UiO : Department of Informatics
University of Oslo

Master Thesis

Risk Assessment Based on CORAS and Fuzzy Logic

Tan Hoang Duy Tran , Supervisor: Ketil Stølen

2nd May 2017

Master Thesis

Tan Hoang Duy Tran Supervisor: Ketil Stølen

2nd May 2017

Contents

List of Figures

List of Tables

Abstract

Risk assessment methodology is a research topic which has been focused and applied in many fields and contexts that consider risk is a part of the system. Currently, the CORAS method is an innovative and effective method to address the issue of risk by providing a comprehensive framework. While fuzzy logic is a mathematical tool which has been adopted to address the issue of uncertainty, imprecision that are attached to the risk analysis. Therefore, a systematic approach to combining CORAS and fuzzy logic has been elaborated in order to address the issue better than the current one based on a selective background of both.

The main building blocks of the approach are the two CORAS rules for reasoning about likelihood that are refined by fuzzy logic. As a result, conditional likelihood and relation scale are defined to facilitate the two fuzzy methods for estimating likelihood. Experimentation has been conducted to test and verify the proposed approach. The outcome of the experimentation gives us an insight into different patterns of results generated by the two fuzzy methods and the two types of membership function.

The tool-supported the approach has been developed in order to simplify fuzzy logic operations that are demanded to adopt the proposed approach. Furthermore, the tool can be considered as a part of the proposed approach.

Keywords: *CORAS*, *fuzzy logic*, *membership function*, *risk assessment*, *likelihood*

Acknowledgements

First of all, I would like to thank Sintef and University of Oslo because they are the institutions that give me the opportunity to study in Norway. In addition, all of my knowledge to complete the thesis was improving with the guidance of the lecturers and my friends who work and study in my same place.

Above all, I would like to thank my parents, my aunt and my sister because they always support and encourage me to complete the thesis. Without their encouragement, my thesis would not be completed in terms of time.

Finally, I would like to give special thanks to Ketil Stølen because of all of your kindness, enthusiasm, motivation to guide me throughout the thesis. Your comments, guidance and support are always worthy for me to further study.

Oslo, May 1st 2017
Tan Hoang Duy Tran

Chapter 1

Introduction

Risk is a general term, and it relates to almost all our everyday activities 'The term risk is used in a variety of context and domain' [25, p.3]. Its presence is unavoidable. However, we can predict, prevent and reduce its consequences by applying analysis techniques and rational decision-making methods [21, 25]. Risk assessment methods require sources of precise data, statistical numbers of unexpected events, etc to estimate and obtain accurate results. In addition to that, risk assessment is a complex task which requires many parameters, and many of those are very difficult to quantify [27, 34]. Therefore, a systematic approach is demanded in order to address the issue.

In this thesis, we proposed an approach which brings CORAS and fuzzy logic closer. While CORAS is a general method for risk analysis, fuzzy logic provides a mathematical tool to model uncertainty, vagueness and imprecision. Our approach therefore inherites the virtues of CORAS to address risk, and the power of fuzzy logic to solve the issue of uncertainty in the risk assessment process.

The approach is an attempt to extend the CORAS by applying fuzzy logic as motivated by Solhaug and Stølen [38]. The approach refines probabilistic and frequency measurements by fuzzy numbers. Additionally, we developed a computer-supported tool which assists risk analysts and stakeholders in using our approach in their situations. Despite our approach have not been evaluated by experts, or applied in practice, we believe that all knowledge which is developed and built from a scientific and systematic approach is worthy. To sum up, this thesis presents an approach to combining CORAS and fuzzy logic.

1.1 Motivation

Risk is considered as something which affects the values we would like to protect [33, p.9]. To address risk, a systematical methodology need to be studied and developed in order to model, predict, estimate as well as provide solutions or making decisions based on those risks. Risk analysis and assessment techniques are diverse, and can be categorized by quantitative, qualitative and hybrid methods [27].

In order to address the issue of uncertainty, there are two main schools in controversies: the probabilistic methodology and the fuzzy logic methodology [12]. The two methods have been applied in many situations and contexts of risk. However, current interest of dealing with uncertainty, vagueness and imprecision in the domain of risk analysis and decision making is supported by fuzzy logic [24].

Fuzzy logic can provide a framework for human interference of inaccurate data to analyze risk [34, 36]. In addition to that, fuzzy logic can be combined with other risk models such as decision trees, artificial neutral network to model complicated issues [36], and uncertainty in fault tree analysis [41].

1.2 Objective

Dealing with uncertainty and vagueness of data is the most difficult task of almost all problems that need to be analyzed. More on that, the domain of risk assessment requires not only skill, experience, knowledge of risk analyst, but also the precision of data which directly affects the outcome. Obtaining and quantifying data demands time and effort, that is a lengthy process. In order to acquire accurate result from analyzing risk, a systematic methodology is demanded. CORAS is a systematic approach for risk analysis that is proved to be effective and comprehensible for analyzing risk of large scale systems [22, 45]. However, it is still dependent on accurate data in order to estimate and calculate risk level. Therefore, combining fuzzy logic and the CORAS method must be a systematic approach that demands acquiring the understanding of both. In this thesis, we pursue a systematic method that combines fuzzy logic and the CORAS method to address the problem above. In addition to that, a tool-supported the method is put forward to assist risk analysts and stakeholders.

1.3 Contribution

The main contribution of the thesis is the extension of the CORAS method by applying fuzzy logic. Our proposed method employs the advantages of the CORAS method for risk analysis and methodology for dealing with imprecision of fuzzy logic to develop an extended version of CORAS which solves the issue of risk analysis that is better than that of the current one.

In summary, the following artifacts have been achieved with respect to success criteria which will be presented later.

- A systematic approach which combines CORAS and fuzzy logic.
- A tool which supports the approach.

1.3.1 Approach to Combining CORAS and Fuzzy Logic

The proposed approach aims to extend the CORAS method by integrating fuzzy logic into two fundamental rules for reasoning about likelihood. Consequently, the approach refines the step of CORAS method by proposing a sub step to construct membership functions. In addition to that, the approach promotes two fuzzy methods for calculation, they are general fuzzy method and simple fuzzy method. While the general fuzzy method refines completely interval scale by applying fuzzy numbers, the simple fuzzy method

With respect to rule for leads-to, the approach refines intervals representing likelihoods by fuzzy numbers. As a consequence, conditional likelihood scale is defined to facilitate fuzzy methods for estimating likelihoods.

With respect to rule for separate, the approach refines CORAS formula and introduces a case of partial separate in the analysis of threat scenarios and unwanted incidents. As a result, the relation scale is define to facilitate the fuzzy methods for calculating.

1.3.2 Tool Supported Method

Applying fuzzy logic operations instead of interval operations strengthens complexity of the risk estimation process. Therefore, the tool supported method has been developed in order to simplify the process of calculation. The current version of the tool supports the rule for leads-to and rule for separate with respect to the approach.

The tool offers three alternatives to estimate likelihood as proposed by the approach with respect to the rule for leads-to, that are general fuzzy method, simple fuzzy method and interval method. While general fuzzy method applies fuzzy numbers to both likelihood and conditional likelihood, simple fuzzy method maintains an interval scale and a fuzzy scale, and interval method is pure CORAS method to reason likelihood.

With respect to rule for separate, general fuzzy method proposes fuzzy scale for likelihoods and their relation, simple fuzzy method remains fuzzy relation scale and apply interval calculation to likelihoods, and CORAS method to calculate rule for separate is the same as the method of rule for leads-to, but it estimates likelihoods in case of completely separate and completely overlap.

1.4 Structure of the Thesis

- **Chapter 1 Introduction**: Chapter 1 introduces requirements for combining CORAS and fuzzy logic. In addition, it presents objectives and contribution of the thesis.

- **Chapter 2 Characterization of needs**: Chapter 2 introduces the success criteria with respect to the artifacts.

- **Chapter 3 Theoretical Background**: Chapter 3 presents some background of CORAS, fuzzy logic and state-of-art which relate to the thesis.

- **Chapter 4 Research Method**: Chapter 4 presents our research method which applies to this thesis.

- **Chapter 5 Approach to Combining Fuzzy Logic with CORAS Method**: Chapter 5 presents our approach which combines CORAS and fuzzy logic step-by-step.
- **Chapter 6 Experimentation**: Chapter 7 presents the steps to conduct experiments.
- **Chapter 7 Implementation of The Tool**: Chapter 6 presents the development of the tool which supports the approach
- **Chapter 8 Evaluation of The Tool**: Chapter 8 presents the evaluation of tool with respect to success criteria.
- **Chapter 9 Discussion**: Chapter 9 discusses issues of our approach with respect to success criteria that are partly addressed and not addressed yet in this thesis.
- **Chapter 10 Conclusion**: Chapter 10 concludes our work and propose further work.

Chapter 2

Characterization of needs

As discussed above, our focus is on risk assessment by employing CORAS and fuzzy logic. To achieve the goals, the theoretical background must be understood in detail, and after that an approach for risk assessment based on fuzzy logic and CORAS will be developed. With respect to the approach, a computer-supported tool for the method should be developed. The purpose of the tool is not only supporting modeling or documenting risk analysis. It should also support risk analysts (or relevant stakeholders) in predicting (or deciding) which threats, risks should be eliminated (based on their likelihoods and consequence), and what solutions for risk (advantages and disadvantages) are preferable.

In summary, the following artifacts must be achieved in conjunction with success criteria which will be presented in the next section.

- A method which combine CORAS method and fuzzy logic.

- A tool that support risk analysis based on the method.

2.1 Stakeholders

This section presents the stakeholders involved and their role in the CORAS risk analysis. The stakeholders include members of risk analysis team and the parties.

The risk analysis team includes one analysis leader, one analysis secretary and analysis members.The analysis leader is responsible for leading the risk analysis tasks and guiding the participants, while the analysis secretary is responsible for documenting and supporting the analysis leader. If the scale of risk analysis is small, then an analysis leader and secretary is sufficient. However, it is better to have additional analysis members such as expert domains, decision makers, evaluators, etc .

The party is usually the customer who hires the risk analysis team to conduct risk analysis. There may be possible to have more than one party in a risk analysis, for instance, shareholders of an organization or a company.

2.2 Theoretical approach to combining CORAS method and fuzzy logic

CORAS is a general method for risk analysis. Therefore, it is sufficient for analyzing risks in almost cases. However, the calculation and reasoning likelihoods in some cases may be complex and difficult due to uncertainty in risk analysis [11]. For that reason, we need to develop a theoretical approach that applies fuzzy logic to support calculation and reasoning of likelihoods in such cases. Furthermore, based on the results, the approach can support us in analyzing as well as predicting risk and relevant factors. In summary, the following success criteria must be achieved for this artifact:

- The method must be general, therefore it can be used in almost situations as in the case of the CORAS approach.
- The method must be sound.
- The method must be comprehensible and applicable, so that it supports the risk analyst and stakeholders to solve issues of risk analysis.
- The method should be effective in comparison with the CORAS method.

2.2.1 The Generality of The Method

CORAS is a general framework for risk analysis [25], and it is can be applied in almost context of defensive risk analysis. Therefore, when bringing fuzzy logic to CORAS to develop an extended version of this framework, the new method should inherit this property of CORAS. For that reason, the new method can be applied in almost situations as the CORAS.

2.2.2 Soundness

Soundness of a logical system is defined as if and only if its interference rules prove only formulas that are valide with respect to its semantics [47]. Therefore, if the method is not sound, then it is useless.

2.2.3 Comprehensible and Applicable

The method is comprehensible means that risk analysis and stakeholders can learn and understand it. Additionally, the method can be applied in practice.

2.2.4 Effectiveness of The Method

By effectiveness, we mean that the method can solve the issue of risk analysis in a more efficient and precise way than the current one, or it has features to address issues which the existing one does not. In our case, the feature of capturing and addressing imprecision of data in risk analysis must be implemented in the method. Therefore, it becomes more effective than the current one by the feature.

2.3 Tool-Supported Approach

The CORAS tool is an open source software, for that reason, we can employ it and extend its functionality. However, the current tool only supports for modeling risks on-the-fly and lack of features for translating the model to sentences and calculating and reasoning likelihoods. Therefore, in order to implement the framework, the following success criteria should be fulfilled:

- Feature for calculating likelihood must be implemented on the tool.
- The tool must support the theoretical framework completely. This means that all the features of the framework must be implemented.
- The tool should be user-friendly, ease to use, and effective.
- Results calculated by the tool must be precise.

2.3.1 Calculating likelihood based on interval and frequency

The tool must support the feature of CORAS approach to calculate the frequency and interval of likelihood. The feature for calculating likelihood is based on rule for leads-to and rule for separate which will be presented in the next section.

2.3.2 The Tool-Supported Method

With respect to the feature of the tool presented above, the tool must implement the features for calculating likelihood based on the proposed method.

2.3.3 Ease-to-use Tool

The tool must be design in a way that it does not require user much effort to figure out all the functions of the tools. Additionally, the tool must be effective so that the processing and interaction of users are not interrupted.

2.3.4 Sound Results

The tool must provide reliable results when it is compared to the approach. By this we means that, results generated by the tool must be as precise as applying the approach manually.

Chapter 3

Research Method

In this chapter, we first present some background of the research methods which are appropriate to our thesis. After that, a strategy to conduct our research will be presented. Additionally, the process of developing the tool will be employed and presented in this chapter.

3.1 Technology Research Process

Research area is categorized by classical research and technology research. While classical research attempts to seek and obtain the knowledge of the world such as nature, society, human, etc. Technology research focuses on creating new artefacts which are better than the existing ones [17]. In this thesis, we elaborate a new method which brings CORAS and fuzzy logic together. Therefore, our research is technology research, because the new method is actually an improvement of the CORAS method.

Technology research is an iterative process and divided into three sub-processes which are problem analysis, innovation and evaluation [17]. These sub-processes will be presented in the following sections. The figure 3.1 is the process which is adapted from [17] and modified to satisfy our objectives.

Problem analysis In this phase, the researcher tries to identify needs from stakeholders to develop new artefacts [17]. By applying this concept to our research domain, we first study the background of the CORAS method and fuzzy logic in order to know the gap between them. Additionally, we identify a set of research questions or success criteria which satisfy the needs. This process is iterative, and the set of research questions will be refined after each iteration.

Innovation The innovation phase involves developing a new artefact based on the set of research questions from the first step [17]. In this phase, we actually try to bridge the gap based from the first step. The artefact is developed by answering the research questions.

Evaluation Based on the research questions, the researcher formulates predictions about the artefact and checks that whether the artefact addresses

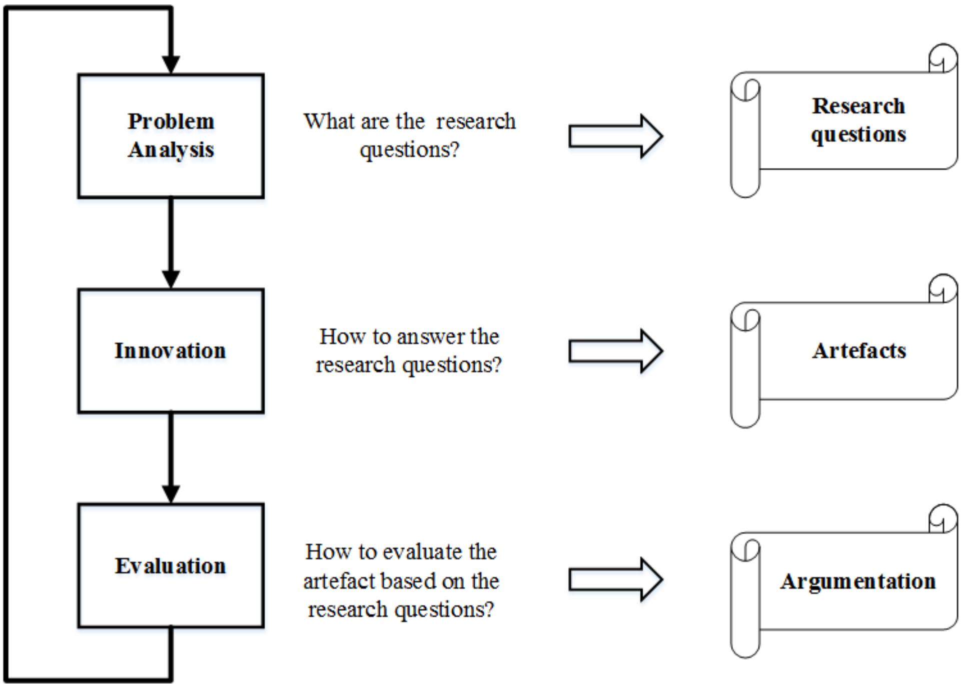

Problem Analysis
What are the research questions?
Research questions
Innovation
How to answer the research questions?
Artefacts
Evaluation
How to evaluate the artefact based on the research questions?
Argumentation

Figure 3.1: The research method process (Adapted and modified from [25])

the need [17]. Actually, research evidence is obtained by the research strategies which will be presented in the following section.

3.2 Technology Development Process

In order to develop the tool supported approach, we adopt the software development methods and disciplines to guarantee that the development of the tool satisfies the plan, objectives and quality.

3.2.1 Rapid Application Development

Rapid application development is an iterative software development method that appeared in response to the weaknesses of the classical software development models [9]. This method incorporates special techniques and computer-aided software engineering tools to fasten the development of software by delivering some portions of the software which is testable after each iteration. User's experience and requirements are incrementally clarified after each iteration, users will have a clear picture of what the current software does offer (functions of the software) and what they really want to have (user's requirements). Therefore, user's feedback is the key factor of the method, and it is an elevator for next iterations. Additionally, computer-aided software engineering tools such as visual programming languages, code generators, third-party libraries, etc are exploited to shorten the development cycle [9].

The figure below presents an adopted version of this method. After the planning and overall analysis phase are conducted, sub-processes include detail analysis, design and implementation phase happens sequentially. Each sub-process delivers a testable software, and the software incrementally matures after each iteration.

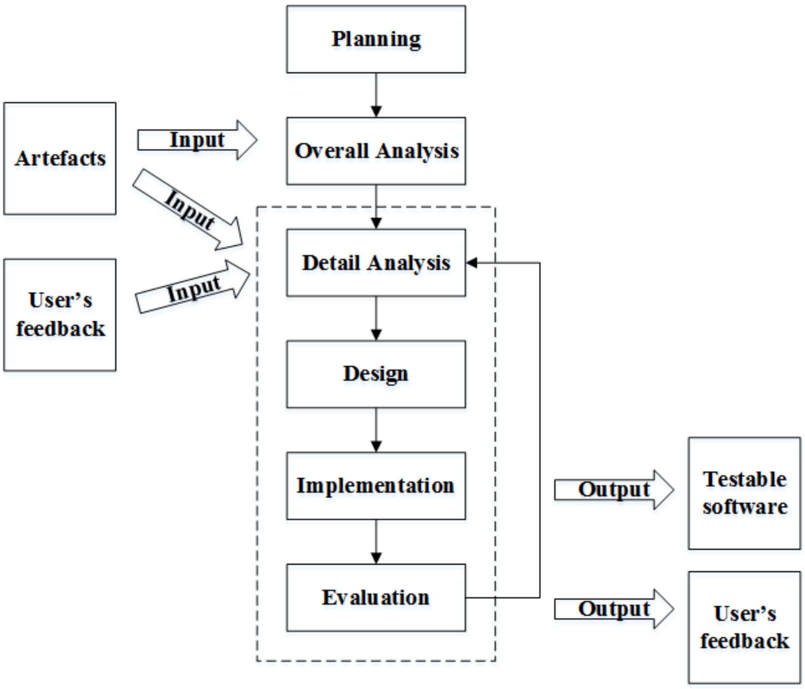


Figure 3.2: The Iterative Software Development Process (Adapted from [9])

- Planning: A plan will be drawn in this phase. It is time plan and necessary tasks.
- Overall Analysis: Requirements and basic functions will be clarified. In addition to that, programming tools, resources, third-party libraries are also selected to support the development of the tool.
- Detail Analysis: Requirements and functions are refined in detail, so that it facilitates the design process.
- Design: The functions and features of the tool will be divided into programmable modules, the components and the relationships will be clarified in this phase.
- Implementation: The modules will be programmed and tested.
- Evaluation: The functions and features of the tool will be test, if they are not verified the reproducing step will be conducted.

This method corresponds to the technology research method presented above in the sense that they are iterative, both of them deliver incremental mature artefacts (software), needs (requirements) are refined with respect to evaluation. Consequently, adopting this method facilitates our research in the

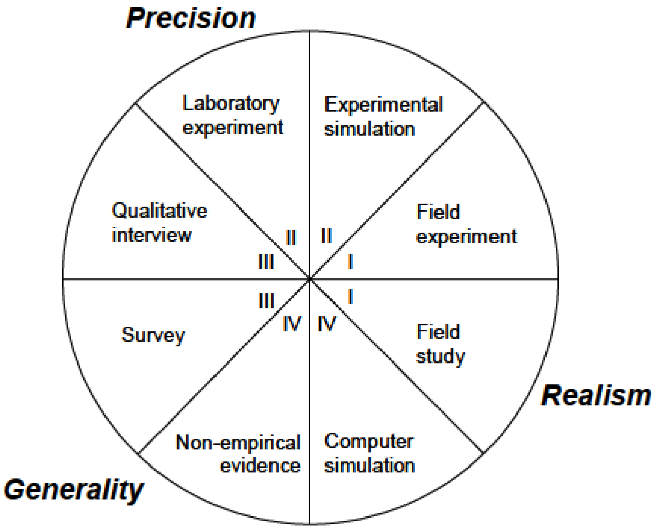


Figure 3.3: The Research Strategies (Adapted from [17])

manner that the artefacts obtained and developed in the technology research process will be input to the analysis phase of the software development process.

3.3 Research Strategies

According to McGrath [30], research strategies have advantages and disadvantages, the selection of the strategies is based on the requirements and the domain of the study, and each strategy has its own methods. There are three factors should be taken into account when obtaining a research evidence, they are generality, precision and realism as described below [30]:

- Generality - The result is valid over populations of actors.
- Precision - The measurement is precise.
- Realism - The result is related to the real situation, or context to which we want it.

In addition to that, there are eight common methods which can be applied in order to obtain the maximum of the three factors as described by the figure 3.3 [17]. In this thesis, we develop a method which combines CORAS and fuzzy logic. Therefore, the method should be general in order to apply in all situations of risk analysis. In addition to that, the method should be sound so that it generates reliable results in the context of risk assessment. As stated by McGrath [30], formal theory (non-empirical evidence [17]) generates maximum of generalizability and keeps much of the precision, but it looses the nature of the context. More on that, laboratory experiment will be applied in order to obtain the maximum of the precision which we want our method can generate reliable results, and we can keep the control of variables and measurement to conduct experimentation.

3.3.1 Formal Theory

Formal theory is the method which employs logical reasoning and mathematics to solve problems [19]. Methods of formal theory and non-empirical research include review existing literature, scholarship, conceptual research, scenario-building, etc [7].The approach to integrating fuzzy logic into CORAS is a systematic approach. Therefore, it requires mathematical foundation, logical reasoning and review of literature to reasonably develop the approach which employs both CORAS and fuzzy logic. In our research, we employ two methods, namely review of literature, and deductive reasoning to develop the approach.

3.3.1.1 Literature Review

In order to develop the approach, a theoretical background of both CORAS and fuzzy logic must be acquired. Fuzzy logic is a mathematical background that is applied to many fields [16, 48]. Therefore, the topics and applications of fuzzy logic are diverse and many of them are not appropriate to our research. Similarly, the CORAS is a general and well-defined framework to conduct risk analysis which includes the process, the rules, the diagrams, etc many aspects that fuzzy logic cannot cover all. Consequently, a selective fuzzy logic background and relevant CORAS aspects should be extracted in order to elaborate the development of the approach.

3.3.1.2 Logical Reasoning

Theoretical background from the review of literature process is the foundation to elicit the approach. However, that is not enough to bring fuzzy logic and CORAS closer. Logical reasoning is adopted to develop a sound approach. There are two basic forms of logical reasoning, they are deductive and inductive reasoning. According to [42], deductive reasoning is a top-down logic that links hypotheses and conclusions by formulating proofs. In other words, if the hypotheses are true, then the conclusions are valid. Our approach are built on a set of hypotheses, we formulate proofs by laws of logical inference to conclude results. While the inductive reasoning is the reverse process of deduction in the sense that hypothesis is constructed from observations [43]. The inductive reasoning is adopted to induce observations from experimentation and construct general patterns of membership functions.

3.3.2 Laboratory Experiment

Laboratory experiment gives us high control of variables and measurements in order to conduct our experiments [17, 30]. Therefore, applying the laboratory experiment in the context of CORAS will facilitate our research in the sense that we can test our approach. However, conducting the whole laboratory experiment process based on the CORAS approach is infeasible in this thesis. Therefore, we assume that the steps of CORAS are reasonably and hypothetically constructed and we come up with a situation where our approach can be applied. In that event, we design, construct and manipulate variables in order to conclude patterns of results.

3.3.3 Tool Evaluation

Tool evaluation is a major step of the development cycle to ensure the quality of the tool, we adopt software testing practices to appraise and verify that the tool conforms the success criteria. Basically, software testing is a set of activities conducted in order to find software errors [32, p.439].

The two fundamental testing methods are white-box and black-box testing. While white-box testing tends to examine internal of software, black-box testing is predicated on external behavior of software [32, p.443-444]. Both of the methods offer a set of techniques. On the other hand, software testing practice defines a set of levels of testing that includes unit testing, integration testing, system testing, and validation testing [32, p.481]. Conducting the whole process is time-consumed and costly. Therefore, we will not conduct the levels of testing instead the black-box method is adopted to facilitate our research. Black-box testing includes a set of techniques such as graph-based technique, equivalence partitioning, boundary value analysis, scenario-based, etc [32, p.460-468]. The method focuses on examining functions of the software by defining a set of predefined input and expected output. The objectives of black-box testing is to verify software functionality and requirements, reveal software bugs by testing valid and invalid inputs. The virtues of black-box testing conform with the success criteria of the tool as we stated.

3.3.3.1 Equivalence partitioning

Attempting to test all cases of a domain can lead to combinatorial explosion. The purpose of this technique is to attempt reducing total of test cases by dividing input data into equivalence classes while test cases still cover all domain input [20, p.99].

Given a program function F that takes two variables x_1, x_2, and $a \leq x_1 \leq b$, $c \leq x_2 \leq d$. The input domain is partitioned into five sub classes as shown in the figure 3.4. There are one valid input domain when $x_1 \in (a, b)$ and $x_2 \in (c, d)$ and four invalid input domains that are when $x_1 < a$, $x_1 > b$, $x_2 < c$ and $x_2 > d$. Therefore, five test cases are conducted when variables x_1 and x_2 are chosen from five classes.

3.3.3.2 Scenario-based Testing

By definition, a scenario is an ordered set of interactions between the system and external actors [35]. The technique concentrates on user's interaction with the system, this means that test case captures user's behavior to uncover interaction errors and incorrect specifications [32, p.641-642]. Test case is a sequence of steps defining interaction of user and the software based on use case specification.

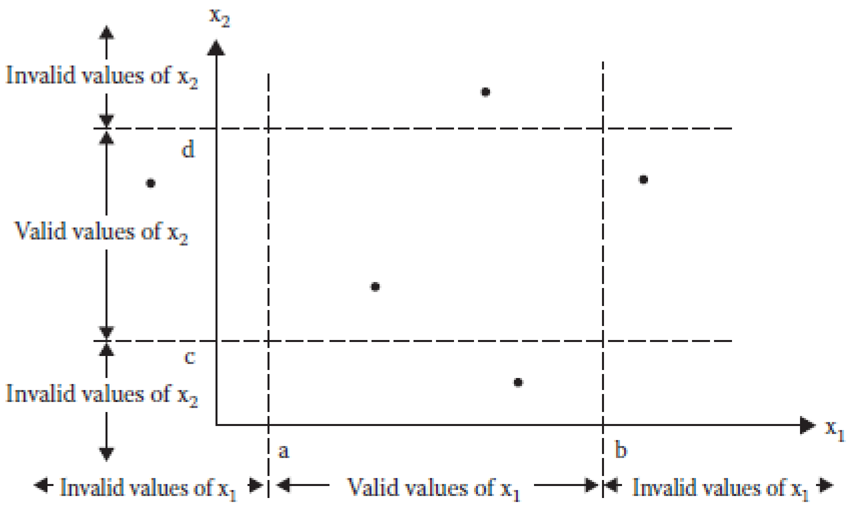


Figure 3.4: Equivalence Testing Technique (Adapted from [20])

Chapter 4

Theoretical Background

In this chapter, we first present some basic terminologies which are related to risk assessment process. Additionally, some background of risk analysis methodology is represented in order to have an overview of the steps conducted in the process of risk analysis. The CORAS approach and related components will be present in an detail manner, and the last part will study Fuzzy logic and relevant subjects of which will be adopted to elaborate our approach.

4.1 Basic Terminologies and Definitions

4.1.1 Asset

Definition 4.1.1. An asset is something to which a party assigns value and hence for which party requires protection.

To identify risks, first we should to know what we want to protect, and we name it an asset. There are two types of asset, that are direct and indirect ones. For example, if confidential information (direct asset) is leaked it will affect the competition of the company (indirect asset). By definition [25, p.55], an indirect asset is the asset that is affected if another asset is harmed, and if not it is a direct asset. The figure 4.1 represents an asset diagram from the example.

4.1.2 Threat and Threat Scenario

Definition 4.1.2. A threat is a potential cause of an unwanted incident.

A threat can be a human with intention (Hacker wants to steal customer's account information of a company) or without intention (Staff in a company incidentally reveals confidential information to externals by accessing social networks), and a threat may be human or non-human such as computer virus, trojan (with intention) or system errors (without intention) such as network connection, reliability of the system, etc. In summary, a threat can be anyone or anything with or without intention to harm the system which we want to protect. Initially, they need to perform a sequence of steps to achieve their goals. It is defined as a threat scenario.

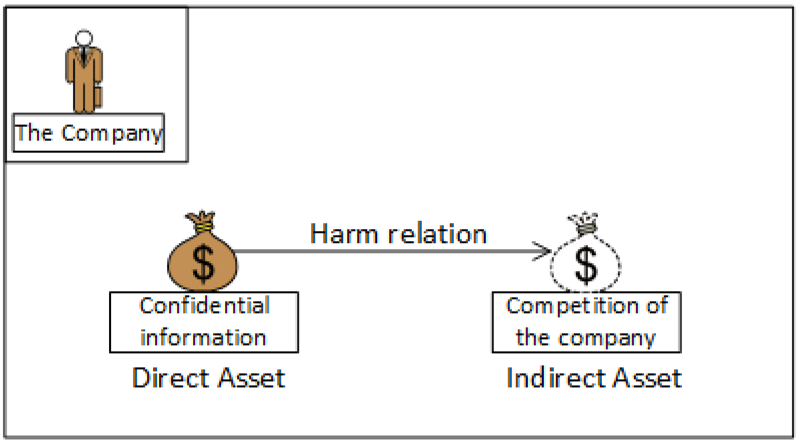


Figure 4.1: Asset Diagram Example

Definition 4.1.3. A threat scenario is a chain or series of events that is initiated by a threat and that may lead to an unwanted incident

The figure 4.2 describes the example of an employee incidentally reveals important information of a company to externals represented by a CORAS threat diagram.

4.1.3 Unwanted Incident

Definition 4.1.4. An unwanted incident is an event that harms or reduces the value of an asset.

Refer to the figure 4.2, when an employee incidentally shares his information, for instance, his company photos on Facebook, he incidentally reveals internal activities of the company to externals despite the fact that this was not his intention, however, it may affects company's competition and it is the unwanted incident.

4.1.4 Vulnerability

Definition 4.1.5. Vulnerability is a weakness, flaw or deficiency that opens for, or may be exploited by, a threat to cause harm to or reduce the value of an asset

A vulnerability may be originated from the system itself, for example a weaknesses of a information system (security issues), or it may be caused by a human such as an employee without sufficient training about company policies. By this way, he or she easily reveals confidential information to others. Figure 4.2 represents the vulnerability is the lack of Internet firewall or filter, for that reason, employee easily gets access to social network pages during working hours.

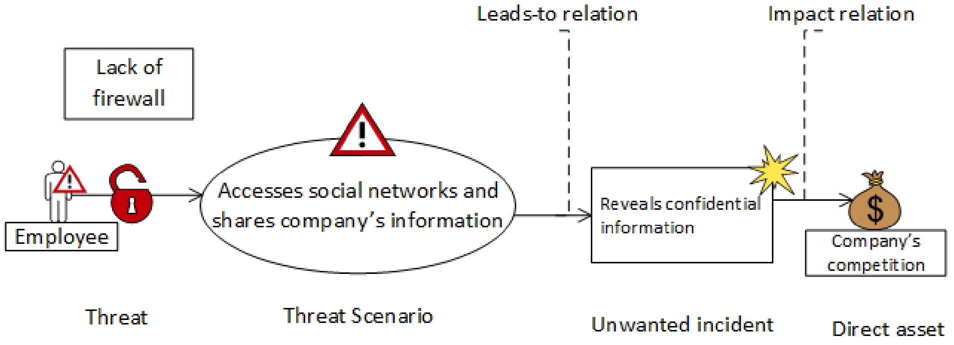


Figure 4.2: CORAS Threat Diagram

4.1.5 Risk

Risk is a general term and there are many definitions which relate to the term *risk*. The definition of risk is dependance on the context [**RISA:RISA433**, 21, 25]. However, we adopt the definition of risk from [25, p.60].

Definition 4.1.6. A risk is a likelihood of an unwanted incident and its consequence for a specific asset.

A risk has two features that are likelihood and consequence. The likelihood describes the frequency of occurrence (how many times does it happen in an interval of time), and the consequence represents 'The impact of an unwanted incident on an asset in terms of harm or reduced asset value'. With respect to the example above, we can see that if the company does not restrict employees to access the internet, they can access to social networks every days, or even every hours when they are working. Therefore, likelihood of the risk will be high and of course it will lead to the leakage of confidential information.

Risk also has a level that describes to what extent a risk can be accepted or denied. In practice, there are many risks in a typical system. Some of them can be accepted because the cost for fixing them may exceed the benefits of the existence of those. For example, in some companies (Bank, Software Development Organizations), it is necessary to have an internet firewall installed on employees' computer for protecting access to the internet. In another companies, when computers are not mandatory for storing and working with confidential information, install and maintain an internet firewall on their network may not be necessary because it may slow down the network access.

The figure 4.3 represents the basic concept of risk and its relationship to related terminologies by UML class diagram. As shown in the diagram, a risk is constituted by an unwanted incident, while unwanted incidents may cause the same risk. An unwanted incident has its own likelihood. An asset is affected by risks, and a party has many assets.

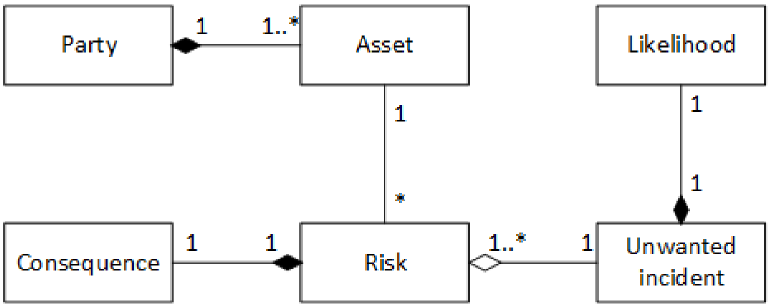


Figure 4.3: The basic risk terminology

4.2 Risk Assessment Process

Risk assessment is a sub-process of the risk management process [39]. As illustrated in figure 4.4, the five sub-processes in the middle of the process constitute to risk analysis process, and the three centered sub-processes (are covered by a dark box) are called risk assessment in which we are interested. The sub-processes in risk analysis happen sequentially in conjunction with Communication and Consult and Monitor and Review. The purpose of other processes (Communication and Consult and Monitor and Review) is to control the risk analysis process. Below is the explanation of the steps above [25, p.16]:

- *Establish the context* is dedicated to identifying assets, vulnerabilities and stakeholders of the system.
- *Identify risks* is to identify threat sources and threat scenarios which constitute risks.
- Estimate risks is the step to estimate likelihoods and consequences from the previous step.
- Evaluate risks is to prioritize risk in order to select what kinds of risks must be taken into account.
- Treat risks is to find appropriate solutions to address risk.

We will not present those processes in detail because CORAS has offered an eight-stepped process which will be presented in the next section. The CORAS steps are correspondent to the sub-processes of the risk analysis defined by International Organization Standard (ISO) [39].

4.3 The CORAS approach

CORAS is a general framework that can be applied to almost all defensive risk analysis situations ('CORAS is a general approach to risk analysis and has been applied to a large variety of risk analysis targets and concerns within numerous domains' [25, p.7]). The CORAS method contains a method, language and tool to assist the risk analysis process. The CORAS method

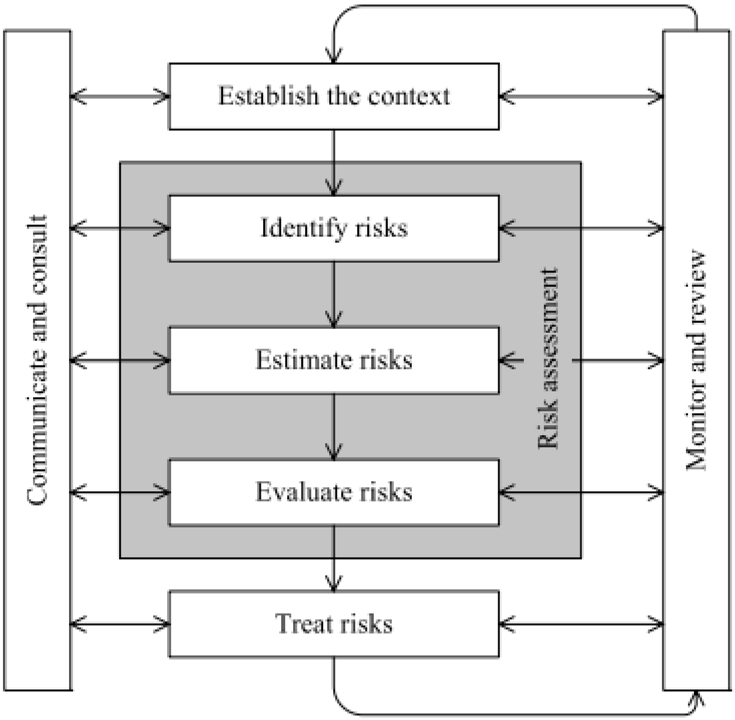


Figure 4.4: The Overall Risk Management Process

is the eight steps to conduct risk analysis and the rules for reasoning about likelihood. The CORAS language employs graphical elements for describing threats, risks, vulnerabilities, etc [25, p.6] (as discussed in section 4.1). The CORAS tool is a software which assists documenting and brainstorming CORAS diagrams.

4.3.1 The CORAS method

4.3.1.1 The eight steps of CORAS

Risk analysis process is conducted by eight steps according to the CORAS approach. The first three steps are involved in Establish the context as in figure 4.4.

Preparation for the analysis This is an initial step to prepare and gather basic information of the system which will be the focus of the analysis team and stakeholders. In addition to that, the analysis team and stakeholders make some agreements before conducting the next step [25, p.73].

Customer presentation of the target In this phase, the analysis team and stakeholders focus on details of the system and parts of the system which they want to protect. More on that, they agree on the CORAS terms and diagrams to model risk situations. Goals, target, focus and scope of the analysis are set out and clarified [25, p.81-94].

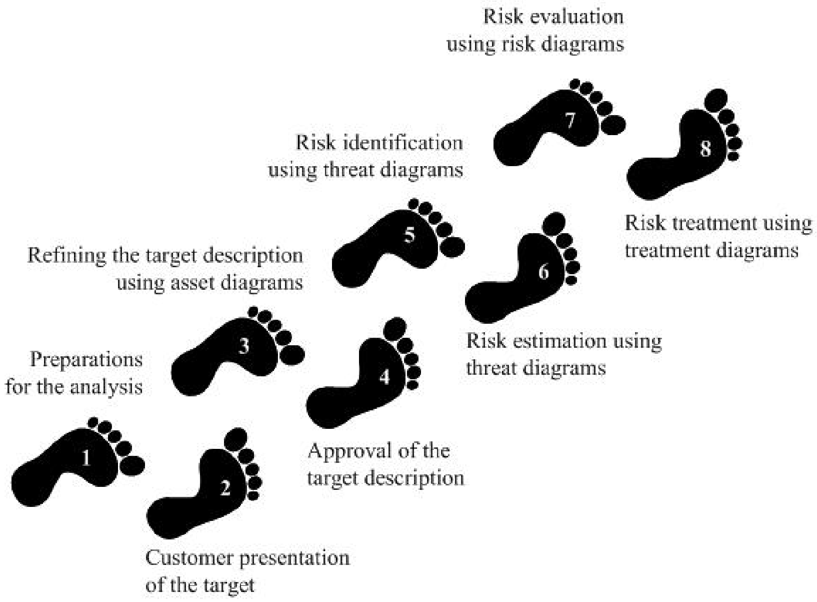

Risk evaluation using risk diagrams
Risk identification using threat diagrams
Risk treatment using treatment diagrams
Refining the target description using asset diagrams
Risk estimation using threat diagrams
Preparations for the analysis
Approval of the target description
Customer presentation of the target
1
2
3
4
5
6
7
8

Figure 4.5: The eight steps of CORAS

Refining the target description using asset diagrams The main purpose of this step is to refine and agree on the focus, targets, scope, etc which are set out in the previous step.

Approval of the Target Description This step is dedicated to obtain approval of which have been agreed on from the previous step. Additionally, the analysis team and stakeholders set up scales of likelihood, consequence and risk function to facilitate the following steps.

Risk Identification using threat diagrams This step includes four main sub-steps, they are Categorizing Threat Diagrams, Identification of Threats and Unwanted Incidents, Identification of Threat Scenarios and Identification of Vulnerabilities . These sub-steps can be conducted in the following manner. First, the analysis team and stakeholders identify threats or threat sources that can harm the assets, and unwanted incidents are also investigated. Next, they analyze threat scenarios which can be conducted by each threat. Based on that, they investigate vulnerabilities of the system from which let threat scenarios conducted. In addion to that, the process is iterative and mutually supported each other in the sense that the risk analysis team and stakeholders brainstorm refining the threat diagrams until risk situations are clearly identified.

Risk Estimation using threat diagrams The objective of this step is to refine the threat diagrams obtained from the previous step. In this step, the analysis team and stakeholders assign likelihoods to threats scenarios, unwanted incidents. They apply rules to reason and calculate likelihoods, risk function to estimate level of impacts on the assets.

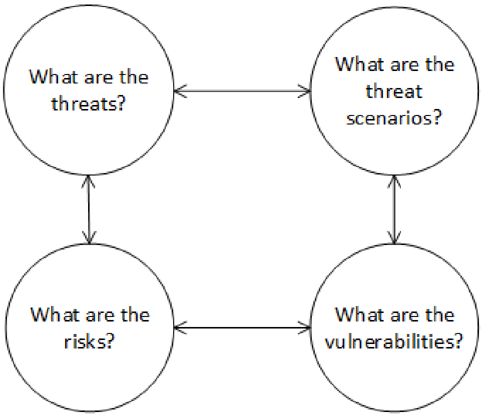


Figure 4.6: Risk Identification Process

In this thesis, we will focus on this step by refining frequency scales by fuzzy logic scales.

Risk Evaluation using diagrams This step is conducted by sub-steps in order to confirm the risk estimation, risk functions from the previous steps to estimate risks. Additionally, they refine risk diagrams from the threat diagrams to have an overview of the identified risks, they aggregate risks to estimate consequence when two or more risks happen simultaneously.

Risk Treatment using treatment diagrams The risk treatment phase consists of three main tasks, the purpose is provide solutions for identified risks.

4.3.1.2 Analyzing likelihood using CORAS diagram

In order to analyze risk, likelihoods must be calculated in detail, based on that treatments for particular risks can be given. There are two kinds of elements and two kinds of relations to which likelihood value can be assigned. The elements are threat and unwanted incident, and the relations are initiates and leads-to relation. CORAS has rules supporting to calculate likelihood based on CORAS diagrams.

Rules for leads-to

$$\frac{H \vdash v_1(f) \qquad H \vdash v_1 \xrightarrow{r} v_2}{H \vdash v_1 \sqcap\!| \, v_2(f.r)} \tag{4.1}$$

Rule for separate

$$\frac{H_1 \vdash v_1(f1) \qquad H_2 \vdash v_2(f2) \qquad s(v_1) \cap s(v_2) = \emptyset}{H \vdash v_1 \sqcup v_2(f1 + f2)} \tag{4.2}$$

4.3.2 The CORAS diagram

CORAS has five kinds of diagram supporting different steps in risk analysis process. The diagrams are defined and exercised by the eight steps of the CORAS consistently. Each subsequent diagram is the refinement of the prior diagram corresponding to the eight steps of the CORAS. The overview of each diagram is given below. The main diagram we focus on is the threat diagram because it involves in the risk assessment process.

4.3.2.1 Asset diagram

Asset diagram is used in the early phase of the process to identify and verify which parts of the system must be protected (assets). Main components of the diagram are party, assets (direct and indirect), and the harm relationship between assets (As shown in the figure 4.1).

4.3.2.2 Threat Diagram

Threat Diagram is the main diagram of the risk assessment process, the purpose of this diagram is to identify and estimate threats, unwanted incidents and risks. It is involved in step 5 of the CORAS approach. Main components of the diagram include threats (human, non-human, deliberate, accidentally), vulnerabilities, threat scenarios, unwanted incidents and assets (direct assets), the relationships between threats and threats scenarios, threats and unwanted incidents are initiates relation, threat scenarios and unwanted incidents or two threat scenarios or two unwanted incidents are leads-to relation, and the relationship between unwanted incidents and assets are impacts relation. The figure 4.2 is an example of threat diagram.

4.3.2.3 Risk Diagram

Risk diagram is the refinement of threat diagram by that it removes threat scenarios and relationships of threat scenarios. The risk diagram represents only threats, unwanted incidents and assets. The relationships between unwanted incidents and assets are impact relation which constitutes to risks.

4.3.2.4 Treatment Diagram

Treatment Diagram is an extension of threat diagram and includes treatment category to treat risk, namely avoid, reduce consequence, reduce likelihood of unwanted incidents, transfer and retain.

4.3.2.5 Treatment Overview Diagram

Treatment Overview Diagram is similar to risk diagram, and is a collapsed version of treatment diagram.

4.3.2.6 Relationship in the CORAS model

In the CORAS model, there are many relationships between the components. Each relationship is related only to a number of components. And the relationships is the way that CORAS model can translate its elements to sentences which is very helpful for a fuzzy approach.

- Harm: The relationship between two assets, asset a is affected then asset b.
- Initiate: The relationship originated from threat to threat scenario or unwanted incident. It means ‘e1 initiate e2 exploits some set of vulnerabilities to initiate e2 with some likelihood’ [25, p.58].
- Leads-to: This is the cause-result relationship. As defined in [25, p.58], ‘e1 leads to e2 with some likelihood, due to some set of vulnerabilities’. The relationship is between a threat scenario to an unwanted incident, or between two threats or two unwanted incidents [25, p.58].
- Impact: The relationship between an asset and an threat scenario or unwanted incident. Treat: The relationship between treatment and risk or vulnerability. It has five categories: avoid, reduce consequence, reduce likelihood , transfer and retain.

4.3.3 The CORAS Tool

The CORAS tool is a graphical user interface tool for describing risk models on-the-fly, and that is the way which we can exploit the tool and the language for risk analysis. The advantage of the tool is that it is developed as an Eclipse plug-in, open source and supports a standard data format, so that it can be extendable as well as developed new functions and features.

4.4 Fuzzy logic

Fuzzy logic is a mathematical tool for addressing issues of uncertainty, imprecise and vagueness in practice [48]. It is a set with value ranges from 0 to 1, therefore it has infinite elements (contrast to classical logic theory, which has only two values, 0 represents false and 1 is true). It has been applied in many fields from engineering, financing, banking, project management and the domain of risk analysis which the topic we pursue in this thesis.

Back to our example above, when the employee are going to a social network page, the frequency (not the probability) that he reveals company’s information to externals is based on what types of information and how much relevant information he is sharing? It is not clear what are the type of information and the amount. If he wrote a sentence which contains a few words relevant to the company, sometimes the frequency was low. However, if he uploaded a photo that revealed confidential information, the the frequency was high. The frequency of doing something, and the type of information can be represented by membership functions. Therefore, fuzzy approach will constitute the methodology for calculating as well as predicting an occurrence of risk and relevant factors. Fuzzy logic is a mathematical theory which includes basic theorems and operations. In this thesis we are only interested in related background and applications of fuzzy which can elaborate the approach.

4.4.1 Fuzzy Set

A fuzzy set is a set where each element belongs to the set by a degree of membership, and a membership function maps every element of the universe of discourse X to the interval $[0,1]$ [16, p. 15]:

$$\mu_A(x) : X \to [0,1] \tag{4.3}$$

A fuzzy set can be represented by a singleton if x is an element of universe of discourse X and A is a fuzzy set defined on X [16, p. 15]:

$$A = \{(x, \mu_A(x))\}, x \in X \tag{4.4}$$

An alternative representation of fuzzy set in case of discrete and continuous is represented as the two equations below respectively:

$$A = \sum_{x_i \in X} \mu_A(x_i)/x_i \tag{4.5}$$

Or

$$A = \int_X \mu_A(x_i)/x_i \tag{4.6}$$

4.4.1.1 Union and Intersection of Fuzzy Sets

The Union of Two Fuzzy Sets is represented as below [16, p.20]:

$$\mu_{A \cup B}(x) = \mu_A(x) \vee \mu_B(x) \tag{4.7}$$

The Intersection of Two Fuzzy Sets can be calculated as [16, p.20]:

$$\mu_{A \cup B}(x) = \mu_A(x) \wedge \mu_B(x) \tag{4.8}$$

And $\wedge, \vee$ are the minimum and maximum operator on fuzzy set.

4.4.2 Membership functions

In the previous section, we represented some basic concepts of fuzzy set, and membership function is a foundation to describe a fuzzy set. This section is dedicated to shed light on the characteristics of membership function in detail.

4.4.2.1 Features of membership function

A membership function has three properties which are: core, support and boundary [37]. And the figure 4.7 is the formal definition of each property.

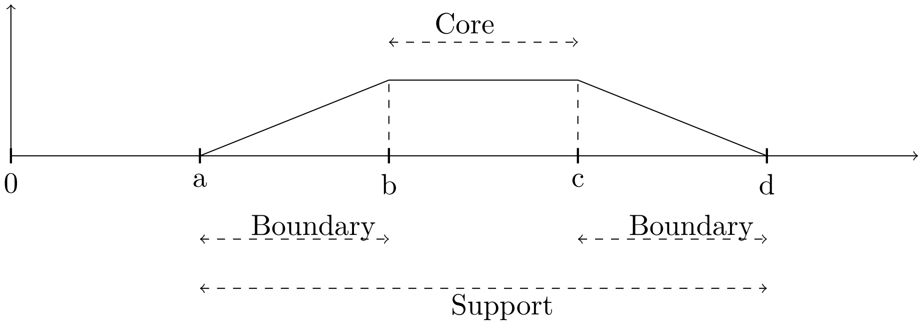


Figure 4.7: Properties of membership function

Core All the elements which their membership function are equal to 1, this means that $\alpha(x) = 1$.

Support All the elements which their membership function are greater than 0, this means $\alpha(x) > 0$.

Boundary All the elements whose membership are between 0 and 1, this means $0 < \alpha(x) < 1$.

4.4.2.2 Types of membership function

There are many types of membership functions which have been widely applied in practice, namely triangular, trapezoidal, Gaussian, Cauchy membership function [31]. However, we will study and apply only triangular and trapezoidal membership function in this thesis. The application of another membership functions to the domain of risk assessment will be the topic for future research. The acquisition of triangular and trapezoidal membership function can be explained theoretically and empirically [3]. Empirical and theoretical fact point out that re-scaling these membership functions is actually linear and direct [23]. Trapezoidal and triangular membership function are the special case of interval scale representing frequency which is adopted by CORAS to calculate likelihoods [38].

The below figures and expressions describe the triangular and trapezoidal membership function respectively.

Triangular membership function The triangular membership function is defined by three parameters: a, b and c as in the expression below [23]:

$$\alpha(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ \frac{c-x}{c-b} & b \leq x \leq c \\ 0 & x \geq c \end{cases} \qquad (4.9)$$

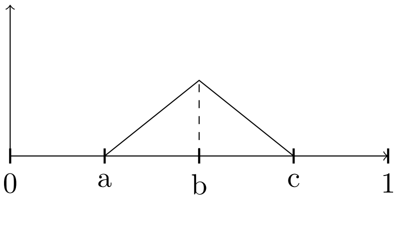


Figure 4.8: Triangular Shape

Trapezoidal membership function Trapezoidal membership function is an extension of triangular membership function, and it is defined by four parameters: a, b, c and d as below:

$$\alpha(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & b \leq x \leq c \\ \frac{d-x}{d-c} & c \leq x \leq d \\ 0 & x \geq d \end{cases} \tag{4.10}$$

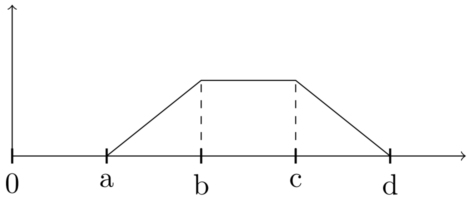


Figure 4.9: Trapezoidal Shape

4.4.3 Classification of Fuzzy Set

A fuzzy set can be classified based on their membership function. There are four types of fuzzy set which are normal, subnormal, convex, and nonconvex [16, 23].

Normal and subnormal If the membership function has at least one element in the universe of discourse whose value is equal to 1, then the set is called normal. Otherwise, if there is not any element in the universe of discourse whose value is equal to 1, then the set is call subnormal.

Convex and nonconvex If the membership function whose elements increasing or decreasing monotonically, or increasing and decreasing monotonically, the set is defined as convex. Or it means that the function does not go up and down more than once [16, p.77,78]. Otherwise, the set is not convex.

4.4.4 Fuzzy Number

A fuzzy number is a fuzzy set with the criteria that it is convex and normal [16, p.77].

4.4.4.1 alpha-cuts

With a fuzzy set A, we can associate a collection of crisp set known as $\alpha - cuts$ or level sets of A. $\alpha - cuts$ of fuzzy set A denoted as A_α is defined as:

$$A_\alpha = \{x \in X | \mu_A(x) \geq \mu\} \tag{4.11}$$

A fuzzy set A can be represented by interval as below.

$$A_\alpha = [a_1^{(\alpha)}, a_2^{(\alpha)}] \tag{4.12}$$

4.4.4.2 Addition

$$A + B = [a_1^\alpha, a_2^\alpha] + [b_1^\alpha, b_2^\alpha] = [a_1^\alpha + b_1^\alpha, a_2^\alpha + b_2^\alpha] \tag{4.13}$$

4.4.4.3 Subtraction

$$A - B = [a_1^\alpha, a_2^\alpha] - [b_1^\alpha, b_2^\alpha \alpha] = [a_1^\alpha - b_1^\alpha, a_2^\alpha - b_2^\alpha] \tag{4.14}$$

4.4.4.4 Multiplication

$$A.B = [a_1^\alpha, a_2^\alpha].[b_1^\alpha, b_2^\alpha] = [a_1^\alpha.b_1^\alpha, a_2^\alpha.b_2^\alpha] \tag{4.15}$$

In addition to that, k is a real number and A is a fuzzy number, we have a multiplication of real number and fuzzy number as below:

$$k.A = k.[a_1^\alpha, a_2^\alpha] = [ka_1^\alpha, ka_2^\alpha] \tag{4.16}$$

4.4.4.5 Division

$$A \div B = [a_1^\alpha, a_2^\alpha] \div [b_1^\alpha, b_2^\alpha] = [a_1^\alpha \div b_1^\alpha, a_2^\alpha \div b_2^\alpha] \tag{4.17}$$

4.4.4.6 Parametric representation of fuzzy number

A triangular and trapezoidal fuzzy number can be represented by parameters that form its geometric shape [14]. The parametric notation of triangular and trapezoidal fuzzy number are adopted and utilized in our experimentation in order to simplify representation of these fuzzy numbers.

Given a triangular and trapezoidal fuzzy number A and B represented by the membership functions that are expressed by the equation 4.9 and 4.10 respectively, parametric notation of triangular and trapezoidal fuzzy number is represented by the equation 4.18 and 4.19 respectively.

$$A = [a, b, c] \tag{4.18}$$

$$B = [a, b, c, d] \tag{4.19}$$

However, we should notice that this notation is for fuzzy number with the criteria of convex and normal.

4.4.5 Fuzzy Relation

4.4.5.1 Cartesian Product of Relation

A Cartesian product of two sets X and Y is defined as [37, p.37]:

$$x \times y = \{(x, y) | x \in X, y \in Y\} \tag{4.20}$$

When $x \neq y$ then $(x, y) \neq (y, x)$

4.4.5.2 Crisp Relation

Crisp relation is defined over the Cartesian Product of two or more sets [16, p.50]. A relation of x and y in X, Y respectively is represented as [37, p.38]:

$$R(x, y) = \begin{cases} 1 & (x, y) \in X \times Y \\ 0 & (x, y) \notin X \times Y \end{cases} \tag{4.21}$$

4.4.5.3 Fuzzy Relation

A pair of x and y in X, Y respectively that is related to a degree is a fuzzy relation. Fuzzy relations are fuzzy sets defined on the Cartesian Product [16, p.52,53].A formal representation of a fuzzy relation is described by the equation below:

$$R(x, y) = \{((x, y), \mu_R(x, y)\} \tag{4.22}$$

When x and y belong to X, Y respectively. And $\mu_R(x, y)$ is a membership function which represents the relationship between x, y.

4.4.6 Interpretation of membership function

A membership function can be interpreted by five different views, they are likelihood view, random set view, similarity view, utility view and measurement view. Each of the interpretation has their own differences and similarities [4] as well as pros and cons. The subject of membership function interpretation and its theoretical background are worth having another research. Therefore, we will not dig into the subject so far because it is out of our scope. However, a simple explanation of the views is given below in order to have an overall understanding about the meanings of membership function.

Given a statement A is F represented by a membership function $\mu_F A = 0.7$. The statement is interpreted by the views below [4].

Likelihood View 70% of a given population declared that A is F.

Random set View 70% of a given population declared that F is an interval containing A.

Similarity View Given a prototypical object which is truly F, A is away from the object to a degree 0.3.

Utility View 0.7 is the utility of asserting that A is F.

Measurement View When compared to others, A is more F than some and can be encoded as 0.7 on some scale.

4.4.7 Elicitation of membership function

So far we have represented membership function as a basic element of fuzzy logic theory, but the construction of membership functions over the domain of interest have not been taken into account yet. There are different ways to elicit membership functions as presented in [4, 23, 37, 44]. As proposed by Verkuilen [44], and Li [23], there are three common manual methods to develop membership functions, namely, direct assignment, indirect assignment and transformation. Each methods has its own pros and cons [23, 44]. We approach these methods in order to provide an overview of constructing membership functions which refines the step of Setting up the scale of the

CORAS. The choice of the methods is the wisdom of the risk analyst who has the perfect knowledge of available sources of data such as statistic numbers, experts, experience, etc. In addition to that, methods for constructing membership functions is empirical work which requires time and effort to conclude obtained results.

4.4.7.1 Direct method

Direct assignment is a subjective method to obtain membership function [4, 23, 44]. It is commonly chosen to almost domain of interest [44]. By acquiring this method, membership functions are constructed by experts' opinion. Each expert is expected to assign a membership function $\mu_A(x)$ to a given element x which belongs to fuzzy set A that according to his or her opinion best captures the meaning of linguistic term represented by fuzzy set A. In case of multiple experts, a procedure is conducted to aggregate the individual expert's opinions and conclude the membership function $\mu_A(x)$. The most common procedure to acquire membership functions from multiple expert's view is the simple weighted average.

$$\mu_A(x) = \sum_{i=1}^{n} c_i \mu_{A_i}(x) \tag{4.23}$$

Where $\mu_{A_i}(x)$ is the membership function given by expert ith, c_i is the weight of the expert ith.

And

$$\sum_{i=1}^{n} c_i = 1$$

4.4.7.2 Indirect method

This method employs both statistical data and expert's opinion to develop membership functions via mathematical modeling [44]. For example, experts are given simple questions to compare elements of universe of discourse with respect to the membership functions describing them. From the results obtained, the membership functions are constructed by procedures. The indirect method is considered to be reliable than the direct method in constructing membership function [23]. However, it is costly and time-consumed than the direct one.

4.4.7.3 Transformation method

In this method, membership functions are constructed from data set or statistical numbers by transformation procedures [23]. Transformation is objective method to elicit membership functions. The figure below shows a histogram graph demonstrating statistical data, the shape of the membership function is estimated and elicited by a procedure.

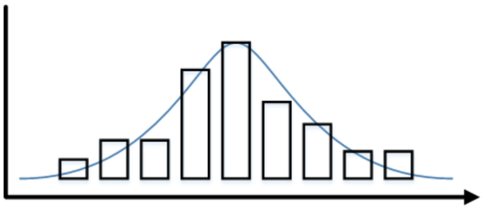


Figure 4.10: Membership function construction based on statistical data

4.4.8 Fuzzification and Defuzzification

4.4.8.1 Fuzzification

Fuzzification is a process of translating a crisp set to a fuzzy set or increase the fuzziness of a fuzzy set [16, p.25]. Actually, this process is involved with the methods of membership function elicitation presented from the previous section.

4.4.8.2 Defuzzification

Defuzzification is a reverse process of Fuzzification that it translates a fuzzy set into a crisp number [16, p.163]. There are many methods of defuzzification and each of them generate different results.

In our thesis, we employ the Center of Area or COA because it is the most prevalent and popular method [16, p.164]. The expression 5.1 is the formula to calculate centroid of function $\mu_A(x)$.

$$u = \frac{\int \mu_A(x).xdx}{\int \mu_A(x)dx} \tag{4.24}$$

The figure below represents different results when applying different methods of defuzzification which are compared to the centroid method.

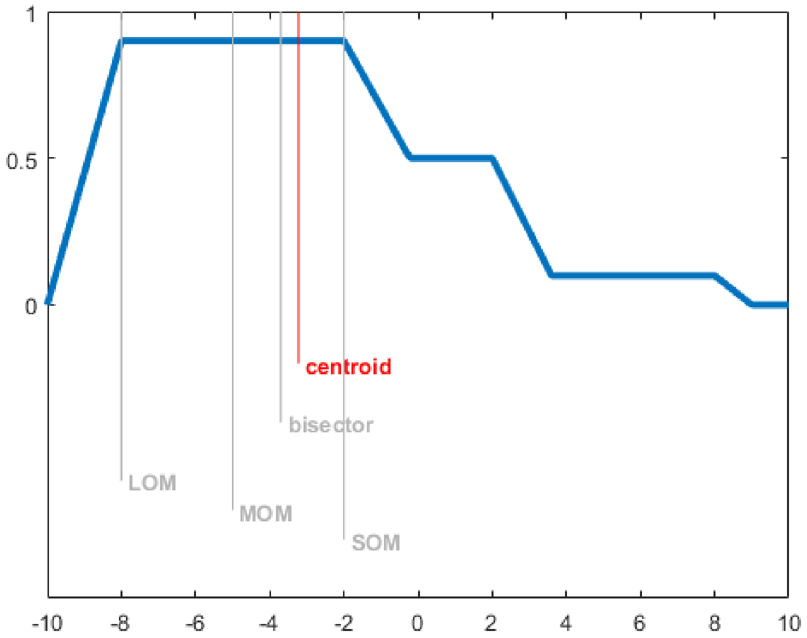


Figure 4.11: Different Results of Applying Different Defuzzification Methods (Adapted from [29])

4.4.8.3 Geometric centroid formulas

Centroid of a trapezoidal and triangular shape can be calculated by geometric formulas. In order to facilitate and simplify computation process, the two centroid formulas of trapezoid and triangular are adopted from []. A centroid of a geometric shape is expressed by two points (C_x, C_y) on the Cartesian coordinate system. However, the Y coordinate represents the degree of membership of an element of a fuzzy set, the X coordinate is the universe of discourse which is the domain of interest. Therefore, the C_x is the result of defuzzification.

Centroid of Trapezoid: Given a trapezoid A represented by four points $[a, b, c, d]$. Centroid of trapezoid A can be calculated by the formula 4.25.

$$C_x = \frac{(c^2 + cd + d^2) - (a^2 + ab + b^2)}{3[(c + d) - (a + b)]} \tag{4.25}$$

Centroid of Triangle: Given a triangular A represented by three points $[a, b, c]$. Centroid of triangular A can be calculated by the formula 5.3.

$$C_x = \frac{a + b + c}{3} \tag{4.26}$$

Centroid of Interval Actually, an interval is represented by a rectangle and a special case of fuzzy logic when all elements belong to the interval having a full degree of membership. Therefore, an interval can be defuzzified by applying geometric formula from [26] to calculate centroid point.

Given an interval (a, b) in the Cartesian coordinate, a centroid point on the X coordinate C_x is calculated by the equation 4.27

$$C_x = \frac{a+b}{2} \tag{4.27}$$

4.4.9 Interval Arithmetic

We present sum and product of two intervals in this section because it is useful for calculating frequencies and approximation of fuzzy logic which will be presented in the next section. In addition to that, centroid of an interval is also presented when it is neccessary to compare results of defuzzification.

Sum of Interval Given two intervals $(a, b), (c, d)$, sum of intervals can be calculated as below expression [15]:

$$(a, b) + (c, d) = (a + c, b + d) \tag{4.28}$$

Product of Interval According to theorem 5 from [15], product of two intervals can be calculated as below:

$$(a, b) \times (c, d) = (min(a \times c, a \times d), max(b \times c, b \times d)) \tag{4.29}$$

Chapter 5

Approach to Combining Fuzzy Logic with CORAS method

In this chapter, we present detail about the approach which combines Fuzzy logic and CORAS based on the rules for reasoning about frequency of likelihood represented in section 4.3.1.2 on page 25.

5.1 Fuzzifying Scales

This step is an extension of the Approval of the Target Description as proposed by CORAS (section 4.3.1.1 on page 23) when the team wants to employ fuzzy logic for the risk assessment process. The risk analyst, experts, and stakeholders of the system need to conduct a meeting to construct membership functions. Basically, membership functions are derived from the original scales which have been approved and refined by the construction methods (section 4.4.7 on page 32), or they can be elicited from scratch by applying the construction methods. However, conducting the construction methods is empirical, time-consumed and out of our scope as mentioned before. Therefore, we derives the membership functions from the original scales and assign them by strategy to construct membership function. Despite the fact that, membership functions should not be assigned arbitrarily, applying the strategy gives us an insight into the effect of rescaling membership functions, and simplifies our research process.

5.1.1 Likelihood Scale

In order to construct membership functions, we first derives original scale for likelihood from [25, p.32]. The likelihood scale describes frequency of occurrences of threat scenario or unwanted incident, and it should be consistent throughout the risk analysis phase [25, p.118].

The likelihood scale is defined as the table 5.1.

Likelihoods	Definition	Description
Rare	Less than once per ten years	[0, 1) : 10y = [0, 0.1) : 1y
Unlikely	Less than once per two years	[1, 5) : 10y = [0.1, 0.5) : 1y
Possible	Less than twice per year	[5, 20) : 10y = [0.5, 2) : 1y
Likely	Two to five times per year	[20, 50) : 10y = [2, 5) : 1y
Certain	Five times or more per year	[50, +∞) : 10y = [5, +∞) : 1y

Table 5.1: Likelihood scale

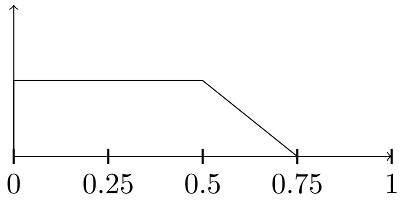


Figure 5.1: Trapezoidal membership function represents the scale of rare

5.1.2 Fuzzifying Extreme Interval

The scales that reach the boundary are extreme values, for instance, the variables rare and certain are extreme intervals. They are extreme value because membership functions assigned to them should be special cases of triangle or trapezoid. The point of view can be explained by adopting one of the views of membership function from section 4.4.6. If we register to the likelihood view, and the interval $(0, 1)$ representing the frequency of rare when an event happens from 0 to 1 time per year. From the likelihood viewpoint, 100% of the population describes that every frequency belongs to $(0, 1)$ is rare. Therefore, the left support of the trapezoidal membership function which represents the scale should not be gradually increased, but it should be as in the figure 5.1, for instance.

We should be careful that the likelihood of certain is defined in original scale being from $[5, \infty)$. For that reason, we can not calculate this value in both CORAS and fuzzy logic. In order to address the problem, the infinite symbol ∞ must be replaced by a concrete number. To what extend the interval of certain can be adjusted? If the finite symbol is replaced by small numbers, it will yield small results and vice verso, and results may be inaccurate.

If we consider the scale of certain representing all frequencies that are greater than 5. Then, the COA defuzzification of the membership function assigned to the scale should be 5 principally.

Given a trapezoidal membership function $A = [a, b, c, d]$ and F is a COA defuzzification function and is defined as:

$$F(A) = r$$

When $r \in \mathbb{R}$.

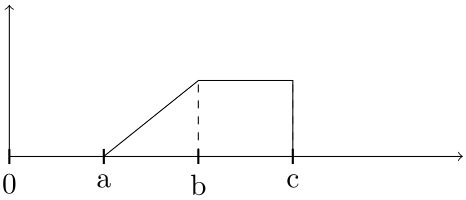
The right support of the function should not be gradually decreased. Therefore, the membership function A can be re-defined as $[a, b, c, c]$. The equation 5.1 is the formula to calculate a centroid point of that type of

Terms	Min	Max
Low	0	0.25
Medium	0.25	0.5
High	0.5	0.75
Critical	0.75	1

Table 5.2: Conditional likelihood scale

trapezoid [28].

$$F(A) = C_x = \frac{3c^2 - (a^2 + ab + b^2)}{3[2c - (a + b)]} \tag{5.1}$$



Actually, we want to find c when $C_x = r$. The equation 5.2 is the solution for the quadratic equation $3c^2 - 6rc - (a^2 + ab + b^2) + 3r(a + b)$ which is transformed from the equation $C_x = r$.

$$c = \frac{6r + \sqrt{(6r)^2 + 4[(a^2 + ab + b^2) - 3r(a + b)]}}{6} \tag{5.2}$$

Similarly, a triangle membership function $B = (a, b, c)$ has the centroid point C_x is calculate by the formula 5.3.

$$C_x = \frac{a + b + c}{3} \tag{5.3}$$

If $C_x = r$, the variable c is calculated as the following:

$$c = 3r - (a + b) \tag{5.4}$$

5.2 Approach to Combining Rule for Leads-to

5.2.1 Conditional Likelihood Scale

Conditional likelihood is likelihood assigned to the leads-to relation, it is treated as a probability or probability interval ranging from $[0, 1]$ [25, p.151]. As argued by Lund, Solhaug and Stølen [25], conditional likelihood is quite difficult to assign. Therefore, we first theoretically divide the conditional likelihood scale from $[0, 1]$ into four levels and they are equivalent sub-intervals as depicted in the table 5.2.

5.2.2 Simple Fuzzy Method

As mentioned above, the conditional likelihood is hard to obtain and estimate, simple fuzzy method proposes the rule that treats conditional likelihoods as fuzzy numbers and retains the original likelihood scale.

Given a likelihood l represents an interval $[i_1, i_2]$, and conditional likelihood c is represented by a trapezoidal fuzzy number $[a, b, c, d]$ and is parameterized by its $\alpha - cut$ level as $[(b-a)\alpha + a, d - (d-c)\alpha]$.

By applying the rule for leads-to, actually the product of two intervals, we have:

$$\begin{aligned} l \times c &= [i_1, i_2] \times [(b-a)\alpha + a, d - (d-c)\alpha] \\ &= [\Big((b-a)\alpha + a\Big)i_1, \Big(d - (d-c)\alpha\Big)i_2] \\ &= [(i_1 b - i_1 a)\alpha + i_1 a, i_2 d - (i_2 d - i_2 c)\alpha] \end{aligned}$$

Therefore, the result is a trapezoidal fuzzy number which is represented by the equation below:

$$\alpha(x) = \begin{cases} 0 & x \leq a i_1 \\ \frac{x - i_1 a}{i_1 b - i_1 a} & i_1 a \leq x \leq i_1 b \\ 1 & i_1 b \leq x \leq i_2 c \\ \frac{i_2 d - x}{i_2 d - i_2 c} & i_2 c \leq x \leq i_2 d \\ 0 & x \geq i_2 d \end{cases} \tag{5.5}$$

Similarly, we obtain the result for the case of triangular fuzzy number $[a, b, c]$ as below:

$$\alpha(x) = \begin{cases} 0 & x \leq i_1 a \\ \frac{x - i_1 a}{i_1 b - i_1 a} & i_1 a \leq x \leq i_1 b \\ \frac{i_2 c - x}{i_2 c - i_2 b} & i_2 b \leq x \leq i_2 c \\ 0 & x \geq i_2 c \end{cases} \tag{5.6}$$

Simple fuzzy approach actually transforms likelihoods represented by intervals into fuzzy numbers. The interval plays a coefficient role adjusting cores and supports of fuzzy numbers.

5.2.3 General Fuzzy Method

General fuzzy method proposes a rule that treats both likelihoods and conditional likelihoods as fuzzy numbers.

Given the conditional likelihood c represented by a trapezoidal fuzzy number $[a_1, b_1, c_1, d_1]$ and the likelihood l represented by a trapezoidal fuzzy number $[a_2, b_2, c_2, d_2]$. The interval representation of the two fuzzy numbers are described as below respectively:

$$c = [(b_1 - a_1)\alpha + a_1, d_1 - (d_1 - c_1)\alpha]$$

And

$$l = [(b_2 - a_2)\alpha + a_2, d_2 - (d_2 - c_2)\alpha]$$

We will not present the step to conduct the product of two fuzzy numbers. However, we derive the result of multiplication of two trapezoidal fuzzy numbers from [2] as below.

$$\alpha(x) = \begin{cases} \frac{-(a_1b_2+b_1a_2-2a_1a_2+\sqrt{(a_1b_2-b_1a_2)^2+4(b_1-a_1)(b_2-a_2)x}}{2(b_1-a_1)(b_2-a_2)} & a_1a_2 \le x \le b_1b_2 \\ \frac{-(d_1c2+c_1d_2-2d_1d_2)+\sqrt{(d_1c_2-c_1d_2)^2+4(c_1-d_1)(c_2-d_2)x}}{2(c_1-d_1)(c_2-d_2)} & c_1c_2 \le x \le d_1d_2 \\ 0 & Otherwise \end{cases} \tag{5.7}$$

Correspondingly, two triangular fuzzy numbers which represent conditional likelihood and likelihood are given as $c = [a_1, b_1, c_1]$ and $l = [a_2, b_2, c_2]$, respectively. The result of the product of two triangular fuzzy numbers is inherited from [13] as below.

$$\alpha(x) = \begin{cases} \frac{-(a_1b_2+b_1a_2-2a_1a_2)+\sqrt{(a_1b_2-b_1a_2)^2+4(b_1-c_1)(b_2-a_2)x}}{2(b_1-a_1)(b_2-a_2)} & a_1a_2 \le x \le b_1b_2 \\ \frac{-(c_1b_2+c_2b_1-2c_1c_2)+\sqrt{(c_1b_2-c_2b_1)^2+4(b_1-c_1)(b_2-c_2)x}}{2(b_1-c_1)(b_2-c_2)} & b_1b_2 \le x \le c_1c_2 \\ 0 & Otherwise \end{cases} \tag{5.8}$$

As shown by the equation 5.7 and 5.8, the result of product of two fuzzy numbers does not maintain their original shapes. The new fuzzy number is represented by two curves.

Similarly, the product of trapezoidal fuzzy number and triangular fuzzy number can be obtained by applying the product of two intervals when two fuzzy numbers are represented by their $\alpha - cuts$.

5.3 Approach to Combining Rule for Separate

5.3.1 Applying fuzzy relation to rule for separate

Two threat scenarios can be treated as separate threat scenarios if they do not overlap in content [25, p.224]. This means that the content of the one can not be a special case of another. Otherwise, they are overlap, or partial overlap. The Venn diagram in the figure 5.2 demonstrates three cases of relation. As shown in the diagram, the likelihood of threat scenario T_1 and T_2 are greater than the case of partial separate. While the likelihood in the case of overlap is the maximum of threat scenario T_1 and T_2.

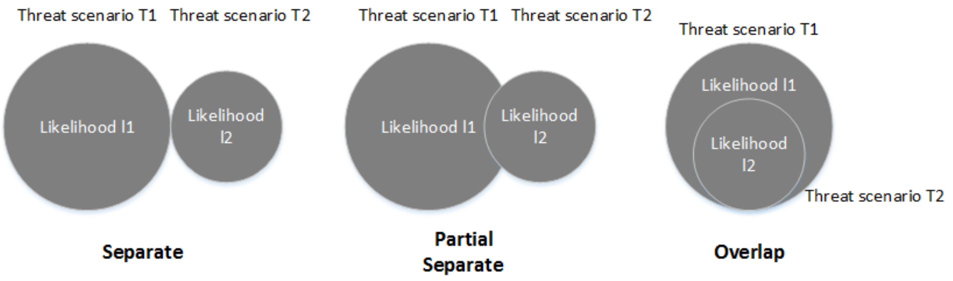


Figure 5.2: Venn diagram demonstrates three cases of relation

Given l_i is the likelihood of threat scenario T_i represented by a fuzzy number, and R is the separate relationship between two threat scenarios

T_1, T_2, a function U representing the result of aggregated T_i as shown in the equation 5.9.

$$U(T_1, T_2, R) = l \tag{5.9}$$

When $l \in L$. L is space of likelihoods.

The relation R representing separate relation of two threat scenarios T_1, T_2:

$$R(T_1, T_2) = r \tag{5.10}$$

And $r \in [0, 1]$. If r is boolean value the relation is crispy. In contrast, if r is represented by the value between $[0, 1]$, it is fuzzy relation.

If $r = 1$, T_1 and T_2 are separate, and if $r = 0$, T_1 and T_2 are overlap. Otherwise, $0 < r < 1$, we have a partial separate (overlap) relation.

By applying rule for separate from the equation 4.2, in case of separate we have:

$$T = l_1 + l_2$$

This means that $R(T_1, T_2) = 1$. In case of overlapping, we have:

$$T = max(l_1, l_2)$$

And $l_1 + l_2 > max(l_1, l_2)$. Therefore in case of separate, the aggregated likelihood must be always greater than that of overlap scenario when applying both CORAS or fuzzy logic principally. In case of partial separate, the aggregated likelihood must be smaller than that of separate, and greater than the case of overlap.

Given that U_S, U_P, U_O are three functions that return aggregated likelihood of two threat scenarios T_1 and T_2 in case of separate, partial separate, and overlap respectively. The criterion above can be expressed by the equation 5.11.

$$U_S > U_P > U_O \tag{5.11}$$

We now apply the concept of separation to fuzzy logic when employing relational value. From the expression 5.9, and rule for separate. The value of r should be 1 in case of completely separate and 0 in case of overlap, because r is actually a coefficient which will decrease size of fuzzy number which represents the likelihood of threat scenario in case of partial separate when $r < 1$.

If $r > 0$ and $r < 1$, we solved the cases of separate and partial separate but not the case of overlap. As we know, the maximum value of two threat scenario when combining together is $l_1 + l_2$ and the minimum is $max(l_1, l_2)$ if l_1, l_2 are likelihood of threat scenarios/unwanted incidents T_1, T_2. Therefore, the expression above 5.9 is redefined as:

$$U(T_1, T_2, R) = max(l_1, l_2) + min(l_1, l_2) \times r \tag{5.12}$$

5.3.2 Relation Scale

The fuzzy relation plays a role as a coefficient which scales the shape of the trapezoids which represents the fuzzy values. We hypothetically defined three intervals which represent the relationship between two threat scenarios. We consider the scale between $[0.75, 1)$ is separate, $[0.25, 0.5)$ is partial separate, and $[0, 0.25)$ is the case of overlap.

The difference between Fuzzy Relation and Fuzzified Relation lies on values which represent the degree of the relationship. In fuzzy relation, the degree of relation is represented by a real number between 0 and 1, which is a crisp number. In case of fuzzified relation, we again apply the concept of fuzzy set to represent the interval of relation.

Chapter 6

Experimentation

6.1 Setting Up the Context

We inherit an example from [25] and hypothetically construct the whole context based on the eight steps of the CORAS. We have not reached the final step of the CORAS method and the experimentation is to focus on the two rules for reasoning about likelihood. Therefore, when the threats diagram of the example is constructed we then conduct different experiments to observe results generated by different scales. We perform fuzzy logic arithmetic by Matlab. Therefore, we do not present details to perform the calculation of the experimentation.

An example is a simple case of an e-commerce system that the customer (the owner of the system) demands the risk analysis team analyzing possible risks that are attached to their system. The followings are necessary steps conducted throughout the process.

6.1.1 Preparations for the Analysis

6.1.1.1 Introduction to the system

Best-One is a company selling electronic items and they are currently interested in selling their products online. They hire a development team from a software company to develop the e-commerce system. After the development of the system is finished and the system are ready to use, they concern about risks that can harm their online business. The e-commerce system is operated by their IT department, they are in charge of every activities that are relevant to the system including business activities.

6.1.1.2 Roles and Stakeholders

From the first glance, the risk analysis team draws attention to relevant parties and stakeholders of the system. They first concern about the Best-one company, the IT department and the software company. In addition to that, clients of the e-commerce system are also relevant to the system. The UML class diagram in the figure 6.1 is adapted from [25, p.75] describing relationship between stakeholders and the risk analysis team. As shown in the

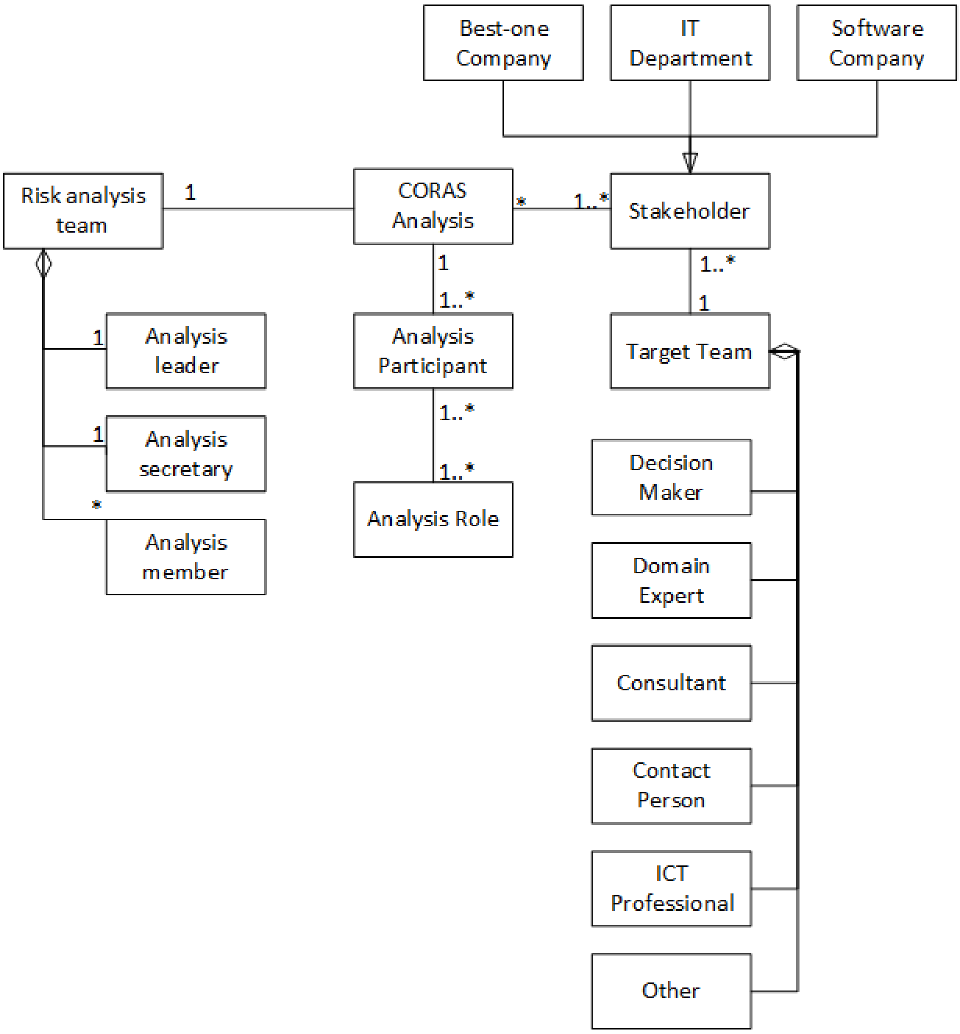


Figure 6.1: The diagram describes the stakeholders of the context

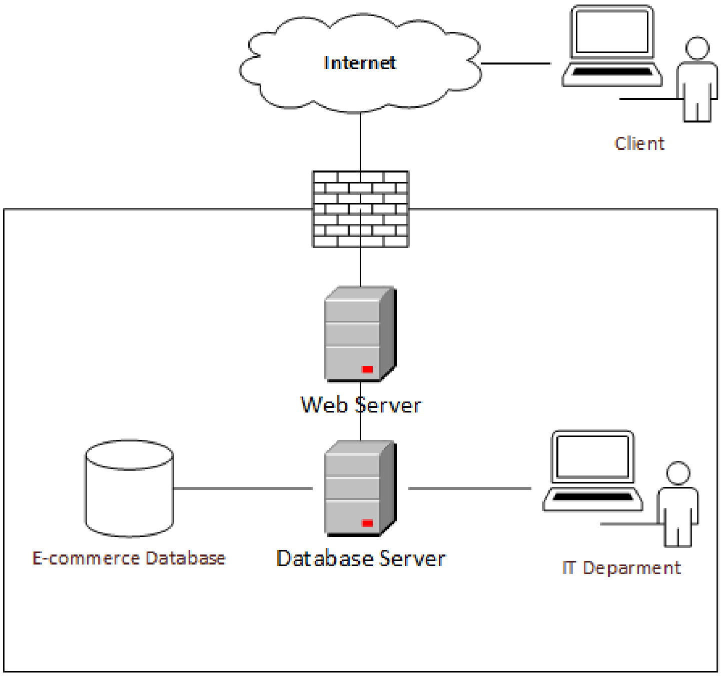


Figure 6.2: Presentation of The Target

figure, a CORAS analysis consists of a risk analysis team and stakeholders. The risk analysis team includes a leader, a secretary and members. The members may be domain experts, decision makers, etc. The stakeholders of the CORAS analysis form a target team that is relevant to the eight steps of CORAS. The stakeholders are the Best-One company, the IT department and the software company. The CORAS analysis defines analysis participants which include analysis roles.

6.1.2 Customer Presentation of The Target

The risk analysis team need to understand the system better, the team conducts a meeting with their stakeholders to obtain more information about the system. The stakeholders of this phase are the Best-one company and the software company. The Best-One company will present the target of the system and the process of online selling, while the presence of the software company is to support the Best-one company to present the technical factors of the system.

The system consists of two servers and a database, the first server is a web server that handles customers' requests from their private devices (computer, mobile, table). The second server is database server with the objective to manipulate the e-commerce database. In addition to that, they have a firewall installed on the web servers to protect the server from hackers' attacks and malware. The figure 6.2 is the target of the system described by the customer. The whole process of the online selling is carried out by the IT department. The IT department is responsible for inputs of the system via the database server. Other business activities are carried out manually.

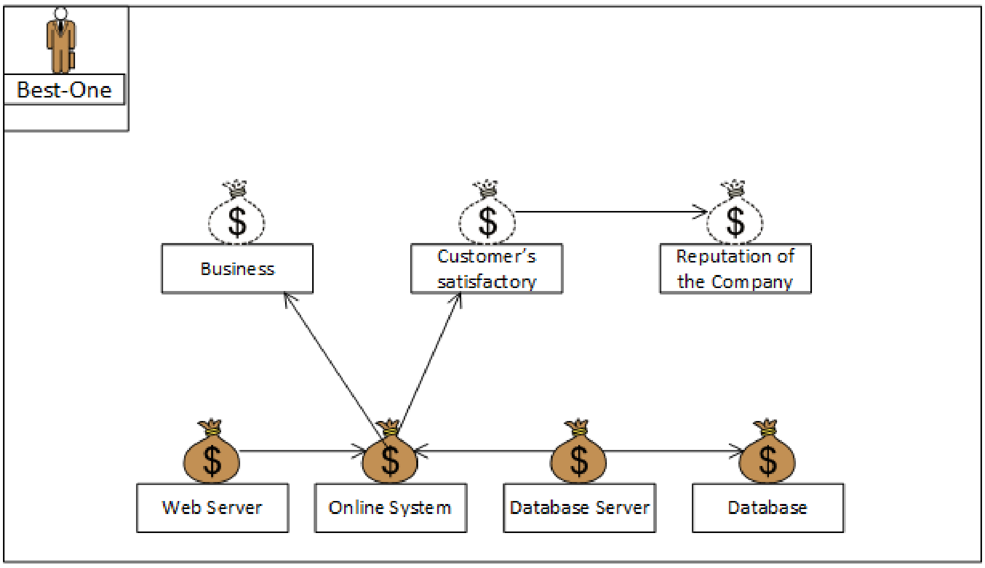


Figure 6.3: Asset diagram

Therefore, the IT department is an only internal factor interacting with the system.

6.1.3 Refining Target Description Using Asset Diagrams

6.1.3.1 Presentation of The Target by the Analysis Team

Based on the target described by the customer from previous step, the analysis team conceptually models the target to ensure that they understand customer's presentation. The UML class diagram below describes the target from the analysis team's point of view.

6.1.3.2 Asset Identification

From the target's description, the analysis team identifies direct assets and indirect assets of the Best-One's online system. At the first time, their customer (Best-One company) want to protect their online business, so that the business is the first asset they identified. They recognized that if the online system (including the servers and database) does not work properly, it absolutely affects the business. Consequently, Best-One's customers will not satisfy with the service of the company and this will affect the reputation of the company. From the point of view above, the analysis team characterizes direct and indirect asset by asset diagram as shown in the figure 6.1.3.2.

6.1.4 Approval of The Target Description

6.1.4.1 Ranking of Assets

The assets from the asset diagram (figure 6.1.3.2) are categorized based on their importance by ranking scale with which values ranges from 1 (very

Asset	Ranking	Type of Asset
Online system	1	Direct
Web server	2	Direct
Database server	2	Direct
Database	2	Direct
Company's business	2	Indirect
Customer's satisfaction	2	Indirect
Company's reputation	3	Indirect

Table 6.1: Assets Ranking

important) to 5 (very low important). Consequently, the online store asset is ranked 1 because it is a core asset that directly affects Best-One's business and customer's satisfactory. In addition to that, database and the servers are ranked 2 because they directly affect the online system. The reputation asset is indicated not important than the others, so its ranking is 3. The table 6.1.4.1 summarizes the ranking of the assets.

6.1.4.2 Setting Up Likelihood Scale

The purpose of establishing likelihood scale is to estimate frequency or probability of threat scenarios and unwanted incidents. The likelihood scale should be consistent throughout the process of risk analysis [25, p.119]. Therefore, we have only one likelihood scale.

As we mentioned above, in order to apply our approach the risk analysis team must construct membership functions by applying construction methods, and this step is dedicated on that task. However, it is outside of our research, for that reason we derives the likelihood scale from [25, p.120] and apply strategy method to obtain membership functions (section 6.2). The table 5.1 is the definition of the likelihood scale which is applied to our experimentation.

6.1.4.3 Setting Up Consequence Scale

Consequence scale is used to facilitate estimation of risk, and there are many instances of consequence scale because it depends on assets [25, p.116].

6.1.4.4 Setting Up Conditional Likelihood Scale

CORAS proposes probability scale ranging from $[0, 1]$ to assign conditional likelihood scale [25, p.151]. This step is a refinement of CORAS step to define conditional likelihood scale. Similar to the likelihood scale, the risk analysis team must assign membership functions. We derive the conditional likelihood scale from the table 5.2 and apply to our experimentation.

6.1.4.5 Setting Up Relation Scale

Similar to likelihood and conditional likelihood scale, this step is an extension of the CORAS approach in order to integrate fuzzy logic into the process of risk assessment. The relation scale is defined in the table **??**.

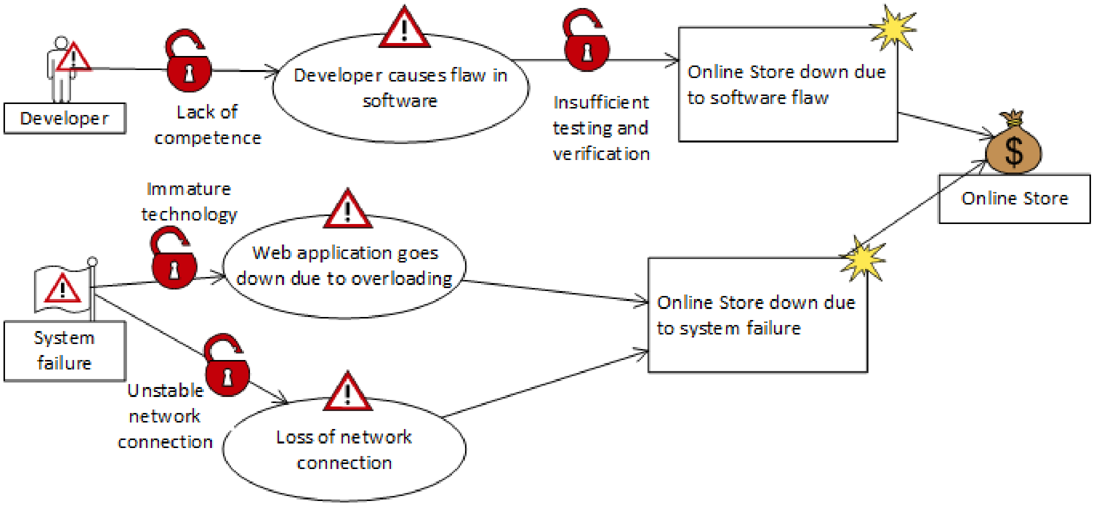

Developer
Lack of competence
Developer causes flaw in software
Insufficient testing and verification
Online Store down due to software flaw
Online Store
Immature technology
Web application goes down due to overloading
System failure
Unstable network connection
Loss of network connection
Online Store down due to system failure

Figure 6.4: Threat diagram

6.1.5 Risk Identification Using Threat Diagrams

The analysis team first focus on the online store system because it is the most important assess in the analysis. Based on that, they identify that there are some major unwanted incidents that relate to the online store system, for instance, the online store can be down due to failure of the servers and databases. The web server can be attacked by external factors such as hacker, while the database server can be failure by internal factors such as insufficient training employees may introduce malware via their personal memory devices. In addition to that, the online store system may have its own issues during its development process such as flaw in software created by incompetence developer, the system does not implement mechanism for dealing with system overloading. The process of risk identification is iterative until the analysis team covers all of possible factors that constitute to risks. The two threat diagrams below summarize the task of this step. The below example is extracted from [25]:

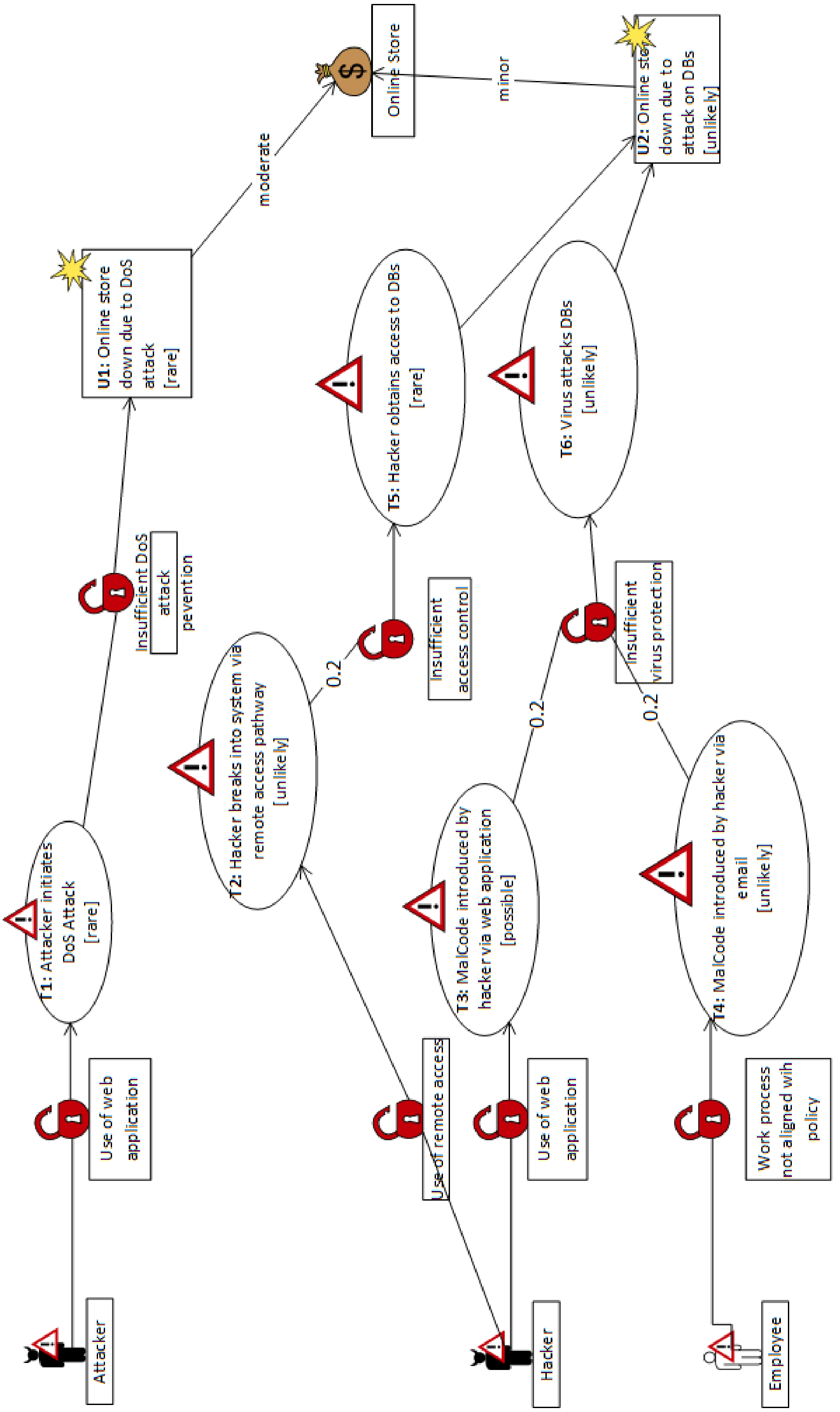

Attacker
Hacker
Employee
Use of web application
Use of remote access
Use of web application
Work process not aligned wih policy
T1: Attacker initiates DoS Attack [rare]
T2: Hacker breaks into system via remote access pathway [unlikely]
T3: MalCode introduced by hacker via web application [possible]
T4: MalCode introduced by hacker via email [unlikely]
Insufficient DoS attack pevention
0.2
Insufficient access control
0.2
0.2
Insufficient virus protection
T5: Hacker obtains access to DBs [rare]
T6: Virus attacks DBs [unlikely]
U1: Online store down due to DoS attack [rare]
U2: Online store down due to attack on DBs [unlikely]
moderate
minor
Online Store

Figure 6.5: Threats diagram

6.2 Strategy to assign membership functions

As mentioned above, membership functions should not be ambiguously assigned, but they should be elicited from the construction methods (section 4.4.7). However, construction methods are out of our scope. Therefore, we derive the original scales from [25] and hypothetically establish membership functions by re-scaling strategies.

6.2.1 Re-scaling supports of membership function

With respect to trapezoidal membership function, supports of intervals are steadily decreased by two parameters d_L and d_R. The parameters d_L, d_R are determined by left neighbor and right neighbor of the interval respectively.

Given three intervals (x_1, x_2), (x_2, x_3), (x_3, x_4) represented by the figure 6.6 and 6.7, and $x_1 < x_2 < x_3 < x_4$.

The left support d_L is determined by the following expression:

$$d_L = \frac{x_2 + x_1}{2}$$

Similarly, the right support of the trapezoidal membership function is calculated as the following:

$$d_R = \frac{x_4 + x_3}{2}$$

After obtaining the two parameters d_L and d_R, the trapezoidal membership function is constructed as the following:

$$A_1 = [d_L, x2, x3, d_R].$$

To obtain a new membership function, we perform a computation as below:

$$A_2 = [\frac{d_L + x_2}{2}, x2, x3, \frac{d_R + x_3}{2}].$$

Similarly, support of triangular membership function are obtained in the same manner. However, the top of the triangular membership function is the middle point of the interval.

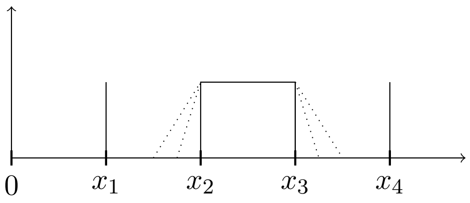


Figure 6.6: Re-scaling trapezoidal membership function

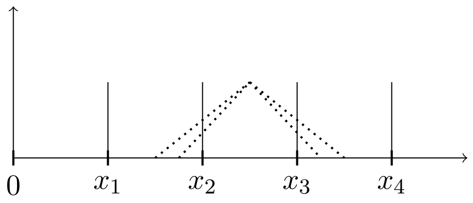


Figure 6.7: Re-scaling of triangular membership function

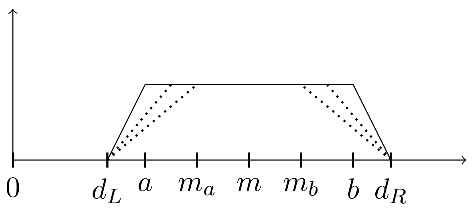


Figure 6.8: Narrowing core of trapezoidal membership function

6.2.2 Re-scaling core of membership function

This strategy applies only to the trapezoidal membership function. Assume that the two parameters d_L and d_R are obtained and trapezoidal membership function is constructed as in the above section.The core of the membership function is determined by a parameter m which is calculated by middle point of interval (a, b).

$$m = \frac{a+b}{2}$$

The parameter m divide the interval (a, b) into two sub intervals (a, m) and (m, b). For each sub interval, we continue to divide it into two sub intervals as below:

$$\begin{cases} m_a = \frac{a+m}{2} \\ m_b = \frac{m+b}{2} \end{cases}$$

If the procedure is repeated, we obtain a set of membership functions that their core are gradually increased as shown in the figure 6.8.

6.3 Risk Estimation Using Threat Diagrams

This is the main step when we actually apply our approach to estimate and reason likelihoods of threat scenarios, unwanted incidents constructed by those steps above. In this step, we will present different experiments to compare results of various membership functions obtained by the strategies to construct membership function.

6.3.1 Applying Rule for Leads-to

At this point, the analysis team has obtained the whole background in order to estimate likelihoods and risk levels. They first focus on the threats from the threat diagram and assign frequencies to each threat. Actually, the relationship between threats and threat scenarios is initiate relation, by applying the rule for reasoning likelihood of initiate relation, they reason that frequency of the threat scenario is also frequency of the threat which initiates the threat scenario. Additionally, they assign conditional likelihoods to relationships between threat scenarios, threat scenarios and unwanted

Threat scenario	Likelihood	Conditional likelihood
T1: Attacker initiates DoS attack	Rare	Medium
T2: Hacker breaks into the system via remote access	Unlikely	Low
T3: Malcode introduced by hacker via web application	Possible	Low
T4: Malcode introduced by hacker via email	Unlikely	Low
T7: Developer causes flaw in software	Possible	High
T8: Web application goes down due to overloading	Possible	Medium
T9: Loss of network connection	Unlikely	Low

Table 6.2: Initial likelihoods and conditional likelihoods from the threat diagrams

incidents. The table 6.2 summarizes the result of this step.

After assigning likelihoods and conditional likelihoods to initial threats and threat scenarios. The next step is to estimate likelihoods of dependent threat scenarios and unwanted incidents. They first apply rule for leads-to to threat scenarios and unwanted incidents that include that type of relation. The results are presented in the following experiments.

6.3.1.1 Experimentation 1: Applying trapezoidal membership function

In this experimentation, we apply trapezoidal membership functions to likelihood and conditional likelihood scale. The experimentation is divided into two groups. Group A involves to the strategy to extend supports of trapezoids, group B is to conduct experiments that narrow the core of trapezoidal membership function. While the conditional likelihood scale are steadily increased by extending their supports or decreased by narrowing their core, the likelihood is remained the same in all experiments.

Extending supports of trapezoidal membership functions The likelihood scale is derived from the table 5.1, by applying the strategy for constructing membership function, each interval is extended supports as shown in the figure 6.9.

In the same manner, the conditional likelihood scale is derived from table 5.2 and fuzzified by applying the strategy to construct membership function. The figure 6.10 is the result of the operation. The table 6.3 describes fuzzified scale in term of trapezoid's representation.

First, the analysis team focus on leads-to relation of threats, threats scenarios and unwanted incidents to estimate likelihood of dependent threat scenarios and unwanted incidents. They apply rule for leads-to to threat scenario T_1, T_2, T_5 to calculate likelihood of U_1, T_2, U_3 respectively. T_3 and

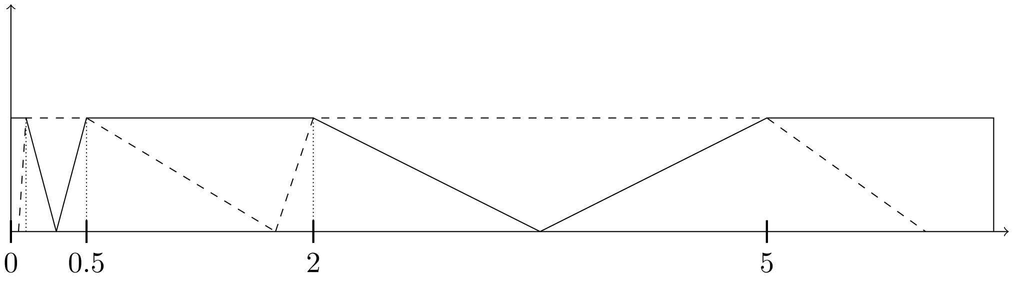


Figure 6.9: Membership functions of likelihoods

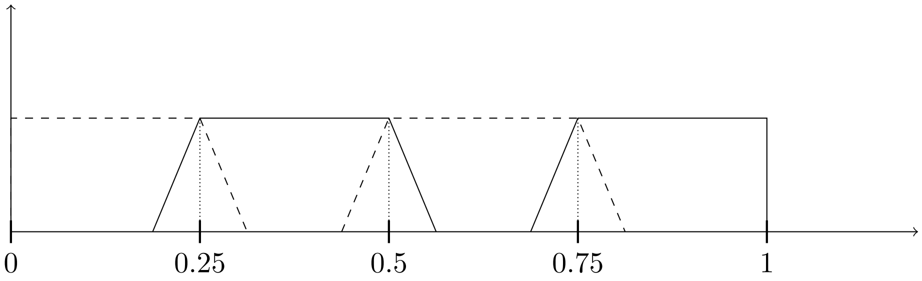


Figure 6.10: Membership functions of conditional likelihoods

Term	Fuzzified Scale
Rare	[0, 0, 0.1, 0.3]
Unlikely	[0.05, 0.1, 0.5, 1.75]
Possible	[0.3, 0.5, 2, 3.5]
Likely	[1.75, 2, 5, 7.5]
Certain	[3.5, 5, 9, 9]

Table 6.3: Trapezoidal membership functions of likelihood

Term	**Experiment A1**	**Experiment A2**
Low	[0, 0, 0.25, 0.3125]	[0, 0, 0.25, 0.375]
Medium	[0.1875, 0.25, 0.5, 0.5625]	[0.125, 0.25, 0.5, 0.625]
High	[0.375, 0.5, 0.75, 0.8125]	[0.3125, 0.5, 0.75, 0.875]
Critical	[0.5625, 0.75, 1, 1]	[0.625, 0.75, 1, 1]

Table 6.4: Conditional likelihood scales for experimentation group A

Input	Likelihood	Conditional Likelihood	CORAS method	Simple Fuzzy method	General Fuzzy method
Experiment A1					
T1, C1	Rare	Low	0.0125	0	0.0083
T2, C2	Unlikely	Low	0.0625	0.0585	0.1705
T3, C3	Possible	High	0.875	0.8974	1.2143
Experiment A2					
T1, C1	Rare	Low	0.0125	0	0.0138
T2, C2	Unlikely	Low	0.0625	0.0693	0.1969
T3, C3	Possible	High	0.875	0.9302	1.2714

Table 6.5: Result of Experimentation Group A

Term	Experiment B1	Experiment B2
Low	[0, 0, 0.1875, 0.375]	[0, 0, 0.125, 0.375]
Medium	[0.125, 0.28125, 0.46875,0.625]	[0.125, 0.3125, 0.4375, 0.625]
High	[0.3125, 0.53125, 0.71875, 0.875]	[0.3125, 0.5625, 0.6875 0.875]
Critical	[0.5625, 0.78125, 1, 1]	[0.625, 0.8125, 1, 1]

Table 6.6: Experimentation Group B

T_4, T_6 and T_7 are treated as the case of rule for separate. The table 6.3.1.1 is the result of applying both CORAS and fuzzy method.

Narrowing core of trapezoidal membership functions In order to conduct the experimentation, the interval of likelihoods are first extended their support and steadily decreased their core while keeping their support constantly. The conditional likelihood scale is of the same as in experimentation group A. The table 6.6 represents the scales applied to this experimentation.

Input	Likelihood	Conditional Likelihood	CORAS method	Simple Fuzzy method	General Fuzzy method
Experiment B1					
T1, C1	Rare	Low	0.0125	0	0.0119
T2, C2	Unlikely	Low	0.0625	0.0636	0.1877
T3, C3	Possible	High	0.875	0.9217	1.2595
Experiment B2					
T1, C1	Rare	Low	0.0125	0	0.0104
T2, C2	Unlikely	Low	0.0625	0.0564	0.1804
T3, C3	Possible	High	0.875	0.9112	1.2483

Table 6.7: Result of Experimentation Group B

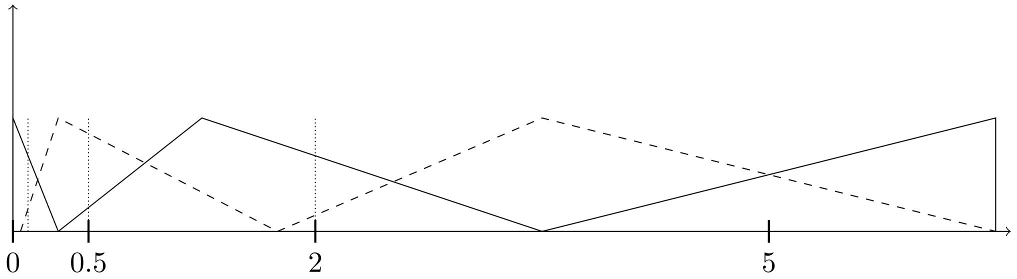


Figure 6.11: Membership functions of likelihoods by triangular membership functions

Term	Fuzzified Scale
Rare	[0, 0, 0.3]
Unlikely	[0.05, 0.3, 1.75]
Possible	[0.3, 1.25, 3.5]
Likely	[1.75, 3.5, 7.5]
Certain	[3.5, 5, 9, 9]

Table 6.8: Triangular membership functions of likelihood

6.3.1.2 Experimentation 2: Applying triangular membership function

In this experimentation, we conduct two experiments that adopt triangular membership function and apply to both likelihood scale and conditional likelihood scale. Top point of triangular membership functions are middle point of interval, support of the function is calculated by strategy to assign membership function. The figure 6.11 represents the result of applying the strategy.

6.3.2 Applying Rule for Separate

The risk analysis team examines the threat diagram and focuses on threat scenarios that lead to the same threat scenarios or unwanted incidents. The threat scenario T_3, T_4 and T_6, T_7 are inspected because they lead to the same threat scenario (unwanted incident). They agree that both T_3 and T_4 are overlapping in their content because they are caused by malcode and hacker. While the threat scenario T_6 and T_7 are clearly separate because web application goes down due to overloading is absolutely not relevant to

Term	**Experiment C1**	**Experiment C2**
Low	[0, 0, 0.3125]	[0, 0, 0.375]
Medium	[0.1875, 0.375, 0.5625]	[0.125, 0.375, 0.625]
High	[0.375, 0.625, 0.8125]	[0.3125, 0.625, 0.875]
Critical	[0.5625, 1, 1]	[0.625, 1, 1]

Table 6.9: Experimentation Group C

Input	Likelihood	Conditional Likelihood	CORAS method	Simple Fuzzy method	General Fuzzy method
Experiment C1					
T1, C1	Rare	Low	0.0125	0	0.0023
T2, C2	Unlikely	Low	0.0625	0.0392	0.1309
T3, C3	Possible	High	0.875	0.8594	1.2103
Experiment C2					
T1, C1	Rare	Low	0.0125	0	0.0072
T2, C2	Unlikely	Low	0.0625	0.0461	0.1347
T3, C3	Possible	High	0.875	0.8930	1.2629

Table 6.10: Result of Experimentation Group C

loss of network connection.

The analysis team is not sure that the threat scenario T_3 and T_4 are completely overlap or partial separate. For that reason, the team estimate the aggregation of the two threat scenarios by dividing it into two cases. The first case is overlap, and the last is partial separate. They apply the rule for separate to calculate the aggregation of the threat scenarios T_3, T_4 and T_5, T_6.

In this experimentation, the likelihood scale is adopted from the experiment A_2 and remained constantly. While the relation scale is adjusted by strategy to assign membership function. Additionally, the two types of membership functions are employed, and we conduct four experiments. The first experiment is apply the trapezoidal membership function, the second one involves applying triangular membership function, the third one is to apply trapezoidal membership function of group D to interval scale, the last one employs triangular membership function of group E to interval scale. For that reason, the objective of the experimentation is to compare the results with CORAS, and verify the criteria of rule for separate with different types of membership functions.

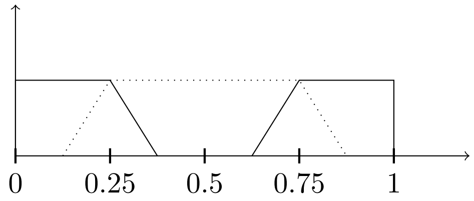


Figure 6.12: Trapezoidal membership functions of relation scale

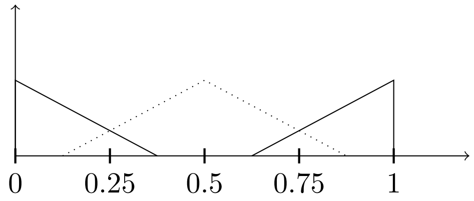


Figure 6.13: Triangular membership function of relation scale

6.4 Comparison

In the previous section, we have conducted five groups of experimentation. While group A, B, C relate to rule for leads-to, group D and E are intended for rule for separate. Overall, the results generated by group A, B, C are not

Term	**Group D**	**Group E**
Overlap	[0, 0, 0.25, 0.375]	[0, 0, 0.375]
Partial Separate	[0.125, 0.25, 0.75, 0.875]	[0.125, 0.5, 0.875]
Separate	[0.625, 0.75, 1, 1]	[0.625, 1, 1]

Table 6.11: Membership functions of relation scale

Input	**Likelihood**	**Overlap**	**Partial separate**	**Separate**
Experiment Group D				
T3, T4	Possible, Unlikely	1.7785	2.1616	2.2980
T8, T9	Possible, Rare	1.6046	1.6576	1.6826
Experiment Group E				
T3, T4	Possible, Unlikely	1.7292	2.1135	2.2980
T8, T9	Possible, Rare	1.5958	1.6434	1.6783
Experiment Group F				
T3, T4	Possible, Unlikely	1.2906	1.5002	1.5697
T8, T9	Possible, Rare	1.2238	1.2592	1.2834
Experiment Group G				
T3, T4	Possible, Unlikely	1.2583	1.4694	1.5697
T8, T9	Possible, Rare	1.2238	1.2511	1.2834

Table 6.12: Result of experimentation of rule for separate

considerably different. Additionally, it is the same as in the case of group D and E. This means that the application of trapezoidal and triangular membership function to the domain of risk assessment is almost equivalent.

6.4.1 Rule for leads-to

In experiment group A, when extending supports of trapezoidal membership function the results are gradually increased corresponding to the increment of the support. In contrast, when narrowing core of trapezoid in group B the results are declined slightly. Similarly to group A, when extending the supports of triangle, the results also increase hardly, but the results of triangular membership functions are insignificantly smaller than that of trapezoidal membership functions.

General fuzzy method always yield greater results than that of the CORAS and simple fuzzy method except for the case of rare. The reason is that rare is a special interval and limited significantly. When applying fuzzy product to the interval of rare, the area of the membership function representing the result is considerably narrowed when comparing with the product of two interval. Therefore, the result of general fuzzy method is slightly smaller that that of the CORAS in case of the rare interval.

Simple fuzzy method yield less significant different results than the CORAS. Generally, it is slightly smaller than the CORAS method. In some cases, for example, the case of possible and high of the experiment C2, it is marginally significant than the CORAS.

6.4.2 Rule for separate

As in the case of rule for leads-to, the use of trapezoidal and triangular membership function in this experimentation is not significantly different. Similar to the case of rule for leads-to, the general fuzzy method yields slightly greater results than that of the simple fuzzy method when comparing experiment group D and F, group E and G.

Our proposed approach satisfies the criterion to estimate likelihood of separate and overlap threat scenarios. In addition to that, the proposed approach presents the case of partial separate while CORAS only offers two formulas for separate and overlapping. By applying CORAS method to calculate likelihood possible and likely in the case of separate and overlapping, the results are 1.55 and 1.25 respectively. Those are significantly smaller than that of our approach.

Chapter 7

Implementation of The Tool

In this chapter, we present the steps which were conducted to implement the tool. First, the requirement analysis is described in order to define the functions of the tool, design alternatives to implement the tool. Next, the design of the tool includes design of class diagram, architecture, sequence diagram and user interface are presented in detail.

7.1 Requirement Analysis

This phase is conducted after each new artefact has been obtained from the technology research process, and to ensure that the implementation of the tool is feasible, the functions of the tool are described in a manner that it is implementable.

7.1.1 Defining the functions

Based on the artefact obtained from the technology research process, the tool should be implemented in the manner that it supports calculating rule for leads-to and rule for separate by offering alternatives to estimate likelihood as proposed by the approach. Additionally, the tool should have a feature that allows user to define likelihood, conditional likelihood and relation scale. However, the tool should not support the user to construct membership functions based on strategies, or fuzzify extreme intervals. The reason is that membership function should be elicited from construction methods, and there are automatic methods to develop membership functions such as neutral network, etc [37]. In addition to that, our proposed method for fuzzifying extreme interval should be proved that it is applicable and intuitive to the domain of risk assessment, and the task is empirical and outside of the scope. As a result, attempting to implement these functions is actually redundant and inefficient. The table 7.1 describes the functions should be offered by the tool.

Function	Description
Setting Up Scales	User sets up likelihood, conditional likelihood and relation scale in order to conduct calculations
Rule for leads-to	Rule for leads-to uses likelihood scale and conditional likelihood scale to reason likelihood of threat scenario or unwanted incident
Rule for separate	Rule for separate uses likelihood scale and relation scale to reason likelihood of two threat scenarios
Alternatives	This option allows user to choose between applying fuzzy approach and CORAS approach

Table 7.1: Description of the Tool's Functions

7.1.2 Supported libraries for development of the tool

7.1.2.1 Fuzzy logic programming library

To facilitate and simplify our development process, we utilize Matlab as an third-party library which provides a foundation for fuzzy logic programming. The drawback of utilizing Matlab is that we need to install Matlab as a prerequisite to use the tool. The reason is that we have not found any fuzzy logic programming library which supports our domain of interest. Some of fuzzy logic libraries facilitate the implementation of fuzzy logic controllers which are not appropriate to develop the tool. Additionally, many of them are underdeveloped, or outdated [6]. However, Matlab provides powerful fuzzy logic library supporting a foundation of fuzzy logic operators, fuzzy set, membership functions and defuzzification methods. Those of them are complicated to implement, and out of our scope. Despite the fact that there are many approximation methods to implement fuzzy numbers [1] and COA defuzzification [18], attempting to implement those is still erroneous and generates unexpected calculations as we first tried to develop our fuzzy logic library.

The use of Matlab in our project is facilitated by a middleware, that is MatlabControl JMI Wrapper [8, 46]. The middleware plays an interface role transferring parameters, commands from the Tool to Matlab and receiving results from Matlab.

7.1.2.2 Visual programming environment

In addition to Matlab, we adopt Eclipse to fasten the development of the tool. Eclipse is an integrated development environment which provides a full-supported java programming framework supporting and managing the development of software projects from small scale to large scale [10].

7.2 Design of The Tool

We adopt Unified Modeling Language 2.0 (UML) to model and represent the architecture and the design of the Tool.

7.2.1 The Architecture of The Tool

To develop a flexible, maintainable and adaptable software, we employ the Model-view-controller (MVC) architecture to structure our tool's components. There are many alternatives to implement the MVC architecture [5]. Basically, the role of each component is described as below [5].

- Model represents the domain of interest, in our case the class diagram from figure 7.2.
- View is basically user interface where user interacts with the system.
- Controller plays the role as the bridge between the model and view. It receives and processes user's commands, updates the views and modifies the model.

The figure 7.1 is our implemented MVC for the tool. As depicted by the figure 7.1, the controller is the bridge between the model and the view, it is also responsible for interacting with external components.

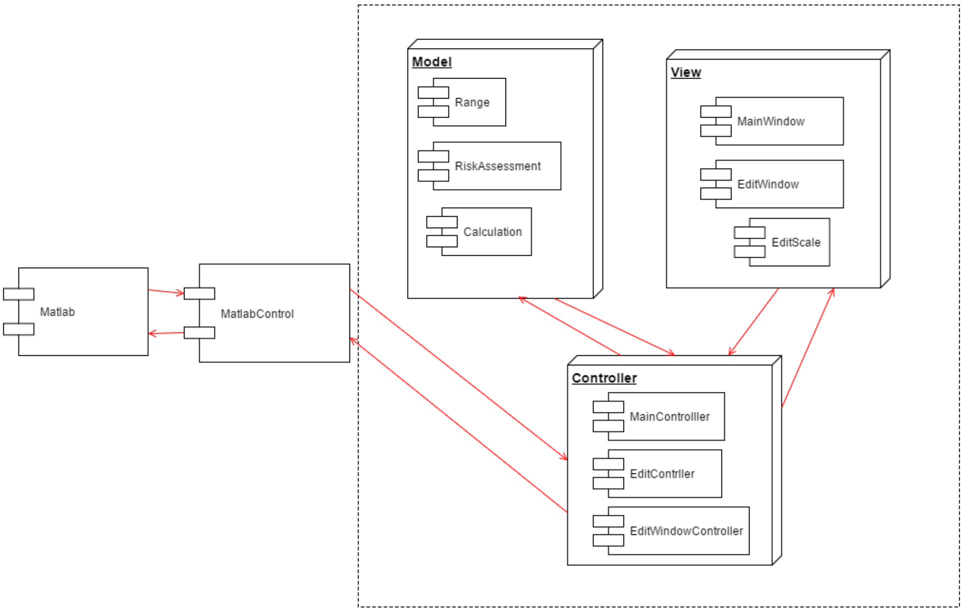


Figure 7.1: The Tool's Architecture

7.2.2 Class Diagram

We first present our domain of interest that is the representation of intervals, likelihoods and methods for calculation, etc. The figure 7.2 is the whole class diagram that represents our domain of interest.

The class *RiskAssessment* is a domain of application which includes ranges and calculation methods. *Range* is a base class representing intervals, trapezoidal and triangular membership function are two sub classes that are inherited from *Range*. *Calculation* is a parent class of *GeneralFuzzyCalculation*, *SimpleFuzzyCalculation* and *IntervalCalculation*, each sub class of *Calculation* presents a typical method to calculate likelihoods.

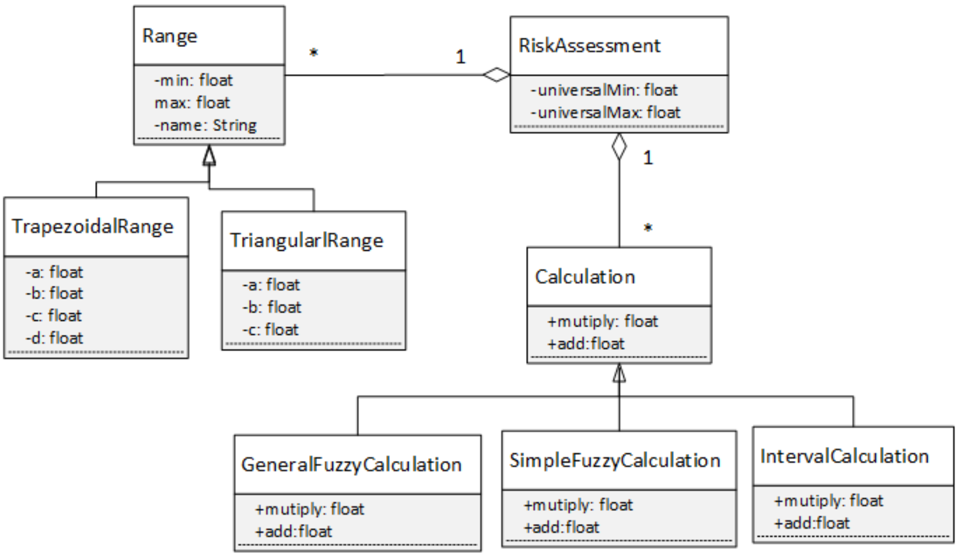


Figure 7.2: Class Diagram for Risk Assessment based on Fuzzy Logic and CORAS

7.2.3 Sequence Diagram

The tool is a simple application that allows users to calculate likelihood based on the approach. This is the first iteration of the software process when our approach proposes fuzzy logic in the rule for leads-to. As shown in the figure 7.3, the user first defines scale of likelihood and conditional likelihood in order to apply the approach. After that, the user calculates likelihoods by choosing the inputs for likelihood and conditional likelihood. The diagram for rule for separate is similar to the rule for leads-to diagram except that it requires user to define relation scale, and it has three parameters to estimate likelihood.

The diagram corresponds to our proposed approach in the manner that the step to define scales should be completed before estimation of likelihoods.

7.2.4 User Interface Design

The tool contains four windows, they are the main window and sub-windows for editing the likelihood, conditional likelihood, and relation scale. The figure 7.4 is the user interface structure of the tool.

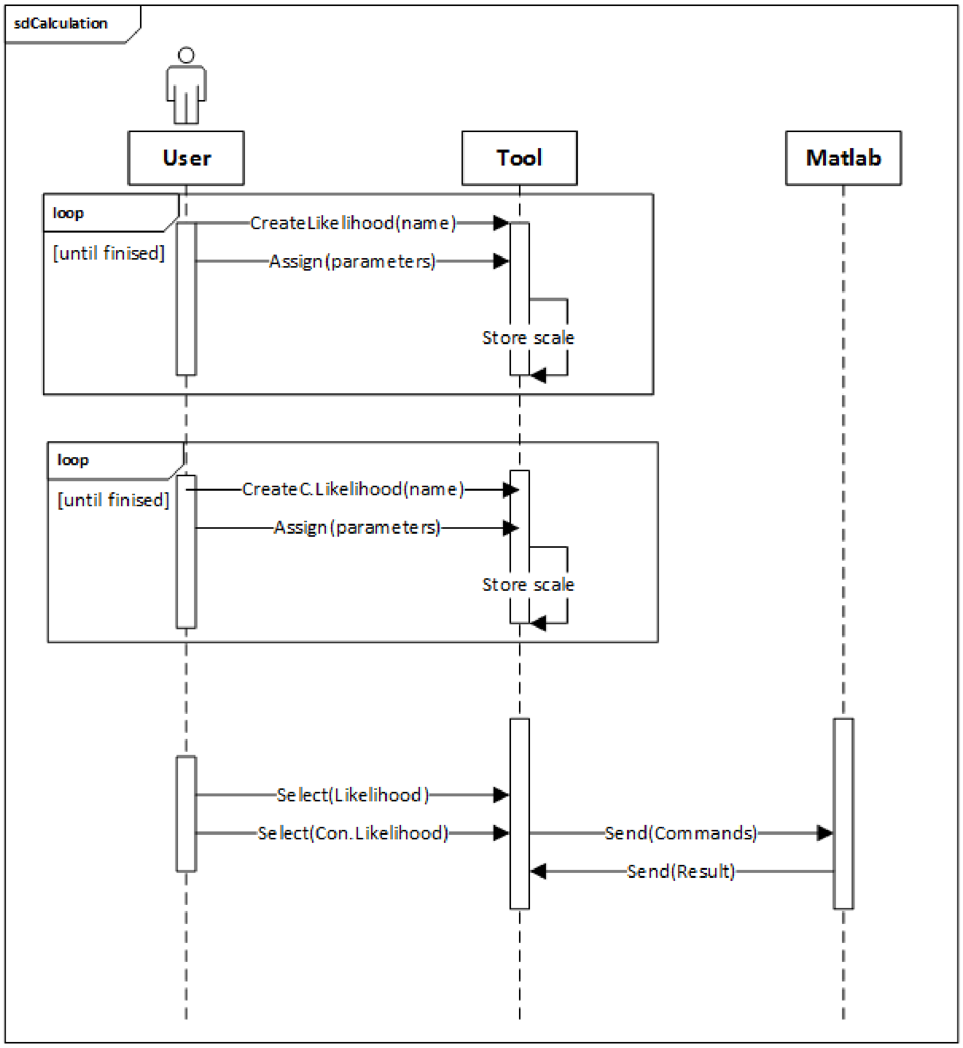


Figure 7.3: Sequence Diagram represents step to estimate likelihood by applyig rule for leads-to

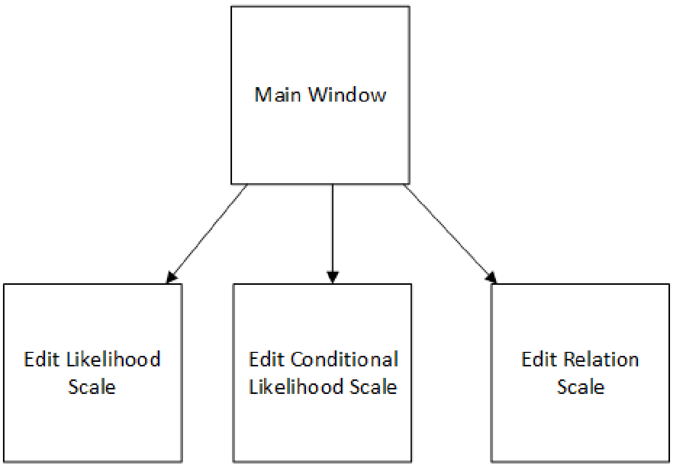
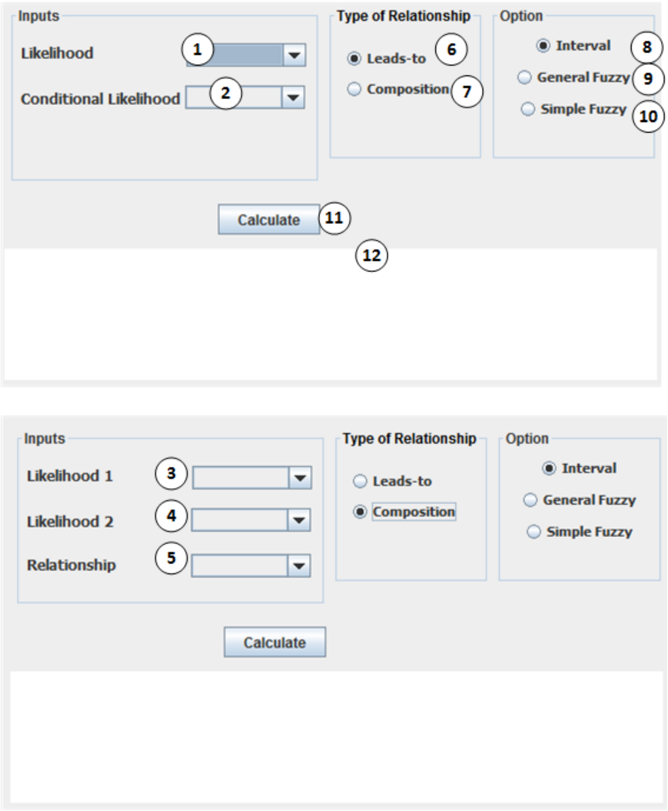


Figure 7.4: User Interface Structure of The Tool



Inputs
Likelihood
Conditional Likelihood

Type of Relationship
● Leads-to
○ Composition

Option
● Interval
○ General Fuzzy
○ Simple Fuzzy

Inputs
Likelihood 1
Likelihood 2
Relationship

Type of Relationship
○ Leads-to
● Composition

Option
● Interval
○ General Fuzzy
○ Simple Fuzzy

Figure 7.5: Main Window of the Tool

7.2.4.1 The main window

The main window includes input panel to select input values, option panel for choosing between fuzzy methods and CORAS method, a relationship panel representing two types of rule for reasoning likelihood, and a screen to display results. When the user selects Types of Relationship, the input panel is changed according to the rule for reasoning likelihood as in the figure 7.5.

The table 7.2 is the specification of the main window.

7.2.4.2 The sub-windows

The sub windows have similar structure, but representing different functions. Each sub window has a panel to define variables (likelihood, conditional likelihood, relation), each variable is represented by its name, original interval and membership function. The table **??** is the specification of the sub windows.

No.	**Control type**	**Label**	**Data type**	**Describe**
1	Combo box	Likelihood	List of likelihoods	Input for likelihood of rule for leads-to
2	Combo box	Conditional likelihood	List of conditional likelihoods	Input for conditional likelihood of rule for leads-to
3	Combo box	Likelihood 1	List of likelihoods	Input for likelihood of rule for separate
4	Combo box	Likelihood 2	List of likelihoods	Input for likelihood of rule for separate
5	Combo box	Relationship	List of relationships	Input for relation
6	Radio button	Leads-to	Boolean	Option for rule for leads-to
7	Radio button	Composition	Boolean	Option for rule for separate
8	Radio button	Interval	Boolean	Option to estimate likelihood by CORAS method
9	Radio button	General Fuzzy	Boolean	Option to estimate likelihood by general fuzzy method
10	Radio button	Simple Fuzzy	Boolean	Option to estimate likelihood by simple fuzzy method
11	Button	Calculate		
12	Text Area		Text	Output for result

Table 7.2: Description of controls for the main window

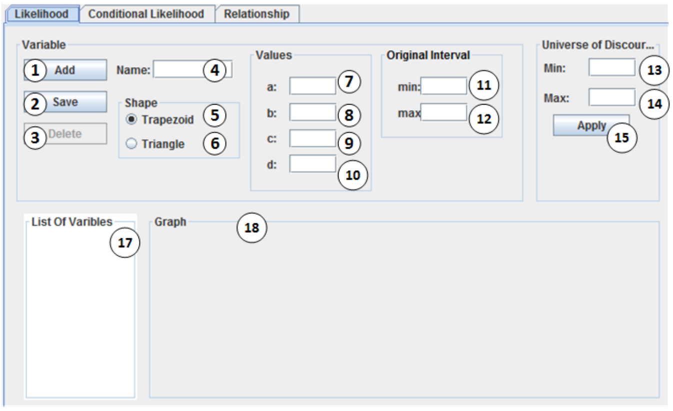


Figure 7.6: Edit Window of the Tool

No.	**Control type**	**Label**	**Data type**	**Describe**
1	Button	Add		Create new range
2	Button	Save		Save range to list of variables
3	Button	Delete		Delete selected range on list of variable
4	Text field	Name	Text	Name of variable
5	Radio button	Trapezoid	Boolean	Option for trapezoidal membership function
6	Radio button	Triangle	Boolean	Option for triangular membership function
7	Text field	a	Boolean	Parameter for membership fucntion
8	Text field	b	Boolean	Parameter for membership fucntion
9	Text field	c	Boolean	Parameter for membership fucntion
10	Text field	d	Boolean	Parameter for membership fucntion
11	Text field	min	Text	Parameter for original interval
12	Text field	max	Text	Parameter for original interval
13	Text field	Min	Text	Minimum of universe of discourse
14	Text field	Max	Text	Maximum of universe of discourse
15	Button	Apply		Save the universe of discourse
16	List box	List of Variables	List of ranges	List of defined ranges
17	Canvas	Graph	Drawing	Display membership functions by graph

Table 7.3: Description of controls for the sub windows

Chapter 8

Evaluation of The Tool

In this chapter, we describe the evaluation of the tool by the testing techniques.

8.1 Equivalence partitioning testing

The equivalence partitioning technique focuses on the input domains of the tool, and it is black-box method. Therefore, test cases are designed based on functional specification of the tool. As shown in the figure 7.5 and 7.6 which describe the user interfaces of the tool. The main window has entries for user to select value of likelihood and conditional likelihood if rule for leads-to is activated, and options for different methods of calculation. However, the likelihood and conditional likelihood are defined in the sub windows. Therefore, those should not to be tested, because they are valid by default if likelihood and conditional likelihood scale are defined reasonably. Consequently, we only need to test only validity of original interval, trapezoidal and triangular membership function defined in the sub-windows.

8.1.1 Test cases for original interval

An interval is presented by two attributes, that are the minimum and maximum. Test cases are divided into two classes, valid intervals where the minimum is less than the maximum, and invalid intervals where the minimum is greater than the maximum.

$I_1 = \{< a, b >: a \leq b\}$

$I_2 = \{< a, b >: a \geq b\}$

The table 8.1 is two test cases that are derived from the equivalence classes I_1, I_2.

Test Case	**Input**	**Expected Output**
I_1	$a = 0.25, b = 0.5$	Valid
I_2	$a = 0.5, b = 0.25$	Invalid

Table 8.1: Test cases for original interval

Test Case	**Input**	**Expected Output**
R_1	$a = 0.25, b = 0.5, c = 0.75$	Valid
R_2	$a = 0.1, b = 0.2, d = 0.3$	Valid
R_3	$a = 0.1, b = 0.4, c = 0.5$	Valid
R_4	$a = 0.4, b = 0.1, c = 0.2$	Invalid

Table 8.2: Test cases for triangular membership function

Test Case	**Input**	**Expected Output**
X_1	$a = 0.25, b = 0.5, c = 0.75, d =$	Valid
X_2	$a = 0.1, b = 0.2, d = 0.3$	Valid
X_3	$a = 0.1, b = 0.4, c = 0.5$	Valid
X_4	$a = 0.4, b = 0.1, c = 0.2$	Invalid
X_5	$a = 0.4, b = 0.1, c = 0.2$	Invalid

Table 8.3: Test cases for triangular membership function

8.1.2 Test cases for triangular membership function

Triangular membership functions have three types, they are equilateral, isosceles and scalene. Therefore, we have four equivalence classes including three types of triangle and not a triangle.

$R_1 = \{< a, b, c >$: a, b, c form equilateral triangle $\}$
$R_2 = \{< a, b, c >$: a, b, c form isosceles triangle $\}$
$R_3 = \{< a, b, c >$: a, b,c form scalene triangle $\}$
$R_4 = \{< a, b, c >$: a, b, c do not form triangle $\}$

The table 8.3 describes four test cases corresponding to the four equivalence classes R_1, R_2, R_3, R_4.

8.1.3 Test cases for trapezoidal membership function

Trapezoid has four geometric shapes which are right side (left side), equilateral, isosceles and scalene trapezoid. Test cases are divided into six equivalence classes that includes four types of trapezoid and one class that is not trapezoid.

$X_1 = \{< a, b, c, d >$: a, b, c form equilateral trapezoid $\}$
$X_2 = \{< a, b, c, d >$: a, b, c form isosceles trapezoid $\}$
$X_3 = \{< a, b, c, d >$: a, b, c form scalene trapezoid $\}$
$X_4 = \{< a, b, c, d >$: a, b, c form right (left) side trapezoid $\}$
$X_5 = \{< a, b, c, d >$: a, b, c do not form trapezoid$\}$

8.2 Scenario-based Testing

We design three test cases that cover use case scenarios of the tool. Three test cases ensure that

Step	Description	Input	Expected output
1	Select Likelihood in the edit menu		The sub window for defining likelihood is opened
2	Select trapezoid in the group box "Shape"		The radio button "Trapezoid" is selected
3	Enter name in the text box "Name"	Name = Rare	
4	Enter value in the text boxes a, b, c, d	a = 0, b = 0.025, c = 0.075, d = 0.3	
5	Select Save button		Rare is shown in the list box "List of Variables"
6	Enter name in the text box "Name"	Name = Unlikely	
7	Enter value in the text boxes a, b	a = 0, b = 0.05	
8	Select Save button		The scale is not valid
9	Turn off the sub window		Rare is shown in the combo box Likelihood

Table 8.4: Scenario to define scales

8.2.1 Test case for defining scales

The purpose of this test case is to verify that the tool supports defining scales as stated in the requirement specification. Additionally, the test case covers some mistakes that the user commonly makes.

8.2.2 Test case for calculating rule for leads-to

This test case is to verify that the tool supports rule for leads-to. Prerequisite to conduct the test case is that likelihood scale and conditional likelihood scale are defined in advance. The scales are employed from the experiment A_2 in the section 8.5.

8.2.3 Test case for calculating rule for separate

This test case is to verify that the tool supports rule for separate. Prerequisite to conduct the test case is that likelihood scale and relation scale are defined in advance. The test case is similar to test case for rule for leads-to in the table 8.5.

Step	**Description**	**Input**	**Expected output**
1	Select Leads-to in the group box Type of Relationship		The radio button Leads-to is selected
2	Select Interval in the group box Option		The radio button Interval is selected
3	Select likelihood from combo box Likelihood	Likelihood = Rare	The combo box Likelihood shows Rare
4	Select likelihood from combo box conditional likelihood	Conditional likelihood = Medium	The combo box conditional likelihood shows Medium
5	Press Calculate button		The screen show result of 0.0125
6	Select General Fuzzy in the group box Option		The radio button General Fuzzy is selected
7	Press Calculate button		The screen show result of 0.0366
8	Select Simple Fuzzy in the group box Option		The radio button Simple Fuzzy is selected
9	Press Calculate button		The screen show result of 0.0205

Table 8.5: Test case for rule for leads-to

Chapter 9

Discussion

In this chapter, we discuss the main results of our research with respect to the success criteria. In addition to that, limitation of the proposed approach and the tool are also taken into account.

9.1 The proposed approach

In chapter 2, we have presented four success criteria for our approach to combining CORAS and fuzzy logic. The first two success criteria can be proved by formulating hypotheses and applying deductive reasoning, but the last two success criteria involve empirical work to verify their validity and it is outside our scope. However, the outcomes obtained from the experimentation can facilitate the comparison of results generated by CORAS and fuzzy logic. Consequently, we can predict the patterns of likelihoods when conducting risk assessment in practice.

9.1.1 Success criterion 1: The approach must be general

The first success criterion states that the approach must be as general as the CORAS. This implies that the approach can be applied in the context of the CORAS. Actually, the proposed approach refines rule for leads-to and rule for separate by fuzzy numbers instead of intervals. Interval is a special case of fuzzy logic [38]. Therefore, our approach can be applied in the context of CORAS.

9.1.2 Success criterion 2: The approach must be sound

Our approach is developed based on the selective theoretical background of fuzzy logic and the CORAS. Bringing fuzzy logic into the two fundamental rules for reasoning about likelihood is elaborated by a mathematical background and logical reasoning as represented in the chapter 5. With respect to the rule for leads-to, results are obtained by fuzzy multiplication and the formula of CORAS. The equation of rule for leads-to has been proved that it is sound by [40]. With respect to rule for separate, we have refined the CORAS equation by the equation [], and it has been justified in the section **??** and verified by the experimentation.

9.1.3 Success criterion 3: The approach should be comprehensible and applicable

The success criterion 3 implies that the progress to adopt and apply the approach in the context of risk assessment is achievable in terms of time and effort. It requires empirical work to justify the characteristic. In this sense, measures and metrics to measure the effort and time to learn the approach should be set up in order to conclude and verify the characteristic. Actually, fuzzy logic is more complex than interval arithmetic in the sense that it requires more computation to obtain final result. However, we developed the tool in order to support fuzzy logic calculation. Although with the assistance of the tool, the risk analyst still needs to acquire membership functions and the construction methods in order to apply the approach.

9.1.4 Success criterion 4: The approach should be effective in comparison with the CORAS

The effectiveness of the approach means that it should generate reliable results and it captures the issues of risk assessment better that that of the CORAS. By introducing fuzzy logic into the rules for reasoning likelihood, the proposed approach derives the virtues of fuzzy logic to model uncertainty by membership functions. With respect to rule for leads-to, the approach proposes conditional likelihood scale to simplify the estimation of conditional likelihood. With respect to rule for separate, the approach proposed relation scale to address the issue of analyzing relationship between threat scenarios (unwanted incidents). In addition to that, the approach captures the case of partial separate while the CORAS proposes the case of separate and overlap. Therefore, proposed approach is an extension of the CORAS.

Despite the fact that the approach is an improvement of the CORAS, results generated by the approach need to be justified in order to conduct in the practical context. The outcome from the experimentation gives us an insight into the patterns of results generated by CORAS and fuzzy methods. Based on the patterns, likelihoods generated by the CORAS are always less significant than that of fuzzy methods generally with respect to rule for leads-to. With respect to rule for separate, likelihoods calculated by fuzzy methods are almost doubled the CORAS. Those imply that final outcome of the risk assessment will certainly be affected in some sense, such as risk level is increased because likelihood of unwanted incident is higher when applying fuzzy methods. However, the reliability of those results need empirical research to justify whether the CORAS or fuzzy logic satisfy the criterion.

9.1.5 Limitation of the approach

The success criteria reflect the limitation of the approach in the manner that the proposed approach needs more empirical work to justify that it can be applied in practice. Our approach inherits the virtues of CORAS and fuzzy logic to address the issue of risk assessment. Contemporaneously, it derives both pros and cons of the methods. With respect to the CORAS method, it is costly and time-consumed to conduct a risk analysis. It requires professional teams and resources in order to conduct the formal steps of the CORAS. Therefore, our research have not reached the final step of the CORAS to conclude risk levels. With respect to the fuzzy logic, the methods for construction of membership functions have not been proposed to our research yet. Additionally, the nature of defuzzification yields significantly diverse outputs depending on the methods. The comparison and selection of defuzzification methods have not been put into our research yet. The COA method is the only defuzzification method that is selected and applied in the research instead.

9.2 The tool-supported the approach

With respect to the tool, the success criteria have focused on the functions and features of the tool. Those are defined and designated by the proposed approach. The first three success criteria have been verified by performing test cases. However, the last success criterion needs empirical work to evaluate.

9.2.1 Success criterion 5: Feature for calculating likelihood based on CORAS

The tool offers the function to calculate likelihood based on the CORAS method. This function is implemented based on interval arithmetic and the COA of interval. It is trivial to implement the function because it is actually arithmetic operators on real numbers.

9.2.2 Success criterion 6: The tool supports the approach

The tool offers the function to calculate likelihood based on the approach which proposes two fuzzy methods. In order to develop the function we utilize Matlab as a fuzzy logic programming library to support the development of the function.

9.2.3 Success criterion 7: Results calculated by the tool are sound

The success criterion implies that results calculated by the tool should be reliable and sound as applying the approach manually. Actually, the calculation of fuzzy logic is done by Matlab. The tool plays a role as a bridge to transfer parameters to and receive results from Matlab. Matlab is a powerful mathematical tool which supports complicated computation of mathematics including fuzzy logic. Therefore, results generated by Matlab are reliable. The test cases are designed to test and verify that the tool transfers right parameters and translates right results.

9.2.4 Success criterion 8: The tool is ease-to-use

The success criterion involves the user interface of the tool. In order to evaluate the user interface, the tool should be tested and experienced by the user and it is empirical work. Therefore, we have not achieved the success criterion.

9.2.5 Limitation of the tool

The limitation of the tool is similar to that of the proposed approach in the sense that it needs practical evidence to verify the success criterion 8. In addition to that, Matlab is a main component of the tool because it undertakes the fuzzy logic operations. For that reason, the tool is not compatible and portable. The function for loading and saving defined scales to file system and the function to visualize membership functions are underdeveloped.

Chapter 10

Conclusion

In this thesis, we have refined the CORAS method by bringing fuzzy logic into the rules for reasoning about likelihood. The objective of combining fuzzy logic and CORAS method is to address the issue of uncertainty in the risk assessment. Consequently, our thesis is designated by the two artefacts:

- The approach to combining fuzzy logic and the CORAS method.
- The tool supported the approach.

Throughout this work we have first presented a selective background of CORAS and fuzzy logic in order to acquire a solid foundation to elaborate the proposed approach. The two fundamental rules for reasoning about likelihood have been extended in the sense that fuzzy numbers refine intervals by the two fuzzy methods for calculating likelihood. As a consequence, conditional likelihood scale and relation scale have been defined to facilitate the approach.

We have conducted experimentation to test and verify the approach by conceptually developed the whole context based on the example from [25]. Additionally, different membership functions generated by strategy to construct membership function have been experimented in order to compare the patterns of the results. The patterns indicate that the use of trapezoidal and triangular membership function is not significantly different. More on that, the greater the support of the membership function the greater the result of defuzification.

We have presented the implementation of the tool through the software development process. The requirements of the tool have been gradually refined by class diagram, sequence diagram, and finally the design of user interface. The tool is evaluated by test cases, and they focus mainly on the functions of the tool to verify that the tool supports the proposed approach. Despite the fact that the tool is an independent artifact, it can be considered as a part of the proposed approach.

The shortage of our research is that it is lack of empirical evidence to justify the success criteria. Additionally, the construction of membership functions and the defuzzification methods have not taken into account yet. With regard to the tool, it is dependent on Matlab because of the lack of fuzzy logic programming library, and some functions are underdeveloped.

Future work indicates more empirical work to verify the success criteria of the approach and the tool. More on that, the underdeveloped functions of the tool should be implemented. To sum up, we close the topic with two theoretical research questions:

- How membership functions are constructed in the domain of CORAS risk assessment that employs fuzzy logic?
- Which defuzzification methods are appropriate in the context of risk assessment?

The two questions designate further research in the sense that they are complementary to the limitation of our proposed approach to achieve a complete approach that can be applied in the context of the CORAS risk assessment.

Bibliography

[1] A.M. Anile, S. Deodato and G. Privitera. ‘Implementing fuzzy arithmetic’. In: *Fuzzy Sets and Systems* 72.2 (1995), pp. 239–250.

[2] A.Teleshian and S.Rezvani. ‘Multiplication Operation on Trapezoidal Fuzzy Numbers’. In: *Journal of Physical Sciences* 15 (2011), pp. 17–26.

[3] Aditi Barua, Lalitha Snigdha Mudunuri and Olga Kosheleva. ‘Why Trapezoidal and Triangular Membership Functions Work So Well: Towards a Theoretical Explanation’. In: 2013.

[4] Taner Bilgic and Burhan Turksen. ‘Fundamentals of Fuzzy Sets’. In: vol. 7. Springer, 2000. Chap. Measurement of Membership Functions: Theoretical and Empirical Work.

[5] Stefano Borini. *Understanding Model-view-controller.* URL: http://forthescience.org/books/modelviewcontroller/index.html (visited on).

[6] Pablo Cingolani and Jesús Alcalá-Fdez. ‘jFuzzyLogic: a Java Library to Design Fuzzy Logic Controllers According to the Standard for Fuzzy Control Programming’. In: (2013), pp. 61–75.

[7] Roger Claker. *Non-empirical Research Techniques.* 2003. URL: https://www.rogerclarke.com/Res/51-NonEmp.ppt.

[8] Google Code. *Matlab Control.* URL: https://code.google.com/archive/p/matlabcontrol/downloads (visited on).

[9] Alan Dennis, Barbara Haley Wixom and Roberta M.Roth. ‘System Analysis and Design’. In: John Wiley & Sons, 2012. Chap. Project Selection and Management.

[10] Eclipse. *Eclipse.* URL: https://eclipse.org/.

[11] M. Elisabeth and Paté-Cornell. ‘Uncertainties in risk analysis: Six levels of treatment’. In: *Reliability Engineering & System Safety* 54 (1996), pp. 95–111.

[12] Brian R. Gaines. ‘Fuzzy and probability uncertainty logics’. In: *Information and Control* 38.2 (1978), pp. 154–169. ISSN: 0019-9958.

[13] Shang Gao, Zaiyue Zhang and Cungen Cao. ‘Multiplication Operation on Fuzzy Numbers’. In: *Journal of Software* 4.4 (2009), pp. 331–338.

[14] Ronald E. Giachetti and Robert E. Young. ‘A parametric representation of fuzzy numbers and their arithmetic operators’. In: *Fuzzy Sets and Systems* 91.2 (1997), pp. 185–202.

[15] T. Hickey, Q. Ju and M. H. Van Emden. 'Interval Arithmetic: From Principles to Implementation'. In: *J. ACM* 48.5 (2001), pp. 1038–1068.

[16] Leftery H.Tsoukalas and Robert E.Uhrig. *Fuzzy and Neural Approaches in Engineering (p11-183).* John Wiley and Sons, 1997.

[17] Solheimk Ida and Ketil Stølen. *Technology Research Explained.* 2007.

[18] National Instrument. *Center of Area (CoA) (PID and Fuzzy Logic Toolkit).* URL: http://zone.ni.com/reference/en-XX/help/370401H-01/lvpidmain/center_of_area/ (visited on).

[19] Guillermina Jasso. 'Handbook of Sociological Theory'. In: Springer, 2006. Chap. Formal Theory.

[20] P.C. Jorgensen. *Software Testing: A Craftsmans Approach, Fourth Edition.* An Auerbach book. Taylor & Francis, 2013.

[21] Stanley Kaplan and B. John Garrick. 'On The Quantitative Definition of Risk'. In: *Risk Analysis* 1.1 (1981), pp. 11–27.

[22] Katsiaryna Labunets et al. 'An Experimental Comparison of Two Risk-Based Security Methods'. In: *2013 ACM / IEEE International Symposium on Empirical Software Engineering and Measurement* 00 (2014), pp. 163–172. ISSN: 1938-6451. DOI: doi.ieeecomputersociety.org/10.1109/ESEM.2013.29.

[23] Qing Li. 'Indirect membership function assignment based on ordinal regression'. In: *Journal of Applied Statistics* 43.3 (2016), pp. 441–460.

[24] A.P.G.J. Lu et al. 'Handbook on Decision Making: Vol 2: Risk Management in Decision Making'. In: Intelligent Systems Reference Library. Springer, 2012. Chap. Computational Intelligence Techniques for Risk Management in Decision Making.

[25] Mass Soldal Lund, Bjørnar Solhaug and Ketil Stølen. *Model-Driven Risk Analysis: The CORAS Approach.* Springer, 2011.

[26] M.M. Malhotra and R. Subramanian. *Textbook in Applied Mechanics.* Wiley Eastern, Limited, 1994. Chap. Centre of Gravity. ISBN: 9788122406450. URL: https://books.google.no/books?id=K3a7t2wzpEsC.

[27] P.K. Marhavilas, D. Koulouriotis and V. Gemeni. 'Risk analysis and assessment methodologies in the work sites: On a review, classification and comparative study of the scientific literature of the period 20002009'. In: *Journal of Loss Prevention in the Process Industries* 24.5 (2011), pp. 477–523.

[28] *Math 251-003 Extra-credit assignment: On GeoMetic Centers and Centroids.* URL: http://mypages.iit.edu/~maslanka/EC.pdf.

[29] Matlab. *Defuzzification Methods.* URL: https://se.mathworks.com/help/fuzzy/examples/defuzzification-methods.html (visited on 01/12/2017).

[30] J.E McGrath. *Groups: Interaction and Performance.* Prentice Hall, 1984.

[31] S. Mitaim and B. Kosko. 'What is the best shape for a fuzzy set in function approximation?' In: *Proceedings of IEEE 5th International Fuzzy Systems.* Vol. 2. 1996, 1237–1243 vol.2.

[32] Roger Pressman. *Software Engineering: A Practitioner's Approach.* 5th ed. McGraw-Hill, 2000.

[33] Atle Refsdal, Bjørnar Solhaug and Ketil Stølen. *Cyber-risk Management.* Springer, 2015.

[34] W.G. de Ru and J.H.P. Eloff. 'Risk analysis modeling with the use of fuzzy logic'. In: *Computers & Security* 15.3 (1996), pp. 239–248. ISSN: 0167-4048.

[35] Johannes Ryser and Martin Glinz. 'A Practical Approach to Validating and Testing Software Systems Using Scenarios'. In: *QWE 99, 3 rd International Software Quality Week Europe.* 1999.

[36] Kailan Shang and Zakir Hossen. *Applying Fuzzy Logic to Risk Assessment and Decision-Making.* URL: https://www.soa.org/Files/Research/Projects/research-2013-fuzzy-logic.pdf (visited on).

[37] S. N. Sivanandam, S. Sumathi and S. N. Deepa. *Introduction to Fuzzy Logic Using MATLAB.* Springer, 2006. ISBN: 3540357807.

[38] Bjørnar Solhaug and Ketil Stølen. 'The CORAS Language – Why it is designed the way it is'. In: CRC Press", 2013.

[39] International Organization Standard. *ISO31000 Risk Management - Principles and Guidelines.* 2009.

[40] Ketil Stølen. *Rules for reasoning about frequencies and consequences in CORAS.* Tech. rep. 2013.

[41] P.V. Suresh, A.K. Babar and V.Venkat Raj. 'Uncertainty in fault tree analysis: A fuzzy approach'. In: *Fuzzy Sets and Systems* 83.2 (1996), pp. 135–141. ISSN: 0165-0114.

[42] Valerie Trifts. *Deductive Inference.* URL: http://penta.ufrgs.br/edu/telelab/3/deductiv.htm (visited on 01/12/2017).

[43] Valerie Trifts. *Inductive Inference.* URL: http://penta.ufrgs.br/edu/telelab/3/inductiv.htm (visited on 01/12/2017).

[44] Jay Verkuilen. 'Assigning Membership in a Fuzzy Set Analysis'. In: *Sociological Methods & Research* 33.4 (2005), pp. 462–496.

[45] J. Viehmann. 'Reusing Risk Analysis Results – An Extension for the CORAS Risk Analysis Method'. In: *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing.* 2012, pp. 742–751.

[46] Kamin Whitehouse. *Calling Matlab from Java.* URL: https://www.cs.virginia.edu/~whitehouse/matlab/JavaMatlab.html (visited on).

[47] Wikipedia. *Soundness.* URL: https://en.wikipedia.org/wiki/Soundness.

[48] Lotfi A. Zadeh. 'Is There a Need for Fuzzy Logic?' In: *Inf. Sci.* 178.13 (2008), pp. 2751–2779.