{L}}} \tag{2.4}$$

and similarly a geometric interpolation is given by:

$$p(r) = p_{\mathrm{L}} \left( \frac{p_{\mathrm{H}}}{p_{\mathrm{L}}} \right)^{\left( \frac{r - r_{\mathrm{L}}}{r_{\mathrm{H}} - r_{\mathrm{L}}} \right)} \tag{2.5}$$

Generally a geometric interpolation is recommended if the range of the parameter variation spans more than one decade.

If we have more than one RIF that is influencing a QRA parameter, we need to develop an interaction model to combine the RIFs.

[1] We here assume that a low value of the RIFs gives a low value of the corresponding QRA parameter to simplify the presentation

## 2.14 Structuring QRA parameters, RIFs and RIs/scores

Ideally we would like to make a clear distinction between QRA parameters, RIFs and RIs. The QRA parameter represent the quantity used in the QRA models which are the starting point of the risk influence modelling. Then RIFs are introduced to represent conditions that are influencing the QRA parameter, where the aim is to give a RIF a theoretical interpretation like competence of maintenance personnel. In order to assess the value of a RIF, the RIs are introduced. In some situations we build a very rigid risk influencing model by formal e.g., a formal BBN model. For example in the Risk_OMT model some 5-8 RIFs were introduced in a hierarchy to show the influence on a particular QRA parameter. In the Risk_OMT the RIFs were more or less identical to variables used in safety audits, and hence there were a one to one relation between the RIF and the RI (where the term score was used rather than a risk indicator). In other situations less effort is made to structure the various RIFs, and at the extreme we may leave out the explicit definition of RIFs and link the risk indicators directly to the QRA parameter, for example by a weighted sum.

## 2.15 Hybrid vs full BBN models for the entire QRA modelling

If the RIFs are treated as random quantities, equation (2) may be used to get a probability distribution over the parameter p. This represents a so-called parameter uncertainty which we in principle may propagate in the QRA model. Since fault- and event trees in principle may be converted to BBNs, and since a BBN is also used to model the distribution over various r values, these BBN models may be combined to give a full BBN model for the total quantitative risk model. Experience from the Risk_OMT project has shown that this is impossible due to memory and time constraints within existing BBN implementations. To create an approximated full BBN model requires careful consideration, and deep knowledge into application of BBN modelling. A hybrid model applies equation (2), but rather than finding a probability distribution over p, the uncertainty regarding the RIFs is being integrated to give an expected value of the QRA parameter. This value could then be used in existing QRA models represented by fault and event tree rather straight forward. Such an approach, is however, not conservative since if the same RIF influences several QRA parameters, we then will ignore the “state of knowledge” dependency in the value of the RIFs, and hence “underestimate” the risk. More research is required, to find ways to improve such hybrid modelling.

References B.A. Gran, R. Bye, O.M. Nyheim, E.H. Okstad, J. Seljelid, S. Sklet, J. Vatn, J.E. Vinnem. 2012. Evaluation of the Risk OMT model for maintenance work on major offshore process equipment. Journal of Loss Prevention in the Process Industries, Volume 25, Issue 3, May 2012, Pages 582-593 J.E. Vinnem, R. Bye, B.A. Gran, T. Kongsvik, O.M. Nyheim, E.H. Okstad, J. Seljelid,

J. Vatn. 2012. Risk modelling of maintenance work on major process equipment on offshore petroleum installations. Journal of Loss Prevention in the Process Industries, Volume 25, Issue 2, March 2012, Pages 274-292 Øien, K. (2001) A framework for the establishment of organizational risk indicators. Reliability Engineering & System Safety, 74, 147-167. Haugen, S., Seljelid, J., Sklet, S., Vinnem, J.E. and Aven, T. (2007) Operational risk analysis: total analysis of physical and non-physical barriers. H3.1, Generalisation Report, Rev. 1 Vatn J (2012). Can we understand complex systems in terms of risk analysis? Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 2012 Sarshar, S., Gran, B.A., Haugen, S and Skjerve, A.B. (2012) Visualisation of risk for hydrocarbon leakages in the planning of maintenance and modification activities on offshore petroleum installations. ESREL/PSAM 2012. Vatn, J and Haugen.S. (2012). On the usefulness of risk analysis in the light of Deepwater Horizon and Gullfaks C. RIO book Hollnagel E, Woods D D and Leveson N (Eds.) (2006) Resilience Engineering – Concepts and Precepts. Ashgate Publishing Company.

# Chapter 3

# Risk_OMT – Hybrid approach

## 3.1 Introduction

This document describes the technical aspects of the hybrid implementation of the Risk_OMT model. The Risk_OMT offers two modelling approaches. The full-BBN (Bayesian Belief Network) approach utilize BBN modelling both for the soft influences between risk influencing factors (RIFs) and the formal probabilistic relations described by fault- and event trees. The hybrid implementation of the Risk_OMT model uses a BBN specification of the relation between the RIFs but uses ordinary processing of the fault- and event trees.

In the fault and event trees failure of an activity is divided into failures of omission and failure of execution. Failure of omission denotes whether or not the prescribed activity is carried out. Failure of execution denotes inadequate actions that may cause failures, e.g., acts performed in a wrong sequence, at wrong time, without required precision etc. Failure of execution is seen as results of *human errors* and *violations*. Human error is further divided into *mistakes* and *slips & lapses*, where mistakes involve actions that are based on failure of interpretation of procedures, and/or failures of judgemental/inferential processes involved in the prescribed activity. This category does not distinguish between whether or not the actions directed by this judgement activities run according to the actor's plan. Typical mistakes are inadequate judgement/conclusion due to intrinsic conditions such as competence, fatigue, mode etc, and extrinsic conditions such as communication, information, work load, time pressure etc. Slips & lapses involve actions that represent unintended deviation from those practiced represented in the formal procedures. This is deviation due to error in execution and/or the storage stage of an action sequence. For our purpose, this category represents only actions where there is no intended violation, failure of interpretation of procedures and judgement failures prior to the action carried out. In the Risk_OMT model separate BBNs are developed for the RIF structure for (i) mistakes, (ii) slips & lapses, and (iii) violations. For failures of omission currently no RIF model is derived.

## 3.2 Risk influencing factors

A Risk Influencing Factor (RIF) represents a condition or a situation that influences the risk in a risk model. In this presentation we always assume that the RIFs are influencing the risk through parameters used in the risk model. In the Risk_OMT we mainly focus on different organisational conditions that have a theoretical and/or empirical grounded influence on the possible deviations from required actions, and hence should be reflected in probability assignment of errors or failures. Further, Risk_OMT operates with 2 levels of RIFs which links the organisational conditions (RIF Level 1) to strategic management decisions (RIF Level 2). In the current Risk_OMT implementation it is only RIFs on level 1 that directly influence the basic event probabilities. Note that the term risk is interpreted differently within the society of risk analysis. In a classical risk analysis framework risk is seen as a property of the system being analysed. Further 2 probabilities in a risk model are considered to represent some true likelihood of e.g., component failures and human errors. With such an interpretation we may think of the RIFs as a way to establish a true causal link between some conditions and the basic event probabilities. In an epistemic interpretation of risk the main focus is on uncertainty. Risk is essential uncertainty regarding the occurrence and severity of undesired events. In such a framework basic event probabilities are not considered as some true values, but are expressions of our uncertainty regarding the occurrence of the basic events. The RIFs will then represent conditions that we take into account when assigning probabilities (expressing uncertainty) to the basic events, but we do not consider any causal link as for the classical interpretation.

The Risk_OMT modelling framework is an extension of the BORA release model (Aven et.al. 2006). There are two major changes in the Risk_OMT model compared to the BORA release model. Whereas the BORA release model combined the RIFs on the same level, the Risk_OMT model introduces a hierarchy between the RIFs. Further the BORA release model considered the RIFs to be known without any uncertainty. In the Risk_OMT model RIFs are still considered to be theoretical constructs that influences the risk, but we do not have exact knowledge regarding the value of the RIFs, and hence they are treated as stochastic variables (random quantities).

Formally we use the term score to denote the summarized information regarding the RIFs form interviews, surveys etc. A score is thus treated as a realization (observation) of the true underlying RIF. In the BBN this corresponds to an arrow from the RIF to the corresponding score. The scoring system is based on characters A to F, where A corresponds to best industry practice, and F corresponds to an unacceptable state with respect to the actual RIF.

Figure 3.1 shows an example of a RIF structure. Level 1 RIFs point to the basic events in the fault tree showing that level 1 RIFs influences the basic events. Level 2 RIFs influence the level 1 RIFs, and there is an arrow from the RIFs to the scores to indicate that the scores are treated as realizations (observations) of the true underlying RIFs.

Examples of level 1 RIFs are *technical documentation* and *time pressure*. Corresponding level

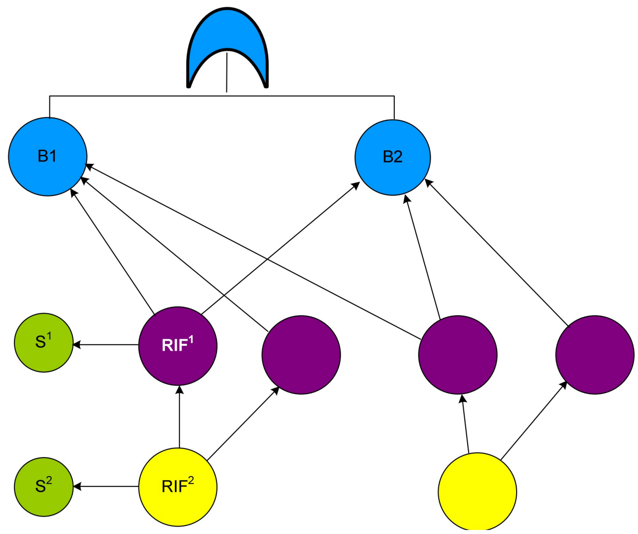


Figure 3.1: RIFs on two levels with scores and relation to basic events in a fault tree

2 RIFs are *management of information* and *management of tasks* respectively.

## 3.3 Impact of the level one RIFs on the basic events

In the hybrid Risk_OMT model the impact of the RIFs are explicitly modelled via the probability of occurrence of a basic event or a barrier in the fault or event tree. We now consider basic event number i. Three quantities span the sample space for the basic event probability for this event:

$q_{i,A}$ = average basic event probability corresponding to average industry practice, i.e., all RIFs equal to the character C.

$q_{i,L}$ = lowest basic event probability corresponding to the best practice in the industry, i.e. all RIFs equal to the character A.

$q_{i,H}$ = highest basic event probability corresponding to the an unacceptable industry practice, i.e. all RIFs equal to the character F. It is not expected to observe RIFs of character F.

Often $q_{i,L}$, and $q_{i,H}$ are specified indirectly by error factors, say $\mathrm{EF}_{i,L} = q_{i,A}/q_{i,L}$ and $\mathrm{EF}_{i,H} = q_{i,H}/q_{i,A}$ respectively. In the modelling it will be convenient to represent the RIFs by numeric values rather than character values. It is convenient to map the character values on the interval [0,1]. Splitting this interval into 6 sub intervals, and mapping the character value into the centre gives the value 1/12 for an A, the value 1/6+1/12 = 3/12 for a B , the value 2/6 + 1/12 = 5/12 for a C up to 11/12 for a value F. etc. In the following we will always use this mapping in the numeric quantifications. In some situations we may use a more differentiated labelling than the pure characters, i.e., we may succeed the characters with extra plusses (+) or minuses (-). For example we may use that A++ corresponds to a value 0, A+ corresponds to a value 1/24 etc.

If we use $r$ as a value of a weighted sum of the RIFs influencing the basic event probability we now introduce $q_i(r)$ to describe the functional relationship between the RIF value ($r$) and the basic event probability. We have that $q_i(0) = q_{i,L}$, $q_i(5/12) = q_{i,A}$ and $q_i(1) = q_{i,H}$. In between these values we may either use linear or geometric interpolation. For high error factors it is recommended to use a geometric interpolation.

We will now assume that there are totally $J$ RIFs that are influencing basic event $i$. Let $\mathbf{R} = [R_1, R_2, \ldots, R_J]$ be a vector of stochastic variables to represent these (standardized) RIFs, and let $p_{\mathbf{R}}(\mathbf{r}) = \Pr(R_1 = r_1, R_2 = r_2, \ldots, R_J = r_J)$ be the joint probability distribution over these RIFs.

If the value of the RIFs are given, we assume there is a relation between the basic event probability and the values of the RIFs:

$$q = q(\mathbf{r}) \tag{3.1}$$

Note the notation here, the upper case $R$ is used for the RIF as a random quantity, and the lowercase $r$ is used for a particular value.

Each RIF might have different weight with respect to the influence on the basic event probability. Let $w_j$ be normalized weight for RIF $j$. A first approximation for the total impact of the RIFs on the basic event probability is given by:

$$q_i = \sum_{\mathbf{r}} q_{\mathrm{L}} \left( \frac{q_{\mathrm{H}}}{q_{\mathrm{L}}} \right)^{w_1 r_1 + w_2 r_2 + \ldots} p_{\mathbf{R}}(\mathbf{r}) \tag{3.2}$$

where $\sum_{\mathbf{r}}$ represents the sum over all possible values of $\mathbf{R}$. Equation (3.2) is then used to establish the basic event probabilities to use in the fault and event tree part of the hybrid risk analysis. It is rather complicated to establsih $p_{\mathbf{R}}(\mathbf{r})$, so it is not straight forward to apply Equation (3.2).

## 3.4 The beta distribution to describe uncertainty regarding the RIFs

A mathematical convenient probability distribution to use for continuous variables on the interval [0,1] is the beta distribution. Although the scoring of the RIFs are on an ordinal level, a continuous ratio scale seems appropriate for the modelling. The probability density function of the beta distribution is given by:

$$f(r) = r^{\alpha-1}(1-r)^{\beta-1}/\mathrm{B}(\alpha, \beta) \tag{3.3}$$

where $\mathrm{B}(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha\beta)$ is the beta function, and $\Gamma()$ is the gamma function. $\alpha$ and $\beta$ are parameters in the distribution.

If $R$ is beta distributed with parameters $\alpha$ and $\beta$ the expected value and variance are given by:

$$\mathrm{E}[R] = \frac{\alpha}{\alpha+\beta} \tag{3.4}$$

$$\mathrm{Var}(R) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{3.5}$$

Note that the beta distribution represents a conjugate prior distribution for the binomial distribution. Thus if the beta distribution is used to describe the parameter $r$ in a binomial distribution with prior parameters $\alpha_0$ and $\beta_0$, then the posterior distribution is also beta distributed with parameters $\alpha_0 + x$ and $\beta_0 + n - x$ where $x$ is the number of successes and $n$ is the number of trials of an experiment provided to update the prior distribution, i.e., the posterior distribution is a beta distribution with parameters $\alpha = \alpha_0 + x$ and $\beta = \beta_0 + n - x$.

## 3.5 Updating the RIF distributions based on the scores

The above results in Equations (3.4) and (3.5) do not apply directly to our situation since we will not get observations from a binomial trial but rather one observation considered to be a realisation of the true RIF.

Let $\alpha_0$ and $\beta_0$ be the parameters in the prior distribution of the RIF prior to observing the score $S$. Given the true value of the RIF, say $r$, it is reasonable to assume that $\mathrm{E}[S|r] = r$, and further we assume that it is possible to specify $\mathrm{Var}(S|r) = V_S$. We now make the following argument: We will use the value of the score, say $s$, by translating the information to a binomial situation, e.g., finding $x$ and $n$. This is done due to the simple result that exists for the binomial situation. Since $X/n$ is an estimator for $r$ in the binomial situation, and the score $S$ is the estimator for $r$ in our situation, it seems reasonable to require:

$$\mathrm{Var}(X/n) = r(1-r)/n = \mathrm{Var}(S) = V_S \tag{3.6}$$

Thus, if we know $V_S$ and replace $r$ with it's estimate $s$, we should have:

$$n = s(1-s)/Var(S) = s(1-s)/V_S \tag{3.7}$$

further since $x/n$ and $s$ both are estimates of $r$, we set $x = s \cdot n$. Utilizing the result from the binomial situation where the posterior distribution is beta distributed with parameters $\alpha_0 + x$ and $\beta_0 + n - x$ we will in our situation approximate the posterior distribution with a beta distribution with parameters:

$$\alpha = \alpha_0 + s^2(1-s)/V_S \tag{3.8}$$
$$\beta = \beta_0 + s(1-s)/V_S - s^2(1-s)/V_S = \beta_0 + s(1-s)^2/V_S \tag{3.9}$$

□

**Exercise 3.1**

Find the expected value and the variance of the posterior distribution with the parameters obtained by Equations (3.8) and (3.9). Compare this result with the expected value and variance of the weighted sum of the prior mean and the score where the reciprocal variances are used as weights. □

## 3.6 Level 2 RIFs

For level 2 RIFs it is straight forward to use the result in Equations(3.8) and (3.9) to find posterior distributions for the RIFs. Various principles may be used for specifying the prior distribution. In order to have a method that is data driven as far as possible, it seems reasonable to apply $\alpha_0 = \beta_0 = 0.5$ corresponding to Jeffreys prior (Jeffreys, 1946).

**Exercise 3.2**
Apply Jeffreys prior together with Equations (3.8) and (3.9) in order to find the expected value and the variance of the posterior distribution for scores corresponding to the characters A, B, ..., F. Present the result in a table for $V_S$ equal to $0.2^2$, $0.1^2$ and $0.05^2$ respectively.

## 3.7 Level 1 RIFs

We will start by assigning the posterior distribution of level 1 RIFs given the value of the parent RIF, i.e., the corresponding level 2 RIF denoted $P$ (i.e., parent). Given the value of the level 2 RIF, say $P = p$, it is reasonable to specify a prior distribution of the RIF, say $R$, with expected value $\mathrm{E}\left[R|P=p\right] = p$. Thus the structural dependencies between the parent RIF and the child RIF is considered to give the same expectation. But how strong is the structural dependency, i.e., the influence of the parent RIF on the child RIF? Such a structural dependency may be expressed by the variance of the child RIF, i.e., $\mathrm{Var}(R|P=p)$. To make the model simple we assume that $\mathrm{Var}(R|P=p) = Var(R) = V_P$ where it is possible to specify $V_P$ independent of the actual value of the parent. $V_P$ may be considered as a measure of the structural dependence in the model. Proposed values for the structural dependency are $V_P = 0.2^2$ (low dependency), $V_P = 0.1^2$ (medium dependency) and $V_P = 0.05^2$ (high dependency). Prior to observing the score, it seems reasonable to express the prior distribution of the RIF with a beta distribution with expected value $p$ and variance $V_P$.

**Exercise 3.3**
Use Equations (3.4) and (3.5) to show that we may obtain the prior parameters in this situation by:

$$\beta_0 = \left(\frac{p(1-p)}{V_p} - 1\right)(1-p) \tag{3.10}$$

$$\alpha_0 = \frac{p\beta_0}{1-p} \tag{3.11}$$

□

Conditional on the value of the parent level 2 RIF; i.e., $P = p$, and the structural influence between these RIFs, the prior distribution may be obtained by applying the parameters in Equations (3.10) and (3.11). Given the score $S = s$ of the level 1 RIF we apply (3.8) and (3.9) to find the conditional posterior distribution, i.e., given the parent value. In order to find the unconditional posterior distribution, we may integrate over the posterior distribution of the parent node. It is, however, important to stress that we do not need the unconditional posterior distribution of the child RIFs, i.e., the level 1 RIFs. In Equation (3.2) we need the joint distribution over the level 1 RIFs that directly influences the basic event probability. From the theory of BBN, we know that the level 1 RIFs are independent *given* their parents, i.e, the level 2 RIFs. This means that we may multiply the conditional posterior distributions for level 1 RIFs to find the required $p(\mathbf{r})$ in Equation (3.2) and then integrate over the joint posterior distribution of the level 2 RIFs.

In Equation (3.2) we did assume that the RIFs were made discrete, i.e., each RIF takes a finite number of values. This is done in order to simplify calculations. Since the scores are measured on six different values, it seems reasonable to use 6 values for each RIF both on level 1 and level 2. Then we may for the posterior distribution of the level 2 RIFs calculate a point probability for each interval, i.e., [0,1/6], [1/6,2/6] etc. Similarly, given the (discrete) values of the level 2 RIFs, we may calculate point probabilities for the level 1 RIFs, and $p(\mathbf{r})$ is then found by applying the law of total probability.

## 3.8 Finding basic event probabilities based on the RIF structure

Up to now we have obtained the following:

1. For the second level RIF's:

- We use Jeffreys prior before we have access to the scores
- When we get the scores, we update the prior to the posterior by applying Equations (3.8) and (3.9)
- The posterior is a Beta-distribution, currently not discretized

2. For the first level RIF's:
   - Given the value of the parent, say $P = p$, the prior distribution parameters are given by Equations (3.10) and (3.11)
   - When we get the scores, we update the prior to the posterior by applying Equations (3.8) and (3.9)
   - The prior is only known for a given value of the parent, i.e., if we know $P = p$. However, we do not know the value of the parent, i.e., second level RIF.
   - The posterior is a Beta-distribution, currently not discretized
3. Equation (3.2) may be used to find the basic event probabilities, but it is not easy to obtain the joint probability distribution over all RIF's.

### 3.8.1 Discretization

Assume that a stochastic variable $R$ is Beta-distributed with parameters $\alpha$ and $\beta$, and that the cumulative distribution function, $F_R(r)$ is known. That is, we have a numerical routine to calculate $F_R(r)$. To discretize the distribution into $K$ values, we use the midpoint for each value, i.e., $p_k = k/K - 1/(2K), k = 1, \ldots, K$. In our case $K = 6$ and the midpoints corresponds to the character scores $A, B, \ldots, F$.

To find the point probabilities we use:

$$\Pr(R = p_k) = F_R(k/K) - F_R((k-1)/K) \tag{3.12}$$

**Exercise 3.4**
Write a simple code (VisualBbasic, Matlab, Python, Fortan or C) to find the point probabilities for each interval [0,1/6], [1/6,2/6], .., given the parameters in the posterior distribution. □

### 3.8.2 Level 2 RIFs

Assume we have $J$ second level RIFs, and let

$$\pi_j(k) = \Pr(R_j^2 = p_k) \tag{3.13}$$

for the $j$th second order RIF, i.e., $R_j^2$. To obtain $\pi_j(k)$ we use Equation (3.12) with the posterior distribution for the second order RIFs. Thus we have established a way to find the probability of each $p-values$ used in the conditional posterior distribution for the first level RIFs.

### 3.8.3 Level 1 RIFs

The basis for the conditional prior distribution for the first level RIFs was to include the structural dependency, $V_P$, and a given value of the parent, say $p_k$ to find the conditional prior distribution parameters $\alpha_0$ and $\beta_0$. When the value of the score $S = s$ is obtain the conditional posterior distribution can be obtained. As for second level RIFs we can also use Equation (3.12) to discretize the level 1 RIF posterior disbributions.

**Exercise 3.5**
Consider a situation with one level 2 RIF, two level 1 RIFs influenced by the level 2 RIF. Write a simple code to find the unconditional distribution over the weighted sum of the level 1 RIFs. Make the code flexible such it is possible to specify the weights, the scores, the structural influences $V_P$'s and the variances of the scores $V_S$'s. □

**Exercise 3.6**
Discuss extension of the model in the previous exercise where there are more than two level 1 RIFs for each level 2 RIF, and where there are more than one level 2 RIF. Hint: Since each level 1 RIF is influenced by one and only one level 2 RIF, the subset of level 1 RIFs with common parent level 2 RIF may be treated separately. Discuss why this will reduce the number of combinations to run through. Also discuss how to implement the solution if this savings should be obtained. □

### 3.8.4 Compiling the results for level 2 and level 1 RIFs

In our case, the value of the parent RIF, $p_k$, is not known, we only have discretized posterior distributions for the second level RIFs. We also reccognize from Bayesian Belief Network theory that first level RIFs are independent only if the second order RIFs are know. This is important when we shall obtain $p_{\mathbf{R}}(\mathbf{r})$ to be used in Equation (3.2).

Actually, in the first place we only need a subset of the RIFs, i.e., $\mathbf{R}^1$ = the first level RIFs. If we fix all level 2 RIFs, i.e., assign a value to $R_j^2$ all children (level 1 RIFs) of level 2 RIF $j$ are independent. Hence, for we can multiply the marginal posterior distributions for the level one RIFs to obtain $p_{\mathbf{R}}^1(\mathbf{r^1})$.

To simplify notation assume that there are $I_j$ level one RIFs that are children of level two RIF $j$. If we only consider these in the joint distribution, $p_{\mathbf{R}}(\mathbf{r})$, we can write the joint distribution as a product of the marginal distributions. If the first level RIFs are discretized similarly as for the second order RIFs, the conditional contribution to $q$ is given by:

$$q_{j,k} = \sum_{\mathbf{r}} q_{\mathrm{L}} \left( \frac{q_{\mathrm{H}}}{q_{\mathrm{L}}} \right)^{\sum_{i=1}^{I_j} w_i p_l^i} \prod_{i=1}^{I_j} \rho_i(l) \tag{3.14}$$

where $\rho_i(l) = \Pr(R_i^1 = p_l), l = 1,\ldots,6$ and where the sum $\sum_{\mathbf{r}}$ is over all possible combination of the first level RIF values, i.e., in total $(I_j)^6$ combinations. Note that $\rho_i(l)$ is found by integrating the conditional posterior PDF over the $l$th interval using Equation (3.12).

$p_l^i$ is used to denote that the $i$'th RIF takes the value $p_l$, e.g., $p_2^1 = 3/12$ if the first level one RIF is an $B$.

To calculate $q_{j,k}$ from equation (3.14) we need to run through all the $(I_j)^6$ combinations of the $\mathbf{r}$-vector. Assume we have written a function `NextComb(r)` that gives the next permutation of the vector `r`, and returns `TRUE` as long as there are more permutations, we may use something like:

```
:
q_jk=0
r=[1,1,1,....,1] ' Initial vector
Do While NextComb(r)
   q_jk = q_jk + new contribution from current r-vector
Loop
:
```

to calculatione $q_{j,k}$.

**Exercise 3.7**

Assume that you have $n$ RIFs where each RIF may take $m$ different values. Propose an algorithm to generate all possible combinations of the $n$ RIFs, i.e., to implement the `NextComb()` function. □

Equation (3.14) represents the "contribution" from RIF number $j$ given it takes the value $p_k$. Since $\pi_j(k)$ is the probability that RIF number $j$ given it takes the value $p_k$ the law of total law of probability may be used to find the *unconditional contribution* to $q$:

$$q_j = \sum_{k=1}^{6} q_{j,k} \pi_j(k) \tag{3.15}$$

Finally, taking all $J$ level two RIFs into account:

$$q = \left(q_{\mathrm{L}}\right)^{-J+1} \prod_{j=1}^{J} q_j \tag{3.16}$$

**Exercise 3.8**
Comment on why the term $\left(q_{\mathrm{L}}\right)^{-J+1}$ is introduced in Equation (3.16).

## 3.9 Interactions between RIFs

Before we start discussing interaction, one should reflect on differences between "interaction (synergy)", "common cause" and "cascading effects".

So far we have assumed that the influence of one RIF on the basic event probabilities is independent of the value of the other RIFs in the basic Risk_OMT model. In many situations it might be reasonable to believe that for example the negative influence of a very bad RIF is higher if the one or more of the other RIFs also have a very bad value compared to more moderate values of these RIFs. Such effects are denoted interaction effects.

In the modelling of interaction effects we introduce sub sets of the total set of RIFs to represent a potential for interaction. In principle we may have several such sub sets. In the presentation we assume that we only consider one subset. Let $\mathscr{I}$ be a sub set of level one RIFs where we will consider interaction effects.

There could be arguments supporting that interaction only comes into play when two or more RIFs are in a very bad condition. We therefore assume that (negative) interaction only applies if the RIFs are w worse than the average value, i.e., a C.

In order to simplify the modelling of interaction effects we assign a weight, $w_{\mathscr{I}}$ of the interaction effect which is relative to the weight of the various RIFs in the interaction sub set, $\mathscr{I}$. For each RIF in the sub set $\mathscr{I}$ we then may find a total weight of the RIF in addition the original weight of the RIF, i.e.,

$$w_{\mathscr{I},i} = w_i w_{\mathscr{I}} f \tag{3.17}$$

where $w_i$ is the original weight of the RIF, and $f$ is a correction factor. If one or more of the RIFs have a value better than the average RIF-value (C) we set $f = 0$.

If all the RIF values in $\mathscr{I}$ have the worst value (F), we set $f = 1$. For values between we apply a linear transformation:

$$f = \frac{\sum_{i\in\mathscr{I}}(r_i/-C)}{\sum_{i\in\mathscr{I}}(F-C)} \tag{3.18}$$

where $C = 5/12$ and $F = 11/12$.

It is then easy to verify that if all RIFs take the value “C” we get $f = 0$, and if all RIFs take the value “F” we get $f = 1$ corresponding to our assumptions. The total impact of the RIFs on the basic event probability is now:

$$r = \sum_i w_i r_i + \sum_{i \in \mathscr{I}} w_{\mathscr{I},i} r_i \tag{3.19}$$

where we have summed the interaction effects for one sub set of interactions.

## 3.10 Common cause failures

Note the difference between interaction between RIFs and common cause failures. In Risk_OMT interaction means that the negative impact of one bad RIF is extra strong if also other RIFs are bad. The interaction is only expressing the impact of bad RIFs, nothing is said regarding the likelihood of bad values.

Common cause failures relate to dependencies, i.e., the likelihood of one failure depend on whether another event has occurred. In our context this also means that interaction effects are used for the RIFs, whereas common cause failures are used for the basic events.

There are many conceptual frameworks for common cause failures. Here we will stick to the $\beta$ factor model. In the $\beta$ factor model we assume that a portion of the failure rate probability represent common cause failures, and will cause two or more basic events to fail simultaneously. This portion is $\beta$. In the fault- and event tree modelling we may introduce extra basic events to represent the situation where two or more event occurs simultaneously due to common cause failures.

The challenge is to assess the numeric value of $\beta$. Podofolini et. al. (2009) summarizes the literature both with respect to factors included, and their importance. They find that the following factors are considered most important:

- Closeness in time
- Similarity of crew/performer(s)
- Stress
- Complexity

To be consistent with the Risk_OMT modelling framework we introduce a quantitative approach where each factor has a weight, say $w_i$ and a score $s_i$, and where the resulting common cause factor is given by:

$$\beta = \beta_0 \prod_i w_i^{s_i} \tag{3.20}$$

Here $\beta_0$ is a baseline common cause factor. The scores are measured on a scale from $s = -1$ representing the best value we can imagine for the factor, and $s = 1$ the worst value we can imagine.

Table 3.1: Weights and principles for setting scores in the CCF model

| **Dependency factor** | **Weight** | **Scores** | |
|---|---|---|---|
| | | **Best = -1** | **Worst = 1** |
| Closeness in time | 2 | The closeness in time is assumed to depend on the type of tasks considered. The following scores are proposed for the relevant situations.<br><br>Control planning \| Planning, $S = 1$<br><br>Control \| Execution, $S = -1$<br><br>Execution \| Execution, $S = ½$<br><br>Note, we assume there are no dependencies between planning activities and execution activities. | |
| Similarity of crew | 3 | Different crew, $S = -1$ | Same crew, $S = 1$ |
| Stress | 2 | The stress level is based on the RIF for *time pressure.* Let $r$ denote the linear mapping of the RIF on the interval (0,1) where A corresponds to 0, and F corresponds to 1. The score of the stress dependency factor is then given by<br>$S = 2(r - ½)$ | |
| Complexity | 1.5 | The Risk_OMT model does not include any RIF explicitly used to describe complexity. The RIF for *design* and *HMI* are considered to be the most relevant RIFs indicating complexity. If the values of these are denoted $r_1$ and $r_2$ respectively, the score of the complexity dependency factor is then given by<br>$S = (r_1 + r_2 - 1)$ | |

Table 3.1 presents numeric values for the weights used in RIsk_OMT. Further principles for assigning a value to the score is given. Some of the scores are determine based on already identified RIFs, whereas other scores are determined by an evaluation of the type of tasks to be executed, or other conditions to be considered related to the actual situation.

The baseline dependency level is set to $\beta_0 = 0.05$ for failure types “violation”, “omission”, and "mistake". For "slips & lapses" the common cause problem is considered slightly lower, and the value $\beta_0 = 0.03$ is reccomended.

There are two feasible ways to include the common cause effects in the modelling. One way is to model explicitly the common cause effects by introducing additional basic events in the fault and event trees. For a full BBN model this is the only way to represent such common cause effects. If the hybrid approach with a mixture of BBN models and event and fault trees is used, we may also introduce common cause failures in the post-processing of the minimal cut sets. The challenge then is to describe the possible dependencies for various classes of basic events, and then add common cause terms when the minimal cut set contributions are calculated. For example if a minimal cut set comprises the following basic events: {P=Planning error, CP=Control Planning error, E=Execution error, CE=Control Execution error} and we introduce $\beta_{\mathrm{CP|P}}$ and $\beta_{\mathrm{CE|E}}$ as common cause factors for controlling the plan, and controlling the execution respectively, we may use the following approximation to find the failure probability contribution from this minimal cut set:

$$Q_j \approx [q_{\mathrm{P}} q_{\mathrm{CP}} + \beta_{\mathrm{CP|P}} \min(q_{\mathrm{P}}, q_{\mathrm{CP}})][q_{\mathrm{P}} q_{\mathrm{CP}} + \beta_{\mathrm{CE|E}} \min(q_{\mathrm{E}}, q_{\mathrm{CE}})] \tag{3.21}$$

## 3.11 Importance measures

Risk importance measures are important in risk management. By having a good understanding of which factors and conditions contribute most to the risk, we may also start evaluating for which conditions risk reducing measures would be most efficient.

A common importance measure is the Birnbaums measure, $I^{\mathrm{B}}(i)$. For a fault tree or a reliability block diagram the measure is essentially a sensitivity measure reflecting the increase in system reliability if a component is improved.

In our Risk_OMT framework we have two challenges

- We deal with fault and event tree, meaning that there is not only one "TOP event", but it will be several end consequences to consider in such a sensitivity analysis
- The objective is to consider the RIFs, and not the basic events. Hence we would like to measure the total risk reduction if the condition of a RIF is improved. However, since the RIFs are described by stochastic variables, it is not straight forward to specify what is meant by improving a RIF

To cope with the situation of several end consequences in the event tree, we may introduce a numerical loss to each end consequences, and calculate the expected loss over all end consequences. The expected loss is then used as our risk measure when establishing our importance measure. In the Risk_OMT project the main focus was to focus on only one critical event in the event trees, hence we can simplify and only treat the frequency of that critical end event, say $F$.

To cope with the fact that the RIFs are only known in terms of the posterior distribution, it is proposed to investigate the impact of a change in the expected value of the RIF. The rationale is now that it might be possible to claim that a proposed measure will improve the condition of the RIF, but the only thing we may say is that the effect of the measure will be a change in the expected value.

Now let $\Delta E_j$ be a small change in the expected value of RIF number $j$ by implementing a risk reduction measures. For example $\Delta E_j = 1/12$ means that we can improve the RIF by a "half mark", for example if we have a "C",we end up with something in the middle between "B" and "C".

Let $\pi_j$ be the posterior distribution of RIF number $j$ before any improvement measure is considered, and let $\pi_j^\Delta$ be the distribution we imagine with the improvement measure. The frequency of the critical end consequence will be a function of the distribution of our RIF under consideration, and the proposed improvement measure is:

$$I_{\mathrm{RIF}}^{\mathrm{B}}(j) = \frac{F(\pi_j^\Delta) - F(\pi_j)}{\Delta E_j} \tag{3.22}$$

The calculation of equation (3.22) requires $\pi_j^\Delta$ and $\pi_j$ to be discretized.

To "change" a RIF is not straight forward. We approach this challenge differently for level one and level two RIFs.

For level one RIFs the posterior distribution is conditional on the corresponding level two RIF. We can therefore not just "change" one of the RIFs without impacting the other RIFs. In the assessment we now handle this by "disconnecting" the level one RIF under consideration. I.e., we calculate the unconditional pdf, say $\pi_j$, and then we assume that this RIF is independent of the other siblings and proceed with $\pi_j$ as a single node in the quantification, i.e., calculation of $F(\pi_j^\Delta)$ and $F(\pi_j)$.

For level two RIFs there will be not very efficient to change only the posterior distribution for this RIF. This is because the children of that RIF has historically been influenced by the "old" RIF, hence the data in terms of scores for the level one RIFs reflects the historical regime, and will have a "moment of inertia". The proposed workaround is just to add $\Delta E_j$ directly to the *score* of all the children of the level two RIF under consideration. $F(\pi_j^\Delta)$ in equation (3.22) is then actually not calculated by a modified posterior distribution, but we have a new calculation regime for $F(\cdot)$ based on "modified" level one scores.

## 3.12 Parameter estimation

The RIM parameters of interest are:

- $q_{\mathrm{L}}$ = Lowest value for the basic event probability, i.e., when all RIFs have state A
- $q_{\mathrm{H}}$ = Highest value for the basic event probability, i.e., when all RIFs have state F
- $w_j$ = Standardized weight of level 1 RIF number $j$ influencing the basic event
- $V_{\mathrm{P},j}$ = Structural importance of parent of level 1 RIF number $j$
- $V_{\mathrm{S},j}$ = Variance of the score of RIF number $j$ given the true underlying RIF value $r$

### 3.12.1 Variances

$V_{\mathrm{P},j}$ and $V_{\mathrm{S},j}$ represent "lack of deterministic" relations that in the end will represent variance in the observed scores. The aim is now to use observations to estimate $V_{\mathrm{P},j}$ and $V_{\mathrm{S},j}$. It might be shown that

$$\mathrm{Var}\left(S_{c,i}|r_{p,i} = s_{p,i}\right) = V_{\mathrm{P},j} + V_{\mathrm{S},j} \tag{3.23}$$

where $S_{c,i}$ is the score of the $i$'th observation of level one RIF number $j$, $r_{p,i}$ is the true underlying value of the corresponding level two RIF (i.e., the parent), and $s_{p,i}$ is the score for the parent RIF for the $i$'th observation. This means that we may find the variance of a level one RIF is we assume that the score of the parent RIF is equal to the true underlying value of the RIF.

Now, assume that we have $n$ observations of child and parent RIFs over some time period. In this period the true underlying values of the RIFs might change. We may now estimate the variance in equation (3.23) by:

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(s_{c,i} - s_{p,i}\right)^2 \tag{3.24}$$

where $s_{c,i}$ is the score of level one RIF number $j$ for the $i$'th observation, and $s_{p,i}$ is the score of the corresponding parent RIF.

We are not able to separate $V_{\mathrm{P},j}$ and $V_{\mathrm{S},j}$. This means that we need to make an expert judgement regarding their relative magnitude. It is proposed to set

$$V_{\mathrm{S},j} = 1/4 V_{\mathrm{P},j} \tag{3.25}$$

### 3.12.2 Estimating $q_\mathrm{L}$, $q_\mathrm{H}$ and $w_j$

Recall, that in Risk_OMT it is only the first level RIFs that influence the failure probability. Therefore, when estimating $q_\mathrm{L}$, $q_\mathrm{H}$ and $w_j$ we only use the scores on the first level RIFs. A better approach would have been to insert the expected values of the underlying RIFs based on a combination of the score and the parent RIF

The functional relation between the RIFs ($r_j$'s) and the basic event probability is given by:

$$q = q_\mathrm{L}\left(\frac{q_\mathrm{H}}{q_\mathrm{L}}\right)^{w_1 r_1 + w_2 r_2 + \ldots} = q_\mathrm{L}\left(\frac{q_\mathrm{H}}{q_\mathrm{L}}\right)^{\sum_j w_j r_j} \tag{3.26}$$

Now, inserting the score values for the RIFs and taking logarithm gives:

$$\ln q = \ln\left(q_\mathrm{L}\left(\frac{q_\mathrm{H}}{q_\mathrm{L}}\right)^{\sum_j w_j s_j}\right) = \ln q_\mathrm{L} + \sum w_j s_j \ln\left(\frac{q_\mathrm{H}}{q_\mathrm{L}}\right) \tag{3.27}$$

By letting $Y = \ln q$, $\beta_0 = q_\mathrm{L}$, $\beta_j = w_j \ln\left(q_\mathrm{H}/q_\mathrm{L}\right)$ and $x_j = s_j$ we get

$$Y = \beta_0 + \sum_j \beta_j x_j = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \tag{3.28}$$

which essential is a multiple linear regression model. Note that in eq. ( 3.28) we use $x$ for the independent variable, but in the estimation we will go back to $s$ to avoid mixing up with the number of errors later on in the estimation procedure.

For one combination of the score vector $\mathbf{s} = [s_1, s_2, \ldots s_r]$, there might be one or more observations. Typically there will be several observations if we have data on the basic event over a period of time where the score vector is assumed to remain constant, typically between surveys executed to assess the scores. We use $i$ as an index to run through the relevant combinations of the score vector, and $s_{ij}$ is the corresponding value of the score for level 1 RIF number $j$. Let $x_i$ be the reported number of failures or human errors, and $n_i$ the number of "trials", i.e., execution of the "basic" event. In the regression model we replace $q$ with it's estimate $\hat{q}_i = x_i/n_i$. However, taking the logarithm on the left hand side will cause problems in case $x_i = 0$.

This might be overcome with an empirical Bayesian approach where the prior distribution over $q$ is found by the sample mean and variance of $\hat{q}_i$. The beta distribution would be a conjugate prior. For the beta distribution with parameters $\alpha_0$ and $\gamma_0$ it might be shown that if we know the mean value, say $\mu$, and standard deviation, say $\sigma$, we have:

$$\gamma_0 = \left(\frac{\mu(1-\mu)}{\sigma^2} - 1\right)(1-\mu) \tag{3.29}$$

$$\alpha_0 = \frac{\mu\gamma_0}{1-\mu} \tag{3.30}$$

In the procedure we will no replace $\mu$ and $\sigma$ in eqs. (3.29) and (3.30) with the sample mean and standard deviation for $\hat{q}_i$.

The complete procedure for the estimation is now

1. Find a prior distribution for $q$ by an empirical Bayesian approach where the parameters in the prior distribution are given eqs. (3.29) and (3.30)

2. For each observation find the Bayes estimate for $q_i$ by calculating $\alpha_i = \alpha_0 + x_i$ and $\gamma_i = \gamma_0 + n_i - x_i$

3. Calculate the dependent variable, $y_i = \ln\left[\alpha_i/(\alpha_i + \gamma_i)\right]$

4. Let $s_{ij}$ be the score of the $j$th first level RIF for observation $i$

5. The standard multiple linear regression model now reads $Y_i = \beta_0 + \beta_1 s_{i1} + \beta_2 s_{i2} + \ldots + \beta_r s_{ir} + \epsilon_i$ Use a standard least square estimation procedure to find the estimates $\hat{\beta}_j$.

6. Find the standardized weights by: $\hat{w}_j = \hat{\beta}_j / \sum_{j=1}^{r} \hat{\beta}_j$

7. Find the range for $q$ by: $\hat{q}_{\text{L}} = e^{\hat{\beta}_0}$ and $\hat{q}_{\text{H}} = \hat{q}_{\text{L}} e^{\sum_{j=1}^{r} \hat{\beta}_j}$.

## Reefs in the sea

Lecture four was started by asking what is the difference:

1. The shipwreck was due to an unknown reef in the sea?

2. A high density of reefs in the sea gives more shipwrecks?

The first situation is relevant if we are analysing data from accident and incident report system, or investigation reports after accidents. The individual events are analysed and causes behind the accident or incident is reported. When analysing investigation reports from shipwreck for example in Norway, we may use such data to find for example the rate of shipwreck caused by hitting a reef:

$$\hat{f}_{\text{Reef}} = \frac{\text{Number of shipwrecks caused by hitting a reef}}{\text{Number of years for which we have data}} \tag{3.31}$$

Obviously have accident rates split on the various failure causes would be important, and this might help us implement risk reducing measure. However, eq. (3.31) will not help us distinguishing between how important the risk influencing factor “reef” is (weight) vs how many reefs there are (score).

Investigation reports might also help us in understanding factors influencing the damage potential of accidents. This is what was demonstrated by the study presented by Bahareh. Let $x_j$ be a factor to consider, and let $Y$ be the severity of an accident. A first order approach to try to understand the relation between the various factors that might affect an accident, and the severity we could establish a standard multiple linear regression model:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots \tag{3.32}$$

For example if $x_1$ is the distance from the bridge to the fire station, we might expect that $\beta_1 > 0$. Such a model can also get insight into the problem, and help us to identify risk reducing measures. For example we can have more dense fire stations, or at least have fire stations close to critical bridges.

Mohaghegh et al. (2009) discusses (i) Regression-based techniques, (ii) Formal probabilistic risk analysis techniques and (iii) BBN = Bayesian belief nets. In our context we emphasize that regression-based techniques could be very valuable, but they are usually not easy to apply in connection with formal probabilistic risk analysis techniques and BBN.

When we approach the question “A high density of reefs in the sea gives more shipwrecks” we would like to have a more explicit risk model utilizing formal probabilistic risk analysis techniques and BBN. For example we might establish fault- and event trees to analyse the seaward approach to ports where we both identify number of reefs as a RIF, but where we also consider use of echo sounder, availability of maps, competence, use of pilot etc as important RIFs.

Note if a reef is an important RIF in a risk model like Risk_OMT, it means that one or more model parameter will become significantly worse as the value of the RIF becomes worse and worse. The *importance* of the RIF is independent of the *value* in this context. In the “failure statistics” approach the RIF is important due to the combination of may reefs and failure to avoid hitting them, i.e., the statistical analysis cannot discriminate between the number of reefs and the mitigating measures we ahve.

In order to estimate model parameters for e.g., Risk_OMT we need accident and incident reports, but also exposure information in terms of description of fairway, technical condition on each ship, log of all ships entering the port etc.

Also in this situation we might establish regression models, but now these regression models are used to establish relation between risk influencing factors and model parameters in our fault- and event tree.

### 3.12.3 Aspects of dynamic risk analysis

The following section relates to the presentation given by Rizza Ardiyanti and some comments I gave. But I will emphasize more, and we could start by asking: What does “dynamic” mean?

a) Wikipedia: The dynamical system concept is a mathematical formalization for any fixed “rule” that describes the time dependence of a point’s position in its ambient space

b) Dynamic risk analysis is the process of updating existing risk analysis in light of new observations, new working practice, new systems installed etc, i.e., it is an update

c) Dynamic risk analysis is about to understand here and now what is happening, and use this to make here and now decisions regarding what is safe to do, is additional temporary barriers required etc.

# Chapter 4

# Bayesian belief network

## 4.1 Introduction

A Bayesian belief network (BBN = Bayesian Belief Network) is a probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed acyclic graph.

- A Bayesian belief network, for example, can represent the probabilistic relationships between diseases and symptoms
- Given symptoms, the BBN can be used to calculate the probability of the presence of various diseases.

The motivation behind the “method” is Bayes’ theorem. We start with the conditional probability rule:

$$\Pr(A|B) = \Pr(A \cap B)/\Pr(B) \text{and} \tag{4.1}$$

$$\Pr(B|A) = \Pr(A \cap B)/\Pr(A) \tag{4.2}$$

which leads to Bayes’ theorem:

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(B|A)\Pr(A)}{\Pr(B)} \tag{4.3}$$

Equation (4.3) says how we can “reverse” the situation. This is useful if, for example, we have a parametric model where Time-to-failure, $T$, is exponential distributed with parameter $\lambda$, i.e., we have $f(t|\lambda) = 1/\lambda e^{-\lambda t}$. If we know $\lambda$, we can make probability statements with respect to $T$ corresponding to $\Pr(B \text{ (i.e.} T)|A \text{ (i.e.} \lambda))$. While if we do not know $\lambda$, but have an observation for the failure time $T$ (a "proof"), we can make a probability statement about the failure rate, at $\Pr(A (\text{i.e. } \lambda)|B \text{ (i.e.} T))$ using Bayes’ theorem. This is actually what is happening in Bayesian statistics, where the probability desnity function for the failure rate $\lambda$ is updated when we get an observation, i.e., a failure time $T = t$.

To determine $\Pr(B)$ in the demonimator, we often use the law of total probability, which yields:

$$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\sum_i \Pr(B|A_i)\Pr(A_i)} \tag{4.4}$$

BBN can both be used to give probability statements about future events, and by "inverting" the equations, i.e., using Bayes' theorem, we can also update model parameters when we get observations (evidence).

The theory behind BBN can be difficult to understand, and especially the technical details. This means that we in limited degree can demonstrate BBN by hand calculation. However, there are computer tools that can make the necessary calculations where the user only specifies dependencies in a graphical tool, as well as probability statements and evidence (observations) at hand.

## 4.2 Definition

A Bayesian network is a directed acyclic graph (DAG = Direct Acyclic Graph) where the nodes represent stochastic variables and the edges (arrows) of the graph represent conditional dependencies.

If there is an arrow from node A to another node B, A is denoted a parent node of B, and B is a child of node A.

## 4.3 Product rule

A Bayesian network is a directed acyclic graph (DAG), while a DAG is only a Bayesian network if the common distribution of node values can be written as a product of local (marginal) distributions to each node and their parent nodes, i.e., we require:

$$\Pr(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \prod_{i=1}^{n} \Pr(X_i = x_i | X_j = x_j \text{ for all parentnodes } j \text{ til node } i) \tag{4.5}$$

If $X_i$ has no parent nodes, then the probability distribution is *unconditional*. Furthermore, if $X_i$ is observed, we say that it is an evident node.

The above product rule can in theory be used to make probability statements in the network, i.e., *inference*. To do this, we need the conditional probabilities. These are structured in so-called conditional probability tables (CPT = Conditional Proability Tables). The challenge with the calculations is that there are a lot of combinations to calculate, and even when using computers, the calculation time will become almost infinitely long, unless you can use clever heuristics or approaches. The commercial tools have good algorithms for such calculations.

## 4.4 Conditional probability tables

To specify a BBN model, we need to specify conditional probabilities. This is done by use of so-called conditional probability tables. Note that a BBN model can have stochastic variables (nodes) that are both discrete and continuous. But in order to be able to carry out calculations it is common to approximate all quantities with discrete stochastic variables. In fact, binary variables are often used to reduce the calculation and specification time of the model.

A conditional probability table (CPT = Conditional Probability Table) is a table for a stochastic variable where conditional probabilities for each value the variable can be specified for all possible combinations for the parent nodes.

If $X_i$ has no parent nodes, then the probability distribution is *unconditional*. Furthermore, if $X_i$ is observed, we say that it is an evident node.

**Example 4.1**

It is reasonable to assume that the probability of a human error depends on whether the work can be carried out indoors or outdoors. We define the following variables:

- $X_1 = T$ (true) if human error, $F$ (false) otherwise
- $X_2 = T$ if the work is carried out indoors, $F$ otherwise

An associated CPT can then be:

Table 4.1: Conditional probability table for human error

| Indors | Human error | |
|---|---|---|
| | T | F |
| T | 0.001 | 0.999 |
| F | 0.005 | 0.995 |

In the example, there is only one parent node, i.e., “Indoors”, while in principle we can have several parent nodes. In this example, it is natural to think of the parent nodes as “influencing factors”, cf. the PSFs in the chapter on human error.

## 4.5 Fault tree represented as a BBN

Any fault tree can be represented as a BBN diagram. This can be useful if the basic events in the fault tree have dependencies that we want to model in more detail. We can then create BBN structures for the basic events, and then connect these together first in the fault tree, and then to convert to a BBN diagram where both the underlying structures and the "logic" of the fault tree are collected in a common BBN model. This is an elegant approach, but it has turned out that the BBN model quickly becomes too large to be handled effectively.

Figure 4.1 illustrates how a fault tree can be represented as a BBN diagram. The fault tree is to the left, and the corresponding BBN diagram is on the right:

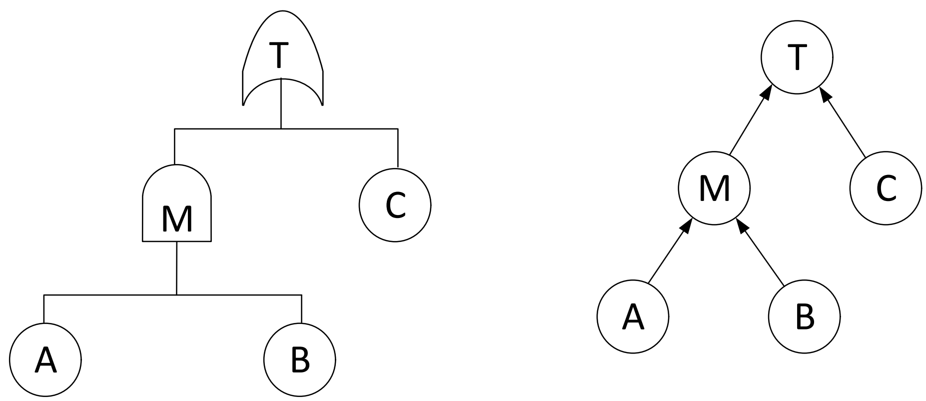


Figure 4.1: Fault tree to the left and BBN diagram to the right.

The TOP event is here indicated as T, M is an AND gate that connects A and B, while the TOP event T is an OR gate. In the corresponding BBN diagram, we do not have gates, and the logical structure will therefore be represented by CPTs, which are referred to here as "truth tables". A truth table is a binary table, where 1 means "True" and 0 means "False". Since this is a fault tree, "True" means that a gate *occurs*, while "False" means that it does not occur.

Figure 4.2 shows truth table for M in the BBN:

Table 4.2: Truth table for M

| A | B | M \| A, B |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |

We see that M is true only if both A and B are true, i.e., that both basic events A and B occur corresponding to the AND gate. Figure 4.3 shows the truth table for T:

Table 4.3: Truth table for T

| C | M | T \| C, M |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 1 |

Note that nodes A, B, and C have no parent nodes. This means that they are stochastically *independent*. The CPTs for these then only have two values, i.e., one for true, and one for false. Here, true will correspond to the probability that the basic event occurs (e.g., a failure), while false is the probability that the basic event is not occuring. Note that the truth tables are *regular* CPTs, so the calculation algorithms can be applied directly by specifying all the tables in a BBN tool.

# Chapter 5

# Human reliability analysis

## 5.1 Introduction

The purpose of a human reliability analysis (HRA) is to analyse human performance of sociotechnical systems. There are several objectives to conduct an HRA analysis such as ensure that humans are able to conduct their main tasks in operation of sociotechnical systems in a safe and reliable manner, to identify measures to improve human performance, and to quantify human errors as part of assessing risk of critical operations. One of the first systematic attempts to quantify human error probabilities (HEPs) was the Technique for Human Error Rate Prediction (THERP, Swain & Gutmann, 1983). A main purpose of the THERP method was to give input with respect to human performance into technical risk analyses. The methodological achievement of HRA techniques has primarily been achieved in the nuclear industry, but HRA methods have also extensively been applied in high risk activities such as aviation and the oil and gas industry. The THERP method is considered to be time consuming from an analysis point of view, hence simplified methods have been proposed. Among these are the Human Error Assessment and Reduction Technique (HEART, Williams, 1992) and more recently the Standardized Plant Analysis Risk (SPAR-H) human reliability analysis method (Gertman et. al., 2005). The SPAR-H methodology will be presented in later sections. Many HRA quantification techniques have been criticized because they treat humans in the same way as components and do not pay sufficient attention to the cognitive performance of humans. Therefore also Cognitive control based techniques have been proposed, among them is the Cognitive Reliability and Error Analysis Method (CREAM, Hollnagel 1998). In 2007 the HSE in UK presented a survey of HRA techniques for which 17 techniques were given a recommendation (Bell and Holroyd, 2009).

The Petro-HRA method was developed in the R&D project "Analysis of human actions as barriers in major accidents in the petroleum industry, applicability of human reliability analysis methods". The project was supported by the Norwegian Research Council and conducted during the period 2012-2016. The main result from the project was a guideline for the Petro-HRA method. The guideline has later been developed and the latest version is from 2022. The Petro-HRA method is mainly based on the SPAR-H method, but adapted to the petroleum industry.

Three important definitions are listed as a basis for the outline of the HRA techniques:

**Human error:** An out-of-tolerance action, or deviation from the norm, where the limits of acceptable performance are defined by the system. These situations can arise from problems in sequencing, timing, knowledge, interfaces, procedures, and other sources (NUREG/CR-6883).

**Human error probability (HEP):** A measure of the likelihood that plant personnel will fail to initiate the correct, required, or specified action or response in a given situation, or by commission will perform the wrong action. The HEP is the probability of the human failure event (ASME RA-S-2002).

**Performance shaping factor (PSF):** A factor that influences human performance and human error probabilities is considered in the HRA portion of risk analysis (NUREG/CR-6883).

In the RiskOMT method the term HEP is not used, rather the failure probability of basic events $q$, in a quantitative risk analysis or a barrier analysis is used. The RiskOMT term risk influencing factor, RIF, is used rather than PSF. Note that PSF's corresponds to level one RIFs in the RiskOMT method.

### 5.1.1 Main steps of an HRA

A typical quantitative HRA comprises the following steps (Rausand, 2011):

1. Identify critical operations where human errors could lead to accidents and/or operational problems. Such operations could be a critical marine operation as erecting the tower of a wind turbine, accessing an offshore energy converter, or conducting maintenance of a safety critical system.

2. Analyse relevant tasks and break down to subtasks and task steps. Various types of task analysis (Kirwan and Ainsworth, 1992) are often used. In subsequent sections the hierarchical and tabular task analysis will be presented.

3. Identify human error modes and, if possible, error causes and performance-influencing factors. Relevant techniques are AEMA, Human HAZOP and the tabular task analysis method presented in subsequent sections.

4. Determine human error probabilities (HEPs) for error modes and for a complete task by combining result for individual tasks with a sequence of tasks analysed by e.g., event tree analysis (ETA).

Task analysis may comprise the following main steps:

1. Breakdown of the task into subtasks and simple task steps.
2. Description of the allocation of task steps between different persons. This description gives an indication of communication needs.
3. Description of the temporal dependencies between subtasks or task steps. This will give input to quantification of so-called common cause failures.
4. Classification of task types (and task step types).
5. Identification of cues and feedback supporting each task step. The cues indicate to the operator that an action can/should be initiated (e.g., a *red traffic light* informs the driver to stop). The feedback informs the operator about the effects of carrying out the action (e.g., the driver will feel the *retardation* knowing that breaking takes place).

In order to get a better understanding of task types Rasmussen (1979) proposed to classify behaviour into the so-called Skill, Rule and Knowledge-based behaviour regime:

- **Skill-based**. Subconscious, automated actions requiring little or no cognitive effort (e.g., speed regulation in curves during ordinary driving).
- **Rule-based**. Actions according to an explicit rule or procedure (e.g., stopping at red light).
- **Knowledge-based**. Coping with unfamiliar situations without a procedure (e.g., diagnosis) requires conscious problem solving and decision-making.

Later Reason (1990) proposed a categorisation of human errors into slips, lapses, mistakes and violations:

- **Slip**. An action that is carried out with a correct intention but a faulty execution.
- **Lapse**. A failure to execute an action due to a lapse of memory or because of a distraction.
- **Mistake**. A correct execution of an incorrect intention.
- **Violation**. A person deliberately applies a rule or procedure that is different from what she knows is required, even though she may often do it with good intent.

Such categories may be of value when analysing tasks and possible human errors. The task analysis of a critical activity is often conducted in two steps. First the relations between tasks are visualised in a hierarchy (HTA = Hierarchical Task Analysis). On the TOP level the main task is shown. Then the necessary sub-task necessary for accomplishing the main task is shown with a plan. The plan shows the sequence of sub-task, typically *do in any order*, or *1 and 2 in order, then IF <condition> DO 3*. Each sub-task may then be further broken down into sub-sub tasks and so on. Figure 5.1 shows an example of a hierarchical task analysis for accessing a wind turbine from a boat.

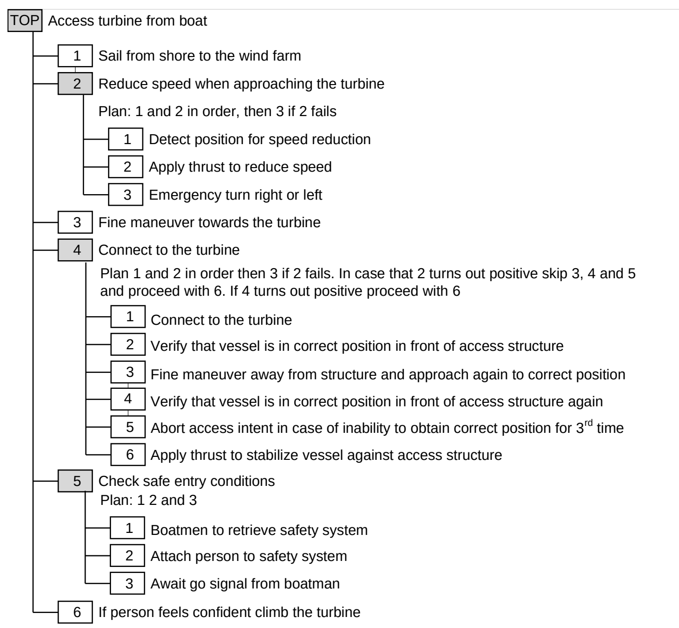


Figure 5.1: Snapshot of a hierarchical task analysis for the task of assessing the wind turbine from a boat

The tasks on the lowest level are now analysed by a tabular task analysis (TTA). Alternatively a Human HAZOP or an AEMA (Action Error Mode Analysis, Rausand, 2011) may be conducted. The purpose of the tabular task analysis is to identify possible human error (HEI = Human Error Identification). To structure the analysis each task from the HTA is analysed in one row in the TTA worksheet. In the following the term *action* is used to denote the task on the bottom level of the hierarchical task breakdown. Standard columns in a TTA comprise:

- ID
- Action
- Cues
- Feedback

- Possible errors (Action errors)
- Error causes
- Comments

The TTA does not usually provide guidewords for identification of possible errors. But we may adopt guidewords from techniques like Human HAZOP or AEMA. Like the standard HAZOP a set of guidewords are used to identify action errors. Typical guidewords are then (Shorrock et al. 2003)

- No action
- More action
- Less action
- Wrong action
- Part of action
- Extra action
- Other action
- More time
- Less time
- Out of sequence
- More information
- Less information
- Wrong information

Combining the action with the guideword gives possible action errors. For example the action “Open valve” combine with “Other action” gives the action error “Open wrong valve”. Note that the error causes will be considered later when assessing the effect of performance shaping factors (PSF) in the quantification part of the analysis.

### 5.1.2 Quantification of HEPs for action and complete tasks

A range of quantification methods have been proposed in order to asses human error probabilities. In the following the SPAR-H method is presented. The SPAR-H method is one of the recommended HRA techniques by HSE in UK(Bell and Holroyd, 2009), and is also partly based on three other methods (THERP, HEART and CREAM). The fact that it is also rather easy to implement and to apply resulted in that this method was chosen for presentation here. Although it is recommended to read the SPAR-H report the following presentation will give an idea how the method works.

#### SPAR-H - Introduction

SPAR-H presents a simple method for assessing the human error probabilities associated with operator and crew actions and decisions in response to initiating and preinitiator events. The method has been developed for use in the nuclear industry, and some adaptions have been made in this presentation. The basic SPAR-H framework comprises the following:

- Decomposes probability into contributions from diagnosis failures and action failures
- Accounts for the context associated with human failure events (HFEs) by using performance-shaping factors (PSFs), and dependency assignment to adjust a base-case HEP
- Uses pre-defined base-case HEPs and PSFs, together with guidance on how to assign the appropriate value of the PSF
- Employs a beta distribution for uncertainty analysis (not covered in this presentation)
- Uses designated worksheets to ensure analyst consistency.

An appealing feature of the SPAR-H method is that only two generic failure modes are considered, i.e., diagnosis failures and action failures. This is in contrast to methods like CREAM and HEART that introduces several generic task types[1]. To compensate for the limited number of task types and failure modes the performance shaping factors therefore have a relative strong impact. Also, compared to other methods, SPAR-H introduces only eight PSFs making the method fast compared to other methods where up to 40 factors are considered.

#### SPAR-H – Generic task types and baseline HEPs

SPAR-H introduces two task types with corresponding error modes:

**Diagnosis failures**: These are failures with respect to correct diagnosis of a situation. The diagnosis failure mode is also used as a failure mode as part of planning activities and verifications.

**Action failures**: These are failures in physical execution of work.

The following baseline (nominal) human error probabilities are given for these failure modes:

- Diagnosis failures: Nominal HEP = $10^{-2}$
- Action failures: Nominal HEP = $10^{-3}$

The baseline HEPs of SPAR-H have been developed for the nuclear industry where tasks are conducted indoor. In situations covered by MARINA most activities will be outdoor activities sometimes conducted under harsh environmental conditions. Since SPAR-H do not provide PSFs for workplace conditions, it is therefore proposed to introduce a more comprehensive set of nominal HEPs (NHEP) shown in Table 5.1:

Table 5.1: Modified nominal HEPs (NHEP) from SPAR-H adjusted for indoor/outdoor applications

| Failure mode | NHEP |
|---|---|
| Diagnosis failure – Outdoor, normal weather | $2 \times 10^{-2}$ |
| Action failure – Outdoor, normal weather | $3 \times 10^{-3}$ |
| Diagnosis failure – Outdoor, bad weather | $5 \times 10^{-2}$ |
| Action failure – Outdoor, bad weather | $10^{-2}$ |
| Diagnosis failure – Indoor | $10^{-2}$ |
| Action failure – Indoor | $10^{-3}$ |

**SPAR-H – Performance shaping factors (PSFs)**

SPAR-H introduces eight PSFs which are:

1. Available time
2. Stress and stressors
3. Experience and training
4. Complexity
5. Ergonomics (including the human-machine interface)
6. Procedures

7. Fitness for duty
8. Work processes

For a detailed description and discussion of these factors reference is made to the SPAR-H report (Gertman et.al., 2005). For each activity for which a HEP is to be assigned, a review of the state of the PSFs is required. A dedicated worksheet is provided, but the essence of this worksheet is shown in Table 5.2 for *diagnosis* failure modes and in Table 5.3 for *action* failure modes.

**SPAR-H - HEP correction based on composite PSF correction factors**

The factors for each PSFs are multiplied to give a composite factor, say $F = f_1 \cdot f_2 \cdot \ldots \cdot f_8$, yielding the adjusted HEP:

$$\mathrm{HEP} = F \cdot \mathrm{NHEP} = f_1 \cdot f_2 \cdot \ldots \cdot f_8 \cdot \mathrm{NHEP} \tag{5.1}$$

In case of more than 3 negative HEPs (factor > 1) an adjustment formula is proposed to avoid probabilities larger than one:

$$\mathrm{HEP} = \frac{F \cdot \mathrm{NHEP}}{(F-1) \cdot \mathrm{NHEP} + 1} \tag{5.2}$$

Note that a correction factor of “Pr(Failure)=1” in Table **8** and Table 9 means that the entire failure mode probability is set to one for this failure mode.

**SPAR-H Dependency calculations**

When activities are conducted in a sequence, it is generally believed that a failure of one activity will lead to a higher failure probability of subsequent activities due to common mode errors. To assess conditional probability for subsequent activities a dependency assessment is carried out. Four factors are considered for activities that may be dependent on previous activities in the course of events:

- Crew (*s*=same or *d*=different)
- Time (*c*=close in time or *nc*=not close in time)
- Location (*s*=same or *d*=different)
- Cues (*a*=additional or *na*=no additional)

Now, let the PSF Without Formal Dependence be denoted HEPw/od and assume that the dependency has been categorised according to Table 5.4. Further let HEPw/d denote the probability of failures With Formal Dependence. HEPw/d is now found form Table 5.5.

Table 5.2: Correction factors for different PSF levels for each PSF for diagnosis

| PSF | PSF level | Factor |
|---|---|---|
| Available Time | Inadequate time | Pr(Fail.)=1 |
| | Barely adequate time (approx. 2/3 x nominal) | 10 |
| | Nominal time | 1 |
| | Extra time (between 1 and 2 x nominal & > 30 min) | 0.1 |
| | Expansive (between 2 and 4 x nominal & > 30 min) | 0.05 |
| | Expansive (> 4 x nominal & > 30 min) | 0.01 |
| | Insufficient Information | 1 |
| Stress/Stressors | Extreme | 5 |
| | High | 2 |
| | Nominal | 1 |
| | Insufficient Information | 1 |
| Complexity | Highly complex | 5 |
| | Moderately complex | 2 |
| | Nominal | 1 |
| | Obvious diagnosis | 0.1 |
| | Insufficient Information | 1 |
| Experience/Training | Low | 10 |
| | Nominal | 1 |
| | High | 0.5 |
| | Insufficient Information | 1 |
| Procedures | Not available | 50 |
| | Incomplete | 20 |
| | Available, but poor | 5 |
| | Nominal | 1 |
| | Diagnostic/symptom oriented | 0.5 |
| | Insufficient Information | 1 |
| Ergonomics/HMI | Missing/Misleading | 50 |
| | Poor | 10 |
| | Nominal | 1 |
| | Good | 0.5 |
| | Insufficient Information | 1 |
| Fitness for Duty | Unfit | Pr(Fail.)=1 |
| | Degraded Fitness | 5 |
| | Nominal | 1 |
| | Insufficient Information | 1 |
| Work Processes | Poor | 2 |
| | Nominal | 1 |
| | Good | 0.8 |
| | Insufficient Information | 1 |

Table 5.3: Correction factors for different PSF levels for each PSF for action

| PSF | PSF level | Factor |
|---|---|---|
| Available Time | Inadequate time | Pr(Fail.)=1 |
| | Time available is approx. the time required | 10 |
| | Nominal time | 1 |
| | Time available >= 5x the time required | 0.1 |
| | Time available is >= 50x the time required | 0.01 |
| | Insufficient Information | 1 |
| Stress/Stressors | Extreme | 5 |
| | High | 2 |
| | Nominal | 1 |
| | Insufficient Information | 1 |
| Complexity | Highly complex | 5 |
| | Moderately complex | 2 |
| | Nominal | 1 |
| | Insufficient Information | 1 |
| Experience/Training | Low | 3 |
| | Nominal | 1 |
| | High | 0.5 |
| | Insufficient Information | 1 |
| Procedures | Not available | 50 |
| | Incomplete | 20 |
| | Available, but poor | 5 |
| | Nominal | 1 |
| | Insufficient Information | 1 |
| Ergonomics/HMI | Missing/Misleading | 50 |
| | Poor | 10 |
| | Nominal | 1 |
| | Good | 0.5 |
| | Insufficient Information | 1 |
| Fitness for Duty | Unfit | Pr(Fail.)=1 |
| | Degraded Fitness | 5 |
| | Nominal | 1 |
| | Insufficient Information | 1 |
| Work Processes | Poor | 5 |
| | Nominal | 1 |
| | Good | 0.5 |
| | Insufficient Information | 1 |

Table 5.4: Assignment of dependency level

| Crew | Time | Location | Cues | Dependency |
|---|---|---|---|---|
| s | c | s | na | complete |
| | | | a | complete |
| | | d | na | high |
| | | | a | high |
| | nc | s | na | high |
| | | | a | moderate |
| | | d | na | moderate |
| | | | a | low |
| d | c | s | na | moderate |
| | | | a | moderate |
| | | d | na | moderate |
| | | | a | moderate |
| | nc | s | na | low |
| | | | a | low |
| | | d | na | low |
| | | | a | low |

**SPAR-H Event Sequence Diagram**

In order to model the sequence of tasks established in the HTA (see Figure 5.1) an event tree, cause consequence diagram or event sequence diagram is used. Figure 5.2 shows a generic example of such a diagram. To find the end event frequencies the frequency of the initiating (hazardous) event is multiplied with the success or failure probabilities along the path leading to the end event in Figure 5.2. Note that conditional dependency probabilities are required, i.e., the HEPw/d-values from Table are required.

For example HEART introduces the following generic task types: 1) Totally familiar, performed at speed with no idea of likely consequences, 2) Shift or restore system to new or original state on a single attempt without supervision or procedures, 3) Complex task requiring a high level of comprehension and skill, 4) Fairly routine task performed rapidly or given scant attention, 5) Routine highly practised, rapid task involving a relatively low level of skill, 6) Restore or shift a system to the original or new state following procedures with some checking, 7) Completely familiar, well-designed, highly practised routine task occurring several times per hour, 8) Respond correctly to system command even when there is an augmented or automated supervisory system.

Table 5.5: Correction formulas for HEPs with formal treatment of dependencies

| Dependency level | Correction formulas for dependencies |
|---|---|
| Complete | HEPw/d = 1. |
| High | HEPw/d = (1+ HEPw/od)/2 |
| Moderate | HEPw/d = (1+6HEPw/od)/7 |
| Low | HEPw/d = (1+19HEPw/od)/20 |
| Zero | HEPw/d = HEPw/od |

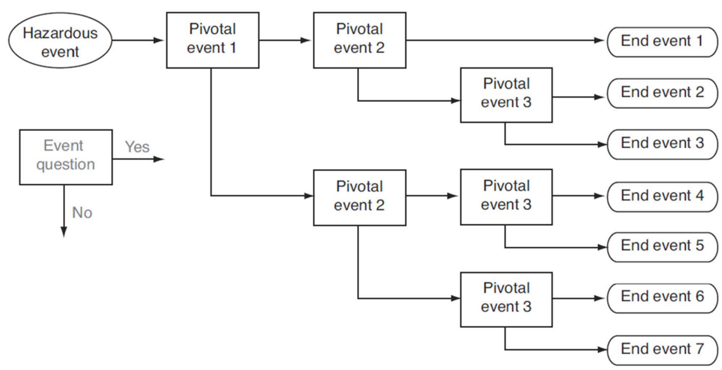


Figure 5.2: Event sequence diagram

# Human error classification and probabilities

This section elaborate more on human error classification and human error probabilties. In particular we review aspects of human error classification used in the Risk_OMT model.

Human error is defined by Reason (1990): *The failure of planned actions to achieve their desired ends - without the intervention of some unforeseeable event.* This definition does not really elaborate on different aspects, so therefore a classification regime is often introduced.

### 5.1.3 Failure of omission and failure of execution

A failure of an activity may further be divided into failure of omission and failure of execution. Failure of omission denotes whether or not the prescribed activity is carried out. Failure of execution denotes inadequate actions that may cause failures, e.g., acts performed in a wrong sequence, at the wrong time, without the required precision, etc.

Failure of execution is seen as results of violations or human errors (Reason, 1990). Violations refer to both deliberate and unintentional omissions of one or several steps within a work task.

Human error is further divided into mistakes and slips & lapses, where mistakes involve actions that are based on failure of interpretation of procedures, and/or failures of judgmental/inferential processes involved in the prescribed activity. Slips & lapses involve actions that represent unintended deviation from those practices represented in the formal procedures. To summarize we have for the failure of an activity:

- Omission failure
- Execution failure
  - Violation failure
  - Human error
    - Slips & lapses
    - Mistakes

Figure 5.3 shows the corresponding fault tree with the RIF structure. Note that in Risk_OMT the RIF structure is only developed for execution failures.

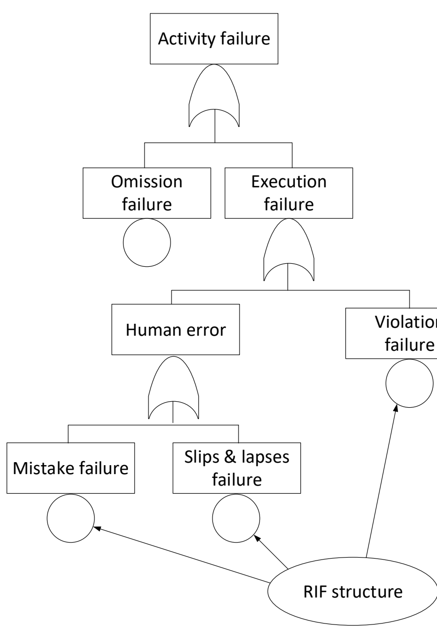


Figure 5.3: Structuring activity failure and RIF structure

### 5.1.4 Generic RIF model for execution and control activities

In Risk_OMT a generic RIF structure is developed. The level two RIFs are management RIFs related to:

- Competence
- Information
- Technical issues
- General issues
- Tasks

For each level two RIF there is one more level one RIFs. In In Risk_OMT there is one generic model for execution and control activities as shown in Figure 5.1.4, and one generic model for planning activities shown in Figure 5.5. This means that the type of activity determines the qualitative nature of the RIF structure.

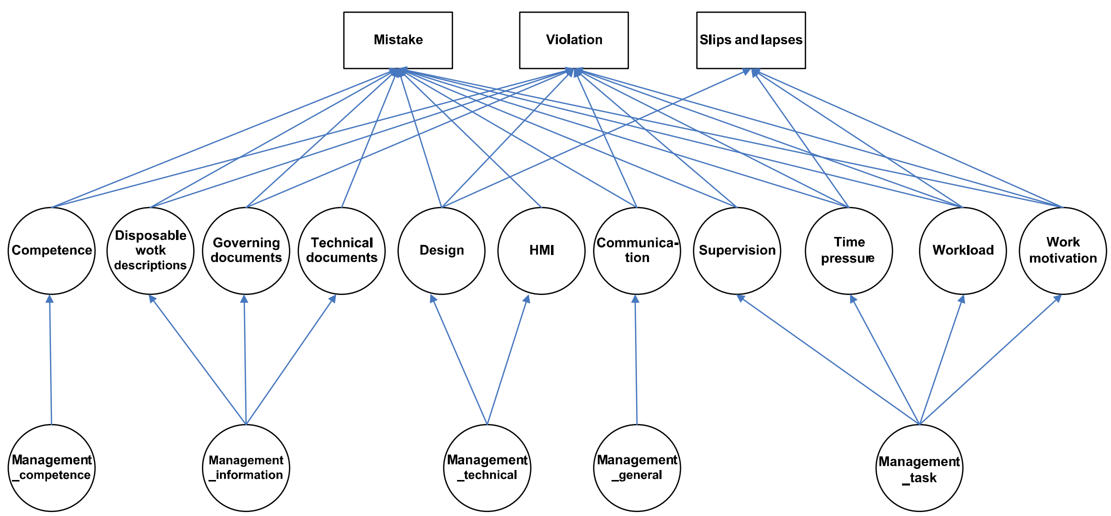


Figure 5.4: Generic RIF model for execution and control activities

### 5.1.5 Weights and "variances"

Note the following:

- In Risk_OMT the basic event failure probabilities depend on the *value* of each RIF, and the relative weight of the RIFs

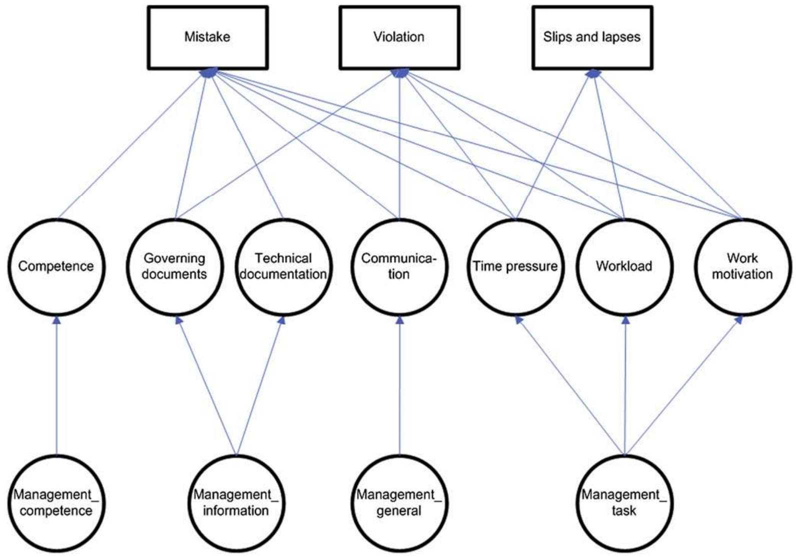


Figure 5.5: Generic RIF model for planning activities

- It is assumed that a particular RIF, e.g., "Technical documentation" is the same for all basic event. This means that we assess the values of the RIFs independent of which basic event is considered.
- The weights of the RIFs could in principle be different for each and every basic event. However, it might be more efficient to give a set of weights for a given type of basic events.
- Referring to Figures 2 and 3, we should as a minimum define 6 set of weights, i.e., for "mistakes", "violations", "slips & lapses" for both "planning" activities and "execution and control" activities.
- In Risk_OMT we define two set of "variances". For all RIFs we need to define the variance of the score given the true underlying RIF. This variance represents how difficult it is to get to know the underlying RIF by the information we have. The second variance is the structural variance, i.e., how much variance it will be in the true value of a first level RIF given a value of the corresponding second level RIF.
- The assessment of these two type of variances are done RIF by RIF, and is common for all basic events.

### 5.1.6 Nominal human error probabilities

In human reliability analysis we distinguish between human error probabilities (HEPs) and nominal HEPs. The HEPs are the probabilities we will use in a given analysis for a given set of the RIFs assessed. The nominal HEPs are the baseline HEPs. In Risk_OMT and BORA these nominal HEPs are those values for the HEPs we will have in the "average" case, i.e., if all RIFs were on their average, corresponding typically to a C.

In the BORA papers numerical values for nominal HEPs are discussed. The proposed numerical values are those that by the authors are considered relevant for oil and gas, and the type of tasks that are involved in typical maintenance and operation activities. It might be relevant to use these numerical values as starting point also for similar activities in maintenance and operation of e.g., infrastructure systems. However, for tasks that are of more "crisis management" and sharp end operation of equipment it might be required to do further investigation.

In the Risk_OMT papers some more considerations are made to take into account the more detailed model for categorizing the activity failures.

### 5.1.7 Error factors

In the Risk_OMT an error factor is used to define the spread of the HEPs relative to the nominal HEPs. The error factors are discussed in the Risk_OMT papers, but no numerical values are given. In the BORA papers the error factors are given for some type of activities, but though not for the different failure categories. Typical values of the error factors are in the order of magnitude 3 to 5.

# Chapter 6

# Siutation awareness

## 6.1 Introduction

Situational awareness or situation awareness (SA) is the understanding of an environment, its elements, and how it changes with respect to a time vector or other factors, is critical for appropriate and optimized decision making in many environments.

Several authors are using the term, and both "situational" and "situation" are used when referring to SA. **?** is the most cited paper, and she defines SA as:

Several authors have criticized the framework proposed by **?** and in a paper from 2015 she claim that most of the critique is based on misunderstanding of her original paper, see **?**.

This memo is to a large extend based on the Wikipedia article on SA.

## 6.2 Endsley's Cognitive Model of SA

The most widely cited and accepted model of SA was developed by Endsley (1995) which has been shown to be largely supported by research findings.

Endsley's model describes the cognitive processes and mechanisms that are used by people to assess situations to develop SA, and the task and environmental factors that also affect their ability to get SA. It describes in detail the three levels of SA formation: perception, comprehension, and projection. Figure 6.1 shows the model:

The next sections describe the core content of the three levels of SA. Some authors have argued that these levels are to be interpreted as linear development of SA, but Endsley argues that this is not the case in a the paper form 2015.

- *Perception (Level 1 SA):* The first step in achieving SA is to perceive the status, attributes, and dynamics of relevant elements in the environment. Thus, Level 1 SA, the most basic level of SA, involves the processes of monitoring, cue detection, and simple recognition,

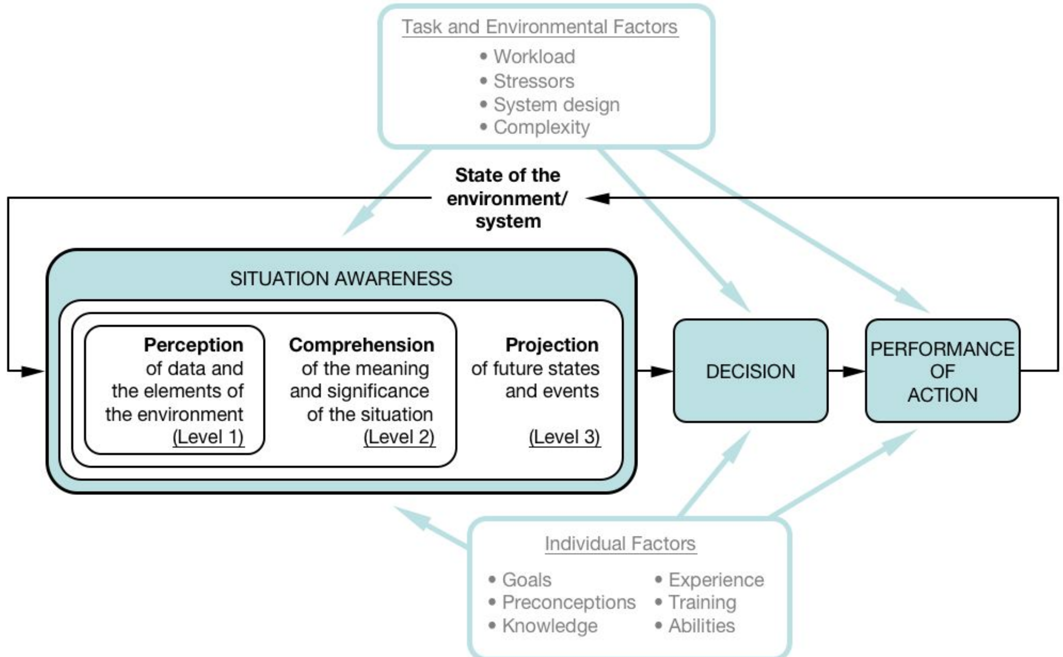


Figure 6.1: Endsley's model of SA

which lead to an awareness of multiple situational elements (objects, events, people, systems, environmental factors) and their current states (locations, conditions, modes, actions).

- *Comprehension (Level 2 SA):* The next step in SA formation involves a synthesis of disjointed Level 1 SA elements through the processes of pattern recognition, interpretation, and evaluation. Level 2 SA requires integrating this information to understand how it will impact upon the individual's goals and objectives. This includes developing a comprehensive picture of the world, or of that portion of the world of concern to the individual.

- *Projection (Level 3 SA):* The third and highest level of SA involves the ability to project the future actions of the elements in the environment. Level 3 SA is achieved through knowledge of the status and dynamics of the elements and comprehension of the situation (Levels 1 and 2 SA), and then extrapolating this information forward in time to determine how it will affect future states of the operational environment.

### 6.2.1 Decision making

Figure 6.1 depict SA as the primary basis for subsequent decision making and performance in the operation of complex, dynamic systems. The feedback loop (State of the environment/system) is important, and a distinction could be made between:

- Decisions to gain more knowledge regarding the state of the environment and system, e.g., collect more information, open a new window on the computer, run simulations etc

- Decisions to change the system state

See Endsley (2015) for further discussions.

## Time and space

SA also involves both a temporal and a spatial component. Time is an important concept in SA since as new inputs enter the system, the individual incorporates them into this mental representation, making changes as necessary in plans and actions in order to achieve the desired goals. SA also involves spatial knowledge about the activities and events occurring in a specific location of interest to the individual.

Endsley's model of SA illustrates several variables that can influence the development and maintenance of SA, including individual, task, and environmental factors.

Key factors that describe the cognitive processes involved in SA are according to Endsley (1995):

- Perception, comprehension, and projection as three levels of SA,

- The role of goals and goal directed processing in directing attention and interpreting the significance of perceived information,

- The role of information salience in "grabbing" attention in a data-driven fashion, and the importance of alternating goal-driven and data-driven processing,

- The role of expectations (fed by the current model of the situation and by long-term memory stores) in directing attention and interpreting information,

- The heavy demands on limited working memory restricting SA for novices and for those in novel situations, but the tremendous advantages of mental models and pattern matching to prototypical schema that largely circumvent these limits,

- The use of mental models for providing a means for integrating different bits of information and comprehending its meaning (relevant to goals) and for allowing people to make useful projections of likely future events and states,

- Pattern matching to schema - prototypical states of the mental model - that provides rapid retrieval of comprehension and projection relevant to the recognized situation and in many cases single-step retrieval of appropriate actions for the situation.

The model also points to a number of features of the task and environment that affect SA:

- The capability of the system and the user interface for conveying important information to the person in a way that is easy to integrate and process.
- Both high workload and stress can negatively affect SA. Information overload is a problem in many situations.
- Underload (vigilance conditions) can also negatively affect SA.
- The complexity of the systems and situations a person is in can negatively affect SA by making it difficult to form accurate mental models.
- Automation is a major factor reducing situation awareness in many environments (e.g. aviation, driving, power operations). See out of the loop performance problems. This is due to it creating situations where people are forced to become monitors which they are poor at (due to vigilance problems), often poor system transparency with needed information not provided, and an overall reduction in the level of cognitive engagement of people with automated systems

Experience and training have a significant impact on people's ability to develop SA, due to its impact on the development of mental models that reduce processing demands and help people to better prioritize their goals.

A note should be made regarding use of *schemata* and automation. In psychology and cognitive science, a *schema* describes a pattern of thought or behaviour that organizes categories of information and the relationships among them. It can also be described as a mental structure of preconceived ideas, a framework representing some aspect of the world, or a system of organizing and perceiving new information, such as a mental schema or conceptual model. Schemata influence attention and the absorption of new knowledge: people are more likely to notice things that fit into their schema, while re-interpreting contradictions to the schema as exceptions or distorting them to fit. Schemata have a tendency to remain unchanged, even in the face of contradictory information. Schemata can help in understanding the world and the rapidly changing environment. People can organize new perceptions into schemata quickly as most situations do not require complex thought when using schema, since automatic thought is all that is required.

### 6.2.2 Relationship of goals and mental models to SA

Endsley (1095) points out that SA is an ongoing dynamic process of gathering and interpreting information to update the situation model and using that situation model to search for information until decisions can be made. Further there is a linkage between goals and mental models that drives the development or selection of plans and scripts for directing actions, the use of the activated mental model to direct attention to the environment to feed into the constantly updated situation model, and use of the situation model in updating the selection of the mental model to be active. Figure 6.2 depicts relationship of goals and mental models to SA.

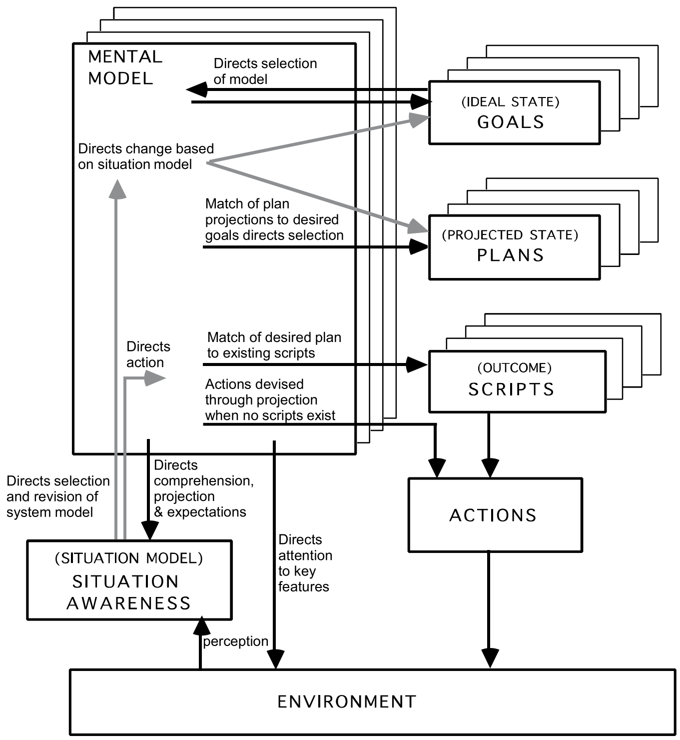


Figure 6.2: Relationship of goals and mental models to SA (Endsley, 1995)

As indicated above, we could distinguish between two types of decisions important to understand the feedback loop in Figure 6.1:

- Decisions to gain more knowledge regarding the state of the environment and system, e.g., collect more information, open a new window on the computer, run simulations etc
- Decisions to change the system state

In Figure 6.2 there are three "feedback" loops:

1. ENVIRONMENT →(perception)→ SIUTATION AWARENESS →(Directs action)→MENTAL MODEL→(Directs attention to key features)→ENVIRONMENT
2. ENVIRONMENT →(perception)→ SIUTATION AWARENESS →(Directs action)→MENTAL MODEL→(Match of desired plan to existing scripts)→SCRIPT→ ACTION →ENVIRONMENT

3. ENVIRONMENT →(perception)→ SIUTATION AWARENESS →(Directs action)→MENTAL MODEL→(Actions devised thorough projection when no scripts exist)→ACTION→ENVIRONMENT

# Chapter 7

# Finding minimal cut sets in combined event- and fault tree systems

## 7.1 Introduction

This memo briefly describes how minimal cut sets may be obtained for combined event- and fault tree systems. It is assumed that the reader is familiar with the definition of cut sets in general, and how to obtain these.

## 7.2 Definitions

A *cut set* in a fault tree is a set of Basic events whose (simultaneous) occurrence ensures that the TOP event occurs. A cut set is said to be *minimal* if the set cannot be reduced without loosing its status as a cut set.

A *path set* in a fault tree is a set of Basic events whose non-occurrence (simultaneously) ensures that the TOP event does not occur. A path set is said to be *minimal* if the set cannot be reduced without loosing its status as a path set.

For small and simple fault trees, it is feasible to identify the minimal cut- and path sets by inspection without any formal procedure/algorithm. For large or complex fault trees we need an efficient algorithm. The MOCUS algorithm (Method for obtaining cut sets) is described in standard FTA textbooks, and an efficient improvement of the algorithm is described by Vatn (1993).

### 7.2.1 Dual fault tree

Let FT be a fault tree with basic events $BE_i$. A dual fault tree to FT, say FT*, is obtained by changing all AND gates in FT to OR gates, and all OR gates in FT to AND gates, and finally the basic events of FT* are the complements of the corresponding basic events in FT.

### 7.2.2 Theorem

The minimal path sets of FT is given by the minimal cut sets of FT* with the complement basic events $BE_i^*$ are replaced with $BE_i$.

## 7.3 Approach

This result will be useful if we have implemented an algorithm to find minimal cut sets, and if we need the minimal path sets. When we combine event- and fault trees this will be the case.

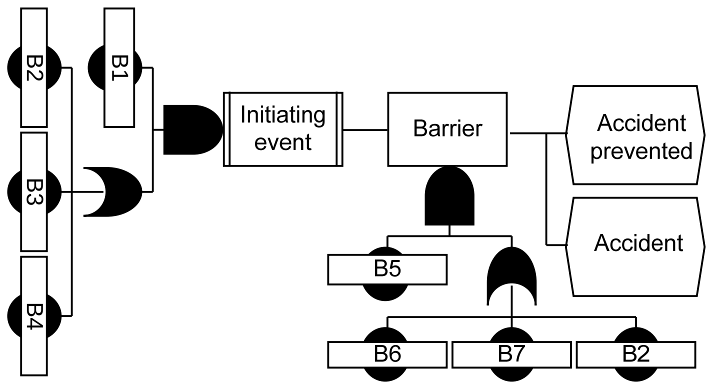


Figure 7.1: Example system combining event and fault tree

Figure 7.1 shows a system of combined fault and event trees. There is only one barrier in the event tree after the initiating event. To the left of the initiating event we have drawn a fault tree representing the combination of events that lead to the initiating event. Further below the barrier we have drawn a fault tree with the TOP event corresponding to the failure of the barrier. Two end consequences are drawn, one for “Accident prevented” corresponding to success of the barrier, and one for “Accident” corresponding to the failure of the barrier, i.e., the occurrence of the TOP event. Minimal cut sets for the left most fault tree are given by {B1,B2},{B1,B3}, and {B1,B4}. Minimal cut sets for the fault tree below the barrier are given by {B5,B6},{B5,B7}, and {B5,B2}.

Figure 7.2 shows the dual fault tree of the rightmost fault tree in Figure 7.1. The minimal cut sets of this dual fault tree is {B5*} and {B6*,B7*,B2*}. Thus the minimal path sets of the original fault tree is given by {B5} and {B6,B7,B2}. We observe that the occurrence of at least one cut set of the dual fault tree will ensure that the outcome of the barrier is a success. For example the occurrence of {B5*} corresponds to the non- occurrence of basic event B5.

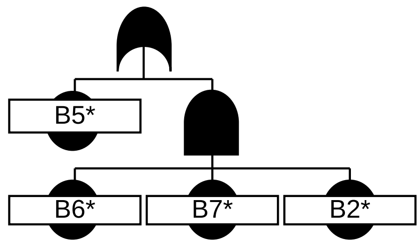


Figure 7.2: Dual fault tree

In order to find the cut sets for each end consequence in the combined fault- and event tree system we need two types of cut sets operators, the & operator, and reduction operators. Let CSs1 and CSs2 be two cut sets, where CSs1 contains the minimal cut sets CS11, CS12, ..,CS1$m$, and CSs2 contains the minimal cut sets CS21,CS22,..,CS2$n$. We now define the &-operator for two set of cut sets:

### &-operator

The &-operator for two set of cut sets is defined such that CSs3 = CSs1 & CSs2 is a new set of cut sets where CSs3 is the set of all combination of minimal cut sets from CSs1 and CSs2. A combination of two cut sets CSa and CSb in this context is the set of all events in CSa and all events in CSb (the union of events in each of them). Note that CSs3 might contain non-minimal cut sets. Also note that each cut set contained in CSs3 may contain repeated evens, and also events that may not occur simultaneously. After applying the &-operator we need post-processing of the result. Three types of reductions are necessary, (i) eliminate repeating events, (ii) remove cut sets with two or more events that may not occur simultaneously, and (iii) eliminate non minimal cut sets. The idea for finding cut sets for each end consequence we collect relevant cut sets along the paths from the initiating event to the various end consequences. Note that if a fault tree is not developed for the initiating event or a barrier the cut sets only contain one cut set, and this one again is only one event, either (i) the initiating event, or (ii) the success, or (iii) failure of the barrier

## 7.4 Procedure

The procedure for finding the minimal cut sets for one end consequence is now as follows where CSs is the running set of cut sets:

1. Start with the initiating event. If a fault tree is developed for the initiating event, let CSs be the corresponding set of minimal cut sets, else the set of cut sets is the initiating event it self.

2. Proceed with the next barrier along the path until the required end consequence is reached.
3. If the barrier along the path we are following corresponds to a failure of this barrier, let CSsB be the set of cut sets for the fault tree of the barrier, or let CSsB be the barrier failure it self if no fault tree is developed. If we are following the success of this barrier let CSsB be the minimal cut set of the dual fault tree, or if no fault tree exists for this barrier, let CSsB be the complement of the barrier failure, i.e., the barrier success.
4. Apply the & operator, i.e., CSs ← CSsB & CSs
5. Remove the second of repeated events in each cut set of CSs
6. Remove cut sets of CSs containing a basic event, say $B_i$, and it's complement $B_i^*$.
7. Remove non-minimal cut sets in CSs
8. GoTo Step 2
9. When the end consequence is reached, CSs now is the set of minimal cut sets for this event

We now demonstrate the procedure for the example in Figure 7.1, and we start with the "accident" end consequence.

In Step 1 we find CSs = { {B1,B2},{B1,B3},{B1,B4} }. In Step 3 we find CSsB = { {B5,B6},{B5,B7}, {B5,B2} }. By applying the & operator we get:

CSs = { {B1,B2,B5,B6}, {B1,B2,B5,B7}, {B1,B2,B5,B2},{B1,B3,B5,B6}, {B1,B3,B5,B7}, {B1,B3,B5,B2}, {B1,B4,B5,B6},{B1,B4, B5,B7},{B1,B4, B5,B2} }.

In Step 5 we remove one occurrence of B2 where it occurs twice, and get:

CSs = { {B1,B2,B5,B6}, {B1,B2,B5,B7}, {B1,B2,B5}, {B1,B3,B5,B6}, {B1,B3,B5,B7}, {B1,B3,B5,B2}, {B1,B4,B5,B6}, {B1,B4,B5,B7}, {B1,B4,B5,B2} }

Neither {B1,B3,B5,B2} nor {B1,B4,B5,B2} are minimal cut sets since {B1,B2,B5} is a minimal cut set, hence these two cut sets are removed in Step 7, and we remain with:

CSs = { {B1,B2,B5,B6}, {B1,B2,B5,B7}, {B1,B2,B5}, {B1,B3,B5,B6}, {B1,B3,B5,B7}, {B1,B4,B5,B6}, {B1,B4,B5,B7}}

These are the minimal cut sets since there are no occurrences of an event and its complement in any of the cut sets, and there are no non-minimal cut sets left.

Proceeding to the “accident prevented” end consequence, we have:

In Step 1 we find CSs = { {B1,B2},{B1,B3},{B1,B4} }. In Step 3 we find for the dual fault tree CSsB = { {B5*},{B6*,B7*,B2*} }. By applying the & operator we get CSs = { {B1,B2,B5*}, {B1,B2,B6*,B7*,B2*}, {B1,B3,B5*}, {B1,B3,B6*,B7*,B2*},{B1,B4,B5*}, {B1,B4,B6*,B7*,B2*} }.

In Step 6 we see that {B1,B2,B6*,B7*,B2*} never will occur since B2 and B2* cannot occur at the same time, hence this cut set is removed, and we remain with:

CSs = { {B1,B2,B5*}, {B1,B3,B5*}, {B1,B3,B6*,B7*,B2*},{B1,B4,B5*}, {B1,B4,B6*,B7*,B2*} }

Note that the minimal cut sets only cover the situation where the initiating event occurs. Finding the dual fault tree to the left most fault tree could also find minimal cut sets for “nothing”, but this is usually not of interest, and hence omitted.

# Appendix A

# Fault tree analysis

## A.1 Introduction

A fault tree is a logic diagram that displays the relationships between a potential critical event (accident) in a system and the reasons for this event. The reasons may be environmental conditions, human errors, normal events (events which are expected to occur during the life span of the system) and specific component failures. A properly constructed fault tree provides a good illustration of the various combinations of failures and other events which can lead to a specified critical event. The fault tree is easy to explain to engineers without prior experience of fault tree analysis.

An advantage with a fault tree analysis is that the analyst is forced to understand the failure possibilities of the system, to a detailed level. A lot of system weaknesses may thus be revealed and corrected during the fault tree construction.

A fault tree is a *static* picture of the combinations of failures and events which can cause the TOP event to occur. Fault tree analysis is thus not a suitable technique for analysing dynamic systems, like switching systems, phased mission systems and systems subject to complex maintenance strategies.

A fault tree analysis may be qualitative, quantitative or both, depending on the objectives of the analysis. Possible results from the analysis may e.g. be:

1. A listing of the possible combinations of environmental factors, human errors, normal events and component failures that can result in a critical event in the system.

2. The probability that the critical event will occur during a specified time interval.

Figure A.1 shows an example fault tree for the bike.

The analysis of a system by the fault tree technique is normally carried out in five steps:

1. Definition of the problem and the boundary conditions.

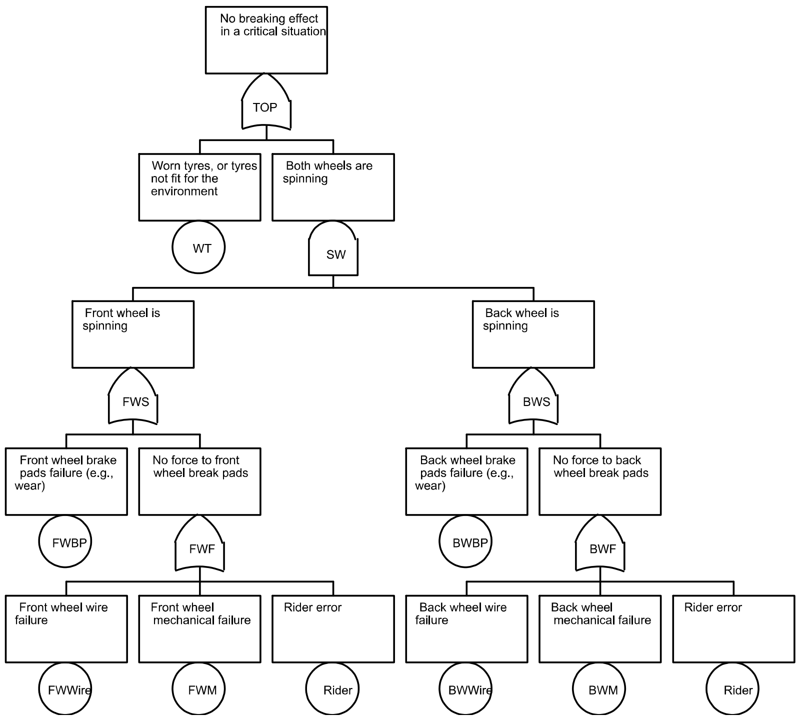


Figure A.1: FTA example for a bike

2. Construction of the fault tree.
3. Identification of minimal cut and/or path sets.
4. Qualitative analysis of the fault tree.
5. Quantitative analysis of the fault tree.

In the following we will present the basic elements of standard fault tree analysis. Then we will conclude this chapter by presenting a numerical example illustrating how the technique could be utilised in relation to maintenance optimisation.

## A.2 Fault tree construction

### A.2.1 Fault tree diagram, symbols and logic

A fault tree is a logic diagram that displays the connections between a potential system failure (TOP event) and the reasons for this event. The reasons (Basic events) may be environmental conditions, human errors, normal events and component failures. The graphical symbols used to illustrate these connections are called “logic gates”. The output from a logic gate is determined by the input events.

The graphical layout of the fault tree symbols are dependent on what standard we choose to follow.

### A.2.2 Definition of the Problem and the Boundary Conditions

This activity consists of:

1. Definition of the critical event (the accident) to be analysed.

2. Definition of the boundary conditions for the analysis.

The critical event (accident) to be analysed is normally called the TOP event. It is very important that the TOP event is given a clear and unambiguous definition. If not, the analysis will often be of limited value. As an example, the event description "Fire in the plant" is far too general and vague. The description of the TOP event should always answer the questions: **What, where** and **when**?

**What:** Describes what type of critical event (accident) is occurring, e.g., collision between two trains.

**Where:** Describes where the critical event occurs, e.g., on a single track section.

**When:** Describes when the critical event occurs, e.g., during normal operation.

A more precise TOP event description is thus: "Collision between two trains on a single track section during normal operation".

1. To get a consistent analysis, it is important that the *boundary conditions* for the analysis are carefully defined. By boundary conditions we mean: **The physical boundaries of the system**. What parts of the system are to be included in the analysis, and what parts are not?

2. **The initial conditions**. What is the operational state of the system when the TOP event is occurring? Is the system running on full/reduced capacity? Which valves are open/closed, which pumps are functioning etc.?

3. **Boundary conditions with respect to external stresses**. What type of external stresses should be included in the analysis? By external stresses we here mean stresses from war, sabotage, earthquake, lightning etc.

4. **The level of resolution**. How far down in detail should we go to identify potential reasons for a failed state? Should we as an example be satisfied when we have identified the reason to be a "valve failure", or should we break it further down to failures in the valve housing, valve stem, actuator etc.? When determining the required level of resolution, we should remember that the detail in the fault tree should be comparable to the detail of the information available

### A.2.3 Construction of the Fault Tree

The fault tree construction always starts with the TOP event. We must thereafter carefully try to identify all fault events which are the immediate, necessary and sufficient causes that result in the TOP event. These causes are connected to the TOP event via a logic gate. It is important that the first level of causes under the TOP event is developed in a structured way. This first level is often referred to as the TOP structure of the fault tree. The TOP structure causes are often taken to be failures in the prime modules of the system, or in the prime functions of the system. We then proceed, level by level, until all fault events have been developed to the required level of resolution. The analysis is in other words deductive and is carried out by repeated asking "What are the reasons for...?"

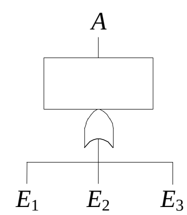


Figure A.2: OR-gate

Figure A.2 shows the OR-gate indicating that the output event $A$ occurs if any of the input events $E_i$ occurs. In relation to the bike example with TOP event "No breaking effects" the two events: "No friction" and "both wheels spinning" are connected by an OR gate since any of these events will lead to the TOP event.

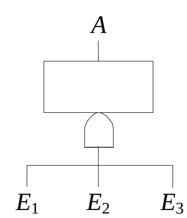


Figure A.3: AND-gate

Figure A.3 shows the AND-gate indicating that the output event $A$ occurs only when all the input events $E_i$ occurs simultaneously. In the bike example, "Front wheel is spinning" and "Rear wheel is spinning" are connected by an AND gate, since both these event have to occur in order to full fill the requirement that both wheels are spinning.

Figure A.3 shows the Basic event representing a basic equipment fault or failure that requires no further development into more basic faults or failures. An example of a basic event in the bike example is "Breakage in break wire".

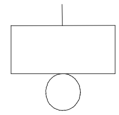

Figure A.4: BASIC-event

## A.3 Identification of Minimal Cut- and Path Sets

A fault tree provides valuable information about possible combinations of fault events which can result in a critical failure (TOP event) of the system. Such a combination of fault events is called a cut set.

*A**cut set** in a fault tree is a set of Basic events whose (simultaneous) occurrence ensures that the TOP event occurs. A cut set is said to be **minimal** if the set cannot be reduced without loosing its status as a cut set.*

*A**path set** in a fault tree is a set of Basic events whose non-occurrence (simultaneously) ensures that the TOP event does not occur. A path set is said to be **minimal** if the set cannot be reduced without loosing its status as a path set.*

In practice only minimal cut sets are used for evaluation of fault trees. To find the minimal cut sets we apply the MOCUS algorithm (Method Of obtaining Cut Sets). The MOCUS algorithm essentially contains the following elements:

1. Start with the TOP event
2. As the algorithm proceeds, the result is stored in a matrix like format of rows and columns
3. AND- and OR-gates are resolved by replacing the gate with it's "children" in the fault tree diagram
4. An AND-gate means that the gate is replaced by new elements for the row(s) it is found
5. An OR-gate means that the gate is replaced by as many rows that the gate has children, where each child is inserted at the position of the OR-gate being replaced
6. When all gates are replaced, we remain with only the basic events, where each row corresponds to a cut set

Note that the the cut sets will not necessarily be be minimal. To make the cut sets minimal we have to:

1. Replace duplicates of one event with only one occurrence of that event in each row
2. If one row is a sub set of another row, then the larger of these two rows (representing non-minimal cut sets) is removed

The MOCUS algorithm is demonstrated here: http://folk.ntnu.no/jvatn/eLearning/TPK4120/Examples/MOCUS.html in relation to the example used in the lectures.

### A.3.1 $k$oo$n$ gate

The $k$oo$n$ gate is something "between" the AND and OR gate. A $k$oo$n$ gate occurs if $k$ out of the $n$ inputs occur. Note that in FTA we focus on fault states, i.e., an event occurring means a failure, hence the "voting" in FTA is different from in RBD. To clarify, the following notation is often used:

- $k$oo$n$ : $G$ is used if we consider the functioning of components (G=Good). The system (block) functions if $k$ or more out of the $n$ components are functioning
- $k$oo$n$ : $F$ is used if we consider the fault of components (F=Fault state). The system (gate/-TOP event) occurs if $k$ or more out of the $n$ inputs are occurring (i.e., in a fault state)

Note the following relation:

$$k\text{oo}n : G = (n–k+1)\text{oo}n : F$$
$$k\text{oo}n : F = (n–k+1)\text{oo}n : G$$

Consider a system with three pumps each having 50% capacity. The system functions if at least 2 of the pumps are functioning. In an RBD we then use the 2oo3 : $G$ block for this system, and for the FTA we use the $k$oo$n$ : $F = n–k+1$oo$n$ : $G = 3-2+1$oo3 = 2oo3 gate.

If we have 4 such pumps, the RBD representation is 2oo4 : $G$, and in FTA we use the $k$oo$n$ : $F = n–k+1$oo$n$ : $G = 4-2+1$oo4 = 3oo4 gate meaning that 3 or more pumps must be in a fault state in order to give a system failure (TOP event).

Computerized FTA programs will offer the $k$oo$n$ : $F$ as part of the drawing palette. For manual construction of a fault tree with a $k$oo$n$ : $F$ gate we can use an OR-gate followed by several AND-gates. Each AND-gate is then a sub-set with $k$ out of the $n$ inputs. There are altogether $\binom{n}{k}$ ways we may choose $k$ inputs out of $n$ inputs, hence we will have $\binom{n}{k}$ AND-gates to put under the OR-gate.

## A.4 Qualitative Evaluation of the Fault Tree

A qualitative evaluation of the fault tree may be carried out on the basis of the minimal cut sets. The importance of a cut set depends obviously on the number of Basic events in the cut set. The number of different Basic events in a minimal cut set is called the *order* of the cut set. A cut set of order one is usually more critical than a cut set of order two, or higher. When we have a cut set with only one Basic event, the TOP event will occur as soon as this Basic event occurs. When a cut set has two Basic events, both of these have to occur at the same time to cause the TOP event to occur.

Another important factor is the type of Basic events in a minimal cut set. We may rank the criticality of the various cut sets according to the following ranking of the Basic events:

1. Human error
2. Failure of active equipment
3. Failure of passive equipment

The ranking is based on the assumption that human errors occur more frequently than active equipment failures, and that active equipment is more failure-prone than passive equipment (an active or running pump is for example more exposed to failures than a passive standby pump).

## A.5 Quantitative analysis

In the quantitative part of a fault tree analysis the main objective is to calculate the following metrics:

- $Q_0(t)$ = Probability that the TOP-event occurs at time $t$
- $F_0(t)$ = Expected number of TOP-event occurrence per unit time at time $t$
- $I(i \mid t)$ = Importance metric for basic event $i$ at time $t$

For the calculations we need the minimal cut set as well as basic event frequencies and probabilities.

### A.5.1 Upper Bound Approximation, $Q_0(t)$

Assume that we have found the minimal cut sets of the fault tree, i.e., $K_j$. Further assume that the minimal cut sets do not contain common components, hence they are independent (also provided that the components are independent). We may now arrange the cut set in a series structure as indicated in Figure A.5: Let $E_j$ denote the event that cut set number $j$ is occurring. The probability that cut set number $j$ is occurring is found by:

$$\Pr(E_j) = \check{Q}_j(t) = \prod_{i \in K_j} q_i(t)$$

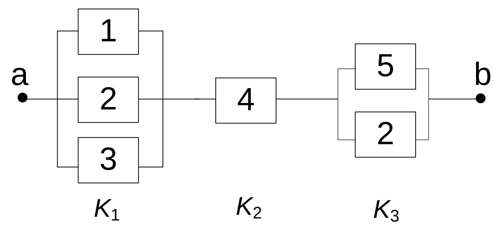


Figure A.5: Example cut set structure

We now have

$$\begin{aligned} Q_0(t) &= \Pr(\text{TOP event occurs at time } t) = 1 - \Pr(\text{TOP event does not occur at time } t) \\ &= 1 - \Pr(\text{No cut set occurs at time } t) \end{aligned}$$

Since the cut sets are independent, and the probability that cut set number $j$ is occurring is given by $\check{Q}_j(t)$, we have:

$$Q_0(t) = 1 - \prod_{j=1}^{k}(1 - \check{Q}_j(t))$$

where

$$\check{Q}_j(t) = \prod_{i \in K_j} q_i(t)$$

Generally there might be some basic events that occur in two or more cut sets, hence the cut sets are *dependent*, and it may be proven that the formula represents an upper bound for the TOP event probability:

$$Q_0(t) \leq 1 - \prod_{j=1}^{k}(1 - \check{Q}_j(t))$$

Hence, we may use:

$$Q_0(t) \approx 1 - \prod_{j=1}^{k}(1 - \check{Q}_j(t))$$

which is referred to as the upper bound approximation and is usually considered to be a good approximation when the $q_i(t)$s are small.

To argue for the less or equal sign we realize that cut sets are “positive dependent” if they have common components. For two cut sets we have

$$\Pr(E_1^C \cap E_2^C) = \Pr(E_1^C | E_2^C)\Pr(E_2^C) > \Pr(E_1^C)\Pr(E_2^C)$$

and

$$Q_0 = 1 - \Pr(E_1^C \cap E_2^C) < 1 - \Pr(E_1^C)\Pr(E_2^C) = 1 - (1 - \check{Q}_1)(1 - \check{Q}_2)$$

and we may give similar arguments for more two or more cut sets.

### A.5.2 The Inclusion-Exclusion Principle, $Q_0(t)$

Referring to Figure A.5 it is also obvious that we may write:

$$Q_0(t) = \Pr(\cup_j E_j)$$

A challenge here is to find the probability of the union of events. For two events $A$ and $B$ we have $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$. For more than two events (cut sets) this becomes more complicated, and we have to use the general addition theorem in probability:

$$Q_0(t) = \Pr(\cup_j E_j) = \sum_j \Pr(E_j) - \sum_{i<j} \Pr(E_i \cap E_j) + \sum_{i<j<k} \Pr(E_i \cap E_j \cap E_k) - \ldots$$

To find $\Pr(E_i \cap E_j)$, $\Pr(E_i \cap E_j \cap E_k)$ is straight forward since these intersections of events are in fact intersection of a set of basic events, and we may multiply the corresponding probabilities as we have done for a single minimal cut set. The challenge is the number of terms we have to calculate. As a starting point we can only take the first sum, i.e., adding the cut set occurrences for each cut set. A slightly better approach would be to subtract the next sum. There are some ways we can optimize the calculations, and finding bounds for the answer to use as a stopping rule, see the textbook. Very often the inclusion-exclusion principle is used by only adding the cut set probabilities:

$$Q_0(t) \approx \sum_{j=1}^{k} \check{Q}_j(t) \tag{A.1}$$

which is faster than the upper bound approximation, but less accurate.

The next challenge is to find the basic event probabilities, $q_i(t)$. Three situations are often considered:

### A.5.3 Non-repairable components

If a component cannot be repaired, the probability that it is in a fault state at time $t$ equals $1 - R(t)$, and provided that the component has an exponentially distributed life time, we therefore have:

$$q_i(t) = 1 - e^{-\lambda_i t} \tag{A.2}$$

where $\lambda_i$ is the constant failure rate of the component.

### A.5.4 Repairable components

To derive $q_i(t)$ for a repairable components we may use Markov analysis. The probability that the component is in a fault state at time $t$ is then shown to be (according to eq. 8.22):

$$q_i(t) = \frac{\lambda_i}{\mu_i + \lambda_i}\left(1 - e^{-(\lambda_i+\mu_i)t}\right) \tag{A.3}$$

where $\lambda_i$ is the constant failure rate of the component, and $\mu_i = 1/\mathrm{MDT}_i$ is the constant repair rate. When $t$ is large compared to $\frac{1}{\lambda_i+\mu_i}$ we have

$$q_i(t) \approx \frac{\lambda_i}{\mu_i + \lambda_i} \approx \lambda_i \mathrm{MDT}_i \tag{A.4}$$

if repair times are short compared to failure times. If this holds, it is safe to use this approximation when $t > 3\mathrm{MDT}_i$, where $\mathrm{MDT}_i$ is the mean time to restoration for the component.

### A.5.5 Periodically tested components

For components with a hidden function, it is usual to perform a functional test at fixed time intervals, say $\tau_i$, to verify that the component is able to carry out it's function. Imay be shown that the (on demand) failure probability of such a component is given by:

$$q_i(t) \approx \lambda_i \tau_i / 2 \tag{A.5}$$

$q_i$ is often referred to as the probability of failure on demand (PFD).

### A.5.6 TOP event frequency, $F_0(t)$

$F_0(t)$ denotes the expected number of occurrences of the TOP event per unit time. In principle we may calculate $F_0(t)$ at various point of times, but usually we focus on the steady state situation, and therefore we omit the time dependency, i.e., we seek $F_0$.

The arguments are as follows:

- We know the minimal cut sets
- If one cut set should be the "contributor" to the TOP event to occur, the other cut sets cannot be occurring
- For a basic event in one cut set to bring the cut set to occur, requires that all other basic events in that cut set are occurring

Let $C_{\mathscr{K}}$ denote a minimal cut set, then the cut set occurrence frequency is given by:

$$\check{w}_{\mathscr{K}} = \sum_{i \in C_{\mathscr{K}}} w_i \prod_{\ell \in C_{\mathscr{K}}, \ell \neq i} q_\ell \tag{A.6}$$

where $w_i$ is the ROCOF of basic event $i$, and $q_l$ is the probability that basic event $l$ is occurring.

The ROCOF is the rate of occurrence of failures. To define the ROCOF we need to have a stochastic process perspective, i.e., we consider what is happening in a time interval rather when things are happening in this interval. Let $N(t)$ be the number of failures that occur in $(0, t]$ and let $W(t) = \mathrm{E}[N(t)]$. The ROCOF at time $t$ is now defined by

$$w(t) = \lim_{\Delta t \to 0} \frac{\mathrm{E}[N(t+\Delta t) - N(t)]}{\Delta t} = \lim_{\Delta t \to 0} \frac{W(t+\Delta t) - W(t)}{\Delta t} = \frac{d}{dt} W(t) \tag{A.7}$$

To obtain the TOP event frequency we may now sum over the $\check{w}_{\mathscr{K}}$'s. However, note that $\check{w}_{\mathscr{K}}$ will not contribute to the TOP event frequency if one of the other cut set is already in a fault state, hence the TOP event frequency is better approximated by:

$$F_0 = w_{\mathrm{TOP}} \approx \sum_{\mathscr{K}=1}^{k} \check{w}_{\mathscr{K}} \prod_{j=1, j\neq \mathscr{K}}^{k} (1 - \check{Q}_j) \approx \sum_{\mathscr{K}=1}^{k} \check{w}_{\mathscr{K}} \frac{1 - Q_0}{1 - \check{Q}_{\mathscr{K}}} \tag{A.8}$$

The formula in Equation (A.8) is the best we can do, but usually $\check{Q}_j$ is rather small, and it will be sufficient to use

$$F_0 \approx \sum_{\mathscr{K}=1}^{k} \check{w}_{\mathscr{K}} = \sum_{\mathscr{K}=1}^{k} \sum_{i \in C_{\mathscr{K}}} w_i \prod_{\ell \in C_{\mathscr{K}}, \ell \neq i} q_\ell \tag{A.9}$$

The ROCOF of the basic events is usually found by the failure rate, say $\lambda_i$. However, a more exact calculation will also take into account the downtime on basic event level, i.e., we may use:

$$w_i = \lambda_i (1 - q_i) \approx \lambda_i \tag{A.10}$$

## A.6 Reliability Importance Metrics

In the literature very many reliability importance metrics are presented. We only focus on the following:

- Birnbaum's metric
- Improvement Potential
- The criticality importance metric
- Fussel-Vesley's metric

In principle a metric is linked to basic events. Very often these basic events are component failures, hence the term component importance is often used. There are many reasons to investigate component importance:

- Considering improving the inherent reliability of critical components
- Establish a preventive maintenance program for the most critical components
- Ensure that we have sufficient spare parts for critical components
- Considering implementing (extra) redundancy at component level for the most critical components
- Given that we have a system failure, which component is the most likely to have caused this?

Several measures are discussed, and the various measures will have their strength and weakness to answer the questions above.

### A.6.1 Birnbaum's Metric of Reliability Importance

Birnbaum's metric of reliability importance of a component is a sensitivity measure expressing the change in system reliability if component $i$ is slightly changed, i.e.,;

$$I^{\mathrm{B}}(i \mid t) = \frac{\partial Q_0(t)}{\partial q_i(t)} \tag{A.11}$$

It follows that a small change $\Delta p_i(t)$ in the component reliability will result in the following change in system reliability:

$$\Delta Q_0(t) = I^{\mathrm{B}}(i \mid t) \Delta q_i(t) \tag{A.12}$$

A disadvantage with Birnbaum's metric is that it is difficult to calculate. If we are able to write down the system reliability function, it should be rather easy to find Birnbaum's measure. But in practice we will not be able to write down the TOP event probability, and hence we cannot derive Birnbaum's metric. In some cases we may utilize that:

$$I^{\mathrm{B}}(i \mid t) = Q_0(t \mid q_i = 1) - Q_0(t \mid q_i = 0)$$

It may be shown that $I^{\mathrm{B}}(i \mid t)$ is the probability that component $i$ is critical at time $t$. This is a valuable result used in maintenance optimization. Often we need to calculate the expected cost of a failure of a specific component. The contribution to downtime depends on whether the system is down or not, and if a failure will cause a system failure. The Birnbaum's metric is exactly what we need, i.e., we should only include downtime cost if the component under consideration is critical, and $I^{\mathrm{B}}(i \mid t)$ is then used for calculating this probability.

### A.6.2 Improvement Potential

The Iimprovement Potential states how much the system reliability will increase if component $i$ is replaced with a perfect component:

$$I^{\mathrm{IP}}(i \mid t) = Q_0(t) - Q_0(t \mid q_i = 0) \tag{A.13}$$

It is easy to show the following relation to Birnbaum's metric:

$$I^{\mathrm{IP}}(i \mid t) = I^{\mathrm{B}}(i \mid t) q_i(t) \tag{A.14}$$

### A.6.3 Criticality Importance

The criticality importance metric $I^{\mathrm{CR}}(i \mid t)$ of component $i$ at time $t$ is the probability that component $i$ is critical for the system and is failed at time $t$, when we know that the system is failed at time $t$. It is easy to show the following relation to Birnbaum's metric:

$$I^{\mathrm{CR}}(i \mid t) = \frac{I^{\mathrm{B}}(i \mid t) \cdot q_i(t)}{Q_0(t)}$$

### Fussell-Vesely's Metric

The Fussell-Vesely's importance metric $I^{\mathrm{FV}}(i \mid t)$ of component $i$ at time $t$ is the probability that at least one minimal cut set that contains component $i$ is failed at time $t$, when we know that the system is failed at time $t$.

In order to calculate $I^{\mathrm{VF}}(i \mid t)$ we need some reasoning. We simplify and skip the index $t$. Now introduce the following notation (we use the terminology "component" whereas the precise word would be "basic event"):

- $D_i$: At least one minimal cut containing component $i$ is failed
- $C$: The system is failed
- $m_i$: Number of minimal cut set containing component $i$
- $E^i_j$: Minimal cut set $j$ containing component $i$ is failed

From the definition we have:

$$I^{\mathrm{FV}}(i) = \Pr(D_i \mid C) = \frac{\Pr(D_i \cap C)}{\Pr(C)} \tag{A.15}$$

Since $D_i$ is a subset of $C$, then $D_i \cap C = D_i$ and we have:

$$I^{\mathrm{FV}}(i) = \frac{\Pr(D_i)}{\Pr(C)} \tag{A.16}$$

To find $\Pr(D_i)$ we use the same approach as for the "upper bound' approximation for $Q_0$. However, note that $D_i = E^i_1 \cup E^i_2 \cup \cdots \cup E^i_{m_i}$ where the union is only taken over minimal cut sets containing component $i$. This gives:

$$\Pr(D_i) = 1 - \Pr({E^i_1}^C \cap {E^i_2}^C \cap \cdots \cap {E^i_{m_i}}^C) \le 1 - \Pr({E^i_1}^C)\Pr({E^i_2}^C)\cdots\Pr({E^i_{m_i}}^C)$$

$\Pr({E^i_j}^C)$ is then obtained by one minus the probability for the event that minimal cut set $j$ is failed, i.e., $\Pr({E^i_j}^C) = 1 - \check{Q}^i_j = 1 - \prod_{l \in K_j} q_l$. The following approximation is usually sufficient to calculate Fussell-Vesely's measure:

$$I^{\mathrm{FV}}(i) \approx \frac{1 - \prod_j^{m_i}(1 - \check{Q}^i_j)}{Q_0}$$

where the product is over minimal cut sets which contain component $i$.

If cut set failure probabilities are small, a faster approximation is given by:

$$I^{\mathrm{FV}}(i) \approx \frac{\sum_j^{m_i} \check{Q}^i_j}{Q_0} \tag{A.17}$$

where the sum is over minimal cut sets which contain component $i$.

By comparing the definition of $I^{\mathrm{CR}}(i)$ and $I^{\mathrm{FV}}(i)$, we see that these measures are rahter close to each other. Thus by assuming $I^{\mathrm{CR}}(i) \approx I^{\mathrm{FV}}(i)$, we could easily get an approximation of Birnbaum's measure from:

$$I^{\mathrm{B}}(i) = \frac{I^{\mathrm{CR}}(i) \cdot Q_0}{q_i} \approx \frac{I^{\mathrm{VF}}(i) \cdot Q_0}{q_i}$$

### A.6.4 System failure frequency obtained by $I^{\mathrm{B}}(i)$

An alternative way to calculate system failure frequency, $F_0$, is to start with Birnbaum's measure. First we recall that $I^{\mathrm{B}}(i)$ is the probability that the system is in such a state that component $i$ is critical. That a component is critical means that the system is in such a state that the system is functioning if component $i$ is functioning, and in a fault state if component $i$ is failed. Then it follows that:

$$F_0 = \sum_i I^{\mathrm{B}}(i)(1 - q_i)\lambda_i \tag{A.18}$$

where $p_i = 1 - q_i$ is the probability that component $i$ is functioning, and $\lambda_i$ is the failure rate of component $i$. Thus, the contribution of component $i$ to $F_0$ is given as the product of:

- The probability that component $i$ is critical, i.e., the state of other components
- The probability that component $i$ is functioning
- The failure rate of component $i$

# Appendix B

# Event Tree Analysis (ETA)

## B.1 Introduction

An event tree is a logical diagram which displays possible event sequences following a specified critical event in a system. An event tree analysis (ETA) is a method for systematic analysis of a system after a critical event has occurred. The result of an ETA is a list of possible event sequences that follows the initiating event. The critical, initiating event may be a technical failure or some human error. In the development of the event sequences, the effects of possible barriers and safety functions, which are designed to prevent the occurrence of the critical event or reduce the consequences of this event, are taken into account. The analysis is both qualitative and quantitative. The qualitative content is primarily a visualisation of different scenarios (the event tree) with corresponding end consequences, while the quantitative analysis gives frequencies for the different end consequences. Figure B.1 shows an ETA example. The initial event could be for example SPAD = Signal passed at danger (obtained from for example an FTA), and then the various barriers are shown as $B_1$, $B_2$ etc. Each barrier has a Y=Yes output and a N=No output.

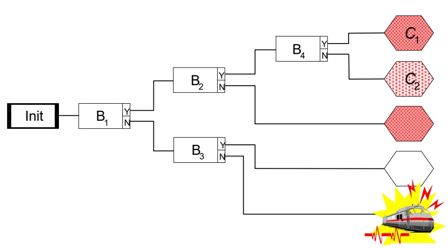


Figure B.1: ETA example

## B.2 Procedure

The event tree analysis is usually carried out in six steps:

1. Identification of a relevant initiating event (which may give rise to unwanted consequences).
2. Identification of the barriers and safety functions which are designed to prevent the occurrence of the initiating event, or to reduce the consequences of this event.
3. Construction of the event tree.
4. Description of the resulting event sequences.
5. Calculation of probabilities/frequencies for the identified consequences.
6. Compilation and presentation of the results from the analysis.

## B.3 Identification of a relevant initiating event

As for the fault tree it is important to define an unambiguous initiating event, use the “what”, “where” and “when” keywords to structure the definition of the initiating event. How early in the course of event the initiating event should be placed depends on the scope and available resources for the analysis. As a starting rule we often define the initiating event as the first significant deviation from normal operations.

## B.4 Identification of the barriers and safety functions

Usually, a number of measures are taken to prevent accidents or limit their consequences These measures are referred to by different names, e.g. “barriers”, “security functions” or “protection layer” (defence in depth). The measures are modelled in the event tree. Other factors related to the physical course are also modelled, e.g.,

- Whether a leak ignites or not
- Whether the fire is large or small
- Whether it’s day or night
- and so forth

## B.5 Construction of the event tree

The event tree is constructed by thinking logical sequences by answering Yes/No questions. The questions should be formulated systematically. Either one uses consistent questions where the “Yes” answer is "success”, or Then the "Yes” answer is systematic failure or error in barrier/safety function The branches corresponding to "Yes” must either systematically go “upwards”, or systematically go “down” in the event tree.

If we adopt the convention that the “No” branch (“barrier fails to hold”) is the down-hand branch from the barrier symbol. The most severe consequences will then normally be located to bottom right corner of the consequence spectrum. Note that in some presentations “Yes” is used to describe that the barrier fails. This will then give a different interpretation of the most critical events.

If we consider a SPAD event, the first barrier, $B_1$, could be {Automatic train protection (ATP) OK}. When constructing the event tree the output from a barrier symbol may lead to another barrier symbol. The development is continued to the resulting consequences, illustrated by consequence symbols, $C_1$, $C_2$ etc in Figure B.1. We should aim at identifying the barriers in the sequence they are expected to be activated. In this way, there will be an implicit time line from left to right. However, in some situations this is demanding because it is not always easy to say which barriers are activated first.

## B.6 Description of the resulting event sequences

The qualitative analysis of the event tree is typically to list the events leading up to the most severe end consequences, and discuss barriers and other circumstances that influence the course of events.

## B.7 Calculation of probabilities/frequencies for the identified end consequences

In order to carry out the quantitative analysis we need the frequency of the initiating event, and the barrier probabilities. During construction of the event tree, we enter the probability that the various barriers fails, i.e., the “No” results. For each barrier, $i$, we need:

- $q_i$ = probability that barrier i fails (“No”), and similarly
- $p_i = 1 - q_i$ probability that barrier i functions as intended (“Yes”)

In addition to the barrier probabilities, we enter the frequency of the initiating event:

- $f$ = frequency of initiating event

When establishing the barrier probabilities and the initiating frequency it might be required to perform separate analyses, e.g., FTA. Also for the barrier probabilities we usually need separate analyses like FTA for the ATP system, failure statistics and "load/strength" methods.

To calculate the frequencies of the various consequences we may multiply the frequency of the initiating event by the barrier probabilities for each barrier along the path leading to the actual consequence . Now, consider consequence $C_j$, and assume that $S$ = is the set of barriers in the path leading to consequence $C_j$, and that represents "success" of the barrier (Yes-terminal), and further $F$ = is the set of those barriers on the path leading to consequence $C_j$, and that represent "the barrier fails" (No-terminal) we have that the frequency of consequence $C_j$ is given by:

$$F_j = f \prod_{i \in S} p_i \prod_{i \in F} q_i$$

This formula is only valid if the barriers are "independent". This is not always the case, and to overcome the problem of "stochastic" dependent barriers, we should in principle specify the barrier probabilities as conditional probabilities given the course of events up to the current barrier. This is not always easy.