Cardiff University
School of Engineering

PhD Thesis

Knowledge-based approach to risk analysis in the customs domain

Panagiotis Loukakos

April 2014

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed .... (candidate) Date ..07/10/2013...

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of ....PhD...............(insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed .... (candidate) Date ...07/10/2013...

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references. The views expressed are my own.

Signed .... (candidate) Date ...07/10/2013

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed .... (candidate) Date ...07/10/2013

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed .... (candidate) Date ..07/10/2013

To my wife,
children and parents
for the countless
family time spent for
the preparation of
this thesis

“Success is
getting what you want.
Happiness is wanting
what you get.”

American Proverb

Acknowledgements

I would like to acknowledge the help of those who made the completion of this work possible. First, I would like especially to express my sincere thanks and gratitude to my supervisor Professor Rossi Setchi not only for her constant assistance, beneficial supervision and amity during this project from the very first till the very last time, but also for her constant support and encouragement to carry out my research. This research wouldn't have been accomplished without her continuous scientific support and guidance. Then, I would like to sincerely thank my second supervisor Mr. Paul Prickett for his beneficial advices during the performance of this thesis.

Finally, I would like to thank from my heart my beloved wife, Eleni, for her patience, motivation and uninterrupted support during all these years to the accomplishment of this study.

Abstract

The aim of this PhD project is to develop a fuzzy knowledge-based approach in support of risk analysis in the Customs domain.

Focusing upon risk management and risk analysis in the Customs domain, this thesis explores the relationship of risk with uncertainty, fuzziness, vagueness, and imprecise knowledge and it analyses state of the art detection techniques for fraud and risk. Special focus is given to fuzzy logic, ontological engineering, and semantic modelling considering aspects such as the importance of human knowledge and semantic knowledge in the context of risk analysis for the Customs domain.

An approach is presented combining the fuzzy modelling and reasoning with semantic modelling and ontologies. Fuzzy modelling and reasoning is explored in the context of risk analysis and detection in order to examine approximate human reasoning based on human knowledge. Ontologies and semantic modelling are explored as an approach to represent domain knowledge and concepts. The purpose is to enable easier communication and understanding as well as interoperability. Risk management is broader, multi-dimensional process involving a number of task, activities, and practises. The presented approach is focused on examining the analysis and detection of the risk, based on the outputs of the risk management process with the use of ontologies and fuzzy rule-based reasoning.

An ontological architecture is developed in the context of the presented approach. It is considered that such architecture is possible to enable modularity, maintainability, re-usability, and extensibility and can also be extended or integrated with other ontologies. In addition, examples are discussed to illustrate representation of concepts at various levels (generic or specific) and the modelling of various semantics.

Furthermore, fuzzy modelling and reasoning are investigated. This investigation consists of literature research and the use of a generic research prototype (examination of Mamdani and Sugeno model types). From theoretical research, fuzzy logic enables the expression of human knowledge with linguistic terms and it could simulate human reasoning in the context of risk analysis and detection. In addition, Hierarchical Fuzzy Systems (HFS) or Hybrid Hierarchical Fuzzy Controllers (HHFC) approaches can be used to manage complexity especially for complex domains. Linguistic fuzzy modelling (LFM) is an aspect that should be considered during fuzzy modelling. From the generic research prototype, fuzzy modelling with the use of ontologies is demonstrated together with their integration in the context of fuzzy rule-based reasoning. It is also considered that Mamdani type of fuzzy models is easier to express human knowledge since the output can be expressed with linguistic terms. However, Sugeno type of fuzzy model could be used from adaptive techniques for optimisation purposes.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference System
BOA	Bi-sector of Area
BPM	Business Process Management
BPMN	Business Process Model and Notation
COA	Center of Area
CREA	Clinical Risk and Error Analysis Method
DL	Description Logic
DMRA	The Decision Matrix Risk-Assessment Technique
EC	European Commission
ETA	Event Tree Analysis
EU	European Union
FL	Fuzzy Logic
FLC	Fuzzy Logic Controller
FMEA	Failure Mode And Effects Analysis
FMECA	Failure Mode Effects and Criticality Analysis
FML	Fuzzy Markup Language
FTA	Fault-Tree Analysis
FuzzyRA	Fuzzy Risk Analysis
HACCP	Hazard Analysis and Critical Control
HAZOP	Hazard and Operability Studies (HAZOP)
HEAT	Human Error Analysis Techniques
HFC	Hierarchical Fuzzy Systems

HFEA	Human Factor Event Analysis
HFS	Hierarchical Fuzzy Systems
HHFC	Hybrid Hierarchical Fuzzy Controller
IRI	Internationalized Resource Identifier
IS	Information System
IT	Information Technology
LFM	Linguistic Fuzzy Modelling
LOM	Largest (Absolute) Value of Maximum
MIMO	Multiple-Input Multiple-Output System
MISO	Multiple-Input Single-Output System
MOM	Mean Value of Maximum
PEA	Predictive, Epistemic Approach Method
PRAT	The Proportional Risk-Assessment Technique
QADS	Quantitative Assessment of Domino Scenarios
QRA	Quantitative Risk-Assessment
RBM	Risk-based Maintenance
RM	Risk Management
RuleML	Rule Markup Language
SOM	Smallest (Absolute) Value of Maximum
STEP	The Sequentially Timed Event Plotting
SWRL	Semantic Web Rule Language
UML	Unified Modelling Language
WRA	The Weighted Risk Analysis
XML	Extensible Markup Language
XSL	Extensible Stylesheet Language
XSLT	XSL Transformations

Chapter 1. Introduction

1.1 Overview

In this chapter, it is provided an overview of the present research. The goal of this chapter is to present and clarify the research problems and areas of this project and integrate it to a general research context.

1.2 Customs Domain

The research is performed in the context of Customs. Therefore, some background information mainly for EU Customs is presented in the following paragraphs prior to the discussion of motivation and objectives of this research.

Customs plays a vital role in the economy, environment, and security. Particularly for European Union (EU), the main role of EU Customs is to facilitate trade and at the same time to protect the interests of the European Union and its citizens. Currently, 28 customs administrations of the EU implement the community customs code (DGTAXUD 2014e). The Customs Union is an important element for the operation of the single market. In EU Customs Union, uniform handling of goods (import, export and transit) is performed by a number of countries and a common set of rules are implemented (DGTAXUD 2014f). Those common rules cover aspects such as common tariff, health and environment controls, protection of economic interests, etc. (DGTAXUD 2014e).

Nowadays, Customs are facing new challenges and in order to achieve these demands, modernization of customs procedures and controls is required as well as cooperation of various services (DGTAXUD 2014e). The Electronic Customs initiative started with the aim of creating a paperless environment in the context of Customs modernisation (DGTAXUD 2014d). The interoperability of Customs IT systems and exchange of information is considered as an essential element so as the EU's economy to continue to compete in a global context (DGTAXUD 2014e).

Customs involves a number of Customs procedures such as release for free circulation, exportation, transit, etc. In EU, those Customs procedures are defined in the Customs legislation into force (DGTAXUD 2014c). The Customs declaration is used by a person to indicate the wish to place the goods under a specific Customs procedure (DGTAXUD 2014b; EEC 1992 Article 4(17); 2008b Article 4(10)).

As stated in (DGTAXUD 2014f), 261 million of customs declarations (Total Extra-EU Trade) performed in 2012, which means 8 declarations a second. This concerns 139 million of customs declarations for import, 105 million of customs declarations for export and 17 million of customs declarations for transit (DGTAXUD 2014f).

In order to ensure correct application of Customs rules and other legislation, Customs Authorities apply Customs Controls in a number of areas (DGTAXUD 2014a; EEC 1992 Article 4 (14)). As it is mentioned in (DGTAXUD 2014a), Customs Authorities must apply various controls in an environment with fast moving of goods and also consistently across the Community. Therefore, Customs controls must be fast, effective and based on risk management techniques. To this context, this research is performed to develop a fuzzy knowledge-based approach in support of risk analysis.

Customs is a complex business domain due to the number and the nature of its processes, the number and complexity of business rules that govern the processes, the number of the actors involved, as well as the number of terms and the concepts that are used in the various procedures. In addition, common understanding of procedures and business rules is required for the performance of customs business. Various entities (e.g. declarations, authorisations, and guarantees) are required for the completion of a customs procedure. A number of actors such as Economic Operators and Customs Authorities are involved in the completion of customs procedures. In the context of electronic customs, Information is exchanged among those actors and hence it is believed that everyone should share the same understanding about the various concepts.

1.3 Motivation

The risk management process has as purpose to identify, analyse and assess factors that may jeopardize the success, operation or function of the assessed context. The results from this analysis and assessment should be used for the establishment of preventive measures and the identification of countermeasures with aim to reduce the probability of these factors from occurring. Therefore, the risk management process has also to treat, monitor, and communicate the risks. The risk management is used in various domains. Another field on which the risk management plays an important role is the Customs domain.

As stated in section 1.2, Customs plays a vital role in the economy, environment, and security. Application of risk management is an important element for modern Customs Administrations (WCO 2011). Risk management is a technique, which used by Customs for setting priorities more effectively and for allocating resources more efficiently with purpose to keep a proper balance between control and facilitation of legitimate trade (DGTAXUD 2013). Furthermore, the "Intelligence-driven risk management" is also a concept where learning from past decisions is utilised in risk related activities. This concept is also enabled with the use and support of IT systems (WCO 2010).

Knowledge is *an understanding of information based on its perceived importance or relevance to problem domain* (Awad 1996, p. 29). Knowledge can be classified into *shallow* or *deep* knowledge. *Shallow* knowledge is minimal understanding of problem domain. Deeper knowledge is required when decision-making is more complex requiring assessment of many parameters. Another way for further classifying knowledge is to *procedural*, *declarative*, *semantic,* or *episodic*. Semantic knowledge is considered as deeper type of knowledge. It is also indicated as “chunked” knowledge residing in long-term memory. Such knowledge requires understanding of various concepts and their interrelationships (Awad 1996). Apparently, this is not an easy task because various relationships exist among the concepts and information. Another important aspect is the concepts and their semantics. For instance, it is considered that Customs contains many concepts and there are complex relationships between the concepts. It is assumed that risk management and assessment is facilitated with the deep understanding of the domain and its concepts as well as the relationships between them. Therefore, the semantic modelling and ontologies would assist on that aspect and formally represent the concepts enabling understanding.

Following literature review, Artificial Intelligence and Expert Systems (Digiampietri et al. 2008; Singh and Sahu 2004; Singh et al. 2003), Neural Networks (Feng et al. 2007; Ye et al. 2007) and Statistical methods (Geourjon et al. 2010; Geourjon and Laporte 2005; Laporte 2011) are some of the techniques that have been applied for detection and risk analysis in this domain.

Human knowledge and expertise is also very important for “intelligence” in Customs. Singh and Sahu (2004) analyse the importance of human intelligence in crime prevention. It is stated that the IT systems are machines and hence devoid of human emotions. Therefore, the systems decide purely based on the data and according to the computer program. The nature of human mind has the capacity to invent ever-new methods to commit crime that cannot be predicted by any computer. Therefore, it is mentioned that *“the emotions and creativity of human mind can be effectively countered only by the intelligence, emotions and creativity of another human mind” (Singh and Sahu 2004, p. 447).* The human knowledge and intelligence must be considered and used effectively during the risk analysis process.

Fuzzy logic and fuzzy inference systems could facilitate this. Application of fuzzy logic enables approximate human reasoning to be applied to knowledge-based systems (Alavala 2008).

The risk analysis shall also consider both the certainty and the uncertainty. The risk is closely related to the uncertainty. The uncertainty is related to the fuzziness since the fuzziness could be an uncertainty occurred from vagueness. Friedlob and Schleifer (1999, p. 127) mentions that L. Zadeh in his *Law of Incompatibility* states that *"as complexity rises, precise statements lose meaning and meaningful statements lose precision"*. The relationship between complexity and uncertainty is proportional i.e. as complexity increases, certainty decreases (Friedlob and Schleifer 1999). Fuzzy logic introduced by Zadeh in 1965 and it is a mathematical tool for dealing with uncertainty (Sivanandam et al. 2007). The concepts of fuzzy sets were used for describing dynamic systems that are too complex and/or ill-defined to synthesize controllers using conventional mathematical modelling techniques (Rao and Saraf 1996).

At the inference engine level, this vagueness or fuzziness could be the reason of the recognised drawback of traditional *Production systems*, which is the partial matching. It is deemed that the partial matching is solved with the use of fuzzy logic during the inference process (Valdez et al. 2007). As far as the knowledge representation is concerned, the uncertainty should also be represented in the knowledge.

1.4 Research Objectives

The key points highlighted in previous section summarise the motivation for this research. The aim of this research is to investigate a fuzzy knowledge-based approach for supporting risk analysis and detection with application in Customs domain. The main characteristics of this will be the combination of fuzzy reasoning and semantic modelling with ontologies. On one hand, fuzzy reasoning is examined for handling imprecise knowledge and vagueness in risk analysis in this complex domain. Also fuzzy inference systems can be used to express knowledge. On the other hand, the semantic modelling is used for representing knowledge and concepts of this domain improving both the communication and understanding. In addition, it offers the flexibility to map different elements with their semantics to ontology concepts. The research objectives are the following four:

1. to develop a conceptual model for fuzzy knowledge-based approach to risk analysis;
2. to develop a high-level ontological architecture for supporting the fuzzy knowledge-based approach to risk analysis, considering the complexity of the domain for knowledge representation and formal representation of concepts;
3. to develop ontology models according to the presented architecture to represent concepts especially specific to the risk analysis with fuzzy logic technique;
4. to investigate fuzzy modelling for risk analysis and assess application of fuzzy logic and approximate reasoning.

1.5 Outline

The structure of research thesis is the following:

Chapter 1 provides an overview of the research motivation and the definition of the research objectives.

Next chapter, Chapter 2, analyses the risk concept and explores risk management process. In addition, it describes the relationship of risk with vagueness and uncertainty. All these are analysed for better understanding of concepts required for this research and the challenges of risk analysis.

Chapter 3 reviews state of the art detection techniques for fraud and risk. It focuses more on fuzzy logic, fuzzy rule-based systems, and Adaptive Neuro-Fuzzy Inference System (ANFIS). In addition, ontologies and ontological engineering are analysed. This chapter gives a ground theoretical background for the practical part of this research.

Chapter 4 presents one of the contributions of this research, which is to develop a conceptual model for fuzzy knowledge-based approach to risk analysis. This model combines semantic modelling with ontologies and fuzzy reasoning.

Chapter 5 presents the semantic modelling work and ontologies developed under this research in order to represent knowledge and concepts. It describes another contribution of this research, which is an architecture of ontologies considering the complexity of the domain for knowledge representation and formal representation of concepts. Finally, individual ontologies are described and discussed with examples as contribution of this research.

Chapter 6 investigates the fuzzy modelling and reasoning following the information presented in the previous chapters. The approach and decisions for this assessment activity is described along with the constraints for that research. The various tools used for this assessment are also mentioned. Finally, the analysis of the results is presented.

Finally, Chapter 7 presents the conclusions drawn from the research findings as well as discusses some ideas for future work.

Chapter 2. Background Information in Risk Analysis

The purpose of this chapter is to analyse the risk concept and the risk management processes in order to gain better understanding of the risk analysis activity. Specifically, it provides the necessary background by unambiguously defining the risk in Customs, which is the domain under study. In addition, the various risk management processes are briefly explored by also making a brief analysis of various activities. Moreover, the relationship of risk with fuzziness, vagueness, and uncertainty is described in order to enable the better comprehension of the concept and the challenges of risk analysis. This background information is considered essential since this research examines a fuzzy knowledge-based approach in support of risk analysis in the Customs domain. Therefore, the various aspects of risk management and risk analysis should be studied.

2.1 Risk

This research examines risk analysis and hence, it is very important to understand what risk means. A number of definitions exist for risk. According to Miller (Miller 2004), *"Risk is a combination of the frequency or probability of a specified hazardous event, and its consequence".* IEEE Standard 1540-2001 (IEEE 2001, p. 3) defines that the risk "*is the likelihood of an event, hazard, threat, or situation occurring and its undesirable consequences".*

In Customs, Customs risk for EU "*means the likelihood that something will prevent the application of Community or national measures concerning the customs treatment of goods" (DGTAXUD 2004, p. 3).*

Examining this non-compliancy with Customs Laws from business perspective, it is translated by Truel (2010) into the following three types of Customs risk: *Regulatory Risk*, *Fiscal Risk* and *Security Risk*.

According to Pearl (1988), the primitive relationships of risk are *Likelihood*, *Conditioning*, *Causation,* and *Relevance*. Understanding of the qualitative relationships of probability language enables the better comprehension of risk. Probability of something or an event to occur is closely related to the risk.

As far as the connection of fraud with risk is concerned, Phua et al. (2005) state in their work that the term *fraud "refers to the abuse of a profit organisation's system without necessarily leading to the direct legal consequences".* Laleh and Azgomi (2009) state that there are several types of fraud providing taxonomy of fraud types. According to this, the main types of fraud are *web network fraud*, *internal fraud (in organisations)*, *insurance fraud*, *credit fraud*, *computer intrusion fraud*, *telecommunication fraud* and *Customs fraud*. Some of those frauds are further decomposed to sub-categories of fraud. It is worth mentioning that the types of fraud mentioned above are only the main types (first level of taxonomy).

In regards to Customs, it is stated by Shao et al. (2002, p. 1241) that "*to avoid administrate regulation or the duties, some lawless persons take the cheating measures while their commodities pass the customs, such as hiding, declaring less or making false reports".*

2.2 Risk Management

2.2.1 Overview

Nowadays, the risk management constitutes an important and integral component of management activities in several fields (e.g. organisational management, project management, etc.) with purpose to achieve the relevant objectives accurately and effectively (Tchankova 2002; Zhu 2008). The trend or recommendation is for a more "risk-based" approach before appropriate decisions or actions have to be taken (Nota 2011).

A number of risk management processes currently exist. For Customs, the WCO Risk Management Process (WCO 2011) is based on ISO 31000:2009. Another risk management process, which is described, is the Standardised Framework for Risk Management in the Customs Administrations of the EU (DGTAXUD 2004).

In terms of definition, risk management for Customs is defined in the Standardised Framework for Risk Management in the Customs Administrations of the EU as follows: "*a technique for the systematic identification and the implementation of all the measures necessary to limit the likelihood of risks occurring. International and national strategies can be effectively implemented by collecting data & information, analysing & assessing risk, prescribing action and monitoring outcomes*" (DGTAXUD 2004, p. 3).

2.2.2 Risk Management Activities

A risk management process is an iterative approach enabling the continuous identification of risks and the improvement of decision-making. Typically, a risk management process consists of the following activities:

- Context
- Risk assessment
  - Risk identification
  - Risk analysis
  - Risk evaluation
- Risk treatment
- Risk monitoring and review
- Risk communication

The various activities are briefly described in the following paragraphs with either generally applied information or specific to the Customs domain. More information can be found in the provided references.

- Context of risk management

The first activity of the risk management process is the context analysis and the definition of the objectives and risk areas of risk management (DGTAXUD 2004; WCO 2011).

- Risk assessment
  - Risk identification

The risk identification is considered an important activity of the risk management process. Since the next stages of risk management are based on the output of this activity, the effectiveness of the risk management process depends on whether this stage will achieve to identify all possible risks (Tchankova 2002). During the risk identification, a comprehensive list of *sources of risks* and *events* that might affect the risk management objectives are defined (Tchankova 2002). For instance, the sources of risks should cover all events that might affect (e.g. prevent or delay) the organisational objectives.

For Customs, possible source of information for this activity could be trade flows, declarations rendered, payments made on time/debt on file and new or changed legislation. Moreover, the experience of operational staff are considered in this activity (DGTAXUD 2004). The identification of risks should be performed with a top-down approach. Some more high-level risks are identified from upper management and then these are refined from the other levels (WCO 2011).

Another important aspect, which must be considered during the risk identification, is the resources that are exposed in risk. Tchankova (2002) provides a categorisation of these resources as it is summarised in Figure 2-1.

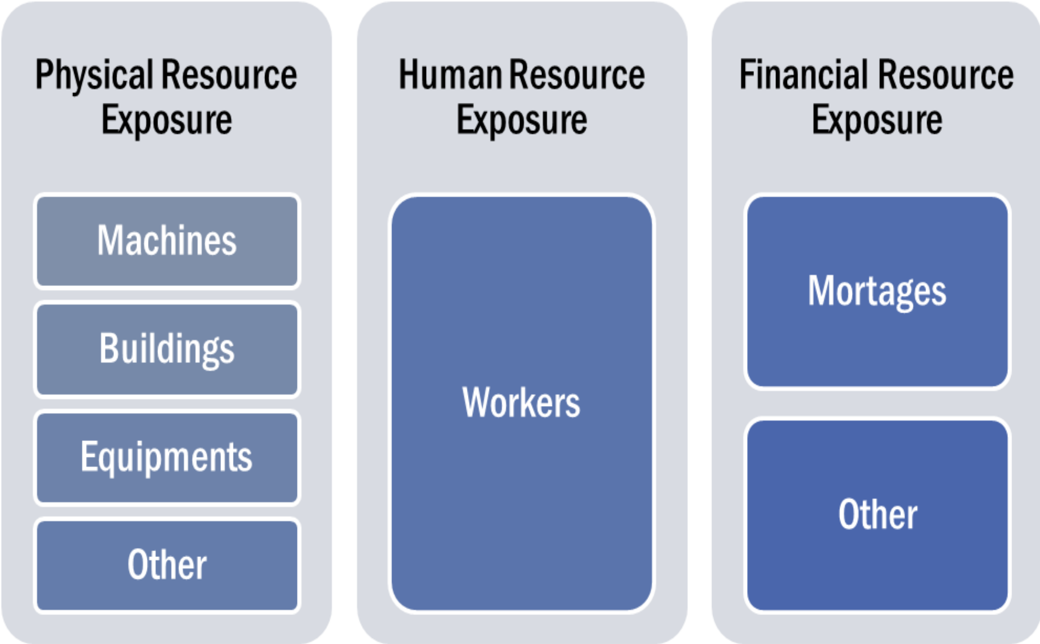

Figure 2-1: Resources exposed to risk (Tchankova 2002)

The risk identification requires well-structured techniques in order to ensure that all risks have been identified and nothing has been forgotten.

Several methods can be used for the risk identification activity. A categorisation of risk identification techniques is based on the form of logic and is provided by Frosdick (1997) where they are classified into:

- ***Intuitive***: include *brainstorming*;
- ***Inductive***: include techniques such as checklists, preliminary hazard analysis, human error analysis, Event-Tree Analysis (ETA) and the most commonly used Hazard and Operability studies (HAZOP) and Failure Mode and Effects Analysis/ Failure Mode Effects and Criticality Analysis (FMEA/FMECA);
- ***Deductive***: are considered the Accident investigation and analysis technique and the Fault-Tree Analysis (FTA).

Another categorisation of techniques is based on the type of technique, i.e. qualitative (e.g. Checklists, HAZOP), quantitative (Quantitative Risk Assessment

(QRA), and hybrid (e.g. FTA) (Marhavilas et al. 2011). However, this kind of categorisation will be further analysed in risk analysis section. It is worth noting at this point that the risk identification is a continuous activity.

- Risk analysis

During risk analysis, the identified risks are analysed. Standardised Framework for Risk Management in the Customs Administrations of the EU states that the following two main categories of risk *Proven risks* and *Potential risks* should be considered during the risk analysis. *Proven* risks are historical facts and have occurred in the past. During risk analysis, those risks can be examined based on current data and if the conditions for these risks do exist then their analysis should be performed by assessing the *Likelihood* and *Consequence* of the risk. *Potential* risks are risks that have not been revealed yet but are suspected. Similar to *Proven* risks, *Potential* risks should be examined based on current data and if the conditions for these risks do exist then as above, their analysis should be performed by assessing the *Likelihood* and *Consequence* of the risk (DGTAXUD 2004).

Risk analysis considers the *Likelihood* and *Consequence (or Seriousness)* of the risk. The *likelihood* denotes the chances (probability) of risk occurring while the *Consequence* indicates the impact and the consequences of this risk when the risk takes effect. Risk analysis estimates the level of risk by assessing the aforementioned defined two factors i.e. the *likelihood* and the *consequence (DGTAXUD 2004; WCO 2011).*

A number of techniques exist for risk analysis. Marhavilas et al. (2011) proposes a classification of techniques into *qualitative, quantitative* and *hybrid. Qualitative* techniques assess the likelihood and consequence and measure the risk based on descriptive scales. *Quantitative* techniques measure likelihood, consequence and level of risk with numerical values (quantify). The *hybrid* techniques include *semi-qualitative* techniques or a combination of *qualitative-quantitative* techniques. This classification is presented in Figure 2-2.

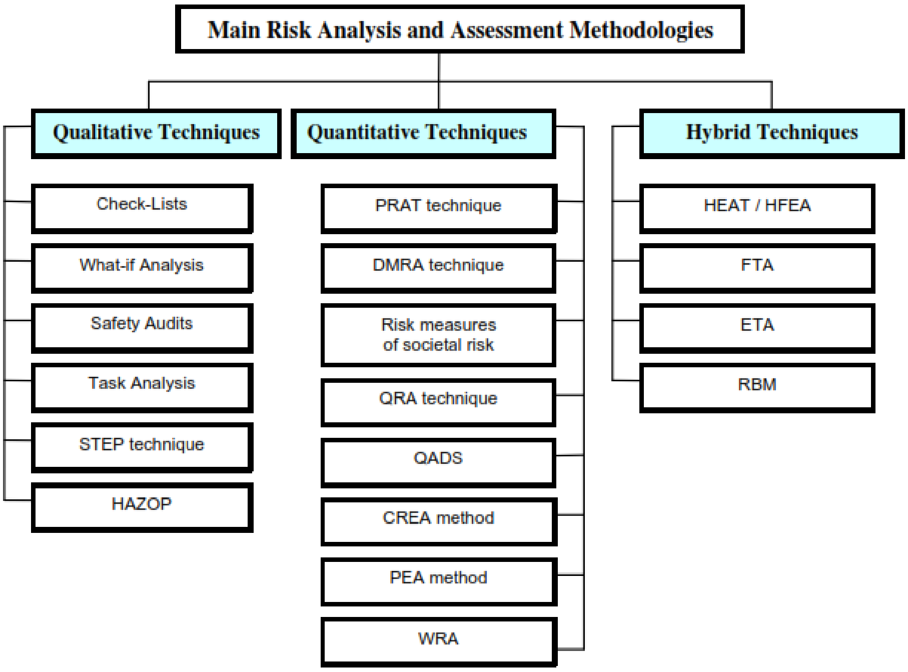

Figure 2-2: Classification of the main Risk Analysis and Assessment (RAA) Methodologies (Marhavilas et al. 2011)

As it shown from Figure 2-2, some of them have already been mentioned in the context of risk identification (e.g. checklists, HAZOP, FTA, ETA, QRA, etc.).

Similar techniques are also indicated by ISO (2009) (as cited in WCO 2011) in the context of Customs Risk Management. For instance, Bayesian statistics and Bayes nets, Decision tree, Markov analysis, and Multi-criteria decision analysis are some techniques indicated for risk analysis. Researches also have been conducted to apply fuzzy logic to FMEA (Braglia et al. 2003; Garcia et al. 2005; Gargama and Chaturvedi 2011; Kai Meng 2009; Tay and Lim 2006; Wang et al. 2009). In general, fuzzy logic has been used for risk analysis purposes (Cameron and Peloso 2005; de Ru and Eloff 1996; Deshmukh and Talluru 1997; Jinting et al. 2009; Liu and Yu 2009).

- Risk evaluation

The outcomes of risk analysis are used to evaluate and prioritise risks.

- **Risk treatment**

Having identified, analysed and evaluated the risks, plans should be developed and actions should be taken for the treatment of risks. As mentioned in Standardised Framework for Risk Management in the Customs Administrations of the EU (DGTAXUD 2004), physical or documentary control actions can be used. Moreover, posteriori control/audits could be performed. A common method for treating risks is to develop risk profiling and targeting (Iordache and Voiculet 2007).

- Risk monitoring and review

A clear distinction is made between the *monitoring* and the *review*. The *monitoring* refers to the evaluation of efficiency of risk management system whereas the *review* refers to the process, which is performed to assess the existing risk profiles and probably to update them accordingly (DGTAXUD 2004).

- Risk communication

All stakeholders involved in the risk management process shall be informed with the later developments on risk management and the taken measures.

2.2.3 Risk Profiling

A risk profile must be concrete and must target specific risk areas. WCO Risk Management Compendium (WCO 2011, p. x (Common Part)) defines the risk profile as *"Description of any set of risks, including a predetermined combination of risk indicators, based on information which has been gathered, analyzed and categorized".*

2.3 Vagueness and Uncertainty

This research focuses on fuzzy knowledge-based approach for risk analysis rather than on using probabilistic theory. Therefore, the concepts such as *vagueness and uncertainty* are studied in relation to risk and risk analysis and are described in the subsequent paragraphs.

In case of *certainty,* there is no *risk* since the outcome of a process or event is predictable. On the contrary, the *uncertainty* creates the risk something to happen. Therefore, it is inferred that the *risk* is closely related to the *uncertainty*. This conclusion is also supported by Friedlob and Schleifer (1999).

The reasoning in realistic domains it is not achieved without made some simplifications. Many exceptions are required in order to create rules to explain some behaviours of real life. Pearl (1988) proposes not to ignore these exceptions but to summarise them, otherwise the reasoning is not valid. All these exceptions prove why we need to bother with uncertainty (Pearl 1988).

Kosko (1992) describes *fuzziness* versus *probability* and explores the use of *fuzziness* as an alternative approach to *randomness* for defining *uncertainty*. In this context, Kosko describes the relationship of *fuzziness, randomness,* and *ambiguity. Fuzziness* describes *event ambiguity* since it shows the degree to which an event occurs. On the other hand, whether an event occurs is the *randomness* because is random. Fuzzy is to what degree this event occurs (Kosko 1992).

Fuzzy sets can be used to model *uncertainty* related to *imprecise information* and *vagueness*. *Uncertainty* is arisen due to ignorance, randomness, lack of knowledge and vagueness. The *set membership* concept has been proposed by Zadeh in order to perform appropriate decision making when uncertainty occurs (Sivanandam et al. 2007).

Lukasiewicz and Straccia (2008) mention a misunderstanding that exists all these years in Artificial Intelligence in regards to the role of *probability/possibility* theory and *vague/fuzzy theory*. Similar to those definitions about *fuzziness* and *randomness* provided by Kosko (1992), Dubois and Prade (2001) highlight the difference between *degrees of uncertainty* and *degrees of truth* by explaining the example of "full" bottle. If someone says that "the bottle is *half-full*", then *full* or *half-full* is more a fuzzy predicate with some *degree of truth*. On the other hand, if we ignore that the bottle is *full* or *empty,* the statement "the *probability* that the bottle is *full* is 0.5" does not mean that the bottle is *half-full*. In this case, we have *degrees of uncertainty*, which according to Dubois and Prade (2001) are clearly a higher level notion, higher than *degrees of truth*. Kosko (1992) claims that "*fuzziness* is a type of deterministic uncertainty" (Kosko 1992, p. 267) while Dubois and Prade (2001) emphasizes the distinction between the handling of vague propositions (*vagueness*) in the presence of complete information and the treatment of *uncertainty* for propositions which are either true or false.

All approaches in which statements are true or false to some *probability* or *possibility* are fallen under the *Uncertainty Theory*. For instance, the example of "it will rain tomorrow" is used to explain this concept. Whether will rain or not tomorrow cannot be certain due to the incomplete knowledge, however, we can estimate the *probability* or *possibility* this to happen (*degree of uncertainty*). Both probability and possibility theories are used to quantify the *degree of uncertainty*, though, they have some conceptual differences because they represent different aspects of the uncertainty (Lukasiewicz and Straccia 2008).

On the other hand, all statements that are true to some degree (taken from the truth space) are fallen under the *vague/fuzzy theory*. When some statements include *vagueness* (imprecise information), it cannot be exactly inferred that those statements are true or false. Another difference is that *vague/fuzzy* statements are "truth-functional" since the *degree of truth* of a statement can be calculated from the *degree of truth* (*uncertainties)* of its parts (Dubois and Prade 2001; Lukasiewicz and Straccia 2008).

If the form of uncertainty happens to arise because of imprecision, ambiguity, or vagueness, then the variable is probably fuzzy and can be represented by a membership function (Virtanen and Helander 2010). *Fuzziness, vagueness* and *imprecise knowledge* are inherent to several real world domains (Bobillo and Straccia 2009). This research investigates *fuzzy logic* and *approximate reasoning* for risk analysis.

2.4 Summary

Risk concept and risk management processes are analysed in this chapter as background information. This is considered necessary for the better understanding on how a fuzzy knowledge-based approach could be applied in risk analysis. It is realised that fuzzy knowledge-based approach should be able to express the knowledge and to be able to analyse and detect the risk. In addition, this risk analysis should consider *fuzziness, vagueness,* and *imprecise knowledge.* Therefore, this research investigates fuzzy logic and approximate reasoning for risk analysis of a *physical entity* in the context of Customs domain.

Chapter 3. State of the Art Review of Detection Techniques and Ontologies

This chapter reviews state of the art detection techniques for fraud and risk. The techniques are classified into the following categories: *Supervised learning*, *Semi-supervised learning, Unsupervised learning and Meta-Learning or Combining Multiple Algorithm*. Those categories are described for gaining an understanding of the characteristics of those techniques. However, this chapter provides special focus on fuzzy logic, fuzzy inference systems, and adaptive network-based fuzzy inference system (ANFIS). Fuzzy logic theory and fuzzy inference systems are examined in more detail since the fuzzy modelling and reasoning are investigated in this research as part of fuzzy knowledge-based approach to risk analysis and detection. Therefore, understanding of the fuzzy inference process is essential since certain decisions must be taken during the development of a fuzzy inference system. Furthermore, the ANFIS technique, which combines Artificial Neural Networks and Fuzzy Logic, is discussed. This technique could assist with fuzzy modelling and the construction or optimisation of adaptive fuzzy inference systems from a given data set for performing reasoning tasks. It is important to understand the ANFIS architecture in order to enable the use of this technique (e.g. select partitioning technique for the generation of a Sugeno type FIS). Finally, this chapter discusses ontological engineering and semantic modelling for knowledge representation in the context of fuzzy knowledge-based approach to risk analysis. This chapter should be considered as a ground theoretical background for the practical part of this research.

3.1 Fraud Detection Systems

Reviewing the literature, it is identified that various researches and works have been conducted for fraud detection systems in various areas (Deshmukh and Talluru 1997; Digiampietri et al. 2008; Farvaresh and Sepehri 2011; Fawcett and Provost 1997; Hilas 2009; Jianhong and Dezhao 2008; Laleh and Azgomi 2009; Ngai et al. 2011; Wei et al. 2008). Studies show that most of intelligent systems techniques are mainly applied in *Telecommunications*, *Insurance*, *Medical Care*, *Auditing*, *Credit Card Transactions* and to *other areas* (Pejic-Bach 2010; Yufeng et al. 2004). Some techniques are *Neural Networks*, *Bayesian Belief Networks*, *Decision Trees*, *Fuzzy Logic*, *Rule-based Systems* and *Data Mining techniques* (Fawcett and Provost 1997; Hilas 2009; Ngai et al. 2011; Pejic-Bach 2010; Roman et al. 2009; Yufeng et al. 2004). According to Pejic-Bach (2010) survey of research articles, Neural Networks are more often used as intelligent systems technique for fraud detection, while few cases are found for Fuzzy Rules and Genetic Algorithms. Considering this, the ANFIS technique is examined in section 3.7 and it is also considered in the fuzzy knowledge-based approach of Chapter 4. ANFIS combines Artificial Neural Networks and Fuzzy Logic and could assist in the construction or optimisation of adaptive fuzzy inference systems from a given data set for performing reasoning tasks.

In general, the fraud detection systems has as purpose to identify or detect general trends of suspicious transactions for fraud (Phua et al. 2005).

Following literature review, various works have been found also for the Customs domain (Digiampietri et al. 2008; Feng et al. 2007; Geourjon et al. 2010; Geourjon and Laporte 2005; Laporte 2011; Liu et al. 2009; Roman et al. 2009; Shao et al. 2002; Singh and Sahu 2004; Singh et al. 2003). Those works mainly examine the application of Neural Networks (Feng et al. 2007), Fuzzy Logic (Singh and Sahu 2004; Singh et al. 2003), Data Mining (Shao et al. 2002), Outlier detection

(Digiampietri et al. 2008) and Statistical methods (Geourjon et al. 2010; Geourjon and Laporte 2005; Laporte 2011) as analysis and detection techniques.

A brief description per category of techniques is provided in the subsequent sections after having examined various areas where researches and works have been conducted for application of fraud detection systems and techniques.

3.2 Supervised Learning Techniques

The supervised techniques are used in most data mining techniques. In supervised techniques a pre-specified target variable must exist and many datasets with the value of target variable must be provided to the supervised technique algorithm in order to be trained and be able to associate the target variable with the predictor variables (Larose 2005). The supervised techniques require both clear data (e.g. legitimate transaction) and fraudulent data for training purposes. The method is trained based on this data and develops models of fraudulent cases and these are used for classifying new cases. Therefore, it is important to ensure the correct classification of trained data (which data is fraudulent and which is legitimate). Moreover, supervised methods are trained with known frauds. They can only be used to detect frauds of the type that have occurred previously (Bolton and Hand 2002). They may fail to detect new strategies of fraud (Roman et al. 2009).

Supervised techniques are mainly classification and regression methods. Neural Networks (Dorronsoro et al. 1997; Feng et al. 2007; Friedlob and Schleifer 1999; Laleh and Azgomi 2009; Larose 2005; Roman et al. 2009), Bayesian Belief Networks (Laleh and Azgomi 2009; Roman et al. 2009), Decision Trees (Laleh and Azgomi 2009; Larose 2005), Statistical Outlier Methodologies of Type 2 (Hodge and Austin 2004), Neural Networks Outlier Methodologies of Type 2 (Hodge and Austin 2004), k-nearest neighbors for classification (Larose 2005), Case-Based Reasoning (Laleh and Azgomi 2009) are some supervised techniques.

3.3 Unsupervised Learning Techniques

In contrast to supervised techniques, the unsupervised techniques do not have a specific target variable. In fact, the algorithm searches the data (all variables) for identifying patterns. Unsupervised techniques include clustering methods (Larose 2005) and unsupervised outlier detection methods (Hodge and Austin 2004) that do not require prior knowledge of data.

3.4 Semi-Supervised Learning Techniques

Traditional classifiers and supervised techniques use only labelled data for training purposes, whereas, unsupervised techniques use only unlabelled data. Labelled data is not easy to be collected or prepared and most of the times is difficult requiring significant effort and time, whereas, unlabelled data can be collected relatively easily. Semi-supervised classification use large amount of unlabelled data together with some labelled data to build the classifier. Semi-supervised techniques are of great interest because consume less effort and has higher accuracy. In case of semi-supervised clustering, clustering is performed with some labelled data in the form of must-links (two points must in the same cluster) and cannot-links (two points cannot in the same cluster) (Zhu 2008).

3.5 Meta-Learning or Combining Multiple Algorithm technique

An alternative technique is to take decision or predict after combining the output for multiple algorithms or models. This technique includes *Bagging (Bootstrap Aggregating)*, *Stacking (Stacked Generalization)* and *Stacking-Bagging*. Normally, such approaches have better performance than using only one model (Laleh and Azgomi 2009).

3.5.1 Bagging (Bootstrap Aggregating)

The technique is used to combine the classifiers or outcome of different models of the same type. In Bagging approach, several models or predictors of a common learning algorithm (e.g. decision trees) are used. All models have equal weight. Each model or predictors is trained with random data set of same size. At the end, the decision is taken by an aggregated predictor. In case of numerical prediction, the aggregation is performed with the average, while for class prediction, the aggregation is the plurality of the vote (Breiman 1994, 1996; Witten and Frank 2005).

3.5.2 Stacking (Stacked Generalisation)

In contrast to *Bagging* that combines the outcome of different predictors of the same algorithm, *Stacking* combines the outcomes/results of various algorithms or base learners/generalisers. For instance, *Stacking* can be used to combine C4.5[1], CART[2], and RIPPER for developing classifiers for a specific data set. According to Wolpert (1992), when the *Stacking* approach is used to combine predictions of multiple generalizers, *Stacking* can be considered as a more sophisticated version of cross-validation. With *Bagging* approach the final decision is taken with voting approach, whereas, *Stacking* uses a *meta-learner* which tries to identify the most trustworthy classifier and use it for prediction (Witten and Frank 2005). *Stacking* combines the base generalisers possibly nonlinearly instead of applying “a winner-takes-all” approach (Wolpert 1992).

[1] C4.5 is an algorithm for generating decision trees
[2] CART is classification and regression trees (CART) method for decision trees.

3.5.3 Boosting

Boosting techniques is also used to combine multiple models with purpose those models to complement one another considering that each model is better to particular case than the other is. *Boosting* has similarities with *Bagging.* First, *Boosting* performs the aggregated prediction (or combination of various outputs) also by using the voting method in case of classification and the average in case of numerical prediction. *Boosting* is also used for combining models of the same type (e.g. decision trees). The main difference is that *Boosting* follows the iterative approach meaning that models are built iteratively and new model is influenced by the previous model. With *Bagging* approach, all models of which will be combined are built individually. Finally, *Boosting* applies the weighted vote or contribution based on the performance of the model, while *Bagging* applies equal weight to all models (Witten and Frank 2005).

3.6 Fuzzy Logic

Fuzzy logic introduced by Zadeh in 1965 and it is a mathematical tool for dealing with uncertainty (Sivanandam et al. 2007). Fuzzy logic is recognised in various fields with various applications. Fuzzy logic is used with success for decision-making and inference purposes. Application of fuzzy logic enables approximate human reasoning to be applied to knowledge-based systems (Alavala 2008).

Section 2.3 provides information about the relationship of concepts such as risk, vagueness, fuzziness, and uncertainty. It is mentioned that if the form of uncertainty happens to arise because of imprecision, ambiguity, or vagueness, then the variable is probably fuzzy and can be represented by a membership function (Virtanen and Helander 2010). As stated by Alavala (2008), the two main characteristics of fuzzy systems that give them better performance in certain applications are (1) in case of uncertain or approximate reasoning and especially where the mathematical model is difficult to be derived and (2) when decision making is performed with estimated values under incomplete or uncertain information.

3.6.1 Fuzzy Sets, Operations and Properties

In *classical (crisp) sets*, an element can have a membership value to either 0 or 1. Therefore, an element either can be a member of the crisp set or not. On the contrary, according to fuzzy set theory, elements of a fuzzy set have a membership function from 0 to 1. Hence, the elements of a fuzzy set have various degrees of membership within a fuzzy set. In addition, the same element can also be a member of another fuzzy set with different degree of membership.

If *x* is an element of the universe, which member of a fuzzy set A then the membership function is defined by the following expression (Sivanandam et al. 2007):

$$\mu_A(\chi) \in [0,1] \qquad (1)$$

A fuzzy set is said to be *normal* when at least one element has membership value equal to one. The height of a fuzzy set can be used to estimate the maximum value of the membership function using the formula **(2)**. If the height (A) is less than one then the fuzzy set is said to be *subnormal (Ross 2010)*:

$$\text{height}(A) = \max\{\mu_A(\chi)\} \qquad (2)$$

In addition, a fuzzy set can be *convex* or *non-convex*. A fuzzy set is *convex* when any element x, y and z in a fuzzy set A with x<y<z (Ross 2010):

$$\mu_A(y) \geq \min[\mu_A(\chi), \mu_A(z)] \qquad (3)$$

An intersection of two convex fuzzy sets always results to a convex fuzzy set (Ross 2010).

As happens with crisp sets, operations also exist for fuzzy sets. Below, some fuzzy sets operations are discussed:

Complement

The complement of a fuzzy set contains all elements that are not in the set. The complement of a fuzzy set consists of elements that have degree of membership 1 minus the degree of membership of the original fuzzy set (Negoita 1985). The complement of fuzzy set $\mu_A(\chi)$ is given by the following equation (Alavala 2008; Klir and Yuan 1995):

$$\mu_{\bar{A}}(\chi) = 1 - \mu_A(\chi) \qquad (4)$$

Union (OR operation)

The union of two fuzzy sets $\mu_A(x)$ and $\mu_B(x)$ is given by the following equation where *max* is the maximum operator (Alavala 2008; Klir and Yuan 1995; Negoita 1985):

$$\mu_{A \cup B}(\chi) = \max[\mu_A(\chi), \mu_B(\chi)] \qquad (5)$$

Intersection (AND operation)

As happens with classical (crisp) sets, the intersection of two sets includes the elements of one set AND the elements of the other set. According to the fuzzy set theory, the intersection of two fuzzy sets $\mu_A(x)$ and $\mu_B(x)$ is given by the following equation where *min* is the minimum operator (Alavala 2008; Klir and Yuan 1995; Negoita 1985):

$$\mu_{A \cap B}(\chi) = \min[\mu_A(\chi), \mu_B(\chi)] \qquad (6)$$

It is also worth noting that from the fuzzy set theoretical point of view the intersection of complemented fuzzy sets is equal to the intersection of original fuzzy sets (Negoita 1985), though the resulted intersections represent different concepts.

$$\mu_{A \cap B}(\chi) = \mu_{\bar{A} \cap \bar{B}}(\chi) \qquad (7)$$

Finally, a number of fuzzy sets properties exist such as *Commutativity*, *Associativity*, *Distributivity*, *Idempotency*, *Identity*, *Absorption*, *Involution* and *de Morgan's Laws* (Negoita 1985).

3.6.2 Fuzzy Rule-Based Systems

A Fuzzy Inference System (FIS) or fuzzy rule-based system or fuzzy model or a fuzzy expert system is a rule-based system or expert system (Shapiro 2004), which uses the fuzzy set theory and fuzzy logic for reasoning tasks. A fuzzy rule-based system consists of a number of membership functions and a set of rules. The set of fuzzy rules constitutes the fuzzy knowledge base or the fuzzy rule base.

A fuzzy rule is defined in the form of 'if-then'. The 'if' part is the antecedent (or premise) of the rule and the 'then' part is the consequent (or conclusion) of the rule. An example of fuzzy rule is shown (8):

$$\text{If } x \text{ is } A \textbf{ and } y \text{ is } B \text{ THEN } z \text{ is } C \tag{8}$$

Where *x* and *y* are fuzzy input variables, while, *A* and *B* represent linguistic values (membership functions) of *x* and *y* respectively used to express the antecedent part of the rule. On the consequent part, *z* is the fuzzy output variable and *C* indicates the linguistic value (membership function) of *z* output variable. Finally, the above example indicates a rule, which have two parts in the antecedent combined with the 'AND' fuzzy operator.

A FIS can be a multiple-input single-output system (MISO) or multiple-input multiple-output system (MIMO). MISO systems returns a single output based on multiple inputs whereas a MIMO system returns more than one (multiple) outputs based on multiple inputs. However, a MIMO system can be considered as multiple MISO systems working in parallel (Khanmohammadi and Jassbi 2012).

Fuzzy Inference is the method for interpreting the input values using Fuzzy Logic and based on the defined fuzzy rules in order to assign to an output. During this fuzzy inference process, concepts of fuzzy logic described before are used such as membership functions, fuzzy set operations, and rules.

The Mamdani and Takagi–Sugeno are the most commonly seen fuzzy inference methods or fuzzy models (Iancu 2012; Sivanandam et al. 2007). Mamdani is the most commonly used fuzzy inference technique. Mamdani fuzzy inference method was proposed initially by Mamdani and Assilian on 1975 attempting to control a steam engine and boiler combination by synthesizing a set of linguistic control rules obtained from experienced human operators (Iancu 2012; Sivanandam et al. 2007).

The Mamdani fuzzy inference process includes four basic activities as depicted in Figure 3-1. This figure illustrates the process using Business Process Model and Notation (BPMN). The goal of the BPMN is to be a notation understandable by all business users including the business analysts who will create the business processes, the technical team who will implement the processes and the business people who will manage and monitor those processes (OMG 2011a). Using the BPMN terminology, the process of Figure 3-1 is a non-executable private process. This means that the process has been modelled to visualise the activities of fuzzy inference process and at modeller-defined level of details. The ADONIS Community Edition (ADONIS 2014) has been used as BPM tool. BPMN provides a number of elements for modelling processes. These elements are grouped mainly in five basic categories: Flow Objects, Data, Connecting Objects, Swimlanes and Artifacts (OMG 2011a). The following paragraphs describe the elements used in Figure 3-1 for modelling the specific process.

At the beginning, the process starts and this is visualised with a start event named “Start of Inference Process”. Start events depicts where the process starts. The first step of fuzzy inference process is to fuzzify the inputs based on the defined membership functions per input variable (Klir and Yuan 1995; Ross 2010; Sivanandam et al. 2007). This step is named “Fuzzify Input” and is modelled using a task (atomic activity) in BPMN (OMG 2011a). This task is an automated task and hence it is modelled as “Service Task”. The Service Task is a special task in BPMN (OMG 2011a) and it is shown with a “Gear” marker in the upper left corner of the shape. Because the “Fuzzify Input” task shall be performed for each input, the task is indicated as "multi-instance" with parallel execution. In BPMN, this is indicated with three vertical lines as a marker in the bottom middle of the shape. The number of instances of the task depends on the number of input items. For that purpose, the "Input value" is modelled as a data object. This data object represents a collection of data and hence it has three vertical lines as a marker in the bottom middle of the shape (OMG 2011a). Finally, the "Fuzzify Input" task requires to retrieve information such as membership functions in order to fuzzify the inputs. This information is retrieved from "Knowledge Base", which is represented as Data Store. The Data Store represents a persistency, which can be used by activities or tasks to retrieve or update information (OMG 2011a). The retrieval is shown with an association having direction from the "Knowledge Base" Data Store to the "Fuzzify Input" task.

After the completion of inputs fuzzification, an intermediate event ("Fuzzification completed") is used to indicate that an important stage of the process completed. In BPMN (OMG 2011a), the intermediate event is shown in the diagram as a circle with double line. The process continues with the rule evaluation atomic activity. This is also modelled as a Service Task with name "Evaluate Rule". During this activity, each rule is evaluated based on the fuzzified inputs. Rules have a form as the one describe in (8). When the antecedent has more than one parts then those parts are linked with fuzzy set operators (AND, OR, Complement). These fuzzy set operators are applied during the antecedent evaluation in order to conclude to a fuzzy value. The fuzzy value from the antecedent part is evaluated based on the consequent part of the rule (membership function) and by applying the implication method. In addition, the weight of rules affects the fuzzy output result for the rule evaluation (Klir and Yuan 1995; Ross 2010; Sivanandam et al. 2007). The "Evaluate Rule" task is executed for each fuzzified input. Therefore, the "Evaluate Rule" task is modelled as multi-instance parallel activity (three vertical lines as a marker in the bottom middle of the shape). The number of instances of the task depends on the input data. The "Fuzzified input" data collection is an output of the "Fuzzify Input" task and input for “Evaluate Rule” task. In addition, the “Evaluate Rule” task requires information such as fuzzy rules, methods for fuzzy set operators (AND, OR) and implication method. The fuzzy rules are retrieved from the “Knowledge Base” Data Store. The methods for fuzzy set operators (AND, OR) and implication method are considered as configuration parameters of fuzzy inference system and therefore are modelled in this case as Data Store named “Configuration Parameters”. This Data Store is also used by the “Evaluate Rule” task in order to retrieve the required information for its execution. At the end of this activity, a fuzzy output for each rule occurs. The intermediate event (“Rule evaluated, one output per rule”) is used to indicate this as an important stage of the process.

Data output of the “Evaluate Rule” task is the “Rule Fuzzy Output” data collection. This is an input for “Aggregate outputs” activity, which aggregates all individual results based on the selected aggregated method (Klir and Yuan 1995; Ross 2010; Sivanandam et al. 2007). This activity is also modelled as a Service Task. The “Aggregate outputs” task requires retrieving information for the aggregation method from the “Configuration Parameters” Data Store. After the completion of this activity, the result is an aggregated fuzzy output. This output is modelled as a data object named “Aggregated Fuzzy Output”. In addition, the intermediate event “Output Aggregated” is used to model this important stage of the process.

The “Aggregated Fuzzy Output” is an input to the “Defuzzify Output” activity. This defuzzification activity is used to determine the output crisp number by applying the selected defuzzification method (Klir and Yuan 1995; Ross 2010; Sivanandam et al. 2007). The latter is considered that it is defined as a configuration parameter of the fuzzy inference system. The “Defuzzify Output” task retrieves this information from “Configuration Parameters” Data Store. Upon the completion of this activity, the defuzzified output value is produced. This is modelled as data output object with name “Output value”. Finally, the completion of the process is modelled as end event with name “Fuzzy Inference completed”. The end event is shown with a circle with thick single line (OMG 2011a). The defuzzification methods are discussed in the following paragraphs.

The most common used defuzzification method for Mamdani's inference is the Centroid or Center of Area or Center of Gravity (COA). Considering the fuzzy conclusion of inference is that *z* is *C* then the output is given by the following equation (Bai and Wang 2006):

$$COA(z) = \frac{\sum_z \mu_c(z) \times z}{\sum_z \mu_c(z)} \tag{9}$$

If *z* is a continuous variable, this defuzzification result is

$$COA(z) = \frac{\int \mu_c(z) z dz}{\int \mu_c(z) dz} \tag{10}$$

Another defuzzification method is the *Mean value Of Maximum* (MOM) (or *Mean–max-membership* or *middle-of-maxima*) which computes the average of those fuzzy outputs that have the higher (maximum) membership. As a limitation of this method could be that, only the maximum (highest) membership values are considered and hence the same result will be produced for membership functions that have different shapes but the same maximum membership values (Bai and Wang 2006). Other defuzzification methods are Bi-sector of area (BOA), Largest (absolute) value Of Maximum (LOM) and Smallest (absolute) value Of Maximum (SOM). BOA is the value at which a vertical line is placed dividing the region into two sub-regions of equal area. BOA is expressed as follows (Naaz et al. 2011):

$$\int_a^{z_{BOA}} \mu_c(z) dz = \int_{z_{BOA}}^{b} \mu_c(z) dz \tag{11}$$

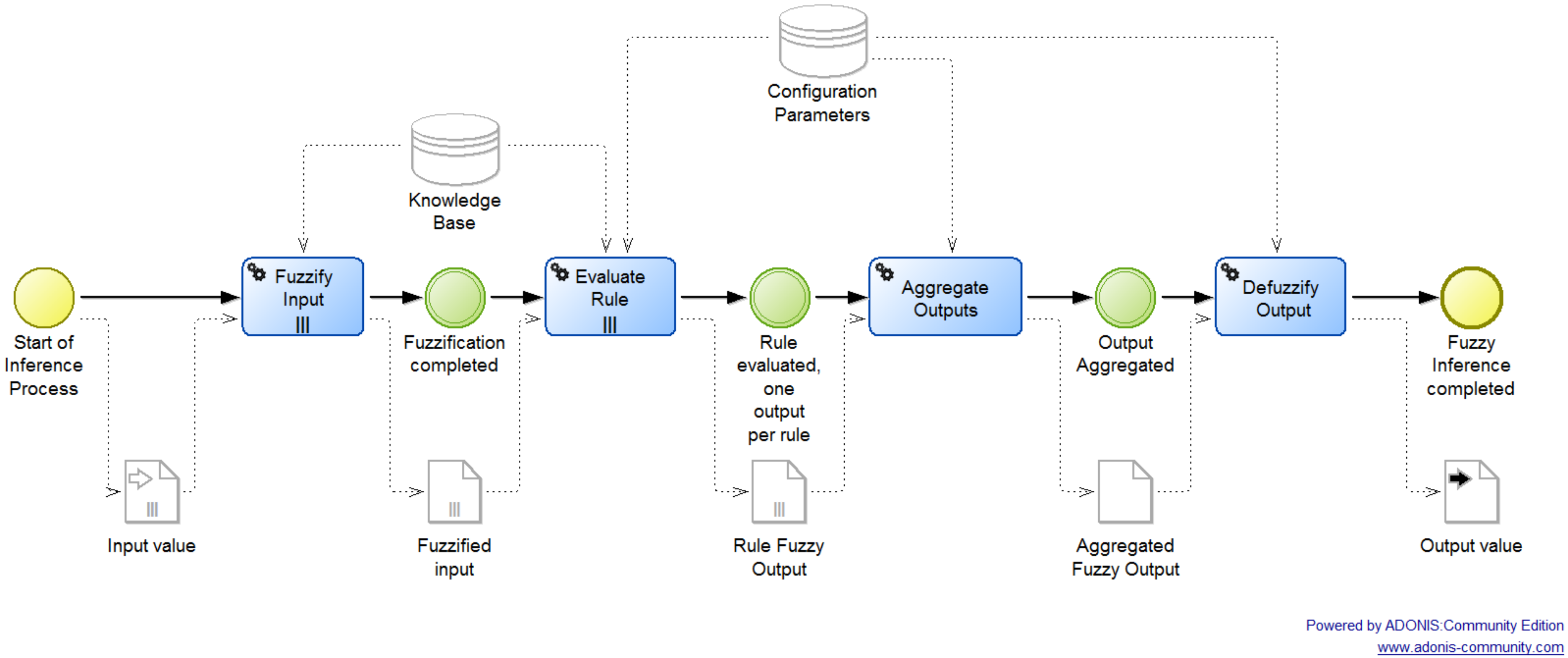

Figure 3-1: Fuzzy Inference Process

SOM and LOM are also referred as First of maxima and Last of maxima respectively. SOM uses the lowest bound of the maximum (highest) membership values in the aggregated membership function output while LOM uses the least upper bound of the maximum (highest) membership values in the aggregated membership function output. Some other defuzzification methods that are not explicitly discussed are *Max membership principle* (or *height method*), *Center of Sums*, *Center of largest area* (Ross 2010).

The TSK or Sugeno fuzzy model was proposed by Takagi, Sugeno, and Kang (TSK) "*in an effort to formalize a system approach to generating fuzzy rules from an input–output data set*" (Sivanandam et al. 2007, p. 123). The main difference with Mamdani method is that the output of Sugeno model is either linear or constant and it is expressed as a function of the input. Therefore, the rules in a Sugeno model take the following form (Sivanandam et al. 2007):

$$\text{If } x \text{ is } A \textbf{ and } y \text{ is } B \text{ THEN } z \text{ is } f(x,y) \tag{12}$$

As above, *x* and *y* are fuzzy input variables, while, *A* and *B* represent linguistic values (membership functions) of *x* and *y* respectively used to express the antecedent part of the rule. In addition, the fuzzy set operator *AND* is used to combine the two parts of the antecedent. On the consequent part, *z* is the fuzzy output variable and *f(x,y)* represents a function which is frequently polynomial or another function defining the relationship with inputs. If the output is a first order polynomial then the Sugeno model is called as first-order Sugeno model. The rule shown in **(12)** takes the following form (Sivanandam et al. 2007):

$$\text{If } x \text{ is } A \textbf{ and } y \text{ is } B \text{ THEN } z = \mathrm{ax} + \mathrm{by} + \mathrm{c} \tag{13}$$

Where *a*, *b* and *c* in the formula **(13)** are constants. In case the *a* and *b* are equal to zero (0) then, the output is defined with a constant crisp value *c* and the Sugeno model is called as *zero-order Sugeno model*. This case can be seen as a special case of Mamdani method where the output is defined as a fuzzy singleton (Ross 2010; Sivanandam et al. 2007).

In terms of fuzzy inference process, the Sugeno fuzzy inference method has similar activities to the Mamdani. One difference is that during rule evaluation, the output of each rule is weighted by the firing strength of the rule. The firing strength is calculated by applying the operator in the antecedent part of the rule (e.g. apply AND or OR operator appropriately). Therefore, an output (w_i) is calculated for each rule weighted by the firing strength. For the rule presented in **(13)** where the AND operator is used to combine the antecedent parts then the is w_i is calculated as follows (Sivanandam et al. 2007):

$$w_i = AND\big(\mu_A(\chi), \mu_B(y)\big) \tag{14}$$

Where $\mu_A(\chi)$ is the membership function A for fuzzy input variable *x* and $\mu_B(y)$ is the membership function B for fuzzy input variable *y*. The AND function is applied depending on the selected method (e.g. minimum)

The aggregation in Sugeno method is the sum of the individual rule outputs. Finally, the defuzzification process is performed by applying the Weighted Average method (WTAVER) as the sum of all weighted average rule outputs (Braglia et al. 2003; Sivanandam et al. 2007):

$$WTAVER = \frac{\sum_{i=1}^{N} w_i \times z_i}{\sum_{i=1}^{N} w_i} \tag{15}$$

Where:

- *WTAVER* is the weighted average of the output result.
- *N* indicates the number of output fuzzy sets;
- z_i symbols at which the i-th membership function reaches its maximum value.

Considering all the above, the main difference between Mamdani and Sugeno methods is that Sugeno output membership functions are linear or constant. Therefore, the consequent part of fuzzy rules for Sugeno fuzzy inference systems are expressed using functions. This fact differentiates the inference process of the

Sugeno method in the aggregation and defuzzification activities as described above. Mamdani method is considered a better approach for expressing human knowledge considering that also outputs can be expressed with fuzzy sets. Therefore, Mamdani is easier for the experts to express knowledge and this is the reason of widely acceptance of this method for decision-making applications using fuzzy logic (Hamam and Georganas 2008; Sivanandam et al. 2007). Moreover, Mamdani method can be used for both MIMO and MISO systems (Hamam and Georganas 2008; Jassbi et al. 2006). On the contrary, Sugeno method fits only for MISO systems (single output) according to Jassbi et al. (2006). Sugeno is considered more efficient computationally since Mamdani's defuzzification process is more complex than Sugeno's, which uses weighted average (Hamam and Georganas 2008; Jassbi et al. 2006; Sivanandam et al. 2007). In addition, Sugeno method is used by adaptive techniques for constructing/optimising fuzzy models, which best models the data (Sivanandam et al. 2007). An example is Adaptive Neuro-Fuzzy Inference System (ANFIS) technique, which is used to construct fuzzy models based on data set (adaptive technique). More information for this technique is provided in section 3.7.

A number of defuzzification methods mentioned previously for both Mamdani and Sugeno fuzzy inference systems. One common question is which method is the best or should be used. Hellendoorn and Thomas (1993) (as cited in Ross 2010) have defined five criteria against which to measure defuzzification methods. These are *continuity, disambiguity, plausibility, computational simplicity,* and *weighting method. Continuity* refers to the fact that a small change in the input of the fuzzy process will not result a big change on the output. *Disambiguity* means that the defuzzification method should always result to a unique defuzzified output value and hence no ambiguity for the output. *Plausibility* refers to whether the output defuzzified value is plausible and in order to be has to lie in the middle of output membership function and also with high membership value. For instance, there are cases where a centroid might result to a value that does not exhibit plausibility because might lie in the middle but not with high membership value. The *computational simplicity* measures how time consuming a method is because this affects a computation system. As an example, methods such as MOM and SOM are computationally simpler (faster) than centroid. Finally, *weighting method* criterion is used to weight the output fuzzy sets (Ross 2010).

3.7 Adaptive Networked-Based Fuzzy Inference System (ANFIS)

The Adaptive Network-Based Fuzzy Inference System (ANFIS) – also called Adaptive Neuro-Fuzzy Inference System – introduced by Jang (1993) by embedding FIS into the framework of Adaptive Networks.

ANFIS consists of five layers each one implementing different node functions for learning and tuning FIS parameters using a hybrid-learning mode (Wei et al. 2007). The output of each adaptive node depends on modifiable parameters applied to these nodes. Those parameters are updated to minimise error based on the learning rule. Figure 3-2 presents a typical ANFIS system with two inputs. The square nodes represent adaptive nodes (node function) whereas circle nodes denote fixed nodes (Güneri et al. 2011).

Layer 1: Every node *i* in layer 1 is a square node with the following function (Jang 1993):

$$O_i^1 = \mu_{A_i}(x) \tag{16}$$

Where: where x is the input to node i, A_i is the linguistic label, and O_i^1 is the membership function of A_i. Parameters in this layer are defined as *premise parameters*.

Layer 2: every node in this layer multiplies the inputs to that node and calculates the product (Jang 1993):

$$w_i \quad = \mu_{A_i}(x) \times \mu_{B_i}(y), \quad i = 1{,}2 \tag{17}$$

This node represents the firing strength of the rule.

Layer 3: Every node *i* calculates the ratio of the *i* rule's firing strength over the sum of all rules' firing strength (Jang 1993):

$$\overline{w_i} = \frac{w_i}{w_1 + w_2}, \quad i = 1{,}2 \tag{18}$$

Layer 4: Every node *i* in layer 4 is a square node with the following function (Jang 1993):

$$O_i^4 = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i) \tag{19}$$

Where: where $\overline{w_i}$ is the output of layer 3. The p_i, q_i, r_i are the parameters, which are defined in this layer as *consequent parameters*.

Layer 5: this node calculates the overall output as the sum of all inputs to this node as follows (Jang 1993):

$$O_i^5 = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{20}$$

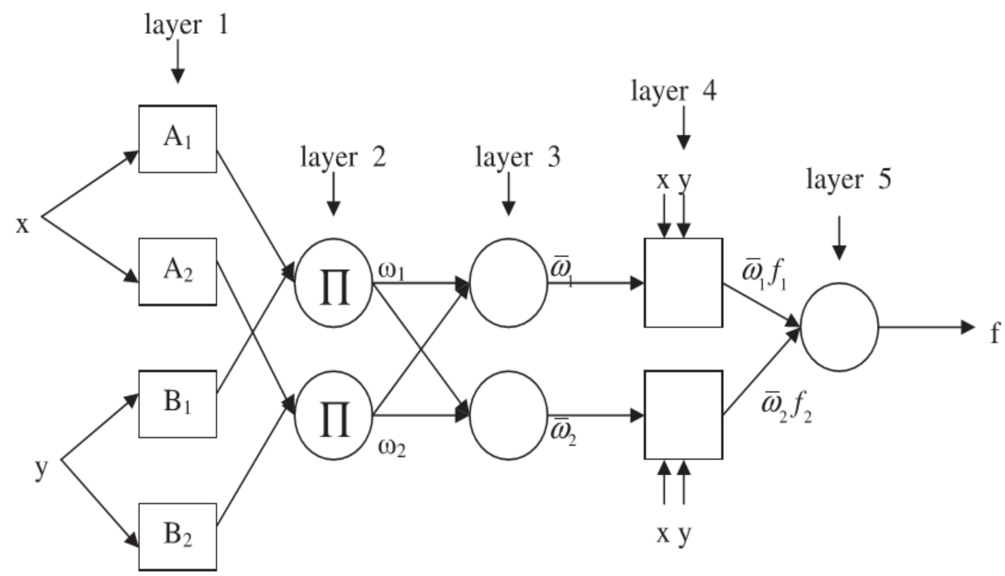

Figure 3-2: Adaptive Neuro-Fuzzy Inference System structure (Güneri et al. 2011)

The identification of FIS (including fuzzy rule base and membership functions) is very important. Two partition techniques that could be used for the generation of a

FIS (Sugeno type) are the *Grid Partitioning* and *Subtractive clustering* techniques. These techniques could be used as the initial model for ANFIS training.

The *Grid Partitioning* is used to generate a MISO Sugeno FIS. The *Grid Partitioning* technique divides the data space into rectangular sub-spaces using axis-paralleled partition based on pre-defined number of membership functions and their types in each dimension. An issue is the so-called *curse of dimensionality* where the number of fuzzy rules is increased exponentially with the increase of the number of inputs (and increase of MF per input). Therefore, it is considered that *Grid Partitioning* is suitable for cases with less than 6 input variables (Wei et al. 2007).

The *Subtractive clustering* is proposed by Chiu (as cited in Wei et al. 2007) by extending the mountain clustering method. This method clusters data points in an unsupervised way by measuring the potential of data points in the feature space. The technique considers each data point as a potential cluster center and calculates the potential for each data point based on the density of surrounding data points. The first cluster center is the data point with highest potential and the data points near to the first cluster center (within the influential radius) are destroyed. Next cluster center is data points with the highest remaining potential and the data points near to the next "new" cluster center (within the influential radius) are destroyed. Therefore, the *influential radius* is very important for the number of clusters. The selection of small *influential radius* results many clusters and hence more rules are defined, and vice versa (Wei et al. 2007).

Following the definition or generation of FIS, the optimisation and training of ANFIS is performed. The hybrid-optimisation algorithm can be used for that purpose by applying the least-squares method (forward pass) and back propagation method (backward pass). In the forward pass, the model is executed with fixed parameters until layer 4 and then the least-squares method is used to identify the consequent parameters. In the backward pass, the consequent parameters are fixed and the error rates propagate back and the premise parameters are updated by applying the gradient descent method. Repeating forward and backward passes will enable the tuning and optimisation of the model. (Jang 1993; Wei et al. 2007).

3.8 Ontologies

Gruber (1993) states that ontology *is an explicit specification of a conceptualisation.* In the context of Artificial Intelligence (AI), the ontology of a program can be defined with a set of representational terms (Gruber 1995). According to Uschold and Gruninger (1996) *Ontology is the term used to refer to the shared understanding of some domain of interest which may be used as a unifying framework to solve the above problems in the above described manner.* The ontologies have various usages. One of them is the *communication* between people, groups or roles with different needs so as ensuring that all parties share the same understanding. Finally, another usage of ontologies is for *interoperability* purposes (Uschold and Gruninger 1996). In reality, ontologies enable the consistent knowledge representation and enable all parties to have at the end the same perception for various concepts.

Obrst (2010) distinguishes *Ontological Architecture* from *Ontology Architecture*. Briefly, *Ontology Architecture* concerns the ontology development, deployment, and maintenance as well as ontology application interaction. On the other hand, *Ontological Architecture "is the architecture that is used to structure the ontologies that are employed by Ontology Architecture".* An *Ontological Architecture* consists mainly of three layers: *upper ontologies*, *middle-level ontologies* and *domain ontologies*. The first two layers are also called "foundational ontologies". *Upper ontologies* describe more basic and general concepts that can be used from *domain specific ontologies*. *Middle-level ontologies* are used to make the bridge between the *upper ontologies* and the *domain ontologies.* Although it is stated that ontologies can be mapped at any level, the *middle-level ontologies* can facilitate the mapping of concepts by representing concepts that are more concrete. As an example, ontologies for *Time* and *Location* are mentioned. Finally, *domain ontologies* define concepts and relationships specific for a domain. The *domain ontologies* might use/mapped to concepts of *middle-level ontologies* or *upper ontologies.* In addition, *domain ontologies* can extend *middle-level ontologies.* This can enable the reusability of existing developed ontologies expressing relevant concepts. The ontologies, which are presented in Chapter 5, concern domain specific ontologies and the described ontological architecture focuses on the *domain ontologies layer*.

Ontology technology has been used in various domains. Various researches and works have been performed. The development of ontology for modelling domain knowledge is a lengthy process involving a number of activities. Researches have examined the automation of ontology model development from existing data models (Albarrak and Sibley 2011).

It is very important to choose the appropriate representation language for expressing the ontology of a specific domain (Shanks et al. 2003; Uschold and Gruninger 1996). The Ontology Web Language (OWL) is a language for representing ontologies. The OWL facilitates greater machine interpretability of the content, and therefore enables the processing of the content from computers. There is the need for an expressive language, which will be used for representing various information of the web. Such language will make feasible the processing of information from machines and the performance of reasoning tasks. The OWL provides the following three sub-languages: OWL-Lite, OWL-DL and OWL-Full. Each sub-language has different expressiveness. The OWL-Lite is the least expressing sub-language while the OWL-Full is the most expressive sub-language. On the other hand, the OWL-DL sub-language is more expressive than OWL-Lite and less expressive than OWL-Full.

There are domains that require representation and reasoning to handle imprecision or vagueness. In such domains, fuzzy or vague concepts are not expressed appropriately by conventional or crisp ontologies (Huang et al. 2011; Loia 2011; Yaguinuma et al. 2013). This research investigates a fuzzy knowledge-based approach to risk analysis in the Customs domain for handling imprecision and vagueness. Many researches and works have been performed on the area of fuzzy ontologies and reasoning. Bragaglia et al. (2010) state that various initiatives have been done in the last few years for implementing fuzzy reasoning either in the context of ontological reasoning or in the context of fuzzy rule-based reasoning. It is also mentioned that researches follow two approaches for the integration of ontological engineering with the fuzzy rule-based reasoning. One is the “tight integration” and the other is the “loose integration”. In case of “tight integration”, a single, unified framework is defined for reasoning tasks. On the contrary, the “loose integration” focuses on combining available technologies in order to satisfy specific requirements (Bragaglia et al. 2010).

Description Logics (DL) is a family of logics for knowledge representation (Bobillo and Straccia 2011). DL defines concepts and role restrictions that can automatically derive classification taxonomies (Davies et al. 2003). Fuzzy DLs are extensions of classical DLs aiming to handle vague or fuzzy concepts. One of them, the fuzzyDL reasoner is proposed by Bobillo and Straccia (2008) aiming to implement fuzzy reasoning in the context of ontological reasoning. Nevertheless, it appears that fuzzyDL supports only LOM, SOM and MOM as defuzzification methods. Bobillo and Straccia (2011) state that there are several researches and implementations on Fuzzy DL reasoners and each one implements its own fuzzy DL language. They propose to represent fuzzy ontologies with the use of OWL2 annotation. This fuzzy ontology representation is integrated also with Fuzzy DL reasoners such as fuzzyDL (Bobillo and Straccia 2008).

Guillaume and Charnomordic (2012) state that the most common types of fuzzy rules are the conjunctive rules and the implicative rules. The conjunctive rules are used in Mamdani and Sugeno FIS. Usually the minimum operator is used for the conjunction of rules. The conjunctive rules are combined disjunctively. The implicative rules use fuzzy implications and they are combined conjunctively (Guillaume and

Charnomordic 2012). According to Yaguinuma et al. (2013), fuzzy ontologies with Fuzzy DLs provide implication operators and hence, they use implicative rules for reasoning. As mentioned before, the implicative rules are combined conjunctively.

On the other hand, there are proposals, which consider crisp ontologies integrated with fuzzy rule-based reasoning. Wlodarczyk et al. (2011) present the SWRL-F, a Fuzzy Logic extension of Semantic Rule Web Language (SWRL). The SWRL (W3C 2004) is syntax for expressing rules using the OWL Knowledge Base. It combines the OWL (DL and Lite) with Rule Markup Language (RuleML). The RuleML is the Rule Markup Language, which developed to express Web rules in XML. According to Wlodarczyk et al. (2011), one of the design decisions for proposing SWRL-F is to follow the principles of fuzzy control systems or fuzzy rule-based systems, i.e. fuzzification, inference and defuzzification. The fuzzy inference with SWRL-F is based only to the rules. The ontology is used for describing the fuzzy knowledge base only. A DL reasoner can be used but it can interpret the ontology based on crisp logic. Object properties have been used for defining the fuzzy assertions in SWRL, which can be interpreted by fuzzy rule reasoners but not from non-fuzzy rule reasoners. The implementation of fuzzy rule engine was based on FuzzyJess. Although the approach above used fuzzy rule-based inference, it is not clear whether the Mamdani inference was used. However, it is understood that FuzzyJess provides the Mean of Maximum (or Average of Maximum) and the Center of Gravity as defuzzification methods (Orchard 2001).

In the context of fuzzy rule-based reasoning, better interoperability is achieved by using a more standardised way to represent fuzzy rules. For that purpose, there are some researches, which exploit the use of Fuzzy Markup Language (FML) and its integration with fuzzy ontologies (Huang et al. 2011; Lee et al. 2009; Yaguinuma et al. 2013). The Fuzzy Markup Language (FML) is an XML-based language for representing Fuzzy Logic Controllers (FLC) or Fuzzy Rule-Based Systems (as defined in 3.6.2). The FML has been proposed by Acampora and Loia

(2005). The FML was initially designed as a middleware between the various platforms. For instance, it could be used to transform a FLC defined in MATLAB (as FIS format) into FML with appropriate transformation. In addition, a FLC defined in FML can be transformed to other platforms depending on the requirements. This transformation can be achieved using Extensible Stylesheet Language (XSL) and XSL Transformations (XSLT) (W3C 2001).

Huang et al. (2011) presents a work for Malware Behaviour Analysis with the integration of fuzzy ontology with FML for knowledge modelling of Malware Behaviours and intelligent decision making for detecting computer anomalies. It is understood that this proposal also use fuzzy ontology for representation of FIS Knowledge Base and focuses on fuzzy rule-based reasoning.

Yaguinuma et al. (2013) proposes an FML-based hybrid reasoner integrating fuzzy ontology and Mamdani reasoning. They mention that supporting the Mamdani FIS, an output value (defuzzified) can be inferred after assessing output of rules and using defuzzification method that consider the shape of the output fuzzy set. They also mention that some proposals use specific formats such as fuzzyDL, Jess and Drools, which are handled by their corresponding inference engines. On the other hand, FML can enable interoperability among platforms. As far as the Mamdani-FIS defuzzification methods, they mention that the Center of Area (COA) and Middle of Maxima (or Mean Value of Maximum) methods were used. In Yaguinuma et al. (2013) work, some interesting experiment comparison results are presented for a certain application scenario. The results from Mamdani FIS are compared with inferences obtained from fuzzyDL reasoner. However, Yaguinuma et al. (2013) mentions that two fuzzyDL approaches are compared with Mamdani-FIS results since fuzzyDL do not provide specific constructors for Mamdani rules representation and reasoning. The first approach is fuzzyDL implicative rules and the second using fuzzyDL concept definitions and defuzzification queries. Regarding the second fuzzyDL approach and as mentioned before in fuzzyDL discussion, it provides the

LOM, SOM, and MOM as defuzzification methods. One conclusion from the experiments of Yaguinuma et al. (2013) work is that for some individuals or instances, the fuzzyDL implicative rules approach cannot conclude inferences. The reason given by the authors of this work is that for some rules consequents do not have intersection and are combined conjunctively (see information provided previously for conjunctive rules and the implicative rules). It is apparent that the Mamdani FIS is considered more appropriate because in most cases an answer or inference is expected. Comparing the second approach of fuzzyDL (using LOM, SOM, and MOM as defuzzification methods) with Mamdani-FIS (COA and MOM as defuzzification methods), Yaguinuma et al. (2013) mention that some instances cannot be distinguished and hence ranked/classified (the same defuzzified output) with defuzzification methods using LOM, SOM, and MOM, however MOM has more appropriate ranking than LOM and SOM.

Following a literature review, some work on ontologies relevant to the Customs under this work, are by Zang et al. (2008) and by Dimakopoulos and Kassis (2008). In the first research, a domain ontology for import and export procedures has been developed to acquire Harmonised System (HS) codes for given products. In fact, the ontology is used for reasoning and particularly for specifying intelligently the HS code of a given product based on its product name. According to the authors of this work, the ontology is intended to be used by the Customs and quarantine departments in order to automate and improve their inspections processes since the HS code can be used to identify the applied policies to the product. Hence, accurate assignment of an HS code to a product implies more efficient and effective inspection. The second work (Dimakopoulos and Kassis 2008) concerns a layered ontology, which includes Customs domain concepts (inward processing and to export customs procedures) and Risk Assessment Ontology. These ontologies developed under RACWeb project, which co-funded by the European Commission under the “Information Society Technology” Programme, Framework Programme 6. The purpose of these ontologies is to store knowledge and then to query this knowledge in the context of risk assessment. Finally, the OWL-DL was used for expressing the ontology.

This research examines the ontologies for modelling concepts at various levels starting from concepts that are more generic and going to more specific to risk analysis. The ontologies are domain specific. Architecture of ontologies is also presented. Ontologies of this project are analysed in Chapter 5. This research follows the approach of "loose integration". It uses crisp ontologies for defining knowledge and focuses on fuzzy rule-based reasoning. For instance, ontology is used to define the concepts of *fuzzy risk model*, which is presented for risk analysis with fuzzy logic technique (5.5.2). This includes the definition of fuzzy variables, fuzzy sets, etc. as concepts. Moreover, it is used for representing the actual knowledge base of a *fuzzy risk model* (or FIS) such as fuzzy variables, membership functions, and fuzzy rules. Therefore, a FIS is constructed for fuzzy rule-based reasoning based on represented knowledge with ontologies. XML-based representation is used for representing this FIS for interoperability purposes. XML representation can also be transformed to FML. Nevertheless, the main reason for selecting in this work the aforementioned approach is the flexibility of using Mamdani-FIS with a number of defuzzification methods such as COA. However, some approaches presented previously used specific defuzzification methods. In this research, five defuzzification methods are examined for Mamdani FIS reasoning. In addition, the Sugeno FIS type can also be examined for fuzzy reasoning and inference. The fuzzy modelling and reasoning is one of the objectives of this study in the area of risk analysis in the Customs Domain. The *fuzzy risk model* should be able to infer an appropriate output depending on the input and based on the defined knowledge base. As discussed before in the review of other researches, this might not be the case for other approaches (e.g. use of implicative rules).

3.9 Summary

A review of state of the art detection techniques for fraud and risk has been performed above. Fuzzy logic, fuzzy inference systems and ANFIS are also analysed taking into consideration that an aim of this research is to examine a fuzzy knowledge-based approach for risk analysis and detection purposes. Fuzzy reasoning is investigated for handling imprecise knowledge and vagueness in risk analysis. The details of inference process for fuzzy reasoning described above have been considered, assessed, and applied in the work performed in the context of this thesis. It is important to understand the various steps of the process because certain decisions must be taken for each step during the development of the model. For instance, a fuzzy inference model type, defuzzification method, aggregation method, etc. shall be selected since the fuzzy reasoning will be performed based on these configuration parameters. Finally, ANFIS technique, which combines Artificial Neural Networks and Fuzzy Logic, could assist in the construction or optimisation of adaptive fuzzy inference systems from a given data set for performing reasoning tasks. Understanding of ANFIS architecture is important for using the technique (e.g. selection of partitioning technique for the generation of a Sugeno type FIS). In this chapter, ontologies are also analysed in the context of fuzzy knowledge-based approach to risk analysis for semantic modelling and knowledge representation. The semantic modelling is used for representing knowledge and concepts of Customs domain enabling both the communication and understanding.

Chapter 4. Fuzzy Knowledge-Based Approach to Risk Analysis

This chapter presents one of the contributions of this research and particular to develop a conceptual model combining fuzzy modelling and reasoning and semantic modelling with ontologies in the context of risk analysis of a *Physical Entity* in Customs domain. At the beginning, the conceptual model is described that would support decision-making based on analysis of risk. After, a high-level architecture of this fuzzy knowledge-based approach to risk analysis is presented and each component of this architecture is described. Finally, some abstract Use Cases and processes are also defined for elaborating this approach.

4.1 Overview

An approach is presented to support the analysis and detection of the risk of a *physical entity* (e.g. consignment) in Customs domain. This approach combines the fuzzy reasoning and the semantic modelling. At this point, it is considered that risk management is broader, multi-dimensional process and involves a number of tasks, activities, and practises. This chapter presents an approach focusing on the analysis of risk of a Customs domain-specific *physical entity* with the definition of fuzzy inference systems. This is based on the outputs of the risk management process. The risk analysis would support decision-making on whether further treatment and actions should be performed accordingly. For instance, a risk analysis of an import consignment can use the declaration data. The outcome of risk analysis of a consignment can be one of the criteria for selecting the consignment to perform inspection (EEC 2008b).

The presented conceptual model (Figure 4-1) uses the ontologies for defining various concepts and for modelling various semantics. Complex domains have many concepts, complex relationships, and semantics and hence such modelling is considered as useful. It enables the unambiguous definition of concepts and the common understanding. In this work, the ontology reasoner is used for checking the consistency of defined ontology definitions. In addition, the ontology reasoner could be used to interpret the ontology based on crisp logic. However, apart from domain specific knowledge, the ontologies are also used in this work to represent FIS related knowledge. For the needs of this research, the inference is investigated with Mamdani and Sugeno fuzzy inference systems. Having examined various approaches in section 3.8, this is a decision for this research with the rationale that it enables a more loosely couple way of integration and there are no restrictions mentioned in section 3.8 (e.g. only specific defuzzification methods can be used). Considering that different type of concepts should be represented with ontologies, an Ontological Architecture is presented. In addition to this, it is considered that this architecture offers modularity, extensibility, maintainability, and re-usability. This is further discussed in section 5.3.

The conceptual model is based on fuzzy logic and fuzzy inference systems with purpose to consider imprecise knowledge and vagueness. This is also discussed in section 2.3. Therefore, the inference or risk analysis is performed based on defined *fuzzy risk models*. The semantics of those *fuzzy risk models* are defined with ontology model (more information can be found in Chapter 5). Fuzzy logic has been applied in various fields and is used with success for decision-making and inference purposes. Application of fuzzy logic enables approximate human reasoning to be applied to knowledge-based systems (Alavala 2008).

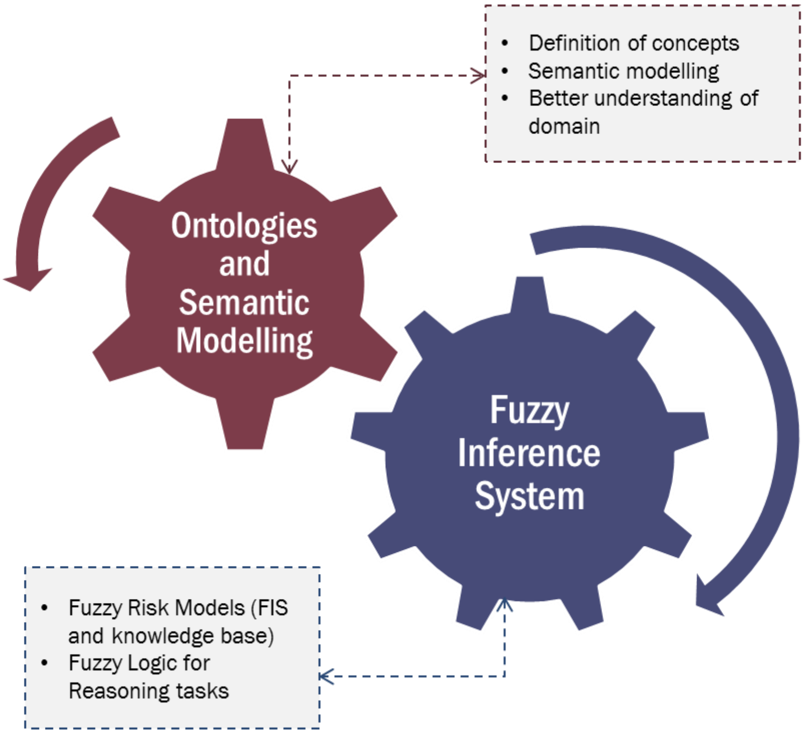

Figure 4-1: Conceptual model for risk analysis with fuzzy knowledge-based approach

The conceptual model (Figure 4-1) is considered generic and it could be explored for other domains. The following sections aim to provide more information for the terms *Physical Entity* and *fuzzy risk model.*

4.1.1 Physical Entity

In this concept and work, the *Physical Entity* term is used to refer to a *physical entity* of the domain, which will be analysed for risk by applying the relevant *fuzzy risk models*. *Physical entity* could be for instance an import consignment. A *Physical Entity* is considered that has various attributes or elements. Some *physical entities* may use the attributes of another *physical entity* or include the definition of other *physical entities.* In this concept, a *physical entity* is defined with ontology, which might refer to the ontologies of other *physical entities*. This is further discussed in sections 5.3 and 5.5.2.

4.1.2 Fuzzy Risk Models

In this approach, a *fuzzy risk model* term is used to represent a *fuzzy model* or a *fuzzy inference system* consisting of fuzzy parameters, membership functions, rules and other attributes of fuzzy inference system. Additionally, the fuzzy parameters or variables of a *fuzzy risk model* are mapped with the attributes of corresponding *physical entity*. The structure of a *fuzzy risk model* is defined by the “Fuzzy Risk Model Ontology”. A *fuzzy risk model* could be considered as a MISO FIS, however, depending on the domain needs *fuzzy risk model* could be a MIMO FIS. According to Torra (2001), two difficulties arise when the application domain of fuzzy knowledge-based systems is a complex system. These are related to the number of variables of the system and the application domain. Usually, the number of variables of the system is large in complex domains. Therefore, the number of required rules is increased exponentially. This is also called “curse of dimensionality”. In addition, the environment changes and hence, those changes cannot be modelled easily with the variables. The Hierarchical Fuzzy Systems (HFS) is a technique for handling the “curse of dimensionality” by decomposing the system into smaller more modular systems connected with input/output variables. The inference is chained among modules of rules. As far as the changing environment is concerned, adaptive intelligent control techniques can be used to handle this Torra (2001). In order to handle changing environment, this approach includes ANFIS as component for assistance/optimisation.

Considering the above, *a fuzzy risk model* might concern a specific area or have specific purpose. If the final output of risk analysis should be estimated with the execution one or more *fuzzy risk model* then this is a matter of definition of risk analysis for the particular *Physical Entity.* It is considered that this definition should specify the hierarchy and the sequence of execution of various models considering that the definition of each *fuzzy risk model* shall describe the input and the output and other details such defuzzification method (Figure 4-2). This is the concept of Hierarchical Fuzzy Logic Controller (Horácek and Binder 1997; Singh et al. 2003) or cascade structure (Jurgutis and Simutis 2011) or stage-wise fuzzy reasoning structure (Dahal et al. 2005) or Hierarchical Hybrid Fuzzy Controllers (HHFC)

(Chiaberge et al. 1995). According to Chiaberge et al. (1995), HHFC differs from traditional HFS since it enables various combinations and can have arbitrary number of outputs. Input to an internal Fuzzy Controller of HHFC can be a state variable or output from other Fuzzy Controller or both. Outputs from Fuzzy Controller at any hierarchical level can coincide with control variables or can be inputs to Fuzzy Controllers of any successive hierarchical levels.

Nevertheless, in the case where the risk analysis of a *Physical Entity* involves more than *fuzzy risk model,* it is necessary to consider and select the appropriate defuzzification method per *fuzzy risk model* otherwise the final output might not be the expected one. As stated by Jurgutis and Simutis (2011) in their work, while the cascade goes downwards, the fuzzy logic systems in the next layers converge towards to the 'center', which means reach to the same conclusion. As a solution in this specific issue, they used 'LOM' as defuzzification method and 'Centroid' in the last fuzzy logic system. Of course, this is also a matter of decision and evaluation of risk analysis of specific *Physical Entity* by also checking the outputs of individual *fuzzy risk models.*

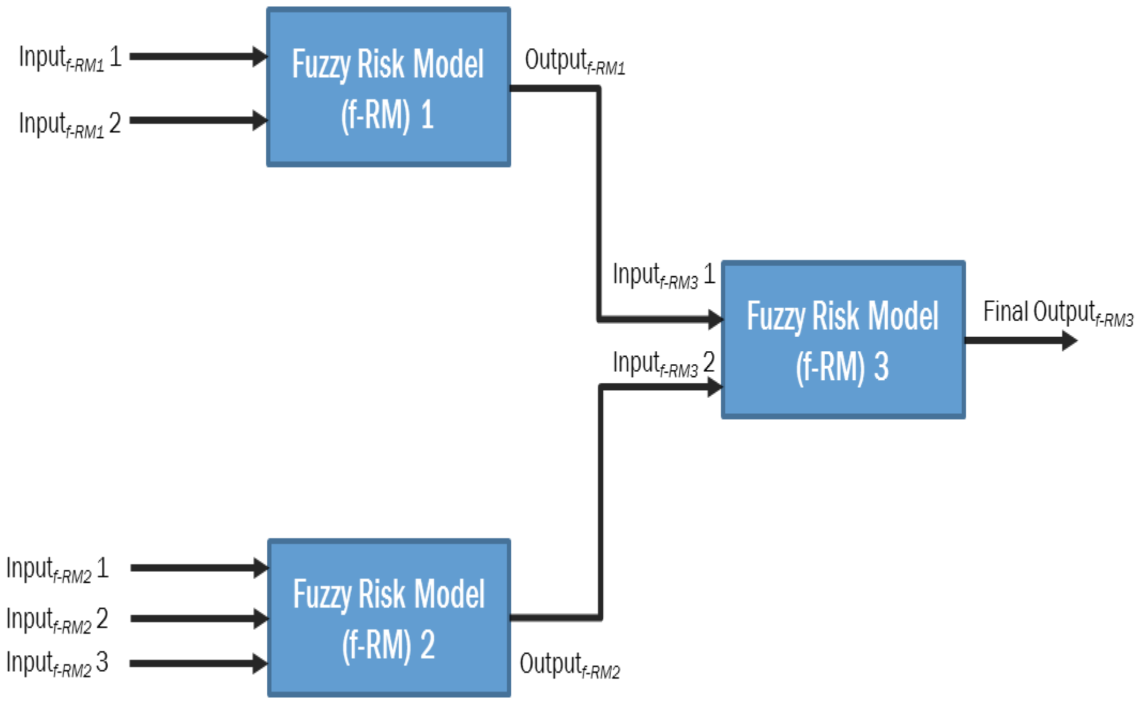

Figure 4-2: Example of concept of risk analysis definition of *Physical Entity* with more than one *fuzzy risk model*

Finally, a business process model for describing the concept of management or development a *fuzzy risk model* is discussed in section 4.3.1. Fuzzy modelling is further discussed in section 6.4.

4.2 High-Level Architecture of the Fuzzy Knowledge-Based Approach for Risk Analysis

A high-level architecture is depicted in Figure 4-3. This architecture is decomposed into three (3) main abstract components, which are described in the subsequent paragraphs:

- Fuzzy Risk Analysis (Fuzzy RA)

This component enables the creation and maintenance of *fuzzy risk models* (fuzzy models) with purpose the risk analysis of various *physical entities*. It is possible more than one model to exist. The final risk analysis result might concern more than one *fuzzy risk models* and according to the risk analysis definition for the particular *Physical Entity*. This component also executes the various fuzzy inference systems according to the risk analysis definition for the particular *Physical Entity*.

- Physical Entity Manager/Editor

The various *fuzzy risk models* have fuzzy variables as input and output. Those fuzzy variables are based on some attributes of the *Physical Entity* is analysed (4.1.1). Ontology is used to define concepts and attributes of each *Physical Entity* according to the approach described in Chapter 5. The purpose is that this component should enable the management of the ontologies of *physical entities* by allowing the addition of new *Physical Entity* or the maintenance of the existing *physical entities*.

- Assistance/Optimisation

This component is defined to assist in the construction or optimisation of fuzzy inference systems. In this work, ANFIS technique could be considered for that purpose. As shown in section 3.7, ANFIS technique can assist in the construction of adaptive fuzzy inference systems from a given data set. ANFIS has the ability to be trained and to learn from the data. ANFIS generates Sugeno-type FIS and hence it is defined with fuzzy variables. Therefore, it could be considered that ANFIS can construct some fuzzy inference systems complementary to those representing human knowledge. Nevertheless, the application of this technique should be done following analysis. This includes selection of the appropriate parameters and consideration of issues such as the *curse of dimensionality* (Wei et al. 2007).

Finally, Figure 4-3 depicts high-level communication links between the User and the various components of fuzzy knowledge-based approach for risk analysis and detection. A number is assigned to each communication link. This number is used in the subsequent paragraphs to explain the purpose of each communication link in the context of each Use Case described is section 4.3.

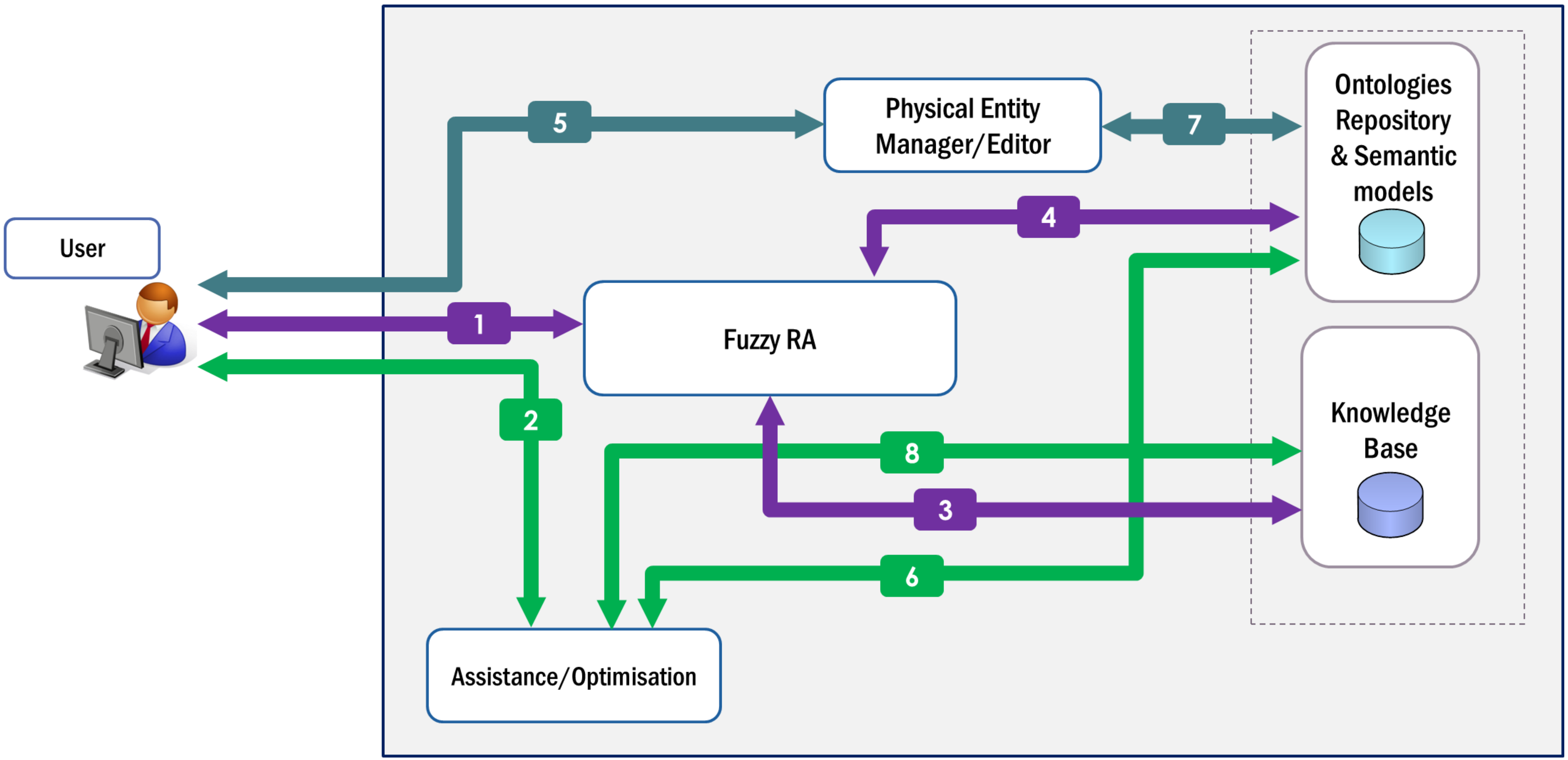

Figure 4-3: High-level (abstract) Architecture of fuzzy knowledge-based approach for risk analysis

4.3 Abstract Use Cases, Processes and Interactions

Main abstract use cases and business process diagrams follow with purpose to elaborate the conceptual model presented in Figure 4-3. In particular, a use case diagram is used to depict in high-level the interactions of various types of users with the components of fuzzy knowledge-based approach for risk analysis. The use case diagram is presented with Unified Modelling Language (UML) (OMG 2011b). UML is the de-facto language for modelling and specification of software systems. In addition, UML can also be used for business modelling and non-software systems. For instance, specific UML Profiles exist tailoring the language to specific areas (e.g. business modelling) (OMG 2013). However, BPMN (OMG 2011a) diagrams (or business process diagrams) are used in this thesis to visualise processes and particular activities in the context of a use case. BPMN notation discussed in section 3.6.2 where the standard fuzzy inference process is described. Finally, the ADONIS Community Edition (ADONIS 2014) has been used for developing the BPMN diagrams and the use case diagrams.

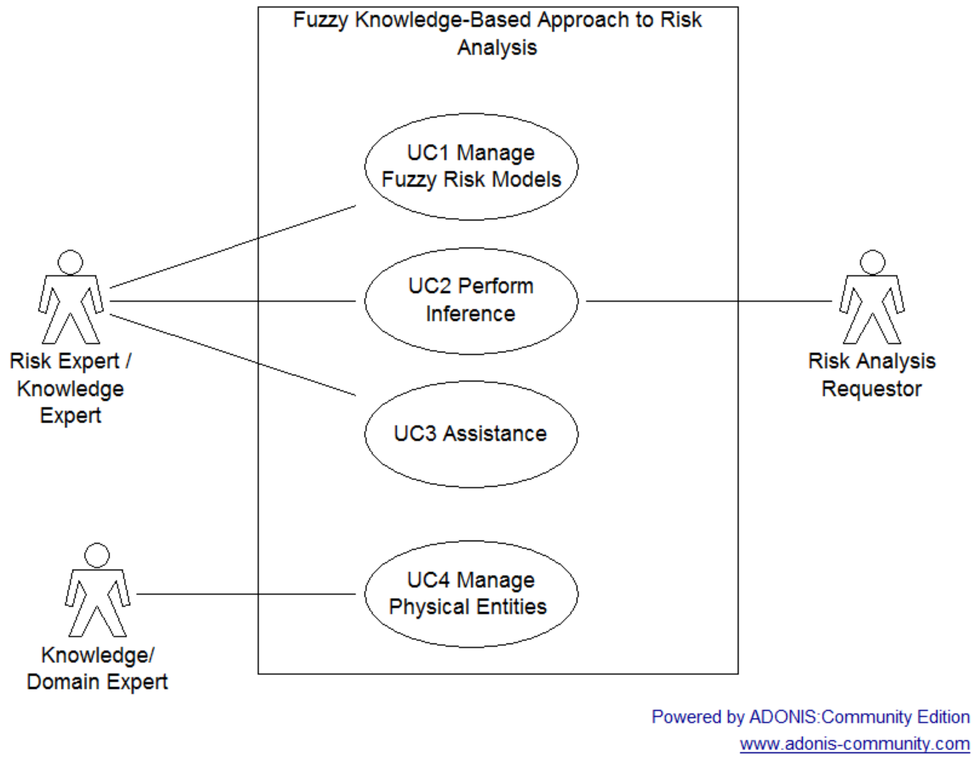

Figure 4-4: Use Case Diagram for interactions of users in fuzzy knowledge-based approach to risk analysis

4.3.1 UC1 Manage Fuzzy Risk Models

Typically, a Use Case defines functional needs and presents the interaction of a user with the system. This Use Case defines that functionally wise the *Risk Expert/Knowledge Expert* (user) shall be able to manage the *fuzzy risk models*. The **Fuzzy Risk Analysis (Fuzzy RA)** is the main component of fuzzy knowledge-based approach for risk analysis, which is involved for the realisation of this Use Case. The interaction of User with this component (activities) is described in the following paragraphs with the aid of Figure 4-5, which illustrates a non-executable private process for the management of *fuzzy risk models*. This designates that the process has been modelled to visualise abstract activities for the management of *fuzzy risk models* at modeller-defined level of details.

The management of *fuzzy risk models* process is realised by the *Fuzzy RA* component. The process starts for either adding new model or modifying/deleting exsting *fuzzy risk models* (e.g. change fuzzy parameters, fuzzy rules). This is reflected with start events "Create new Fuzzy Risk Model" and "Re-assess/Update Fuzzy Risk Model" respectively. In the first case ("Create new Fuzzy Risk Model"), the start event is defined as conditional. In BPMN (OMG 2011a), conditional events denote that the event is triggered as soon as the guard condition of the event is true. In this specific case, the condition is that the related "Physical Entity Ontology" should exist. This also modelled in the process as note associated with the conditional start event. For the second case ("Re-assess/Update Fuzzy Risk Model"), the start event is also defined as conditional requiring at least one model to exist before the update of *fuzzy risk model*.

For the case of new model, it is considered that, the "Fuzzy Risk Model Ontology" should be loaded at the beginning from the *Ontologies Repository and Semantic Models*. The latter holds the definition of the ontology and the *fuzzy risk models* instances (Knowledge Base). The "Load Fuzzy Risk Model Ontology" activity is modelled using a BPMN task (atomic activity) (OMG 2011a). This task is an automated task and hence it is modelled as Service Task. As mentioned in section 3.6.2, the "Service Task" is a special task in BPMN (OMG 2011a) and it is shown with a "Gear" marker in the upper left corner of the shape. In addition, the *Ontologies Repository and Semantic Models* repository is represented as Data Store with BPMN (OMG 2011a) notation. The retrieval of *fuzzy risk models* instances from the *Ontologies Repository and Semantic Models* repository is shown with a data association between the "Load Fuzzy Risk Model Ontology" Service Task and the Data Store representing the *Ontologies Repository and Semantic Models* repository. Having loaded the "Fuzzy Risk Model Ontology", then User creates a *fuzzy risk model* instance. This is shown in the process with the "Create Fuzzy Risk Model instance" User Task. The "User Task" is a special task in BPMN (OMG 2011a) and it is shown with a "Human" marker in the upper left corner of the shape. The User Task denotes a task of a process, which is performed by a User with the assistance of a system (OMG 2011a). Hence, the User Tasks indicate the interaction of users with a process. The created *fuzzy risk model* instance is stored in the *Ontologies Repository and Semantic Models* repository. This is modelled in the process of Figure 4-5 with data association between the mentioned User Task and the *Ontologies Repository and Semantic Models* repository (Data Store).

In case the process starts for the update or re-assessment of existing model, it is presented in the process (Figure 4-5) that the User shall select a *fuzzy risk model* ("Select Fuzzy Risk Model" User Task). For that task, retrieval of *fuzzy risk models* is required from the *Ontologies Repository and Semantic Models* repository (Data Store). This is modelled with data association between the task and the *Ontologies Repository and Semantic Models* repository (Data Store). Following the selection of particular *fuzzy risk model*, the process continues with the retrieval of details of selected *fuzzy risk model* ("Retrieve fuzzy risk model instance details").

In either case (new model or update of existing one), the process continues by updating or adding the *fuzzy risk model* details (parameters and rules) and/or the *fuzzy risk model* attributes (mapping with Physical Entity, main FIS attributes). This is illustrated in the process of Figure 4-5 with an Inclusive Gateway (decision point). The Inclusive

Gateway is used in BPMN as an OR Gateway and has the marker "O" inside the Gateway (diamond shape). This means that all outgoing flows might be executed based on the conditions assigned per path. If management of parameters and rules is selected, then the management of fuzzy variables is deemed as the first step. This is considered as a sub-process involving addition, modification or deletion of fuzzy variables instances ("Manage Fuzzy Variables instances" sub-process) and their mapping with the *fuzzy risk model*. This sub-process also includes addition or update of fuzzy variables' attributes including mapping with *physical entity* attribute. This sub-process needs to communicate with *Ontologies Repository and Semantic Models* repository (Data Store) in order first for retrieving any existing fuzzy variable and second for creating, updating, or deleting fuzzy variables depending on the actions selected. This is shown in the process as data associations between the sub-process and the mentioned Data Stores.

Subsequently, the process continues with the management of membership functions, which is also considered as a sub-process ("Manage Membership Functions instances"). This sub-process needs to communicate with *Ontologies Repository and Semantic Models* repository (Data Store) in order first for retrieving any existing membership functions per fuzzy variable and second for creating, updating, or deleting membership functions of a fuzzy variable depending on the actions selected. Such interactions of *Fuzzy RA* component with *Ontologies Repository and Semantic Models* repository (Data Store) are modelled as data associations between the sub-process and the pertinent Data Store. Finally, the management of parameters and rules is completed with the management of Fuzzy Rules using the Fuzzy Variables and Membership Functions previously defined in the process. The management of Fuzzy Rules is also considered as a sub-process ("Manage Fuzzy Rules"). This sub-process also interacts with *Ontologies Repository and Semantic Models* repository (Data Store) for retrieving any existing Fuzzy Rules for the model and for creating, updating, or deleting Fuzzy Rules depending on the actions selected. Those interactions are modelled as data association between the sub-process and the *Ontologies Repository and Semantic Models* repository (Data Store).

For the management of *fuzzy risk model* attributes, the “Manage Fuzzy Risk Model attributes” sub-process is triggered. This sub-process involves User interaction and also interacts with the *Ontologies Repository and Semantic Models* repository (Data Store) for retrieving existing *fuzzy risk model* attributes and for creating or updating existing *fuzzy risk model* attributes. This is modelled with Data Associations between the sub-process and the Data Store.

As a final point, the process uses again an Inclusive Gateway for merging and synchronising the flow before the End Event, which is “Model creation/update completed”. It is worth noting that the process described above shall be considered as an iterative process.

Finally, the purpose of communication links 1 and 4 in Figure 4-3 has been described in the context of UC1 Manage Fuzzy Risk Models. In particular, the communication link 1 in Figure 4-3 is depicted with the User Tasks analysed above for the management of *fuzzy risk models* process (Figure 4-5). The communication link 4 in Figure 4-3 enables the interaction of Service Tasks of the process with the *Ontologies Repository and Semantic Models* repository (Data Store).

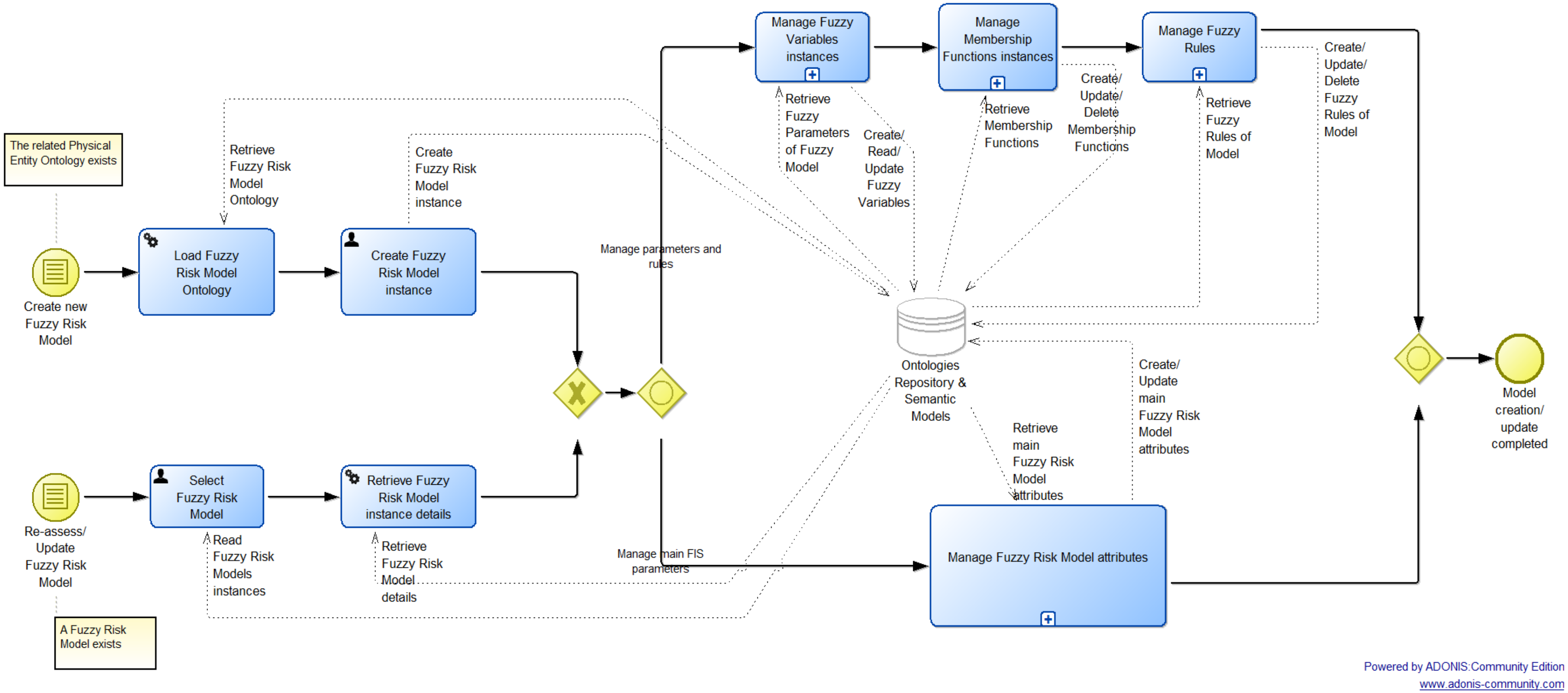

Figure 4-5: Manage Fuzzy Risk Models process

4.3.2 UC2 Perform Inference

This Use Case describes the functional need for performing risk analysis and fuzzy inference by executing the applicable fuzzy models. In the example of an import consignment, this Use Case is executed for analysing the risk. The Actor interacting with that Use Case might be the *Risk Expert / Knowledge Expert* or any other Actor (*Risk Analysis Requestor*) requiring to analyse for risk a *physical entity* (e.g. import consignment). This is illustrated in Figure 4-4.

The main component, which is involved for the realisation of this Use Case, is the **Fuzzy Risk Analysis (Fuzzy RA)**. Figure 4-6 illustrates the interactions of Actors with this component as well as the various activities for the realisation of this Use Case. The User Tasks of the process indicates interactions of the Actor while Service Tasks denotes automated actions by the **Fuzzy Risk Analysis (Fuzzy RA)** component. Finally, the process of Figure 4-6 is a non-executable private process. This designates that the process has been modelled to visualise abstract activities for performing Inference or risk analysis at modeller-defined level of details.

The process starts when there is the need to perform risk analysis. In BPMN (OMG 2011a), this is denoted with the start event "Risk Analysis for a specific Physical Entity has been selected". As a first step in this process, it is considered that the Actor has to select the *physical entity* for inference (risk analysis). This interaction is modelled with a User Task named "Select Physical Entity for Inference". Following that, the definition of risk analysis for the particular *physical entity* is retrieved from *Ontologies Repository and Semantic Models* repository (Data Store). As discussed in section 4.1.2, this might concern the execution of more than one *fuzzy risk model*. This activity is modelled with the Service Task "Retrieve the analysis definition" and the pull of information with a data association having direction from the *Ontologies Repository and Semantic Models* repository (Data Store) to the Service Task.

The process continues with two parallel activities (“Provide required inputs values” and “Transform Fuzzy Risk Model to FIS”). This is illustrated in the process of Figure 4-6 with a Parallel Gateway. The Parallel Gateway is used in BPMN as an AND Gateway and has the marker “+” inside the Gateway (diamond shape). This means that all outgoing flows are executed in parallel. The Parallel Gateway can also be used for synchronising or joining flows. The outgoing flow from this Gateway is executed only if all incoming flows have been completed. The User Task “Provide required inputs values” illustrates that the User provides the required input values for the execution of *fuzzy risk model(s)*. A collection of input values is an output of this task and input for the “Fuzzy Inference Process” sub-process. This sub-process follows for each *fuzzy risk model* that shall be executed according to the risk analysis definition. In the example of Figure 4-2, the $\text{Input}_{f\text{-}RM1}1$, $\text{Input}_{f\text{-}RM1}2$, $\text{Input}_{f\text{-}RM2}1$, $\text{Input}_{f\text{-}RM2}2$, and $\text{Input}_{f\text{-}RM2}3$ are the input values that are provided by the User. For the same example, the *fuzzy risk model* (f-RM1) and the *fuzzy risk model* (f-RM2) should be executed in parallel and the output of each *fuzzy risk model* should be input to *fuzzy risk model* (f-RM3), which is executed at the end. The “Fuzzy Inference Process” sub-process is used for the execution of each *fuzzy risk model*. This process discussed in section 3.6.2 and it is depicted in Figure 3-1.

The second parallel activity is the Service Task “Transform Fuzzy Risk Model to FIS”. This task transforms the *fuzzy risk model* into the FIS format that will be executed. The “Transformed Fuzzy Risk Model to FIS” is an output of this task and it is stored in the *Knowledge Base* Data Store. This is modelled with the relevant data association.

Both parallel activities are joined with a Parallel Gateway as shown in the process of Figure 4-6. After the execution of the *fuzzy risk model(s)*, an output value is produced as the outcome of risk analysis (inference process). Then, this can be stored to *Knowledge Base* persistency. This is modelled with a data association having direction from “Fuzzy Inference Process” sub-process to *Knowledge Base* Data Store. The output of fuzzy inference is presented to the User. This is depicted in the process with User Task “Present results”. Possibly, the output of this process might lead to fine tune the *fuzzy risk model* if this analysis is performed in the context of evaluation. If there is a need to update the *fuzzy risk model* parameters, then the User can do it as described in section 4.3.1.

Finally, the need of communication link 1 in Figure 4-3 is depicted with the User Tasks of the process for performing inference (Figure 4-6). The communication link 3 in Figure 4-3 enables the interaction of Service Tasks of the process with *Knowledge Base* Data Store for accomplishing the inference process and the risk analysis. The communication link 4 in Figure 4-3 facilitates the interaction of Service Tasks of the process with the *Ontologies Repository and Semantic Models* repository (Data Store).

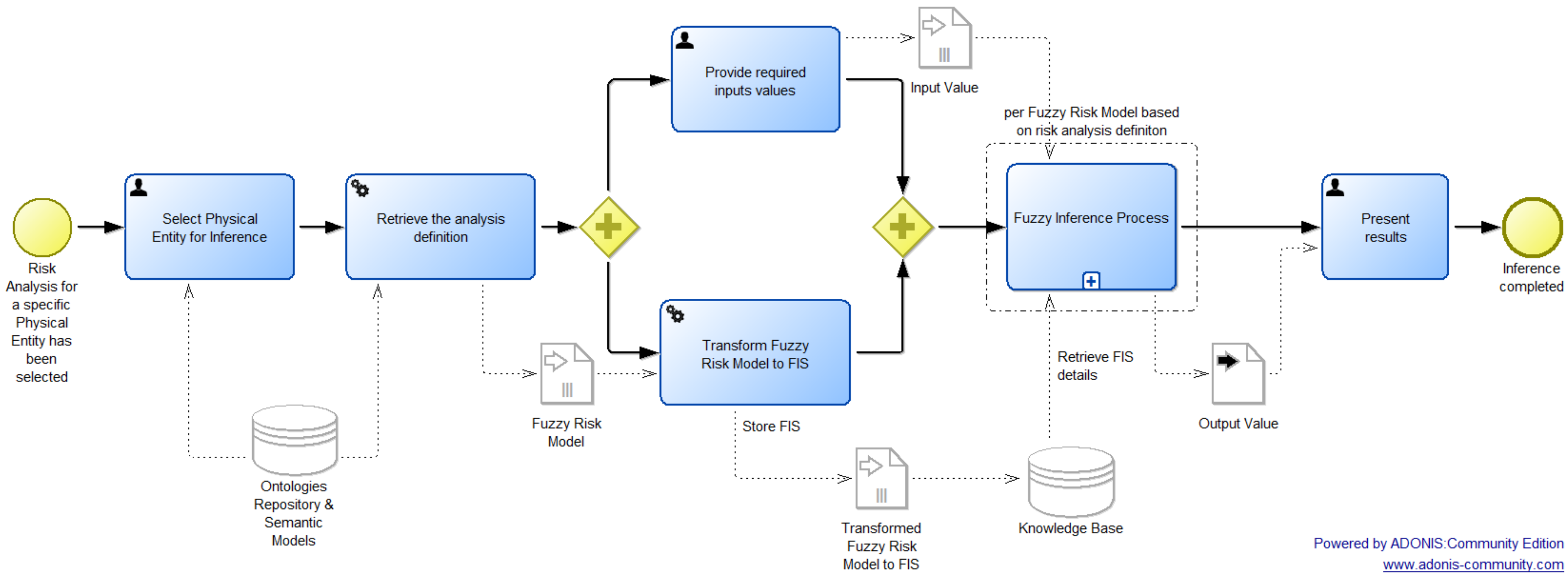

Figure 4-6: Perform Inference Process

4.3.3 UC3 Assistance

This Use Case describes in abstract form, the interaction when the User wants to construct or optimise a model using adaptive techniques. This process is very complex. Depending on the selected technique, this process should be adapted accordingly. Attributes to be used as input for that process are already configured. Training, test, and checking data set should also be defined. In the context of fuzzy knowledge-based approach for risk analysis, the ANFIS technique could be considered for that purpose. The Actor (User) interacting with that Use Case is the *Risk Expert / Knowledge Expert* requiring to construct new fuzzy models or optimising fuzzy models using ANFIS technique. This is illustrated in Figure 4-4. However, it is considered that the *Risk Expert / Knowledge Expert* have the knowledge to apply the ANFIS technique.

The main component of fuzzy knowledge-based approach for risk analysis, which is involved for the realisation of this Use Case, is the **Assistance/Optimisation**.

The User interacts with the **Assistance/Optimisation** component for constructing new fuzzy models or optimising fuzzy models using advanced techniques (e.g. ANFIS). The process is complex and this is the reason for not providing specific process with interactions as happens for other Use Cases. The **Assistance/Optimisation** component will enable the User to develop new *fuzzy risk models* based on pre-selected advanced techniques (e.g. ANFIS). Input attributes and output are configured along with the training dataset. Therefore, it is considered that the **Assistance/Optimisation** component retrieves the definition of the concerned *physical entity* from *Ontologies Repository and Semantic models* and it could possibly retrieve information from *Knowledge Base* persistency. Following that, **Assistance/Optimisation** component constructs or optimises a fuzzy inference system based on the provided data sets and displays the output to the User. The User might select to evaluate the generated fuzzy model. This may happen via the UC2 Perform Inference. Finally, the need of communication link 2 in Figure 4-3 is described above with the User interactions for the construction or optimisation of a model using adaptive techniques. The communication link 8 in Figure 4-3 enables the interaction of **Assistance/Optimisation** component with *Knowledge Base* Data Store as described before in the Use Case description. The communication link 6 in Figure 4-3 facilitates the interaction of **Assistance/Optimisation** component with the *Ontologies Repository and Semantic Models* repository (Data Store) for the tasks described previously in the Use Case description.

4.3.4 UC4 Manage Physical Entities

The Use Case "Manage Physical Entities" specifies the need for managing the ontologies and semantic models of *physical entities*. The term *physical entity* is described in section 4.1.1. The example of import consignment mentioned. The primary Actor of this Use Case is the Knowledge Expert. It is assumed that the Knowledge Expert has the knowledge to create/maintain ontologies based on the knowledge acquired. This is illustrated in Figure 4-4.

The **Physical Entity Manager/Editor** is the component of fuzzy knowledge-based approach to risk analysis, which is involved for the realisation of this Use Case. The interactions of Actors with this component are depicted in Figure 4-7 using User Tasks of BPMN (OMG 2011a). In addition, BPMN's (OMG 2011a) Service Tasks are used in Figure 4-7 in order to model the automated activities. Finally, the process of Figure 4-7 is a non-executable private process. This designates that the process has been modelled to visualise abstract activities for managing *physical entities* at modeller-defined level of details.

The process starts when there is the need to create a new ontology or update an existing one. Those two events are modelled with BPMN (OMG 2011a) as "Create new Physical Entity ontology" and "Update a Physical Entity ontology" start events. In the case of new ontology, the User has to interact in order to create it. This is modelled with User Task "Create Physical Entity Ontology". The latter shall also store the output from this task to *Ontologies Repository and Semantic Models* Data Store. In Figure 4-7, this is visualised with data association having source the User Task and destination the concerned Data Store. If an update of existing ontology is selected, then the User first has to select the "Physical Entity Ontology" ("Select Physical Entity Ontology"). Then, the details of this "Physical Entity Ontology" are automatically retrieved. This is visualised with a Service Task "Retrieve Physical Entity Ontology details" and its association with *Ontologies Repository and Semantic Models* Data Store.

In both cases (new or existing ontology), the process continues with the addition, modification or deletion of concepts/classes of "Physical Entity Ontology". This is considered a separated Sub-process ("Manage Concepts/Classes of Ontology"). The classes of "Physical Entity Ontology" can model complex elements of represented Physical Entity. The "Manage Concepts/Classes of Ontology" Sub-process stores changes on Concepts/Classes to *Ontologies Repository and Semantic Models* Data Store.

Then the "Physical Entity Ontology" is elaborated by defining Object Properties and/or Data Properties (and Data Types if required). Both activities are represented in process of Figure 4-7 as individual Sub-processes. Outputs from these sub-processes are stored in the *Ontologies Repository and Semantic Models* Data Store (e.g. new Data Properties). Moreover, it is depicted that one of the Sub-processes or both can be performed. This is modelled with an Inclusive Gateway (decision point) in the process of Figure 4-7. The Inclusive Gateway is used in BPMN (OMG 2011a) as an OR Gateway and has the marker "O" inside the Gateway (diamond shape). This means that all outgoing flows might be executed based on the conditions assigned per path.

As a final point, the process uses again an Inclusive Gateway for merging and synchronising the flow. Following that, it modelled that a decision should be taken whether the ontology is considered as complete using Exclusive-OR Gateway. In BPMN (OMG 2011a), the Exclusive-OR Gateway has the marker "X" inside the Gateway (diamond shape). If the answer is "Yes", then the process is completed with BPMN (OMG 2011a) End Event "Ontology defined". In the case that ontology needs more update, the flow returns at the point the process of actual maintenance of "Physical Entity Ontology" starts (before the "Manage Concepts/Classes of Ontology" sub-process). It is worth noting that the process described above shall be considered as an iterative process.

Changes on *physical entities* might have impact on models and fuzzy inference systems since the fuzzy parameters of the latter are based on the *physical entities*. Therefore, the user might need to manage the models as described in UC1 Manage Fuzzy Risk Models. It is also considered that similar process of Figure 4-7 is used to manage the other ontologies.

Finally, the need of communication link 5 in Figure 4-3 is illustrated with the User Tasks of the process of Figure 4-7. The communication link 7 in Figure 4-3 allows the interaction of Service Tasks of the process with *Ontologies Repository and Semantic Models* Data Store for accomplishing the management of *physical entities*.

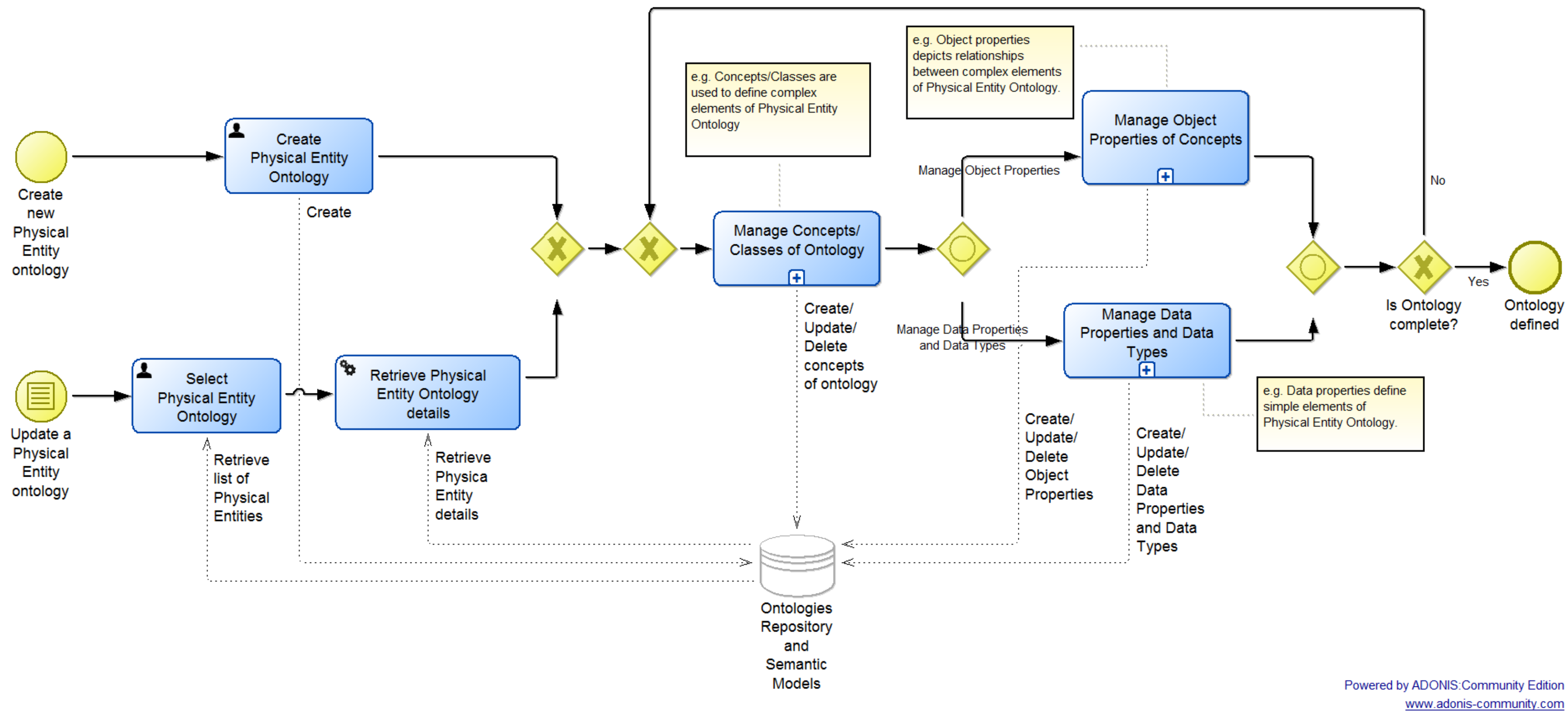

Figure 4-7: Manage Physical Entities process

4.4 Summary

Having considered information presented in previous chapters, this chapter presented the conceptual model for risk analysis with fuzzy logic technique and using semantic modelling. It is considered that addresses the first objective of this research, which is the development of a conceptual model for supporting risk analysis with fuzzy knowledge-based approach. The two main parts for the presented concept is the use of fuzzy logic for expressing *fuzzy risk models* and performing inference tasks for risk analysis purposes and the use of ontologies for semantic modelling of generic and specific concepts related to risk analysis. This concept is elaborated with three components described above.

Chapter 5. Semantic Modelling and Ontologies

This chapter presents the semantic modelling work and ontologies developed under this research in order to represent knowledge and concepts. At the beginning, the role of ontologies is mentioned in relation to the conceptual model presented in Chapter 4. Following that information about the development of ontologies is given. The chapter continues with the presentation of conceptual architecture of ontologies considering the complexity of the domain for knowledge representation and formal representation of concepts. Finally, individual ontologies are described and discussed through examples.

It is worth noting that the scope of this activity is mainly to demonstrate the modelling of concepts with ontologies rather than to develop complete ontologies for Customs, etc. It can only be considered as a research activity to represent some concepts of this domain with formal representation and to model the complex relationships that exist in this domain. This activity also helps to examine the use of ontological engineering as a tool to represent and share various concepts and knowledge in this domain. Ontologies are used for modelling knowledge in the context of fuzzy knowledge-based approach to risk analysis in the Customs domain. Continuous refinement, review, and formal validation are very important for the ontology's effectiveness. Such ontologies require a significant effort in order to fully represent the knowledge of such complex domain and consider the ontology complete. Generally, the Knowledge Acquisition phase is a challenging task for such complex business domains. Finally, formal validation of the ontology and continuous refinement is very important for the ontology's effectiveness.

5.1 Role of Ontologies in this Concept

As discussed in section 3.8, Ontologies can improve the communication and understanding as well as the interoperability. This work explored the domain of Customs, which is a complex domain. As mentioned in section 4.1, the ontologies are used in this work for defining various concepts and for modelling semantics. Therefore, the role of ontologies in this research is for representing semantics and modelling complex relationships. Some examples are provided in the next paragraphs during the discussion of ontologies. Ontologies have different levels of detail with more generic and more specific concepts. Finally, the ontology reasoner is used for consistency checks of the ontologies. In addition, the ontology reasoner can be used to interpret the ontology based on crisp logic. In fuzzy knowledge-based approach to risk analysis, the fuzzy reasoning is investigated to be performed with Mamdani and Sugeno type fuzzy inference systems as it is explained in section 4.1. The rationale of adopting this approach was discussed in section 3.8.

5.2 Ontological Engineering Approach

The approach towards ontological engineering proposed by Uschold and Gruninger (1996) has been used in this work. The activities of ontological engineering approach are briefly described in the next paragraphs as applied in this work:

- Ontology Capture

The concepts of ontology were captured using and assessing sources of information (libraries, online resources, existing knowledge, etc.) such as (DGTAXUD 2004, 2010; EEC 1992, 1993, 2008a, b; EUROSTAT 2005; ISO 2006). In addition, acquired knowledge from literature review presented in Chapter 3 has been used for the development of ontologies. During this step, a number of concepts were defined, a hierarchy of concepts was built, descriptions were added for concepts, and finally the relationships among them were defined. As it is mentioned at the beginning, the ontology focuses more on risk analysis aspects in the Customs domain.

- Ontology Coding

It is very important to choose the appropriate representation language for expressing the ontology of a specific domain. (Shanks et al. 2003; Uschold and Gruninger 1996). The Ontology Web Language (OWL) and particularly the OWL-DL sub-language has been used for coding the ontologies being discussed in this work. OWL facilitates greater machine interpretability of the content and therefore enables the processing of the content from computers as well as the performance of reasoning tasks (Smith et al. 10 February 2004). The complex nature of the domain requires quite expressive language for describing the various concepts and the relationships among them. Therefore, the OWL-Lite could not be used for the ontologies due to its simplicity in terms of expressiveness. On the other hand, the OWL-Full is very expressive sub-language; however, it is undecidable because it does not include restrictions on the use of transitive properties, which are required for decidability (Antoniou and Harmelen 2004; Horrocks and Patel-Schneider 2004). As a conclusion, the OWL-DL has been selected for coding the ontologies. The ontologies are not used in this research for reasoning since the latter is performed with fuzzy rule-based reasoning. The motivation is to represent the knowledge with formal language with extensibility. Therefore, this version of the ontology could be considered as starting point for further work. Nevertheless, the ontology reasoner is mainly used for consistency checks.

The ontology coding was performed using the Protégé v4.3.0 (build 304) (Protégé). The generated OWL code is compliant to OWL 2.0 and has been generated with OWL API (version 3.4.2) provided with Protégé tool. Moreover, the HermiT (v 1.3.8) has been used as ontology reasoner.

The Protégé tool visualises the ontologies with specific forms. The *OntoGraf* plugin (Falconer 2010) of the Protégé tool has been mainly used for visualising the ontology classes and their relationships. In order to visualise also some constraints and data properties of concepts, the OWLGrEd Protégé Plugin have been used, which integrates with OWLGrEd tool (OWLGrEd 2013).

- **Integrating existing ontologies**

Integration of ontologies is discussed in the architecture of ontologies (section 5.3). Nevertheless, the ontologies of this work could be integrated or mapped with more generic ontologies (upper ontologies and/or middle-level ontologies) as discussed in section 3.8 in order to use already defined concepts. This could be considered in a future work.

5.3 Architecture of Ontologies

Considering that different type of concepts (e.g. generic or specific) should be represented with ontologies, an architecture of ontologies is presented in Figure 5-1 based on the Ontological Architecture principles discussed in section 3.8. However, this architecture focuses on the domain ontology layer.

A modular architecture is followed for developing the pertinent ontologies under this project. The architecture is currently decomposed into three main components/ontologies. The first one is a “Generic Customs Ontology” defining the various Customs concepts focusing more to risk management and risk analysis concepts. The second component is the “Ontologies of Physical Entities”. In fact, this contains one or more ontologies each one modelling a *physical entity* (4.1.1). *Physical entities* could be a document or any other entity. An example is mentioned in section 4.1.1 and it is further discussed in section 5.5.2. It is considered that a *physical entity* might re-use concepts of other *physical entity(ies)* enabling re-usability and manageability. Therefore, it might be a dependency between the various ontologies of *physical entities* in terms of concepts and attributes. Concepts of the “Physical Entity Ontology” might refer to or associate with concepts of an ontology modelling another *physical entity*. This can be achieved with one ontology importing another ontology using the “import” statement of OWL. However, the OWL principles for import of ontologies should be considered for that.

The third component is the “Fuzzy Risk Model Ontology” defining the main concepts of *fuzzy risk model* (4.1.2). The architecture of the ontology is illustrated in Figure 5-1.

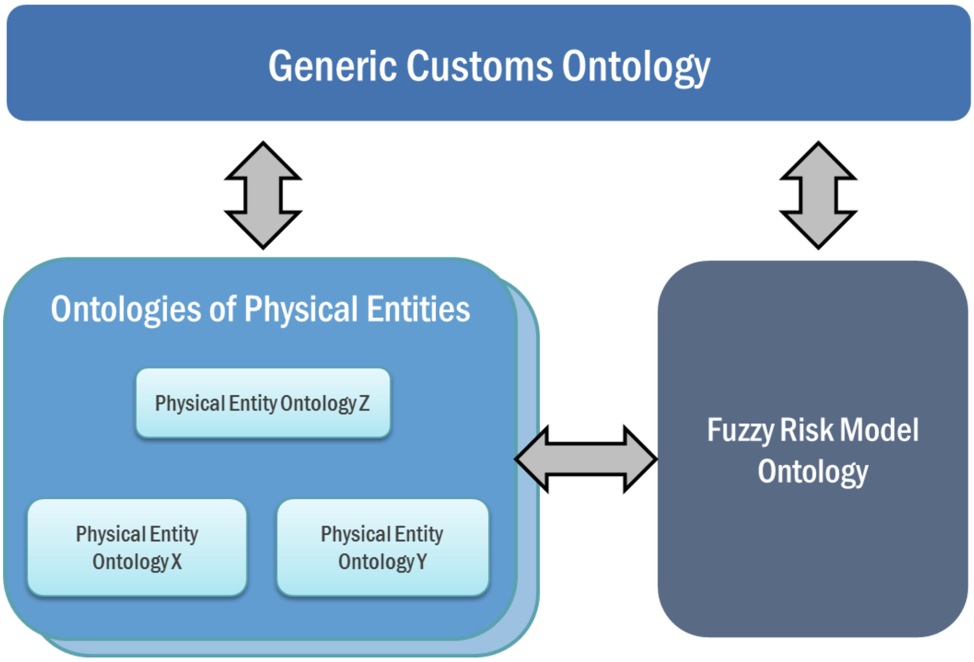

Figure 5-1: Conceptual Architecture of Ontologies

The following benefits are considered having such architecture:

- **Modularity**: Instead of having a single unmanageable Ontology, this is decomposed in other ontologies with concrete purpose and concepts. The decomposition is a matter of decision and organising the ontologies. The current decomposition is described above although some further decomposition could have been performed for “Generic Customs Ontology” or integration with other upper ontologies and/or middle-level ontologies.
- **Maintainability**: Having smaller ontologies it is considered that they can be maintained easier. This facilitates the change management especially if such ontologies are used by systems. Normally, other linked ontologies might be impacted from a change of an Ontology, however, this change is

considered more controllable. Moreover, it is considered that makes easier and consistent the maintenance of ontologies of *physical entities* because changes on the 'imported' ontologies are automatically propagated to the 'importing' ontologies when the latter use concepts that changed in the 'imported' Ontology. Nevertheless, maintenance requires a concrete change management procedure.

- **Re-usability:** Ontologies are re-used in order to construct ontologies of *physical entities* instead of replicating the concepts. Moreover, concepts from other ontologies are used to create relationships and model information. For instance, a concept from "Generic Customs Ontology" might be used from a "Physical Entity Ontology" or concepts from "Physical Entity Ontology" are used to the "Fuzzy Risk Model Ontology". Finally, as an improvement more re-usability could be achieved by integrated those ontologies with other upper ontologies and/or middle-level ontologies.
- **Extensibility:** More ontologies of *physical entities* can be added or existing ontologies can be extended with more concepts. For instance, another ontology of *physical entity* could be added with new concepts without affecting the other relationships.

5.4 Ontology Evaluation

The development of Ontology is an iterative process. Therefore, the Ontology is continuously updated and verified. According to Gomez-Perez (1995), Ontology can be evaluated by the development team, by other development teams, and by end users or experts. Each actor validates it from different perspective. Normally, the development team focus the evaluation on the technical properties of the concepts whereas end users evaluate the actual value and correctness of defined concepts within a given organization or domain. Some technical validation has been performed. The HermiT reasoner has also been used for consistency check of the Ontology. However, it is worth noting that the ontologies have not been validated from any official body or any organisation. They have been developed based on available sources mentioned previously. Hence, this is the frame of reference for the technical evaluation. As mentioned previously, the purpose of the development of ontologies is for research purposes and for exploring their benefits for communication, common understanding, and interoperability in complex domains. The ontologies cannot be considered as complete or validated by end-users. It includes some indicative concepts to indicate the above benefits. Besides, the ontologies must always be enriched which implies that the continuous evaluation and formal validation of the ontologies is required.

5.5 Knowledge Modelling

5.5.1 Generic Customs Ontology

The *Generic Customs Ontology* has taken advantage of all OWL components for representing various concepts of Customs business with special focus on the risk management and risk analysis. The modelled information in this Ontology has been captured from the knowledge sources indicated in section 5.2.

Before discussing the Ontology, it is mentioned that the Protégé tool (Protégé) has been used as the Ontology editor for this *Generic Customs Ontology*. The *OntoGraf* plugin (Falconer 2010) of the Protégé tool has been used for visualising the Ontology and presenting some graphs in the subsequent sections. In an OntoGraf graph, the rectangles with yellow circle in the top left corner represent the classes of the ontology. The rectangles with purple diamond shape in the top left corner represent the individuals of the ontology. The solid purple line between two classes indicate hierarchy relationship (*has subclass*) while the dashed line between two classes indicate relationship due to the existence of object property.

A number of classes have been defined to represent various Customs concepts or entities. Three annotation types have been used as attributes for the definition of classes. These are the *label*, *comment* and *source* attributes. The *comment* annotation has been used to provide a small description about the specific class and hence the user of the Ontology to be able to understand the various business concepts. An example for the "Customs Declaration" class (Figure 5-2) is the *comment* annotation '*means the act whereby a person indicates in the prescribed form and manner a wish to place goods under a given customs procedure, with an indication, where appropriate, of any specific arrangements to be applied'*". The *source* annotation mainly indicates the knowledge source from which this class captured. For the specific example, the '*Article 4(10) of Modernised Customs Code No 450/2008'* was the source for the "Customs Declaration" class. Some illustrative examples from the developed Ontology are presented and discussed in the following paragraphs.

```
<AnnotationAssertion>
   <AnnotationProperty abbreviatedIRI="rdfs:comment"/>
   <IRI>#Customs_declaration</IRI>
   <Literal datatypeIRI="&xsd;string">means the act whereby a person indicates in the prescribed form and manner a wish to place goods under a given customs procedure, with an indication, where appropriate, of any specific arrangements to be applied</Literal>
</AnnotationAssertion>
```

Figure 5-2: Example of OWL syntax for comment annotation on "Customs Declaration" class

The ontology can be used to present the hierarchical structure of various entities. Focusing on the "Risk Management Framework" and specifically on the "EU Risk Management Framework" class, it is shown that the "EU Risk Management Framework" (DGTAXUD 2004) consists of some activities and that manages the Customs Risk by using the OWL component object property. The object properties are used to express the various relationships between the concepts/classes of the Ontology. In this case, the "EU Risk Management Framework" class has the object properties *'consists_of_activities'* and *'is_used_to_manage_risks'*. The first one specifies that the "EU Risk Management

Framework" consist of a number of activities, which are defined by the entity "Risk Management activities". The second object property defines that the "EU Risk Management Framework" is used to manage the "Customs Risk" entity (see Figure 5-3). In addition to the above, the "Risk Management Framework" is a disjoint class with "Risk Management Activities" class. The latter class aims to represent the various activities that shall be performed in the context of a risk management.

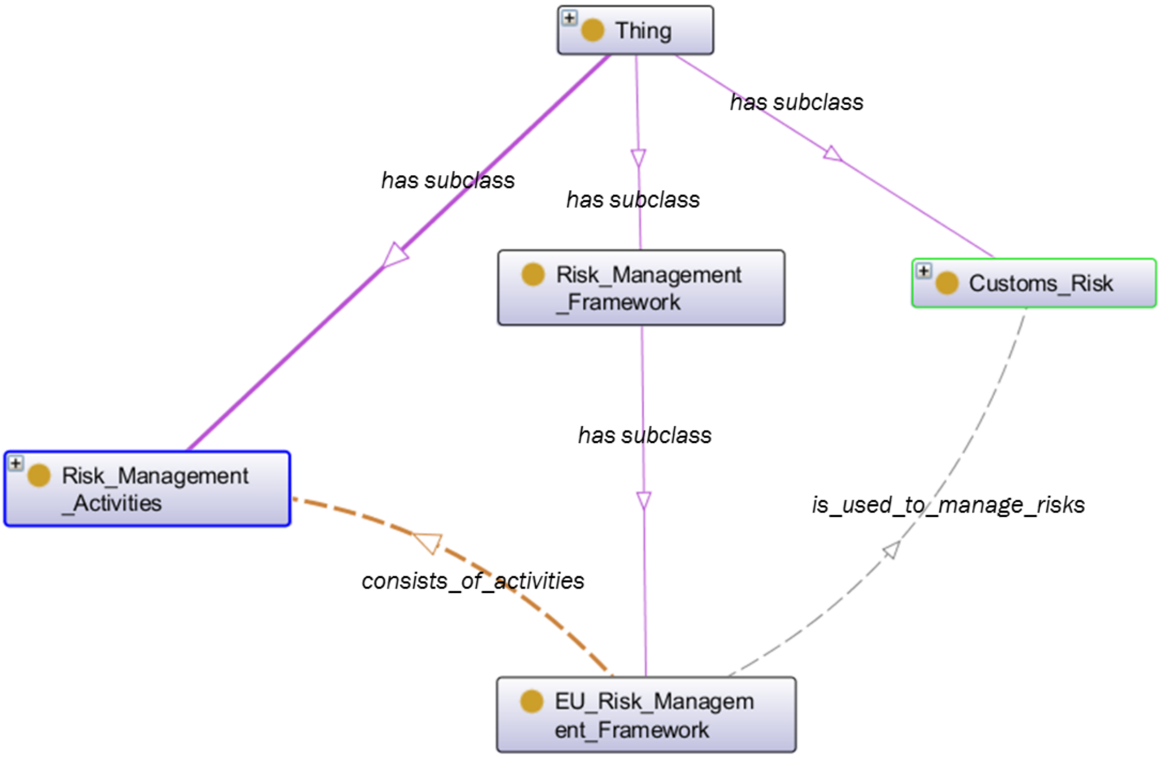

Figure 5-3: Example with Relationships of "EU Risk Management Framework" class with other concepts (OntoGraf Graph)

As it is shown in Figure 5-4, the ontology models the main activities of risk management as defined in (DGTAXUD 2004). The "Risk Analysis" activity consists of the "Identify Risk Data", "Analyse Risks", and "Weigh Risks" sub-activities. As it is stated in (DGTAXUD 2004), ranking of assessment of risk into "High", "Medium" and "Low" is widespread. Therefore, the example of Figure 5-5 depicts that the "Weigh Risks" sub-class has been defined as a value partition of "High Risk", "Medium Risk" and "Low Risk" aiming to model this. The value partition is considered as a design pattern (Jupp et al. 2007). This pattern has been used to restrict the values of "Weigh Risks" and indicate that it has equivalent class the "High Risk or Low Risk or Medium Risk".

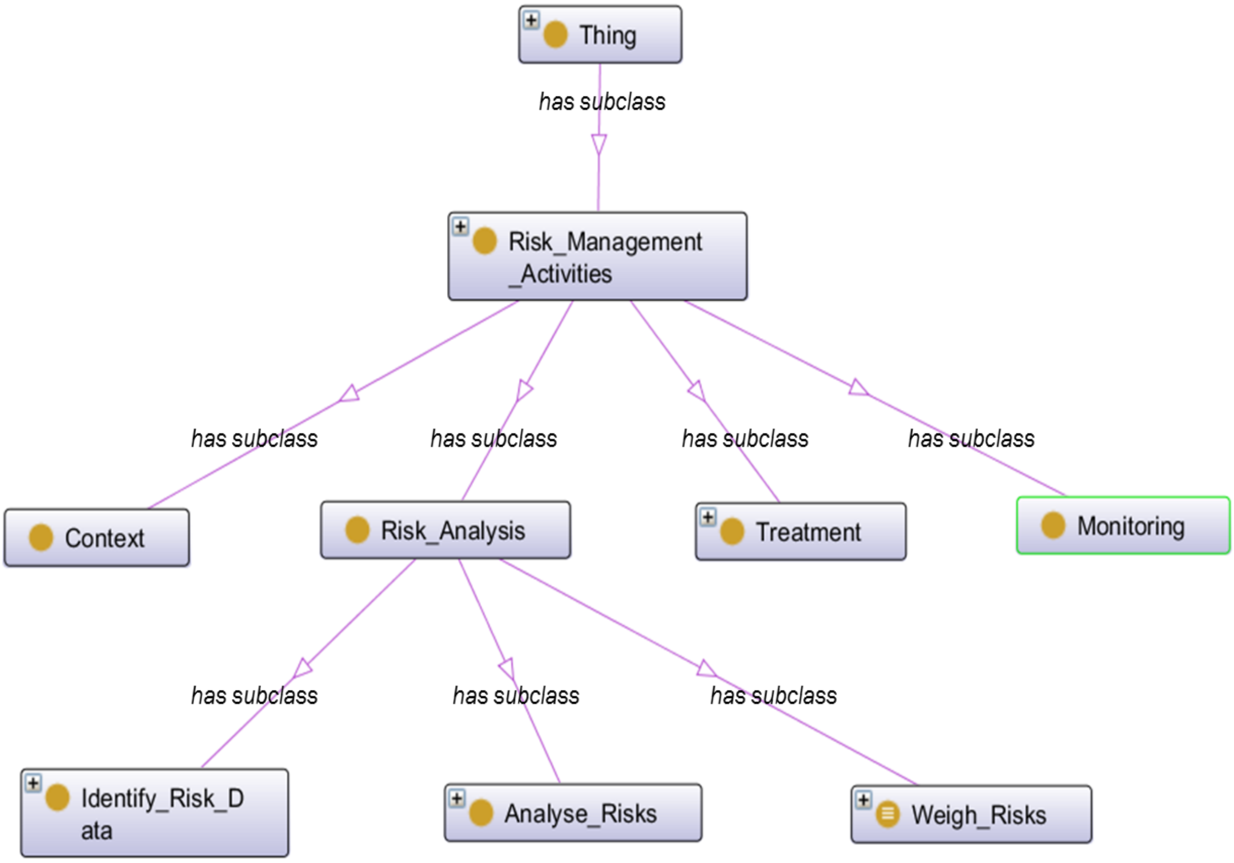

Figure 5-4: “Risk Management Activities” class hierarchy with two levels of children (OntoGraf Graph)

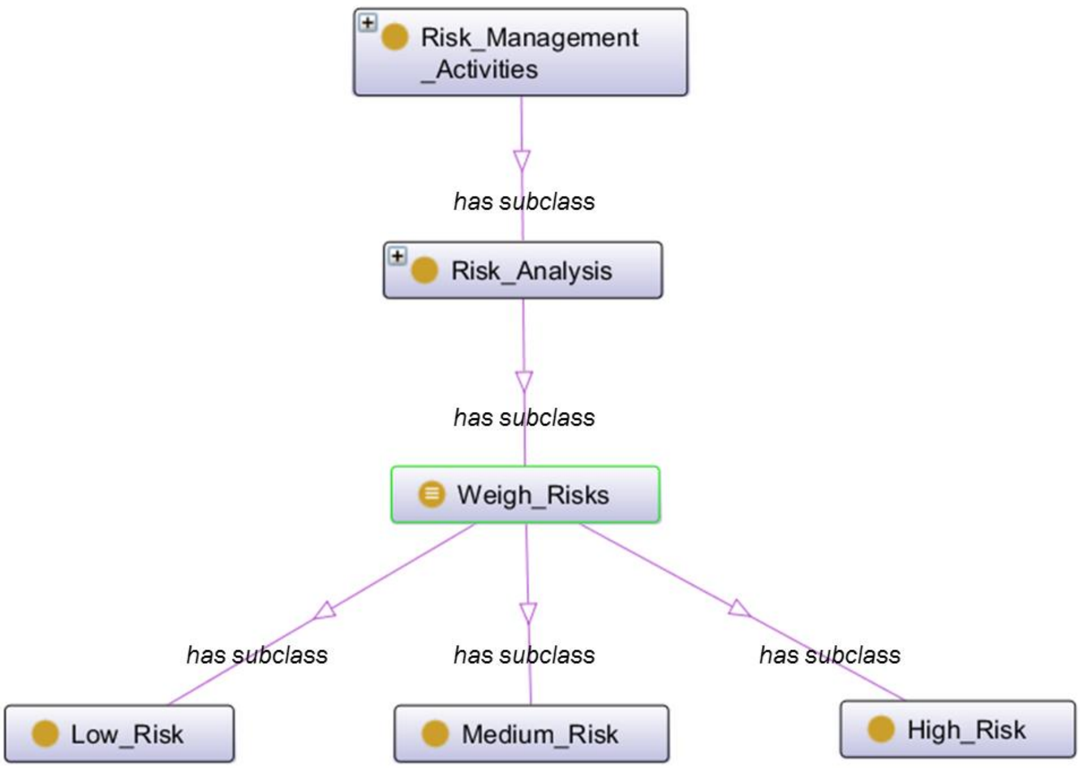

Figure 5-5: "Weigh Risks" class hierarchy (OntoGraf Graph)

Moreover, the hierarchy of "Analyse Risks" class depicts that there are two types of analysis; analysis on proven risks and on potential risks (Figure 5-6). This is also reflected by looking at the "Customs Risk" concept, which classifies the risks into potential and proven risks (DGTAXUD 2004) (Figure 5-6). The Ontology presents the fact that the "Analyse Proven Risks" and "Analyse Potential Risks" activities are used to analyse the "Proven Risks" and "Potential Risks" respectively through OWL object properties. In this particular case, the object property *'is_used_to_analyse_potential_risk'* relates the "Analyse Potential Risks" class with the "Potential Risk" class. The same applies for the object property *'is_used_to_analyse_proven_risks'*, which links the "Analyse Proven Risks" class with the "Proven Risk" class.

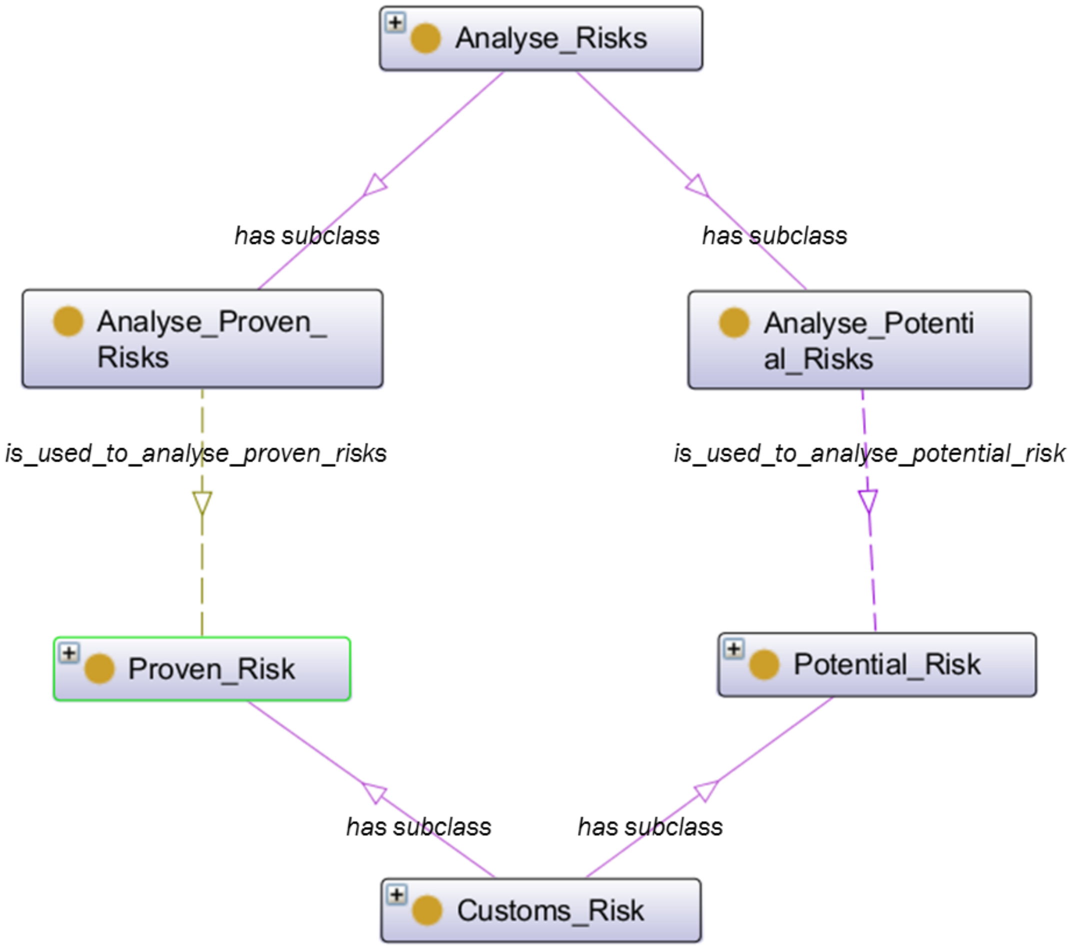

Figure 5-6: Relationships between "Analyse Risks" and "Customs Risk" classes (OntoGraf Graph)

The “Analyse Risk of consignment” class denotes the activity, which is performed for analyzing the risk of a consignment using the declaration data (DGTAXUD 2004; EEC 1993). This is modelled in the Ontology with two object properties. The first one (*‘analyse_risk_of_consignment’*) relates the “Analyse Risk of consignment” class with the “Consignment” class while the second one (*‘is_based_for_analysing_the_risk_of_consignment_on’*) relates the “Analyse Risk of consignment” class to “Customs Declaration” or “Summary Declaration” classes. Finally, the object property *‘is_considered_for_the_control_decision’* between the “Analyse Risk of consignment” class and the “Control consignment” class verifies the fact that outcome of risk analysis of consignment will be one of the criteria for selecting to perform movement inspection (EEC 2008b). The “Control consignment” activity is related to the Treatment activity of risk management (DGTAXUD 2004). This has been modelled in the Ontology with the object property *‘is_part_of_treatment’* between the “Control consignment” class and the “Treatment” class (Figure 5-7).

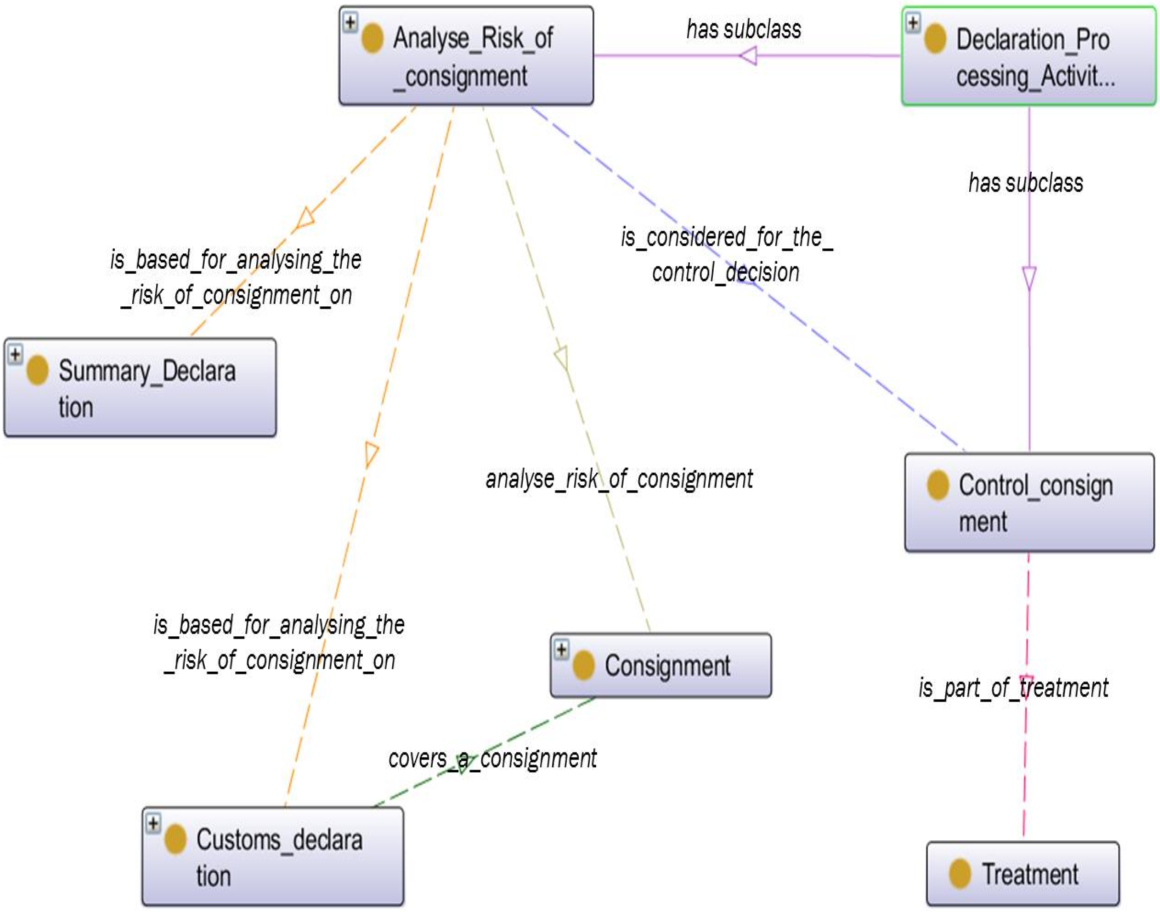

Figure 5-7: Relationships of Analyse Risk of Consignment class with other concepts (OntoGraf Graph)

5.5.2 Physical Entity Ontologies

UC4 Manage Physical Entities describes in high-level the management of the ontologies of *physical entities*. It is considered that a "Physical Entity Ontology" defines the attributes of the *physical entity* using object and data properties. Figure 5-8 depicts an example of such Ontology for the "Physical Entity 1". The "Physical Entity 1" could represent the import consignment considering the example mentioned in section 4.1.1. The "Import Consignment" can be considered a sub-class of "Consignment" class presented in Figure 5-7 (Generic Customs Ontology).

The Protégé tool (Protégé) has been used as the Ontology editor for this "Physical Entity Ontology". The model of Figure 5-8 has been generated by OWLGrEd tool (OWLGrEd 2013) via the OWLGrEd Protégé Plugin. The OWLGrEd tool uses a specific notation for representation of an OWL Ontology. This representation has been used because it visualises various constraints of the Ontology, including data properties and object properties. The rectangles of Figure 5-8 represent the classes of the Ontology. The relationships of sub classes with super class are presented in Figure 5-8 with purple colour ending with an arrow. The object properties along with cardinality restrictions between two classes are shown with red associations in Figure 5-8. Finally, complex or equivalent classes are also shown as rectangles starting with equal symbol.

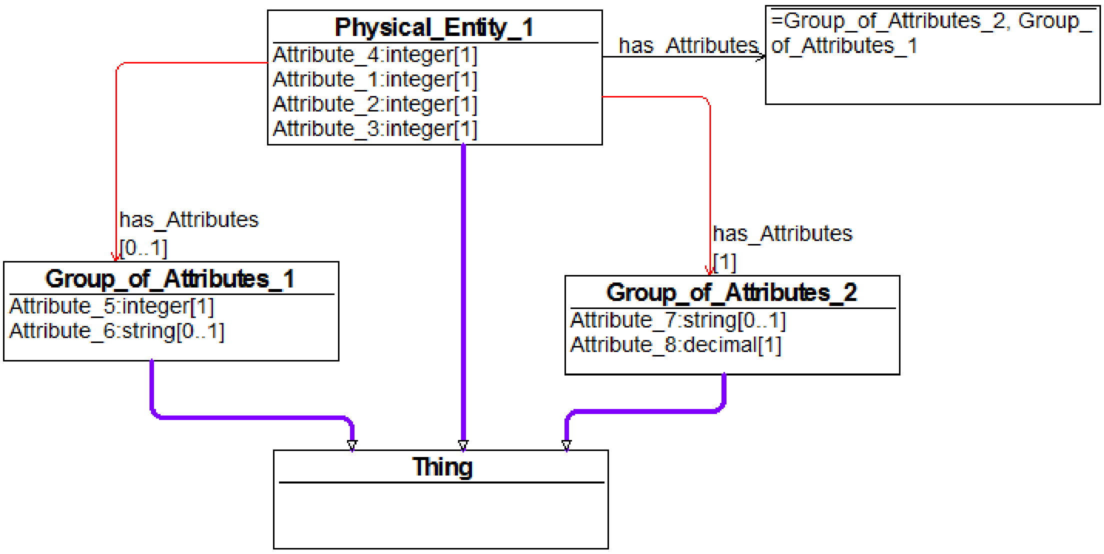

Figure 5-8: Example of "Physical Entity Ontology" (with OWLGrEd notation)

The "Physical Entity 1" class of Figure 5-8 models the root concept of Physical Entity. As it is shown in the example, this Physical Entity has four data properties and two object properties. Classes "Group of Attributes 1" and "Group of Attributes 2" models the grouping of some data properties. The '*has_attributes*' object property has been used to relate the "Physical Entity 1" class with "Group of Attributes 1" and "Group of Attributes 2" classes. Therefore, the object property '*has_attributes*' is used to relate a Physical Entity class with the classes representing a group of attributes. In the "Physical Entity Ontology", the domain of the object property '*has_attributes*' is the "Physical Entity 1" class and the ranges are defined as "Group of Attributes 1" or "Group of Attributes 2". Figure 5-8 presents this as a complex or equivalent class "=Group of Attributes 1 or Group of Attributes 2", which is linked with the "Physical Entity 1" class though the object property '*has_attributes*'. The definition of object property '*has_attributes*' with OWL syntax is shown in Figure 5-9.

```
<ObjectPropertyDomain>
    <ObjectProperty IRI="#has_Attributes"/>
    <Class IRI="#Physical_Entity_1"/>
</ObjectPropertyDomain>
<ObjectPropertyRange>
    <ObjectProperty IRI="#has_Attributes"/>
    <Class IRI="#Group_of_Attributes_1"/>
</ObjectPropertyRange>
<ObjectPropertyRange>
    <ObjectProperty IRI="#has_Attributes"/>
    <Class IRI="#Group_of_Attributes_2"/>
</ObjectPropertyRange>
```

Figure 5-9: Definition of object property 'has_attributes' with OWL syntax

The group of attributes represent a group of information for the Physical Entity. One group of attributes might have more than one instance in a Physical Entity (repetition) or might be optional. Such cases can be modelled with cardinality restrictions of the object property '*has_attributes*'. For instance, the example of Figure 5-8 defines that the "Physical Entity 1" class might have maximum 1 "Group of Attributes 1" and exactly 1 "Group of Attributes 2". The first means that "Group of Attributes 1" might be present in the "Physical Entity 1", while the second denotes that "Physical Entity 1" must have one instance of "Group of Attributes 2". The OWL syntax for such cardinality restrictions of object properties is shown in Figure 5-10.

Attribute 1, Attribute 2, Attribute 3 and Attribute 4 can be considered as the main attributes of "Physical Entity 1" and they are modelled as data properties of the class *('Attribute_1'*, *'Attribute_2'*, *'Attribute_3'* and *'Attribute_4'* respectively). Figure 5-8 shows that all data properties are required. Moreover, the format or data type of each data property is shown in Figure 5-8 after the attribute name followed by the symbol ":". The OWL syntax for cardinality restrictions of data properties of "Physical Entity 1" is shown in Figure 5-10.

```
<SubClassOf>
    <Class IRI="#Physical_Entity_1"/>
    <ObjectExactCardinality cardinality="1">
        <ObjectProperty IRI="#has_Attributes"/>
        <Class IRI="#Group_of_Attributes_2"/>
```

```
        </ObjectExactCardinality>
    </SubClassOf>
    <SubClassOf>
        <Class IRI="#Physical_Entity_1"/>
        <ObjectMaxCardinality cardinality="1">
            <ObjectProperty IRI="#has_Attributes"/>
            <Class IRI="#Group_of_Attributes_1"/>
        </ObjectMaxCardinality>
    </SubClassOf>
    <SubClassOf>
        <Class IRI="#Physical_Entity_1"/>
        <DataExactCardinality cardinality="1">
            <DataProperty IRI="#Attribute_1"/>
            <Datatype abbreviatedIRI="xsd:integer"/>
        </DataExactCardinality>
    </SubClassOf>
    <SubClassOf>
        <Class IRI="#Physical_Entity_1"/>
        <DataExactCardinality cardinality="1">
            <DataProperty IRI="#Attribute_2"/>
            <Datatype abbreviatedIRI="xsd:integer"/>
        </DataExactCardinality>
    </SubClassOf>
    <SubClassOf>
        <Class IRI="#Physical_Entity_1"/>
        <DataExactCardinality cardinality="1">
            <DataProperty IRI="#Attribute_3"/>
            <Datatype abbreviatedIRI="xsd:integer"/>
        </DataExactCardinality>
    </SubClassOf>
    <SubClassOf>
        <Class IRI="#Physical_Entity_1"/>
        <DataExactCardinality cardinality="1">
            <DataProperty IRI="#Attribute_4"/>
            <Datatype abbreviatedIRI="xsd:integer"/>
        </DataExactCardinality>
    </SubClassOf>
```

Figure 5-10: OWL syntax for object properties restrictions on Physical Entity 1 class

An ontology and its elements (e.g. classes, object properties, data properties) are identified with the Internationalized Resource Identifiers (IRIs) (W3C 2012). The IRIs of classes and data properties of “Physical Entity Ontology” are used in “Fuzzy Risk Model Ontology” (section 5.5.2) in order to refer to those classes and data properties depending on the need. The IRIs are used to uniquely identify the referenced resources.

The example of “Physical Entity Ontology” in Figure 5-8 is also used in the “Fuzzy Risk Model Ontology” (section 5.5.3) and in the example presented in Chapter 6.

5.5.3 Fuzzy Risk Model Ontology

The “Fuzzy Risk Model Ontology” defines concepts of *fuzzy risk model*, which is used for risk analysis with fuzzy logic technique (please see section 4.1.2). UC1 Manage Fuzzy Risk Models (4.3.1) described in high-level the management of *fuzzy risk models.* These models are based on the “Fuzzy Risk Model Ontology”. Various OWL components have been used for representing concepts and relationships of “Fuzzy Risk Model Ontology”. The modelling of concepts has been based on the fuzzy logic principles and fuzzy inference systems discussed in section 3.6.

Figure 5-13 illustrates the main class of the ontology, which is the “Fuzzy Risk Model”. As it is shown, the *fuzzy risk model* is related to a number of classes of the ontology, which represent some concepts. The *fuzzy risk model* analyses a specific *physical entity* (section 4.1.1). This logical relationship is modelled with the object property *‘analyse_Physical_Entity’* between the “Fuzzy Risk Model” class and the “Physical Entity” class. The constraint *‘analyse_Physical_Entity exactly 1 Physical Entity’* is added in the class definition in order to restrict and specify that the *fuzzy risk model* concern only one *physical entity*. The OWL syntax of this restriction is shown in Figure 5-11.

In section 5.5.2, it was described that a *Physical Entity* is defined with a “Physical Entity Ontology”. As mentioned before, a *fuzzy risk model* is mapped to a specific *physical entity* class definition (e.g. “Physical Entity 1”) of a particular “Physical Entity Ontology”. In that example, the “Physical Entity 1” class (Figure 5-8) is the root concept of “Physical Entity Ontology”. The data property *‘Analyse_Physical_Enity_IRI’* of “Fuzzy Risk Model” class is used for mapping the *fuzzy risk model* with a specific *physical entity* class definition (e.g. “Physical Entity 1”) of a particular “Physical Entity Ontology”. This data property (*‘Analyse_Physical_Enity_IRI’)* must have 1 *Physical Entity* IRI *(‘Analyse_Physical_Enity_IRI exactly 1 anyURI’*). Therefore, the value of this data property is the corresponding Physical Entity IRI as stated in section 5.5.2. This is also shown in

Figure 5-19. An example is also presented in section 6.4 when fuzzy modelling is further discussed.

```
<SubClassOf>
    <Class IRI="#Fuzzy_Risk_Model"/>
    <ObjectMinCardinality cardinality="2">
        <ObjectProperty IRI="#has_Input_variables"/>
        <Class IRI="#Fuzzy_Input_Variable"/>
    </ObjectMinCardinality>
</SubClassOf>
<SubClassOf>
    <Class IRI="#Fuzzy_Risk_Model"/>
    <ObjectExactCardinality cardinality="1">
        <ObjectProperty IRI="#analyse_Physical_Entity"/>
        <Class IRI="#Physical_Entity"/>
    </ObjectExactCardinality>
</SubClassOf>
<SubClassOf>
    <Class IRI="#Fuzzy_Risk_Model"/>
    <ObjectExactCardinality cardinality="1">
        <ObjectProperty IRI="#has_Output_variables"/>
        <Class IRI="#Fuzzy_Output_Variable"/>
    </ObjectExactCardinality>
</SubClassOf>
```

Figure 5-11: OWL syntax for some restrictions on "Fuzzy Risk Model" class

Figure 5-13 shows that a *fuzzy risk model* has fuzzy rules with the object property *'has_Fuzzy_Rules'* between the class "Fuzzy Risk Model" and "Fuzzy Rule". It is also depicted that the "Fuzzy Risk Model" has "Fuzzy Input Variable" and "Fuzzy Output Variable" and these are modelled with object properties: *'has_Input_variable'* and *'has_Output_variable'* respectively. However, it is considered that the "Fuzzy Risk Model" has minimum 2 "Fuzzy Input Variables" and 1 "Fuzzy Output Variables" (MISO). In the "Fuzzy Risk Model Ontology", this is specified with cardinality restrictions "min" and "exactly" accordingly. This is not visible in Figure 5-13 but it can be seen with OWL syntax in Figure 5-11.

In addition, Figure 5-13 shows that "Fuzzy Input Variable" and "Fuzzy Output Variable" are both "Fuzzy variable" using the *'has subclass'* relationship. Moreover, it is modelled that "Fuzzy variable" has "Membership Function" with the *'has_MF'* object property. In addition, the restriction that "Fuzzy variable" has minimum 1 membership function (*'has_MF min 1 Membership_Function'*) is defined in the "Fuzzy variable" class definition. The cardinality restriction of minimum 1 membership function is inherited by both "Fuzzy Input Variable" and "Fuzzy Output Variable" classes because they are sub-classes of the "Fuzzy variable". This is also visualised in Figure 5-12.

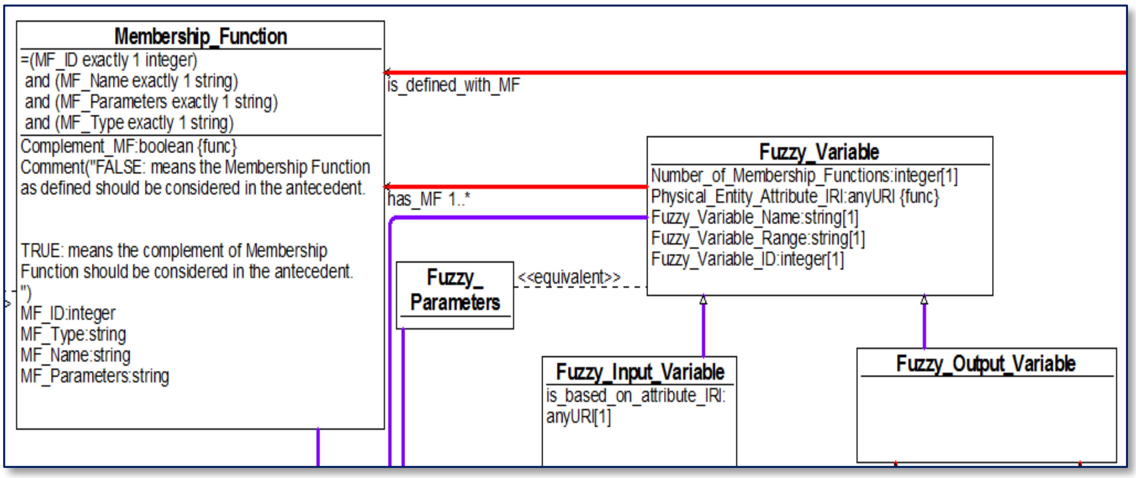

Figure 5-12: Relationships between Fuzzy Variable, Membership Function, Fuzzy Input Variable and Fuzzy Output Variable (with OWLGrEd notation)

Moreover, Figure 5-12 indicates that the "Fuzzy Input Variable" class has a data property *'is_based_on_attribute_IRI'* with cardinality restriction exactly 1. This data property can be used to indicate the mapping of "Fuzzy Input Variable" of the *fuzzy risk model* with a specific attribute of *Physical Entity* (e.g. *'Attribute_1'* of "Physical Entity Ontology" discussed in 5.5.2). A "Fuzzy Input Variable" applies to an attribute of *physical entity* so the mapping is necessary for the risk analysis of an instance of the particular *physical entity.* The value of this data property is the IRI of the relevant *physical entity* data property (attribute). An example is also presented in section 6.4 when fuzzy modelling is further discussed. This is also shown in Figure 5-19.

Finally, the concept "Fuzzy Parameter" it is defined as equivalent to "Fuzzy Variable" in this Ontology. The equivalent class feature has been used to indicate whether one concept is the same with another although they have different names.

Figure 5-13 depicts the relationship (*'has_Fuzzy_Rules'* object property) between the "Fuzzy Risk Model" and the "Fuzzy Rule". In this Ontology, the various parts of a rule antecedent and consequence have been defined as concepts. Figure 5-15 visualises how the Ontology specifies that a “Fuzzy Rule” has antecedent and has consequent. This is done through the corresponding object properties ‘has_antecedent’ and ‘has_consequent’ of “Fuzzy Rule” class with “Fuzzy Rule Antecedent” and “Fuzzy Rule Consequent” classes respectively. However, it is defined that a “Fuzzy Rule” has exactly 1 “Fuzzy Rule Antecedent” and exactly 1 “Fuzzy Rule Consequent”. The antecedent of a rule might have various parts combined with fuzzy operators. For that purpose, the object property *‘antecedent_consists_of’* is defined to indicate that the “Fuzzy Rule Antecedent” consists of “Fuzzy Rule Antecedent Component”. However, the definition is more specific and the class definition contains that the “Fuzzy Rule Antecedent” consists of either exactly 1 “Fuzzy Rule Antecedent Component” or minimum 2 “Fuzzy Rule Antecedent Component” and a fuzzy operator (AND or OR). The expression defined in the ontology is the following: *“((antecedent_consists_of min 2 Fuzzy_Rule_Antecedent_Component) and (Fuzzy_Operator exactly 1 {"AND" , "OR"})) or (antecedent_consists_of exactly 1 Fuzzy_Rule_Antecedent_Component).* Finally, the OWL syntax of the “Fuzzy Rule Antecedent” class definition is shown in Figure 5-14.

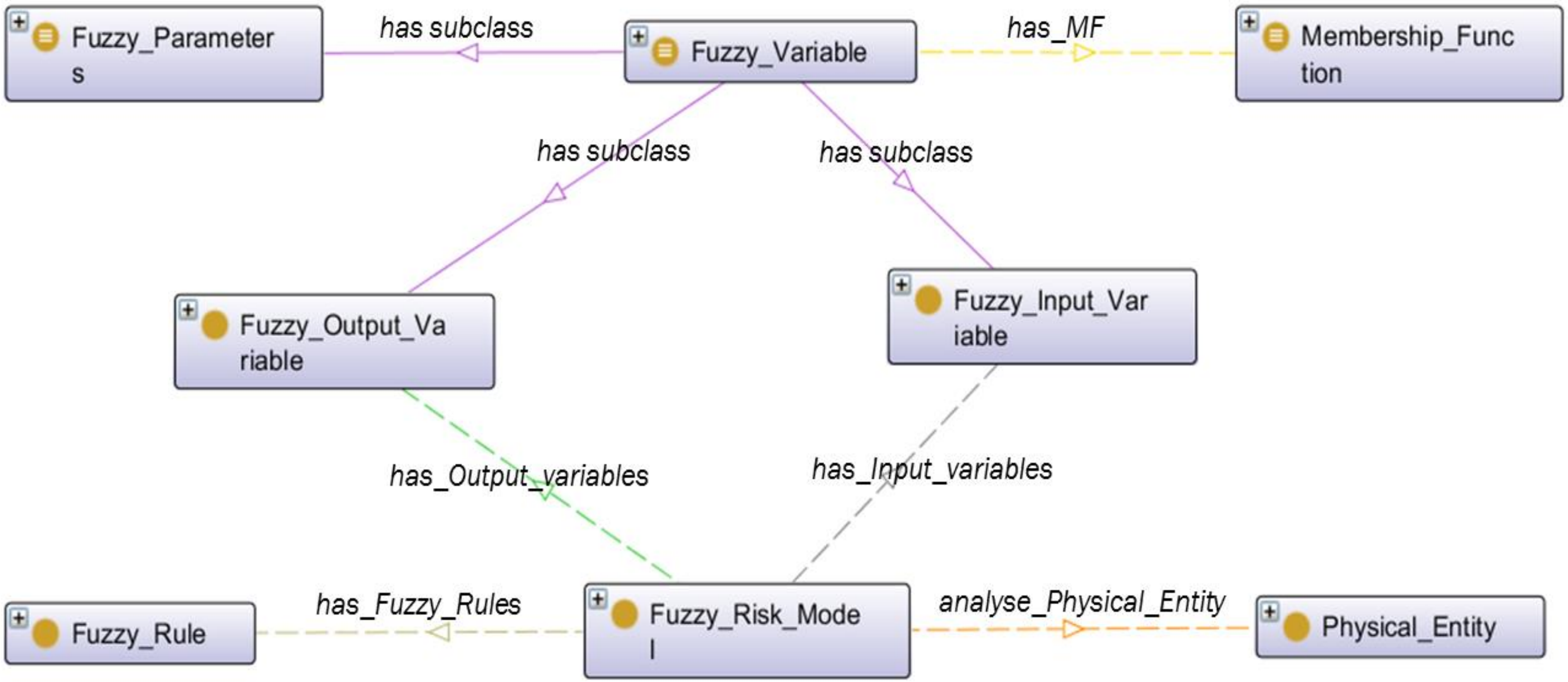

Figure 5-13: Top Level Diagram for "Fuzzy Risk Model Ontology" (OntoGraf Graph)

```
    <SubClassOf>
        <Class IRI="#Fuzzy_Rule_Antecedent"/>
        <ObjectUnionOf>
            <ObjectIntersectionOf>
                <ObjectMinCardinality cardinality="2">
                    <ObjectProperty IRI="#antecedent_consists_of"/>
                    <Class IRI="#Fuzzy_Rule_Antecedent_Component"/>
                </ObjectMinCardinality>
                <DataExactCardinality cardinality="1">
                    <DataProperty IRI="#Fuzzy_Operator"/>
                    <DataOneOf>
                        <Literal
datatypeIRI="&rdf;PlainLiteral">AND</Literal>
                        <Literal
datatypeIRI="&rdf;PlainLiteral">OR</Literal>
                    </DataOneOf>
                </DataExactCardinality>
            </ObjectIntersectionOf>
            <ObjectExactCardinality cardinality="1">
                <ObjectProperty IRI="#antecedent_consists_of"/>
                <Class IRI="#Fuzzy_Rule_Antecedent_Component"/>
            </ObjectExactCardinality>
        </ObjectUnionOf>
    </SubClassOf>
```

Figure 5-14: OWL syntax for restrictions in “Fuzzy Rule Antecedent” class

It is specified that the “Fuzzy Rule Antecedent Component” is based on a “Fuzzy Input Variable” (object property *‘antecedent_is_based*_on’) and is defined with a “Membership Function” (linguistic variable) (object property *‘is_defined_with_MF’)*. However, in order to be more specific on the definition of the “Fuzzy Rule Antecedent Component”, the definition contains that the “Fuzzy Rule Antecedent Component” consist of exactly 1 “Fuzzy Input Variable”, exactly 1 membership function (“Membership Function” class) and whether the complement of membership function (data property of “Fuzzy Rule Antecedent Component”) is applicable. This is illustrated in the example of OWL syntax of the “Fuzzy Rule Antecedent Component” class definition (Figure 5-16).

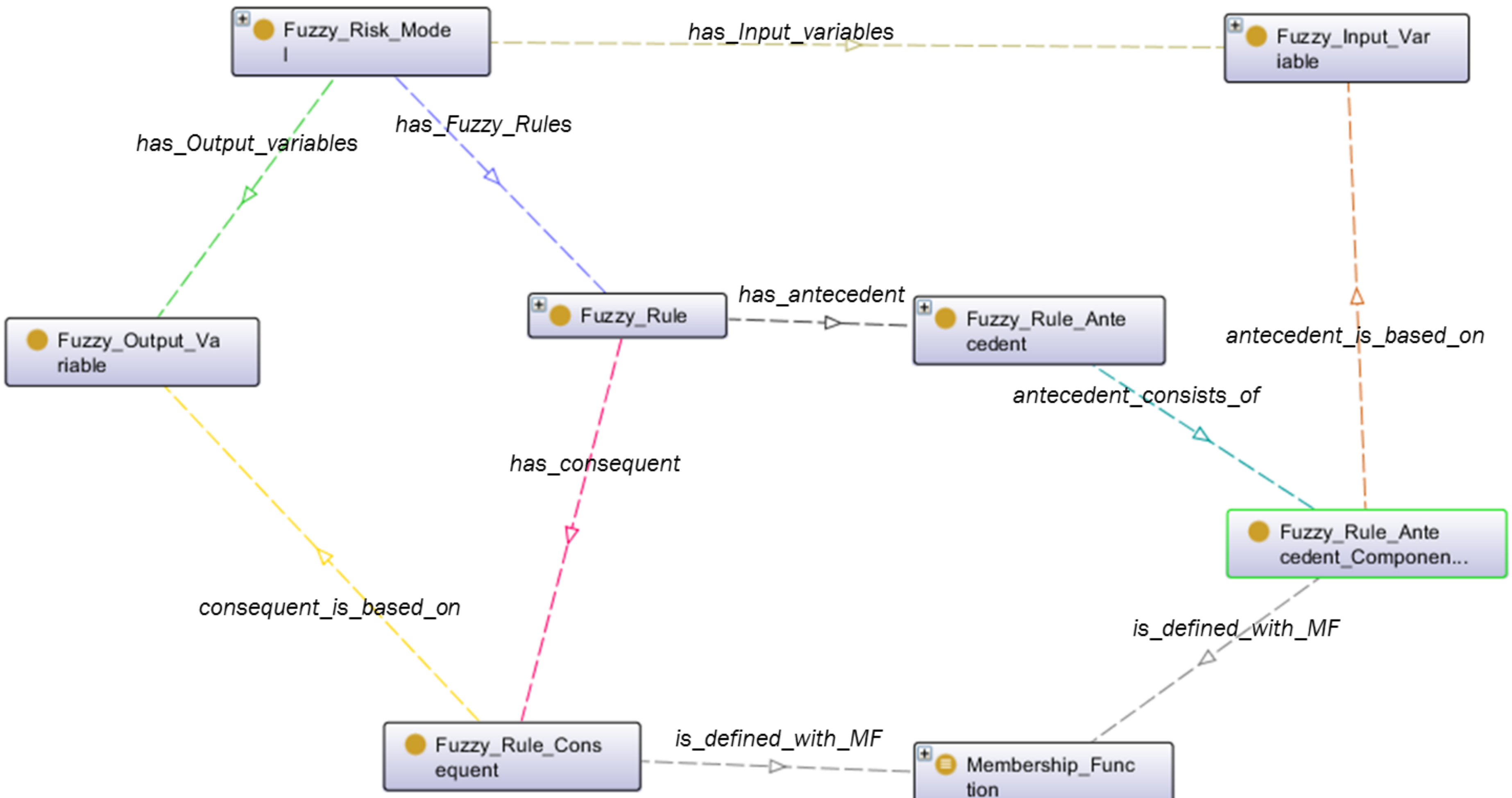

Figure 5-15: Relationships of “Fuzzy Rule” concept with other concepts (OntoGraf Graph)

```
    <SubClassOf>
        <Class IRI="#Fuzzy_Rule_Antecedent_Component"/>
        <ObjectIntersectionOf>
            <ObjectIntersectionOf>
                <ObjectExactCardinality cardinality="1">
                    <ObjectProperty IRI="#is_defined_with_MF"/>
                    <Class IRI="#Membership_Function"/>
                </ObjectExactCardinality>
                <DataExactCardinality cardinality="1">
                    <DataProperty IRI="#Complement_MF"/>
                    <Datatype abbreviatedIRI="xsd:boolean"/>
                </DataExactCardinality>
            </ObjectIntersectionOf>
            <ObjectExactCardinality cardinality="1">
                <ObjectProperty IRI="#antecedent_is_based_on"/>
                <Class IRI="#Fuzzy_Input_Variable"/>
            </ObjectExactCardinality>
        </ObjectIntersectionOf>
    </SubClassOf>
```

Figure 5-16: OWL syntax of definition of “Fuzzy Rule Antecedent Component” class

Similarly, the Ontology expresses that the “Fuzzy Rule Consequent” of a “Fuzzy Rule” is based on a “Fuzzy Output Variable” with the object property *‘consequent_is_based_on’* (Figure 5-17) and that is defined with a membership function (linguistic variable) with the object property *‘is_defined_with_MF’*.

```
<Declaration>
        <ObjectProperty IRI="#consequent_is_based_on"/>
</Declaration>

<ObjectPropertyDomain>
    <ObjectProperty IRI="#consequent_is_based_on"/>
    <Class IRI="#Fuzzy_Rule_Consequent"/>
</ObjectPropertyDomain>

<ObjectPropertyRange>
    <ObjectProperty IRI="#consequent_is_based_on"/>
    <Class IRI="#Fuzzy_Output_Variable"/>
</ObjectPropertyRange>
```

Figure 5-17: OWL code with parts of definition of *‘consequent_is_based_on’* object property

More specifically, it is specified that the “Fuzzy Rule Consequent” consists of exactly 1 “Fuzzy Output Variable”, exactly 1 membership function (“Membership Function” class) and whether the complement of membership function (data property of “Fuzzy Rule

Consequent") is applicable. This is visible in the example of OWL syntax of the "Fuzzy Rule Consequent" class definition shown in Figure 5-18.

```
<SubClassOf>
    <Class IRI="#Fuzzy_Rule_Consequent"/>
    <ObjectIntersectionOf>
        <ObjectExactCardinality cardinality="1">
            <ObjectProperty IRI="#consequent_is_based_on"/>
            <Class IRI="#Fuzzy_Output_Variable"/>
        </ObjectExactCardinality>
        <ObjectExactCardinality cardinality="1">
            <ObjectProperty IRI="#is_defined_with_MF"/>
            <Class IRI="#Membership_Function"/>
        </ObjectExactCardinality>
        <DataExactCardinality cardinality="1">
            <DataProperty IRI="#Complement_MF"/>
            <Datatype abbreviatedIRI="xsd:boolean"/>
        </DataExactCardinality>
    </ObjectIntersectionOf>
</SubClassOf>
```

Figure 5-18: OWL syntax of definition of "Fuzzy Rule Consequent" class

Apart from classes and object properties, which indicate the various concepts of *fuzzy risk model* as well as their complex relationships, various data properties have also been defined in various Ontology classes. These data properties indicate attributes of the classes but also they can be used in various restrictions as described in the examples above. Some other examples of data properties usages follow.

For instance, the "Fuzzy Risk Model" class has data properties such as Fuzzy Model Name (*'Fuzzy_Model_Name'*), FIS type (*'FIS_Type'*), Aggregation Method (*'Aggregation_Method'*), Implication Method (*'Implication_Method'*), method for AND operator (*'And_Method'*), method for OR operator (*'Or_Method'*), etc. Some of them are required for the execution of the model and the fuzzy reasoning according to the fuzzy logic principles. In addition, restrictions to the value of some data properties have been defined. An example is FIS type (*'FIS_Type'*) data property, which can take only the following two values: (Mamdani or Sugeno). The Data Property Range of this data property is defined with exact (*DataOneOf*) literal values *{"Mamdani" , "Sugeno"}*.

Below, an overview of the *Fuzzy Risk Ontology* is provided (Figure 5-19):

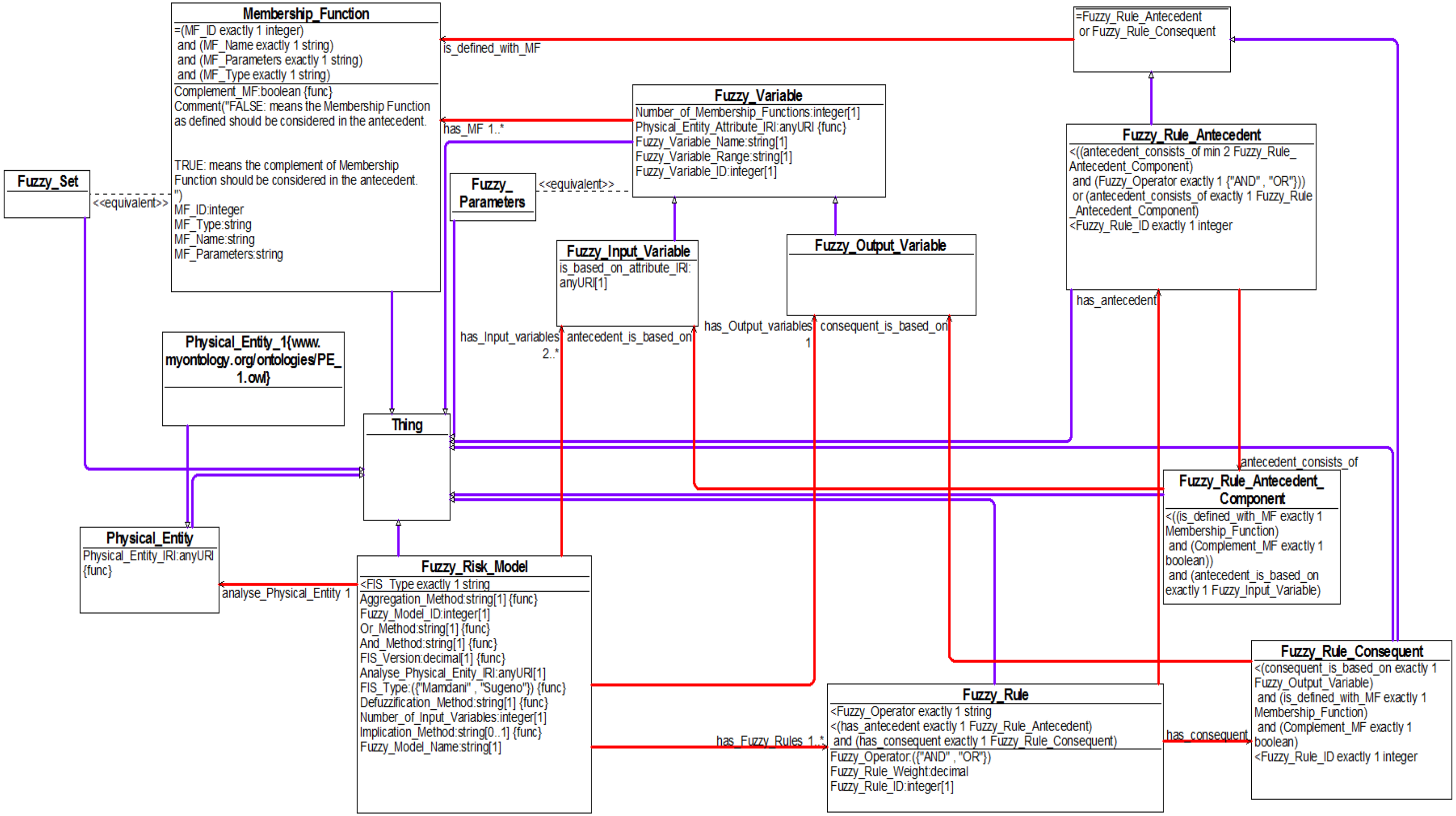

Figure 5-19: "Fuzzy Risk Model Ontology" – Class Diagram View (with OWLGrEd notation)

5.6 Summary

Following the description of the concept in Chapter 4, the semantic modelling and ontologies part discussed in this chapter. The role of ontologies in this concept clarified in 5.1. Due to the complexity of the domain, an architecture of ontologies presented with aim to manage better the various concepts. Benefits such modularity, maintainability, re-usability and extensibility discussed. In addition, the linking of ontologies mentioned for cases where there is the need to elaborate to concepts or to re-use concepts. It is worth noting that the decomposition is a matter of decision for organising the ontologies. The current decomposition described in detail, however, some further decomposition could have been performed for “Generic Customs Ontology” or integration with other upper ontologies and/or middle-level ontologies for re-using existing concepts from existing published ontologies.

Some illustrative examples presented above to demonstrate the representation of concepts at various levels (e.g. domain concepts or concepts related to Fuzzy Risk Ontology based on the principles of the fuzzy logic). In these examples, various OWL components were used. They used to depict the hierarchical structure of entities and the modelling of equality between concepts using equivalent classes (modelling same concepts (synonymous) with different names). Examples presented the usage of ontologies syntax for the definition of enumeration in data properties restricting the allowed data (e.g. fuzzy operator data property). In addition, object properties used for modelling the complex relationships among concepts. Property restrictions (e.g. cardinality restrictions) added in object properties to enrich definitions. The example of value partition presented for modelling specific concepts (e.g. example of Weigh Risk defined as the union of “High Risk”, “Medium Risk”, and “Low Risk”). As a conclusion, it is considered that this chapter addresses the second and third objectives of this research as expressed in section 1.4. A conceptual architecture of ontologies developed supporting the concept presented in Chapter 4. In addition, ontology models developed based on this architecture to represent concepts especially specific to the risk analysis with fuzzy logic technique. Those ontologies will be further used for fuzzy rule-based reasoning. Both fuzzy modelling and fuzzy reasoning are discussed in Chapter 6.

Finally, it is emphasized once more that this is a research activity for examining the use of ontologies in the Customs domain as a tool for modelling concepts and semantics in order to facilitate the communication, understanding, and interoperability in the context of risk analysis. This is also discussed in section 5.4.

Chapter 6. Fuzzy Modelling and Reasoning

This chapter investigates the fuzzy modelling and reasoning following the information presented in the previous chapters. For that purpose, fuzzy inference systems MAMDANI and SUGENO are analysed. The assessment approach is described along with the constraints for that research. Integration of ontologies (*fuzzy risk model*) with fuzzy rule-based reasoning is also presented. In addition, it is specified which types of FIS are used in this investigation and various decisions. The Linguistic Fuzzy Modelling (LFM) practises and other principles are discussed during this assessment. The various tools used for this assessment are also mentioned. At the end, analysis of the results is presented.

6.1 Overview

The fuzzy modelling and reasoning investigated through literature research and the development of generic research prototype. This prototype consists of six MISO fuzzy inference systems. Five of the fuzzy models are of Mamdani type and one is of Sugeno type. Those types of fuzzy inference systems have been discussed in section 3.6. In addition, the five Mamdani type fuzzy inference systems have different defuzzification methods.

The assessment focuses in the fuzzy reasoning. In the context of this generic research prototype, the ontologies are used for specifying the knowledge base with *fuzzy risk models* (e.g. fuzzy variables and membership functions). All *fuzzy risk models* are based on the principles of “Fuzzy Risk Model Ontology” discussed in section 5.5.3. In addition, it is considered that in this hypothetical scenario the *fuzzy risk models* created in the context of *Physical Entity* described in “Physical Entity Ontology” of section 5.5.2.

It is worth noting that one of the constraints of this research, and which had to be addressed for the development of the research prototype and the development of those fuzzy inference systems, was the approach and data to be used for that activity. Considering that previous sections focus on Customs domain, it is clarified that for the purpose of this activity no information and knowledge for real scenarios and real data are available, known, or found in order to be used for the development of fuzzy inference system and for evaluation purposes, due to the sensitive nature of such information. Therefore, this research has not access to such kind of information.

Consequently, the above constraint was an important factor for deciding to investigate fuzzy modelling and fuzzy reasoning (apart from bibliographic research) through the development of the aforementioned fuzzy inference systems based on randomly machine-generated evaluation data and not real classes of data. The evaluation data is randomly generated data, is not real and not specific to a real scenario. In fact, this fuzzy inference system is considered a generic FIS and it examines fuzzy reasoning from engineering point of view. Classes of data prepared for this testing and considering the constraints of this research described above. This data is designed in such a way to expect a specific output result based on the input vector in order to be able to compare the expected result and the actual result from fuzzy reasoning. The membership functions and fuzzy rules have been designed based on the classes of data and the evaluation data set in order to check the output of fuzzy inference per model. A set of evaluation data is randomly generated based on the classes of data defined in Table 6-2. The purpose is to evaluate the developed prototype. The research prototype FIS is evaluated against this randomly generated evaluation data set in order to analyse the various outputs of the system. More information about evaluation data used in this prototype is provided in section 6.2.

The fuzzy modelling and reasoning are discussed in the subsequent sections following literature research and by including information from the research prototype activity.

Finally, the activities described in the Use Cases UC1 Manage Fuzzy Risk Models (4.3.1), UC2 Perform Inference(4.3.2) and UC4 Manage Physical Entities(4.3.4) are required and performed for this prototype.

6.2 Dataset preparation for Research Prototype

For investigating the fuzzy modelling and fuzzy reasoning, evaluation data is used for testing and evaluation purposes of the research prototype. As stated in section 6.1, the evaluation data used in this research prototype is randomly generated data and is not real as discussed in the constraints of this research. A set of evaluation data is randomly generated based on the classes of data as shown in Table 6-2. Similarly, the classes of data shown in Table 6-2 have also been prepared for this testing and considering the constraints described in section 6.1. The classes are designed in such a way that based on specific input vector to expect a specific output result within the expected output range. Hence, this enables the comparison of the expected result with the actual result following the execution of the models (inference process). Each parameter or variable can take a specific range of values per class. The input values are generated randomly based on the defined range per class. The classes are built with the two-value logic and will be represented with fuzzy logic. For instance, the input parameter 4 (P4) is defined for Class A that takes values between '0' to '9'. The value '10' for P4 does not belong to Class A. The number of evaluation data records generated per Class for this prototype is shown in Table 6-1. The 61% of data represents class A as it is illustrated in the table. The 92% of data is of basic classes A, B, C, and D. As it is shown in Table 6-2, the remaining 8% is distributed to the classes that group set of data, which cannot be mapped to the basic classes A, B, C, and D based on the two-value logic. For instance, the class indicated as 'P4 NOT Class A and NOT Class B' cannot be considered neither as Class A nor as Class B because values for P4 are not in the range of those classes with the two-value logic. For the example of class 'P4 NOT Class A and NOT Class B', the input parameter 4 (P4) takes values '20' to

'30'. As stated above, the values of P4 in class A are between '0' to '9'; this range of data cannot be indicated as Class A with two-value logic. For the example of class 'P4 NOT Class A and NOT Class B', it means that members of this class cannot belong in either Class A or Class B. However, this can be considered in fuzzy logic with the definition of fuzzy sets accordingly. As it shown in Figure 6-10, values of this range have different degree of membership in the membership functions of P4.

Table 6-1: Number of evaluation data records generated per Class

Classes of Data	% of Total Records	# of Records
Class A	61,00%	1220
Class B	17,00%	340
Class C	10,00%	200
Class D	4,00%	80
P4 NOT Class A and NOT Class B (with P4 neither fully MF_1 nor fully MF_2)	2,00%	40
P3 NOT Class A and NOT Class C (with P3 neither fully MF_1 nor fully MF_2)	2,00%	40
P3 NOT Class B and NOT Class C (with P3 neither fully MF_1 nor fully MF_2)	2,00%	20
P4 NOT Class C and NOT Class D (with P4 neither fully MF_2 nor fully MF_3)	1,00%	40
P3 NOT Class C and NOT Class D (with P3 neither fully MF_2 nor fully MF_3)	1,00%	20
Total	**100%**	**2000**

Finally, the Expected Output MIN and the Expected Output MAX should be defined for the classes other than A, B, C, and D in order to be able to compare the actual output for those classes. Therefore, for this research prototype, the Expected Output MIN for the classes other than A, B, C, and D is defined as the *midrange* of Class with lowest values. The Expected Output MAX for the classes other than A, B, C, and D is defined as the *midrange* of the class with highest values. Therefore, in the case of 'P4 NOT Class A and NOT Class B' the Expected Output MIN and Expected Output MAX are calculated for this research prototype as shown in Figure 6-1. In this particular example, Class A has the lowest values and hence the Expected Output MIN is calculated based on the *midrange* of this class.

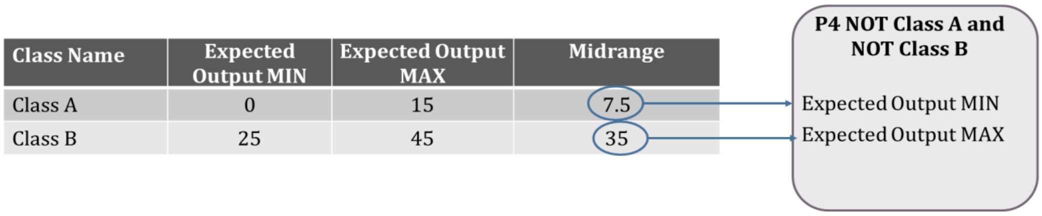

Figure 6-1: Example of MIN and MAX calculation of 'P4 NOT Class A and NOT Class B' class

Table 6-2: Classes of evaluation data[3]

	Expected Output		
Class Name	Expected Output MIN	Expected Output MAX	Midrange
Class A	0	15	7.5
Class B	25	45	35
Class C	55	75	65
Class D	85	100	92.5
P4 NOT Class A and NOT Class B (with P4 neither fully MF_1 nor fully MF_2)	7.5	35	21.25
P3 NOT Class A and NOT Class C (with P3 neither fully MF_1 nor fully MF_2)	7.5	65	36.25
P3 NOT Class B and NOT Class C (with P3 neither fully MF_1 nor fully MF_2)	35	65	50
P4 NOT Class C and NOT Class D (with P4 neither fully MF_2 nor fully MF_3)	65	92.5	78.75
P3 NOT Class C and NOT Class D (with P3 neither fully MF_2 nor fully MF_3)	65	92.5	78.75

6.3 Tools

Two tools have been primarily used for the purposes of this research prototype. These are the *Protégé* and the *MATLAB software*. In this prototype, the *Protégé* tool (Protégé) has been used as the *Physical Entity Manager/Editor* component and also as part of *Fuzzy Risk Analysis (Fuzzy RA)* component for the management of *fuzzy risk models* (section 4.2). Therefore, the *Protégé* has been used as a tool to define the *fuzzy risk models* (FIS) in the "Fuzzy Risk Model Ontology" (UC1 Manage Fuzzy Risk Models). In particular, the "Fuzzy Risk Model Ontology" defines the main parameters of *fuzzy risk*

[3] Classes of evaluation data have been prepared for this testing and considering the constraints. Please refer to section 6.1.

models (e.g. Defuzzification method), the fuzzy variables of each fuzzy model, the membership functions per fuzzy variable and the fuzzy rules. This is further elaborated in section 6.5.

The *MATLAB Fuzzy Toolbox* has been used in this prototype for performing inference as part of *Fuzzy Risk Analysis (Fuzzy RA)* component. The defined *fuzzy risk models* of the "Fuzzy Risk Model Ontology" are imported to *MATLAB.* This import is performed following a transformation of instances of the "Fuzzy Risk Model Ontology" (rendered in RDF) to the specific FIS format of MATLAB. The *MATLAB Fuzzy Logic Toolbox Graphical User Interface (GUI) Tools* can be used for visualising the membership functions and the rules (Rule Base) of the *fuzzy risk models* defined in the "Fuzzy Risk Model Ontology". Moreover, the evaluation data is generated based on classes of data shown in Table 6-2 (please see more information in section 6.2). The generated evaluated data is imported to MATLAB software as data input for analysis. The analysis is executed once and all individual results per model of Table 6-3 are recorded. This has been achieved by writing M-files with commands or functions for executing the analysis via the command line of MATLAB Fuzzy Toolbox. Inference and execution is discussed in section 6.5.

Figure 6-2 illustrates diagrammatically in high-level the tools used for this research prototype.

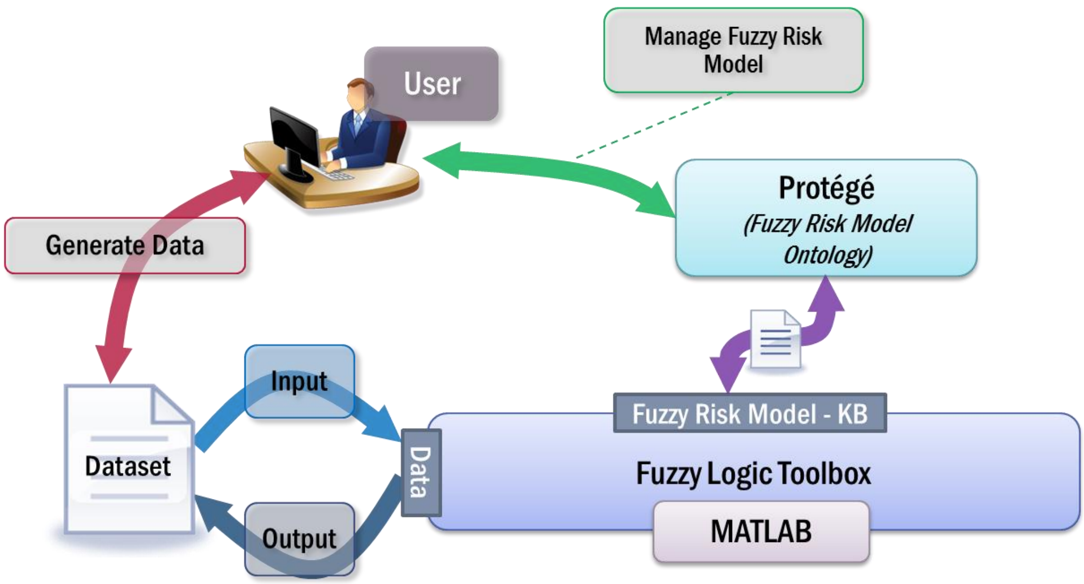

Figure 6-2: Overview of Environment – Research Prototype

6.4 Fuzzy Modelling

This section discusses the fuzzy modelling performed in the context of this research prototype. This activity is performed in the context of UC1 Manage Fuzzy Risk Models (4.3.1). It discusses the fuzzy modelling with the use of ontologies analysed in section 5.5. Particularly, it presents the use of "Fuzzy Risk Model Ontology" for modelling *fuzzy risk models*. Those *fuzzy risk models* will be used for fuzzy rule-based reasoning. In addition, it provides examples of fuzzy modelling for Customs domain and best practises from research that can be applied in this activity (e.g. LFM).

As mentioned before, six *fuzzy risk models* of two FIS types developed and assessed in the context of the research prototype. One *fuzzy risk model* is based on Sugeno FIS and five *fuzzy risk models* are Mamdani FIS. The main difference among the Mamdani FIS models is the defuzzification method. In fact, the following defuzzification methods were used:

1. **Centroid:** centroid of area
2. **Bisector**: bisector of area
3. **MOM**: mean value of maximum
4. **SOM**: smallest (absolute) value of maximum
5. **LOM**: largest (absolute) value of maximum

The above defuzzification methods have been described in section 3.6.2.

The same fuzzy inputs (including membership functions), same rule base (Fuzzy rules) and same evaluation dataset were used in order the models of this prototype to be comparable. The following table (Table 6-3) summarises the *fuzzy risk models* assessed under this research prototype.

Table 6-3: Fuzzy Models used for Research Prototype

Model	Type	# Input	# Output	Defuzzification Method
1	MAMDANI	4	1	centroid of area (centroid)

2	MAMDANI	4	1	bisector of area (bisector)
3	MAMDANI	4	1	mean value of maximum (MOM)
4	MAMDANI	4	1	smallest (absolute) value of maximum (SOM)
5	MAMDANI	4	1	Largest (absolute) value of maximum (LOM)
6	SUGENO	4	1	weighted average (wtaver)

Initially, the fuzzy variables of the models are defined. Apparently, the selection of correct input affects also the expected output of the FIS. The fuzzy variables of the model are stored in the "Knowledge Base". Gacto et al. (2011) present an overview of interpretability measures and techniques with purpose to have more interpretable linguistic fuzzy rule-based systems. Their work focuses on Linguistic fuzzy modelling (LFM). The *number of features or variables* is stated as one measure, which is used for controlling the *complexity at the level of fuzzy partitions*. The readability of knowledge base is improved with the reduction of the number of features. Finally, Gacto et al. (2011) discusses methods and works related to feature reduction. Torra (2001) states one of the two difficulties in complex domains is the fact that the number of variables is many. Therefore, this increases exponentially the required rules (it is also called 'curse of dimensionality'). The Hierarchical Fuzzy Systems are mentioned as a technique to handle this 'curse of dimensionality'. This is also discussed in section 4.1.2 to address such issues.

Singh and Sahu (2004) presents a decision support system for Customs examination, which incorporates human intelligence and experience of officers using linguistic terms and fuzzy logic based expert system. It is stated that fuzzy logic enables the use of linguistic variables for risk analysis. It is also mentioned that in reality, several factors affect the risk and therefore the overall risk is calculated from all risk factors. An example is provided in their work for the risk analysis of an import consignment.

In addition, Singh et al. (2003) presents an decision support system for Customs assessment to detect Valuation frauds. This system uses the expertise of officers to determine the sensitivity of the import. As stated by Singh et al. (2003), the sensitivity refers to the belief of import being sensitive to under-valuation. Moreover, it is mentioned by Singh et al. (2003) that the number of variables is very high and therefore the HHFC was used to reduce the rule base. The sensitivity of each input is given after defuzzification and based on expert rules. As stated by Singh et al. (2003), the Mamdani FIS has been used with "Centroid" defuzzification method. Four modules are defined with different input variables depending on the type. The sensitivity per type and the overall sensitivity are calculated based on an algorithm (Singh et al. 2003).

In this research prototype, all fuzzy models are defined with four (4) input parameters (P1, P2, P3, and P4) and a single output (MISO). Figure 6-3 illustrates the generic structure of fuzzy inference system for this research prototype, which is either Mamdani or Sugeno type as defined in Table 6-3 above. The number of inputs has been selected to have a moderate number of variables for the purposes of this research prototype requiring less number of rules for fuzzy reasoning.

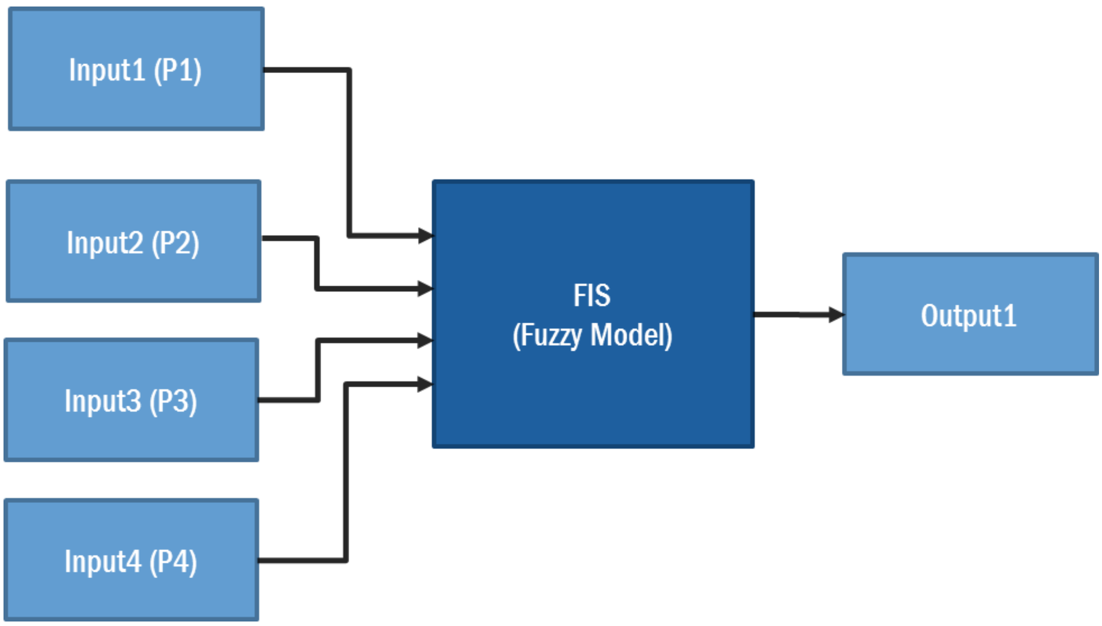

Figure 6-3: Structure of Fuzzy Inference System

In this research prototype, all models (Table 6-3) have the same four fuzzy input parameters. The membership functions and fuzzy rules of this model are built on those fuzzy parameters or variables and based on the classes of Table 6-2. The output of the fuzzy model indicates the risk analysis result. Therefore, all models have one output parameter (MISO) as shown in Figure 6-3. The "Fuzzy Risk Model Ontology" defines the *fuzzy risk models* of this research prototype mentioned above. This is visualised in Figure 6-4. The OntoGraf graph notation is explained in section 5.5.1. As it is shown in Figure 6-4, six individuals are defined in the ontology representing the *fuzzy risk models*. These are individuals of "Fuzzy Risk Model" class. In addition, four individuals are defined as individuals of "Fuzzy Input Variable" class representing the input parameters of the models for this prototype. Finally, two individuals of type "Fuzzy Output Variable" are defined to denote the output of the fuzzy model (risk analysis result).

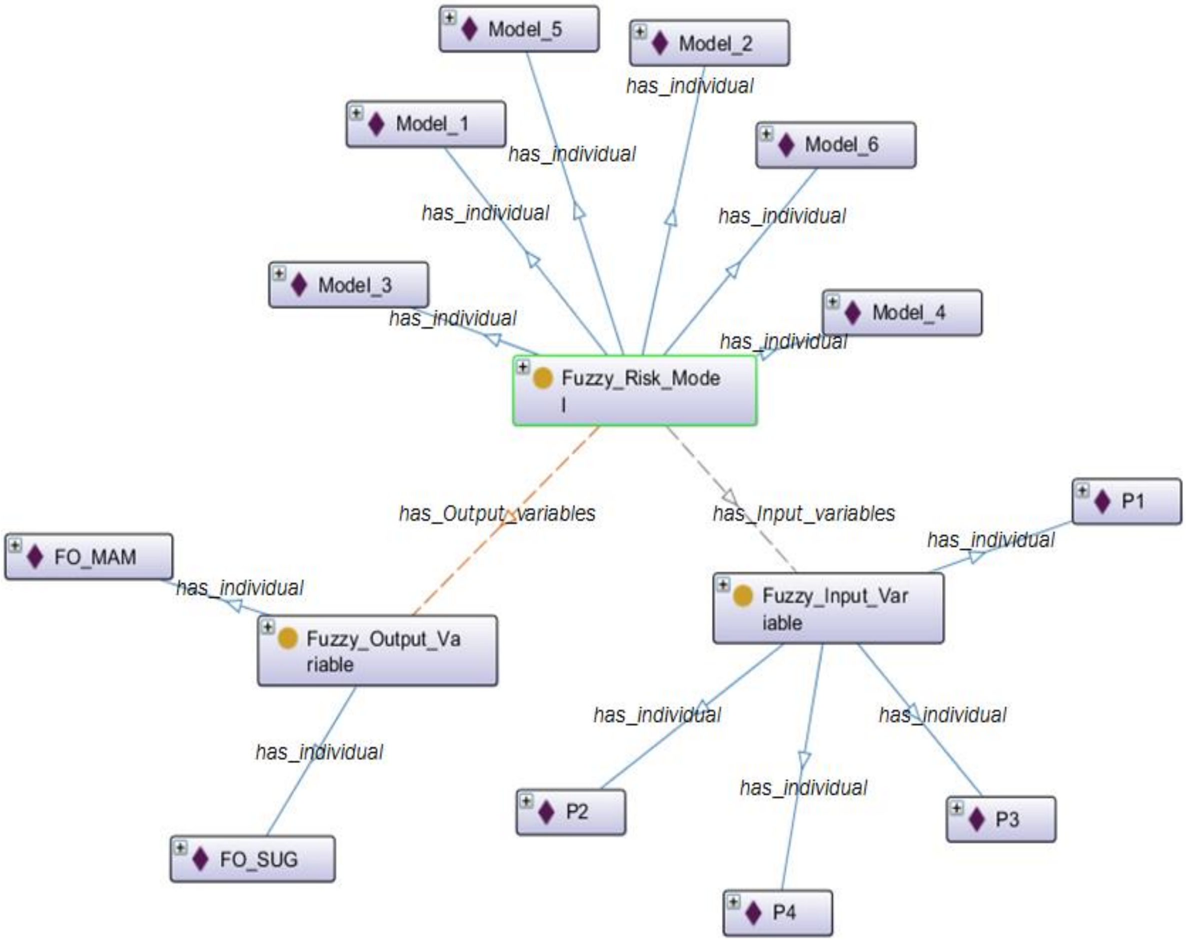

Figure 6-4: Fuzzy risk models of prototype in the "Fuzzy Risk Model Ontology" (OntoGraf Graph)

Nevertheless, Figure 6-4 do not show the relationships between individuals, which are inherited from their classes. The relationships are shown in the examples of Figure 6-5 and Figure 6-6. Those figures present the input parameters and output parameters of "Model 1" and "Model 6" respectively. These are defined using the object properties '*has_Input_variables*' and '*has_Output_variables*', which discussed in section 5.5.3. It is apparent that the same individuals of "Fuzzy Input Variable" type are used to specify the input parameters of "Model 1" and "Model 6" *fuzzy risk models*. This enables the re-usability if a parameter is the same in more than one model. Finally, Figure 6-5 shows that "Model 1" has output parameter the Mamdani fuzzy output (FO_MAM). While Figure 6-6 depicts that "Model 6" has a Sugeno fuzzy output (FO_SUG). This is in line with Table 6-3.

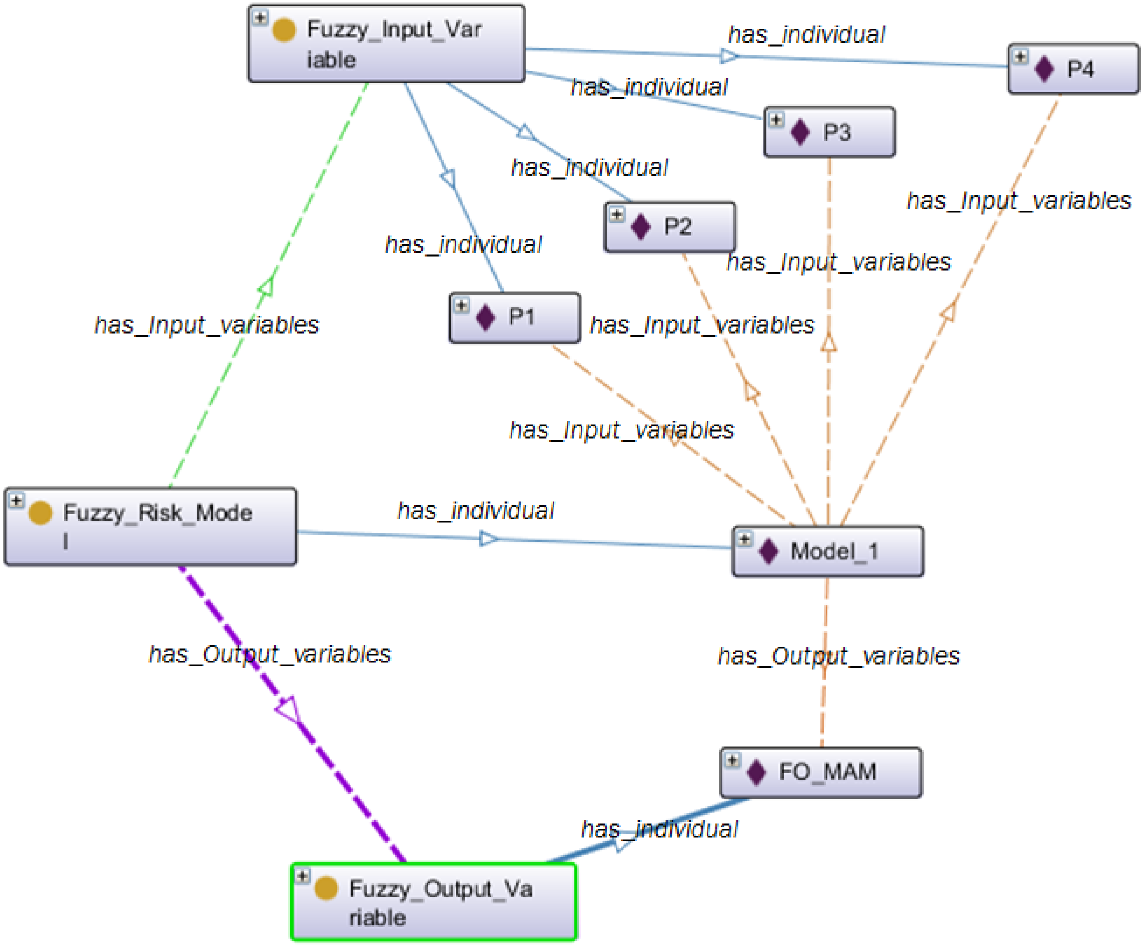

Figure 6-5: Relationships of "Model 1" with Fuzzy Input and Fuzzy Output variables individuals (OntoGraf Graph)

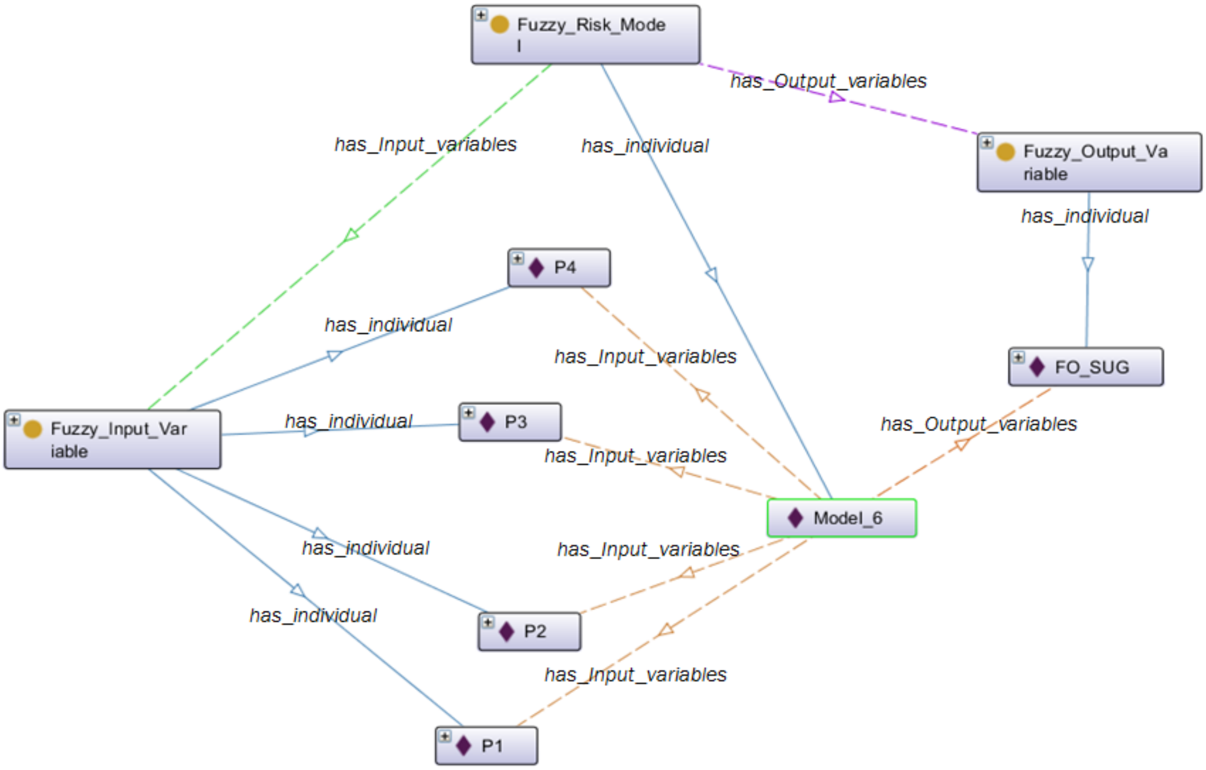

Figure 6-6: Relationships of "Model 6" with Fuzzy Input and Fuzzy Output variables individuals (OntoGraf Graph)

The membership functions are defined for each fuzzy parameter (input and output) in order to define how the values should be interpreted by the fuzzy model and how the output result should be inferred. In fact, the membership function of input fuzzy parameter is used to identify the degree of membership of crisp values to each fuzzy set. The latter is part of the fuzzification activity of the fuzzy inference process. The membership functions of output parameter defines the fuzzy sets that will be used for the expressing the consequent part of fuzzy rules. Those membership functions are also used during the implication method or rule evaluation for evaluating each fuzzy rule.

The *number of membership functions* is mentioned by Gacto et al. (2011) as another measure used for controlling the *complexity at the level of fuzzy partitions*. As stated, the number of membership functions should be moderate. An increase on the number of membership functions may increase precision of the system but also decrease its relevance. In addition, the number of membership functions should not exceed the principle 7 ± 2 since it is the number of conceptual entities a human being can handle

Gacto et al. (2011). Finally, Gacto et al. (2011) discusses methods and works related for decreasing number of membership functions.

A membership function has a linguistic term such as "low", "medium", "high", "few", "some", etc. representing the fuzzy variable. The fuzzy logic enables to use fuzzy sets and express information with linguistic terms. Additionally, the same element can also have different degree of membership to different fuzzy sets in contrast to classical (crisp) sets. Therefore, the linguistic terms can be used to express information for the input variables and the output of the risk analysis. Nevertheless, the relationship of input and output is defined with fuzzy rules. In the context of Decision Support System for Customs, some examples are also provided by Singh and Sahu (2004) indicating the use of fuzzy sets for risk analysis in this domain.

Analysing the aspect of semantics interpretability at fuzzy partition level, Gacto et al. (2011) state that complex fuzzy partitions (huge overlapping between membership functions) reduces the semantic interpretability. Some of the properties that are mentioned regarding semantics interpretability at fuzzy partition level are *completeness or coverage, normalisation*, *distinguishability* and *complementarity.* For instance, *distinguishability* requires that a membership function should represent a linguistic term with clear semantics and distinguishable from the other membership functions of this fuzzy variable. In addition, *normalisation* specifies that at least one data point in the membership function with membership value equal to one (1).

The membership functions for this research prototype defined in order to model the classes of data shown in Table 6-2. Input1 (P1), Input2 (P2), Input3 (P3) and Input4 (P4) have 4, 3, 4 and 3 membership functions respectively (Figure 6-10). These membership functions are defined in the "Fuzzy Risk Model Ontology" as individuals of type "Membership Function". Figure 6-7 depicts those individuals. In addition, the *'has_MF'* object property is used in order to define in the "Fuzzy Risk Model Ontology" the membership functions, which belong to each fuzzy parameter. Figure 6-8 presents an example for "Model 1" and "Model 6" of this prototype. Some of the properties at fuzzy partition level discussed in the paragraph above are shown in Figure 6-10 and Figure 6-11. Nevertheless, Gacto et al. (2011) mentions that it is not always possible to have or impose *strong fuzzy partitions* (satisfy the properties mentioned before) because if the system is based on experts knowledge then different fuzzy partition might be applied appropriate to the problem. Such properties could be considered during the definition of fuzzy partitions as semantic interpretability measures by also taking into consideration the particular problem.

Methods for assigning the membership values are intuition, inference, rank ordering, angular fuzzy sets, neural networks, genetic algorithms and inductive reasoning. Briefly, intuition concerns definition of membership functions based on human intelligence and understanding. It is also stated that this involves contextual and semantic modelling about an issue. In addition, it is stated for inductive reasoning that the entropy minimisation principle is used for the induction, which clusters the parameters corresponding to the output classes. It is also mentioned that a well-defined database for input-output relationships is needed for inductive reasoning method. However, it is mentioned that this method suites for complex systems with plenty and static data but not for cases with dynamic data. In the latter case, it does not suit because the membership functions changes continuously with time (Ross 2010; Sivanandam et al. 2007).

The membership functions of input fuzzy parameters of both Mamdani and Sugeno type models are shown in Figure 6-10. In this research prototype, the membership functions of input fuzzy parameters are the same of all six models defined in Table 6-3. Furthermore, the example of Figure 6-8 shows the re-usability of fuzzy parameters and membership functions by the *fuzzy risk models.*

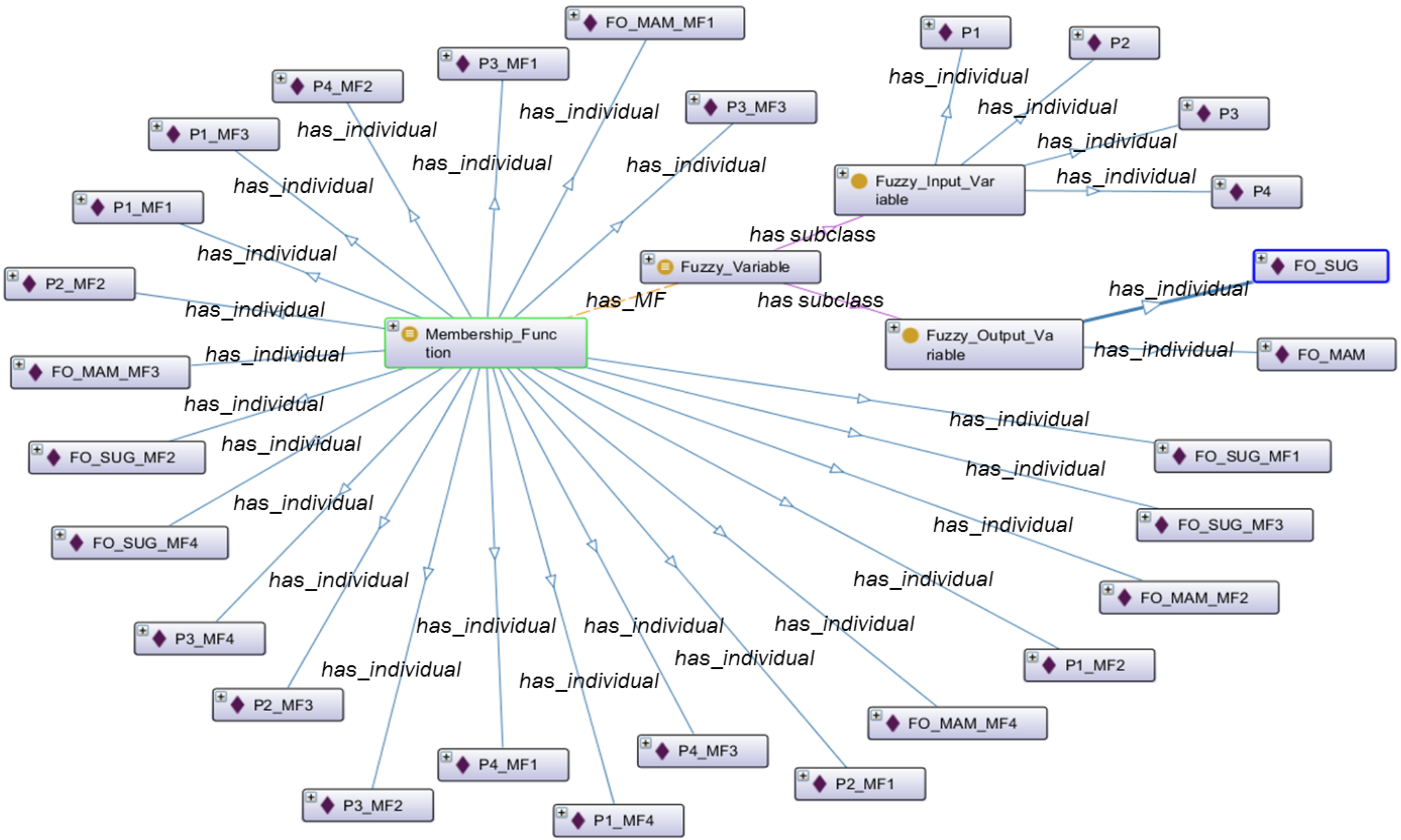

Figure 6-7:Fuzzy Variables and Membership Functions (OntoGraf Graph)

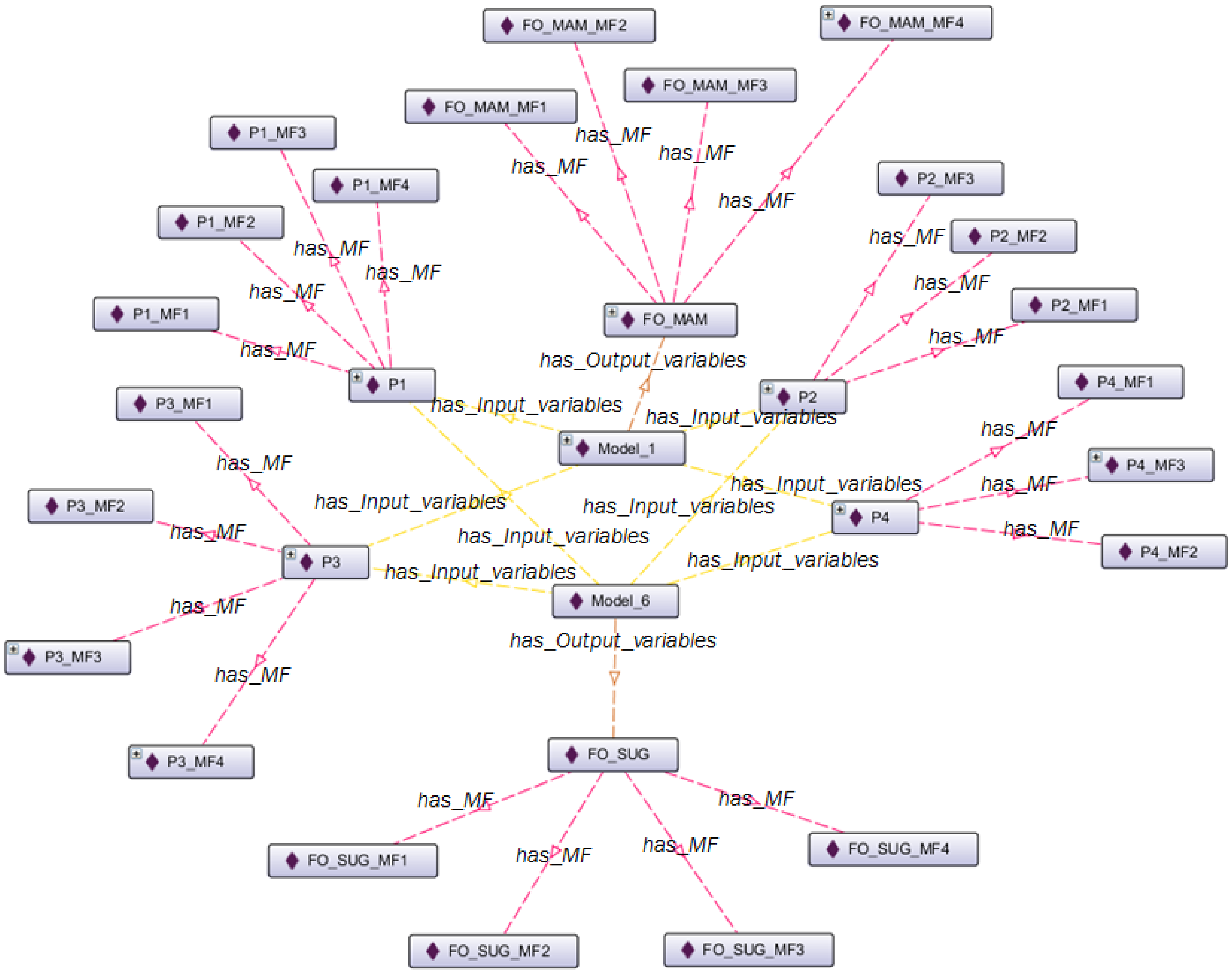

Figure 6-8: Example with “Model 1” and “Model 6” Fuzzy Variables and Membership Functions (OntoGraf Graph)

The membership functions of output fuzzy parameter are four for the fuzzy models of Mamdani type (please see example in Figure 6-8). In this particular example of research prototype, four fuzzy sets are defined for the output. The linguistic term class A, class B, class C and class D have been given in this example based on Table 6-2. The consequent part of fuzzy rules is constructed based on the aforementioned membership functions. The individuals FO_MAM_MF1, FO_MAM_MF2, FO_MAM_MF3 and FO_MAM_MF4 represents the linguistic terms class A, class B, class C and class D respectively in the “Fuzzy Risk Model Ontology”. Figure 6-9 presents an example for defining the membership function name (linguistic term) of FO_MAM_MF1 as class A. A *Data Property Assertion* is used for the data property *‘MF_Name’*.

```
    <DataPropertyAssertion>
        <DataProperty IRI="#MF_Name"/>
        <NamedIndividual IRI="#FO_MAM_MF1"/>
        <Literal datatypeIRI="&xsd;string">A</Literal>
    </DataPropertyAssertion>
```

Figure 6-9: OWL syntax for MF Name (linguistic term) of FO_MAM_MF1 in the “Fuzzy Risk Model Ontology”

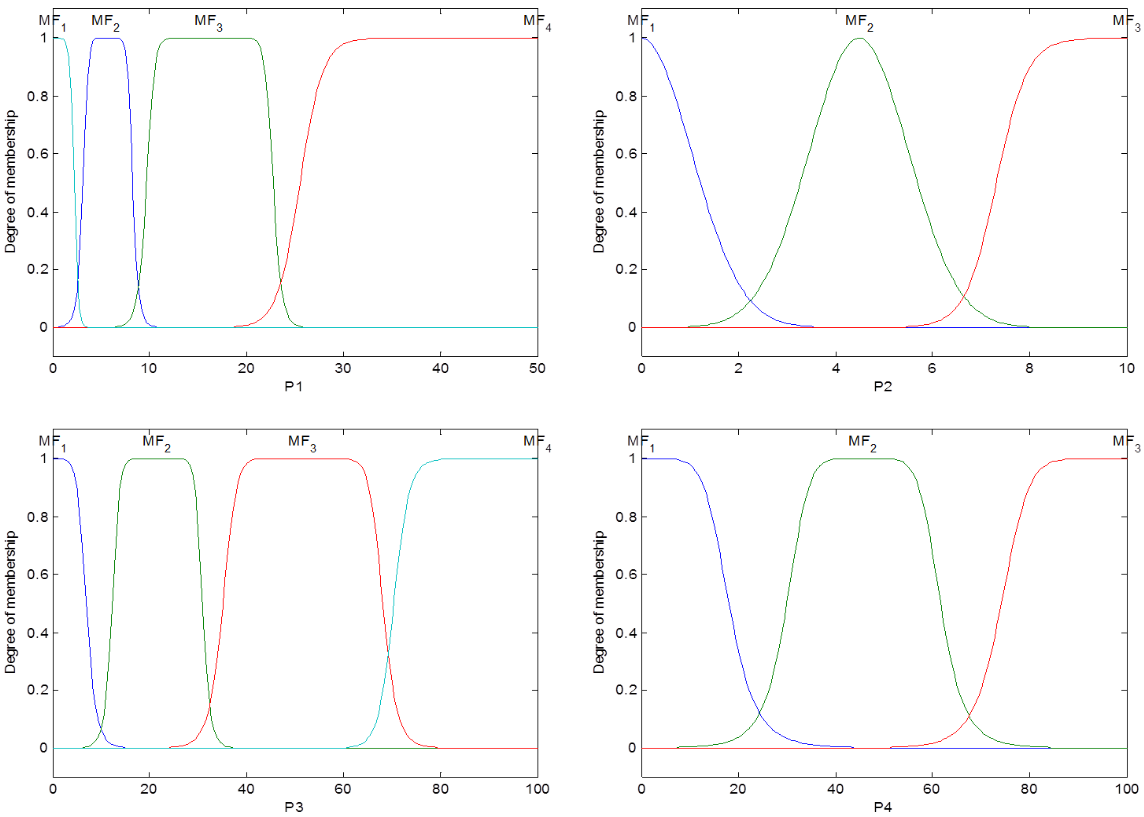

Figure 6-10: Fuzzy Input Variables (MATLAB)

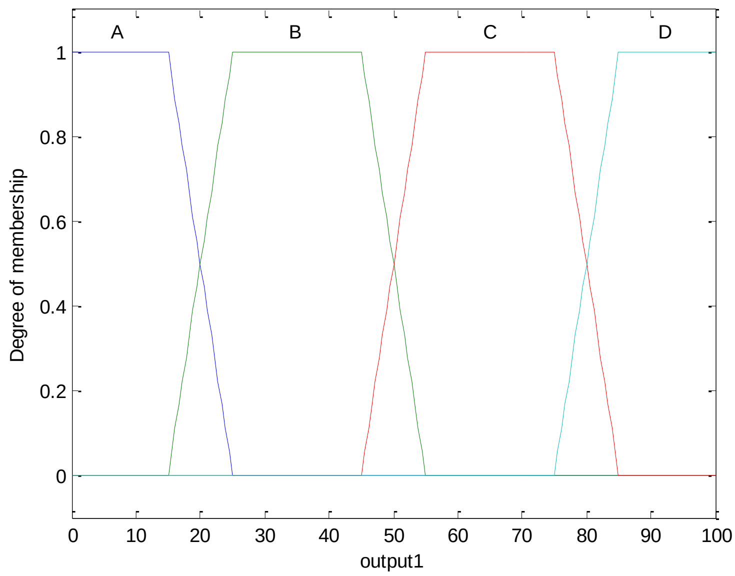

Figure 6-11: Membership Functions of Mamdani type Fuzzy Output Variable (MATLAB)

In Sugeno fuzzy inference system, the output membership functions can be either linear or constant. In this research prototype, the Model 6 is Sugeno type and the output membership functions are defined as constant. Therefore, the fuzzy model of Sugeno type (Model 6) is a zero-order Sugeno model as discussed in section 3.6.2. The definition of output for Sugeno requires to knowing the relationship with output. For this research prototype, the membership functions of output fuzzy parameter of Model 6 (Sugeno-type) have been constructed based on membership functions of output fuzzy parameter of models of Mamdani-type (models 1-5). It is also considered that this will enable to have outputs of fuzzy models comparable to the maximum extent. In particularly, four membership functions were defined for Model 6 (Sugeno type) with type constant. The constant value of each membership function is calculated from the Center of Gravity (centroid) of the corresponding membership function of Mamdani-type models (similar to Jassbi et al. (2006)). The membership functions of Mamdani type models are of trapezoidal form. The MATLAB Fuzzy Toolbox has been used for this research prototype as it is mentioned in section 6.3. Functions of MATLAB Fuzzy Toolbox are used as follows for defining the membership functions of (constant) output variable of Sugeno model. The trapezoidal membership function in Mamdani system is defined in MATLAB Fuzzy Toolbox as a vector with four scalar parameters MFp(1), MFp(2), MFp(3) and MFp(4), which are used to draw trapezoidal curve as depicted in Figure 6-12. Those scalar parameters of vector of trapezoidal membership function are used to estimate the centroid. The procedure is automated based on the definition of Mamdani fuzzy system. Initially, the output membership functions of Mamdani model are retrieved. Then, a trapezoid curve is defined per membership function based on the above four scalar parameters MFp(1), MFp(2), MFp(3) and MFp(4). For each output MF, the '*defuzz*' function is used to find the centroid of the specific trapezoid MF as described in Figure 6-13. The procedure is performed for each of the four membership functions. It is worth noting that the definition of output membership functions of Sugeno model based on the Mamdani fuzzy model is an approach selected for the purpose of this research prototype. For instance, the output membership functions (constant) could be defined with intuition or other logic.

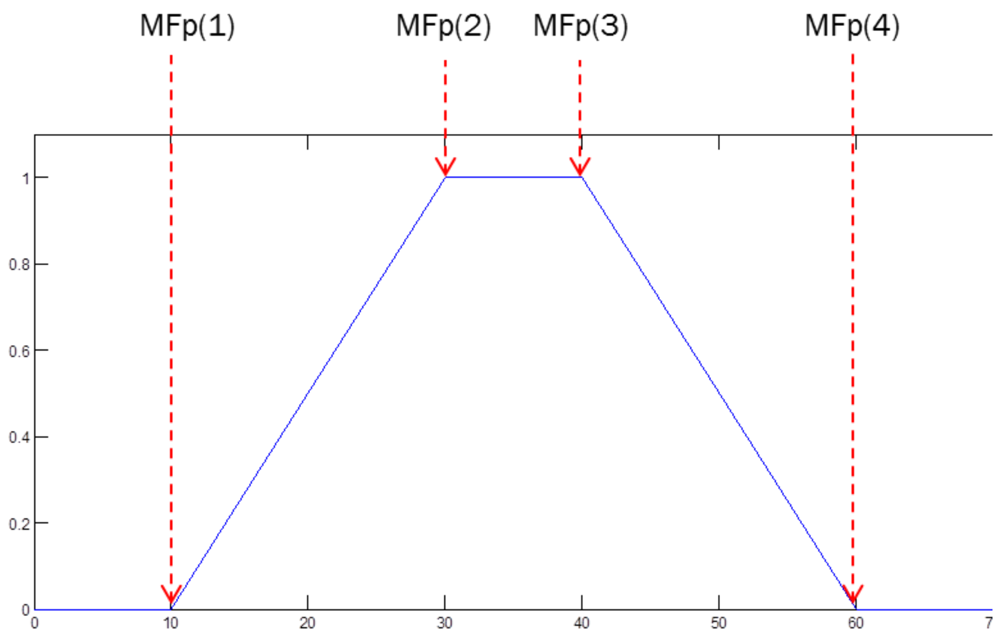

Figure 6-12: Example of trapezoidal membership function

```
…

% Read output membership function A of Mamdani model
out_mf_A = getfis (RA_FIS_MAMDANI, 'output', 1, 'mf', 1, 'params');

…

% Define range
x = 0:10:100;
```

```
% Calculate Centroid for Class A

mf_A = trapmf(x,[out_mf_A(1) out_mf_A(2) out_mf_A(3) out_mf_A(4)]);
Cx_mf_A = defuzz(x, mf_A,'centroid')


…
```

Figure 6-13: Example of Code for generation of Output MF 1 of Sugeno model from Output MFs of Mamdani type – MATLAB file

Subsequently, the fuzzy rules of the model are specified for the identified fuzzy parameters or variables and based on the defined membership functions. The fuzzy rules are used in this approach as the knowledge for risk analysing a particular instance of a *physical entity* and define the relationship of the input with the output using linguistic terms. This is very useful for expressing human knowledge especially for Mamdani systems where the outputs are also defined with membership functions.

For this research prototype, the fuzzy rules defined based on the classes shown in Table 6-2. The fuzzy rules of Mamdani fuzzy inference systems have the form shown in equation (8). The fuzzy rules of Sugeno fuzzy inference systems have the form shown in equation (13) where *a* and *b* are equal to zero (0) since it is a *zero-order Sugeno model.*

The fuzzy operators 'AND' and 'OR' are used to combine the various input fuzzy parameters in case the antecedent of fuzzy rule is defined with more than one input fuzzy parameter. As an example, the antecedent of the following fuzzy rule uses the 'AND' fuzzy operator to combine the input fuzzy parameters *P3* and *P4*: "***IF (P3 is MF3) and (P4 is MF3) THEN..***". This example is visualised in Figure 6-16 using the "Fuzzy Risk Model Ontology". The fuzzy rules of various models of this prototype modelled using the structure of "Fuzzy Risk Model Ontology" (section 5.5.3). As explained in that section, each rule has an antecedent and a consequent. For "Rule_1" (first rule in Table 6-4), this is defined with the "FRA_1" individual representing the "Fuzzy Rule Antecedent" and the "FRC_FO_MAM_MF4" individual representing the "Fuzzy Rule Consequent". The definitions of "FRA_1" and "FRC_FO_MAM_MF4" individuals with RDF syntax are shown in Figure 6-14 and Figure 6-15 respectively.

```
<!-- http://www.myontology.org/ontologies/FRO.owl#FRA_1 -->

    <NamedIndividual rdf:about="&FRO;FRA_1">
        <rdf:type rdf:resource="&FRO;Fuzzy_Rule_Antecedent"/>
        <FRO:Fuzzy_Operator>AND</FRO:Fuzzy_Operator>
        <FRO:antecedent_consists_of
rdf:resource="&FRO;FRA_P3_MF3_COMPL_FALSE"/>
        <FRO:antecedent_consists_of
rdf:resource="&FRO;FRA_P4_MF3_COMPL_FALSE"/>
    </NamedIndividual>
```

Figure 6-14: RDF syntax for FRA_1 Individual (Fuzzy Rule Antecedent) in the "Fuzzy Risk Model Ontology"

As it is defined in the "Fuzzy Risk Model Ontology", a "Fuzzy Rule Antecedent" consists of one or more "Fuzzy_Rule_Antecedent_Component". In this example of "Rule_1", the "FRA_1" (Figure 6-14) consists of the antecedent components "FRA_P3_MF3_COMPL_FALSE" and "FRA_P4_MF3_COMPL_FALSE". Those relationships are illustrated in Figure 6-16. Finally, the data property *'Fuzzy_Operator'* of "FRA_1" defines that the 'AND' fuzzy operator is used to combine the input fuzzy parameters *P3* and *P4* in the antecedent part of the rule (Figure 6-14).

Figure 6-15 shows that the "FRC_FO_MAM_MF4" rule consequent applies to output variable "FO_MAM" of "Model_1" and it is defined with membership function "FO_MAM_MF4". Those relationships are also visualised in Figure 6-16.

```
    <!-- http://www.myontology.org/ontologies/FRO.owl#FRC_FO_MAM_MF4 -
->

    <NamedIndividual rdf:about="&FRO;FRC_FO_MAM_MF4">
        <rdf:type rdf:resource="&FRO;Fuzzy_Rule_Consequent"/>
        <FRO:Complement_MF
rdf:datatype="&xsd;boolean">false</FRO:Complement_MF>
        <FRO:consequent_is_based_on rdf:resource="&FRO;FO_MAM"/>
        <FRO:is_defined_with_MF rdf:resource="&FRO;FO_MAM_MF4"/>
    </NamedIndividual>
```

Figure 6-15: RDF syntax for FRC_FO_MAM_MF4 Individual (Fuzzy Rule Consequent) in the "Fuzzy Risk Model Ontology"

The *number of rules* and the *number of conditions* are measures, which is used for controlling the *complexity at the level of rule base*. The number of fuzzy rules should be reduced but without affecting the system performance, which shall remain at satisfactory level. Similarly, the *number of conditions* in the antecedent part of the rule should be reduced and consider the principle 7 ± 2 (referring to the number of conceptual entities a human being can handle), but also without affecting the system performance, which shall remain at satisfactory level (Gacto et al. 2011). In addition, the “don’t care” approach (Ishibuchi et al. 1998) reduces the “conditions” in the antecedent part and the complexity at rule base. The example mentioned above of the following fuzzy rule of this prototype “***IF (P3 is MF3) and (P4 is MF3) THEN…”*** considers the “don’t care” approach since P1 and P2 are “don’t care” conditions. Some indicative rules of the rule base of this prototype are presented in Table 6-4.

Table 6-4: Indicative rules of the Rule Base of this Prototype

If (P3 is MF_3) and (P4 is MF_3) then (output1 is D) (1)
…
If (P3 is MF_1) and (P4 is MF_1) then (output1 is A) (1)
…
If (P1 is MF_3) and (P3 is MF_1) and (P4 is MF_2) then (output1 is B) (1)
…
If (P1 is MF_3) and (P3 is MF_2) and (P4 is MF_2) then (output1 is C) (1)
…
If (P1 is MF_4) and (P2 is MF_2) and (P3 is MF_2) and (P4 is MF_1) then (output1 is B) (1)
…
If (P1 is MF_4) and (P2 is MF_2) and (P3 is MF_1) and (P4 is MF_3) then (output1 is B) (1)

During the inference process, the fuzzy operators are applied after the fuzzification of input parameters in order to combine the individual inputs of the antecedent and determine one number, which is used in the next step of the inference process where the implication method is applied. For the fuzzy models of Table 6-3, the Minimum (min) method is used for ‘AND’ operator and the Maximum (max) for ‘OR’ operator. Please refer to fuzzy operations in section 3.6.1 and specifically to equations (6) and (5) respectively.

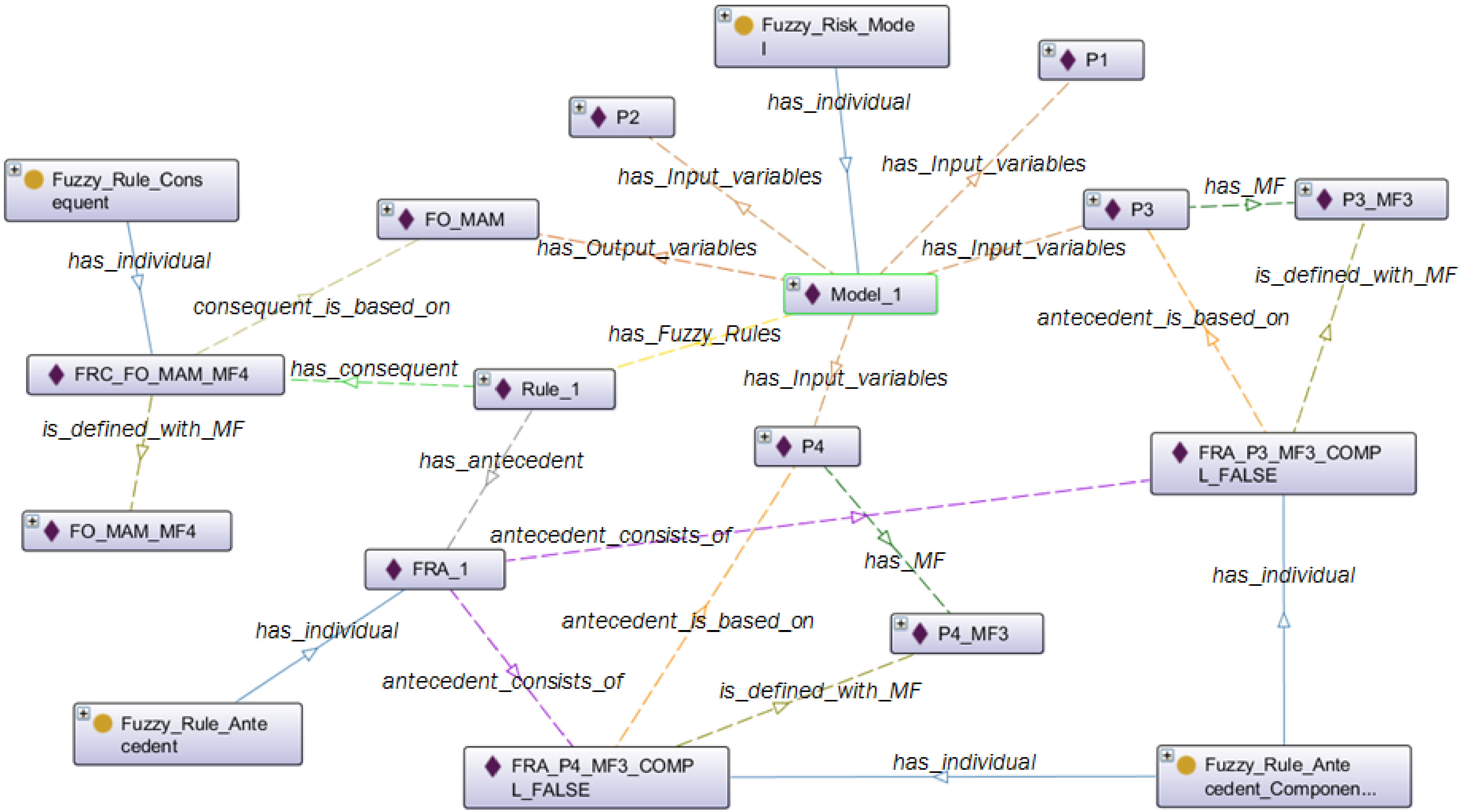

Figure 6-16: Example of "Rule_1" for "Model_1" (OntoGraf Graph)

Finally, an activity parallel to the definition of fuzzy parameters, membership functions, and fuzzy rules is to specify the global parameters or attributes of *fuzzy risk model* (e.g. Defuzzification Method). However, these parameters or attributes are used in specific steps of the inference process. For instance, it is stated above that the ‘AND’ and ‘OR’ fuzzy operators are used for combining (evaluating) the antecedents of a fuzzy rule. Table 6-3 summarises the various methods used per model for ‘AND’ and ‘OR’ fuzzy operators. It is noted that for Sugeno type model, the *Algebraic Product (prod)* (21) and the *Algebraic Sum* or *Probabilistic OR* (22) were used as methods for ‘AND’ and ‘OR’ fuzzy operators respectively.

$$\mu_{A\cap B}(\chi) = \mu_A(\chi)\,\mu_B(\chi) \qquad (21)$$

$$\mu_{A\cup B}(\chi) = \mu_A(\chi) + \mu_B(\chi) - \mu_A(\chi)\,\mu_B(\chi) \qquad (22)$$

In addition, the Mamdani type fuzzy models of Table 6-3 use the *Product* method as implication method, which scales the output fuzzy set based on the number of fuzzy rule antecedent. For Sugeno model, the output of each rule is weighted by the firing strength of the rule (see equation (14)). Another parameter of fuzzy inference system is the Aggregation method, which is used for aggregating all the results from individual rule evaluation and produces the aggregated output fuzzy set. The fuzzy models of Mamdani type use the *Maximum (max)* method for that purpose. In Sugeno model, the aggregation is the sum of the individual rule outputs as mentioned in section 3.6.2. Finally, the defuzzification method takes as input the aggregated output fuzzy set and based on the method, the output is defuzzified into a crisp number. As it was discussed, different defuzzification method is applied to each model of Table 6-3. The various defuzzification methods used for the fuzzy models are discussed in section 3.6.2. Table 6-5 summarises the parameters of *fuzzy risk models* used in this particular prototype.

Table 6-5: Global parameters per model

Model	Type	# Input	# Output	‘AND’ method	‘OR’ method	Implication Method	Aggregartion Method	Defuzzification Method
1	MAMDANI	4	1	min	max	prod	max	COA
2	MAMDANI	4	1	min	max	prod	max	BOA

Model	Type	# Input	# Output	'AND' method	'OR' method	Implication Method	Aggregartion Method	Defuzzification Method
3	MAMDANI	4	1	min	max	prod	max	MOM
4	MAMDANI	4	1	min	max	prod	max	SOM
5	MAMDANI	4	1	min	max	prod	max	LOM
6	SUGENO	4	1	prod	probor			Wtaver

Figure 6-17 and Figure 6-18 illustrate part of the definition of "Model_1" and "Model_6" individuals respectively in the "Fuzzy Risk Model Ontology". These definitions presents how the global parameters or attributes per *fuzzy risk model*, which described in Table 6-5, are defined in the ontology. As mentioned in section 5.5.3, the data properties of "Fuzzy Risk Model" class are used for that purpose. "Model_1" and "Model_6" are individuals of type "Fuzzy Risk Model". For instance, the data property *'FIS_Type'* specifies that "Model_1" is 'Mamdani' (Figure 6-17) while 'Model_6' is 'Sugeno' (Figure 6-18).

```
<!-- http://www.myontology.org/ontologies/FRO.owl#Model_1 -->

    <NamedIndividual rdf:about="&FRO;Model_1">
        <rdf:type rdf:resource="&FRO;Fuzzy_Risk_Model"/>
        <FRO:Fuzzy_Model_ID
         rdf:datatype="&xsd;integer">1</FRO:Fuzzy_Model_ID>
        <FRO:FIS_Version
         rdf:datatype="&xsd;decimal">2.0</FRO:FIS_Version>
        <FRO:Number_of_Input_Variables
         rdf:datatype="&xsd;integer">4</FRO:Number_of_Input_Variables>
        <FRO:FIS_Type
         rdf:datatype="&xsd;string">Mamdani</FRO:FIS_Type>
        <FRO:Fuzzy_Model_Name  rdf:datatype="&xsd;string">Model  1  -
        Mamdani - COA</FRO:Fuzzy_Model_Name>
        <FRO:Defuzzification_Method
         rdf:datatype="&xsd;string">centroid</FRO:Defuzzification_Meth
         od>
        <FRO:Analyse_Physical_Enity_IRI
         rdf:datatype="&xsd;anyURI">http://www.myontology.org/ontologi
         es/PE_1.owl#Physical_Entity_1</FRO:Analyse_Physical_Enity_IRI
         >
        <FRO:Aggregation_Method
         rdf:datatype="&xsd;string">max</FRO:Aggregation_Method>
        <FRO:Or_Method rdf:datatype="&xsd;string">max</FRO:Or_Method>
        <FRO:And_Method
         rdf:datatype="&xsd;string">min</FRO:And_Method>
        <FRO:Implication_Method
         rdf:datatype="&xsd;string">prod</FRO:Implication_Method>

    .........

    </NamedIndividual>
```

Figure 6-17: RDF syntax with part of definition of "Model_1" in the "Fuzzy Risk Model Ontology"

```
<NamedIndividual rdf:about="&FRO;Model_6">
        <rdf:type rdf:resource="&FRO;Fuzzy_Risk_Model"/>
        <FRO:FIS_Version
        rdf:datatype="&xsd;decimal">2.0</FRO:FIS_Version>
        <FRO:Number_of_Input_Variables
        rdf:datatype="&xsd;integer">4</FRO:Number_of_Input_Variables>
        <FRO:Fuzzy_Model_ID
        rdf:datatype="&xsd;integer">6</FRO:Fuzzy_Model_ID>
        <FRO:Fuzzy_Model_Name   rdf:datatype="&xsd;string">Model   6   -
        Sugeno</FRO:Fuzzy_Model_Name>
        <FRO:FIS_Type rdf:datatype="&xsd;string">Sugeno</FRO:FIS_Type>
        <FRO:Or_Method
        rdf:datatype="&xsd;string">probor</FRO:Or_Method>
        <FRO:And_Method
        rdf:datatype="&xsd;string">prod</FRO:And_Method>
        <FRO:Aggregation_Method
        rdf:datatype="&xsd;string">sum</FRO:Aggregation_Method>
        <FRO:Defuzzification_Method
        rdf:datatype="&xsd;string">wtaver</FRO:Defuzzification_Method
        >

    .........

    </NamedIndividual>
```

Figure 6-18: RDF syntax with part of definition of “Model_6” in the “Fuzzy Risk Model Ontology”

6.5 Fuzzy Inference

During this activity, the various fuzzy models (section 6.4) of research prototype are executed against the prepared datasets (section 6.2). This refers to the UC2 Perform Inference described in section 4.3.2.

In section 6.3, the tools for this research prototype are described. As it is explained, the MATLAB is used in this prototype for performing inference. The *fuzzy risk models* defined during fuzzy modelling are imported to *MATLAB.* A necessary task for this is the transformation of *fuzzy risk models* individuals defined in “Fuzzy Risk Model Ontology” (rendered in RDF syntax) to the specific FIS format of MATLAB. This performed mainly in two stages. The first stage is to transform the *fuzzy risk models* expressed in RDF into XML format. The second stage is to transform the *fuzzy risk models* in XML syntax to FIS format of MATLAB.

Figure 6-19 presents an example of extracting the basic details of *fuzzy risk model.* It includes the fuzzy variables list and fuzzy rule list of the specific model. The ‘*FuzzyVariableIRI’* element refers to the IRI of the specific individual of the “Fuzzy Risk

Model Ontology". The list of variables and rules per *fuzzy risk model* are transformed in XML by combining the various object properties of individuals (e.g. '*has_Input_variables*' and '*has_Output_variables*') in "Fuzzy Risk Model Ontology". For instance, the relationships of "Model 1" with Fuzzy Input and Fuzzy Output variables individuals illustrated in Figure 6-5 are expressed in transformed XML as shown in the example of Figure 6-19 (Fuzzy Variable List).

```
<?xml version="1.0" encoding="UTF-8"?>
<FuzzyRiskModel xsi:noNamespaceSchemaLocation="FuzzyRiskModel.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
      <ModelName>http://www.myontology.org/ontologies/FRO.owl#Model_1<
/ModelName>
      <FuzzyVariableList>

      <FuzzyVariableIRI>http://www.myontology.org/ontologies/FRO.owl#P
1</FuzzyVariableIRI>
      </FuzzyVariableList>
      <FuzzyVariableList>

      <FuzzyVariableIRI>http://www.myontology.org/ontologies/FRO.owl#P
2</FuzzyVariableIRI>
      </FuzzyVariableList>
      <FuzzyVariableList>

      <FuzzyVariableIRI>http://www.myontology.org/ontologies/FRO.owl#P
3</FuzzyVariableIRI>
      </FuzzyVariableList>
      <FuzzyVariableList>

      <FuzzyVariableIRI>http://www.myontology.org/ontologies/FRO.owl#P
4</FuzzyVariableIRI>
      </FuzzyVariableList>
      <FuzzyVariableList>

      <FuzzyVariableIRI>http://www.myontology.org/ontologies/FRO.owl#F
O_MAM</FuzzyVariableIRI>
      </FuzzyVariableList>
      <FuzzyRuleList>

      <FuzzyRuleIRI>http://www.myontology.org/ontologies/FRO.owl#Rule_
1</FuzzyRuleIRI>
      </FuzzyRuleList>
      <FuzzyRuleList>

      <FuzzyRuleIRI>http://www.myontology.org/ontologies/FRO.owl#Rule_
2</FuzzyRuleIRI>
      </FuzzyRuleList>

.........


</FuzzyRiskModel>
```

Figure 6-19: Example of XML syntax for Model_1 (list of fuzzy variables and rules)

The rules per *fuzzy risk model* also transformed in XML syntax (Rule Base) based on the “Fuzzy Risk Model Ontology”. Figure 6-20 presents an example for fuzzy rule list for “Model_1”. This example shows the “Rule_1”, which also presented in Figure 6-16. Again the relevant object properties where used for this transformation.


```
<?xml version="1.0" encoding="UTF-8"?>
<FuzzyRulesBase xsi:noNamespaceSchemaLocation="FuzzyRules.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
      <FuzzyRulesList>
      <FuzzyRuleIRI>http://www.myontology.org/ontologies/FRO.owl#Rule_
1</FuzzyRuleIRI>
      <FuzzyModel>http://www.myontology.org/ontologies/FRO.owl#Model_1
</FuzzyModel>
            <FuzzyRuleID>1</FuzzyRuleID>
            <FuzzyRuleAntecedent>
      <FuzzyRuleAntecedentIRI>http://www.myontology.org/ontologies/FRO
.owl#FRA_1</FuzzyRuleAntecedentIRI>
                  <FuzzyRuleAntecedentComponent>
      <FuzzyInputVariableIRI>http://www.myontology.org/ontologies/FRO.
owl#P1</FuzzyInputVariableIRI>
                  <FuzzyInputVariableID>1</FuzzyInputVariableID>
                  <FuzzyInputMFID>0</FuzzyInputMFID>
                  </FuzzyRuleAntecedentComponent>
                  <FuzzyRuleAntecedentComponent>
      <FuzzyInputVariableIRI>http://www.myontology.org/ontologies/FRO.
owl#P2</FuzzyInputVariableIRI>
                        <FuzzyInputVariableID>2</FuzzyInputVariableID>
                        <FuzzyInputMFID>0</FuzzyInputMFID>
                  </FuzzyRuleAntecedentComponent>
                  <FuzzyRuleAntecedentComponent>

      <FuzzyInputVariableIRI>http://www.myontology.org/ontologies/FRO.
owl#P3</FuzzyInputVariableIRI>
                        <FuzzyInputVariableID>3</FuzzyInputVariableID>
                        <FuzzyInputMFID>3</FuzzyInputMFID>
                  </FuzzyRuleAntecedentComponent>
                  <FuzzyRuleAntecedentComponent>

      <FuzzyInputVariableIRI>http://www.myontology.org/ontologies/FRO.
owl#P4</FuzzyInputVariableIRI>
                        <FuzzyInputVariableID>4</FuzzyInputVariableID>
                        <FuzzyInputMFID>3</FuzzyInputMFID>
                  </FuzzyRuleAntecedentComponent>
                  <FuzzyOperator>AND</FuzzyOperator>
            </FuzzyRuleAntecedent>
            <FuzzyRuleConsequent>

      <FuzzyRuleConsequentIRI>http://www.myontology.org/ontologies/FRO
.owl#FRC_FO_MAM_MF4</FuzzyRuleConsequentIRI>

      <FuzzyOutputVariableIRI>http://www.myontology.org/ontologies/FRO
.owl#FO_MAM</FuzzyOutputVariableIRI>
                  <FuzzyOutputVariableID>1</FuzzyOutputVariableID>

      <FuzzyOutputMFIRI>http://www.myontology.org/ontologies/FRO.owl#F
O_MAM_MF4</FuzzyOutputMFIRI>
                  <FuzzyOutputMFID>4</FuzzyOutputMFID>
```

```
                    <ComplementMF>false</ComplementMF>
              </FuzzyRuleConsequent>
        </FuzzyRulesList>

. . . . .
</FuzzyRulesBase>
```

Figure 6-20: Example of XML syntax for of Rule 1 for Model 1

The examples of Figure 6-19 and Figure 6-20 presents some of the information transformed from the "Fuzzy Risk Model Ontology" (rendered in RDF) into XML. The Extensible Stylesheet Language (XSL) and XSL Transformations (XSLT) (W3C 2001) can be used for the transformations. In addition, *fuzzy risk models* in XML could be transformed similarly into FML syntax if there is the need. FML is discussed in section 3.8. Such transformation would enable interoperability.

Following the final transformation of *fuzzy risk model* into FIS format of MATLAB, the fuzzy inference can be executed. This execution of models in MATLAB performed as follows. For Mamdani type fuzzy models, a MATLAB file defined for the execution of the models with instructions, functions, and commands. The definition of Mamdani type fuzzy models has been analysed in section 6.4. An example is shown in Figure 6-21 with a MATLAB file used for the evaluation of "Model_1". Similarly, Sugeno fuzzy model is defined and executed. A MATLAB file defined for the execution of the model with instructions, functions, and commands. This execution file also includes the estimation of four output membership functions of Sugeno type as described in 6.4 and shown in Figure 6-13. The output results of fuzzy inference are analysed in section 6.6.

```
% *** ANALYSIS MAMDANI (Defuzz Method CENTROID) ***

.........

%setfis(FIS_MAMDANI,'defuzzmethod','newfisprop')
RA_FIS_MAMDANI_CENTROID = setfis
(RA_FIS_MAMDANI,'defuzzmethod','centroid');
display (RA_FIS_MAMDANI_CENTROID)

display (' ** MAMDANI Analysis of input data started **')

%Evaluate Fuzzy Inference System again Input Data
RA_FIS_MAMDANI_CENTROID_Result =
evalfis(ANALYSIS_DATA_INPUT,RA_FIS_MAMDANI_CENTROID);
```

```
. . . . . . . . .
```

Figure 6-21: Code example of Mamdani type Fuzzy Model evaluation with ‘centroid’ (COA) Defuzzification method (Model 1 of Table 6-5)

6.6 Analysis of Results from Research Prototype

This section presents an analysis of results following the evaluation or execution of various fuzzy models of this research prototype. All output of analysis have been assessed and classified as “OK” or “NOK” by comparing the actual output of analysis and the expected output as defined per class in Table 6-2. In particular, if the actual output of analysis is within the range of class expected minimum and maximum, then the result is considered as “OK”. Else, the result is considered as “NOK”.

As it is shown in Table 6-6, all evaluation data of classes A, B, C and D are indicated as “OK” (100%) from all models. In regards to evaluation data belonging to classes other than A, B, C or D, it is shown that Model 2 and Model 3 have better results with 99.38% evaluation data classified as “OK”. In addition, the 92.5% of evaluation data indicated as “OK” with Model 6. While, the 90.63% of evaluation data indicated as “OK” with Model 1.

Checking Table 6-7, it is observed that both models (1 and 6) have most of evaluation data as “NOK” in class “*P3 NOT Class B and NOT Class C (with P3 neither fully MF_1 nor fully MF_2)*”. Further analysing the results for this, it is shown that with Model 1, the 14 instances of vector input data indicated as “NOK” for this class, has output very close to the expected MIN of this class (the average percentage difference of Model 1 output from the expected MIN of the class is 0.04%). Similarly for Model 6, the data indicated as “NOK” are 11 and the average percentage difference of Model 6 output from expected MIN of the class is 0.18%. The difference is considered as marginal for both models. Although bigger percentage of data are indicated as “NOK” (not within the range) for Model 1 (35.00%) compared to Model 6 (27.50%), the average percentage difference of Model 1 output from expected MIN of the class (0.04%) is less than the average percentage difference of Model 6 output from expected MIN of the class (0.18%).

Table 6-6: Output results summary – Group of Classes

Class		MAMDANI - CENTROID	MAMDANI - Bisector	MAMDANI MOM	MAMDANI SOM	MAMDANI LOM	SUGENO
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Total (Class A+B+C+D)	*OK*	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
	NOK	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Total (Other - Not Class A, B, C or D)	*OK*	90.63%	99.38%	99.38%	31.25%	71.88%	92.50%
	NOK	9.38%	0.63%	0.63%	68.75%	28.13%	7.50%

Table 6-7: Output results – Other Classes - Not Class A, B, C or D

Class		MAMDANI - CENTROID	MAMDANI - Bisector	MAMDANI MOM	MAMDANI SOM	MAMDANI LOM	SUGENO
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
P3 NOT Class A and NOT Class C (with P3 neither fully MF_1 nor fully MF_2)	*OK*	100.00%	100.00%	100.00%	37.50%	77.50%	100.00%
	NOK	0.00%	0.00%	0.00%	62.50%	22.50%	0.00%
P3 NOT Class B and NOT Class C (with P3 neither fully MF_1 nor fully MF_2)	*OK*	65.00%	100.00%	100.00%	7.50%	92.50%	72.50%
	NOK	35.00%	0.00%	0.00%	92.50%	7.50%	27.50%
P3 NOT Class C and NOT Class D (with P3 neither fully MF_2 nor fully MF_3)	*OK*	100.00%	100.00%	100.00%	80.00%	20.00%	100.00%
	NOK	0.00%	0.00%	0.00%	20.00%	80.00%	0.00%
P4 NOT Class A and NOT Class B (with P4 neither fully MF_1 nor fully MF_2)	*OK*	97.50%	97.50%	97.50%	40.00%	57.50%	97.50%
	NOK	2.50%	2.50%	2.50%	60.00%	42.50%	2.50%
P4 NOT Class C and NOT Class D (with P4 neither fully MF_2 nor fully MF_3)	*OK*	100.00%	100.00%	100.00%	0.00%	100.00%	100.00%
	NOK	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%

Another observation of the analysis results for Class A, B, C and D is that Model 3 has output results equals to the *midrange* value of the class (e.g. for Class D the midrange is 92.5 and the output is 92.5). This could be justified since Model 3 uses *mean value of maximum (MOM)* defuzzification method. Similarly, it is observed that the output for Model 4 and Model 5 is equal to the Expected MIN and to the Expected MAX of the class respectively. This could be explained considering that Model 4 uses the defuzzification method *smallest (absolute) value of maximum (SOM)* and the Model 5 the defuzzification method *largest (absolute) value of maximum (LOM)*. Finally, it is also observed that the output of Model 2 is almost always constant per class for classes A, B, C and D.

In addition, the analysis shows that the average Sugeno fuzzy model output results are relative close to the midrange of the classes A, B, C and D above (see Table 6-8).

Table 6-8: Output results – Percentage (%) difference (Model 1 and 6)

Class				Percentage (%) difference between Model 1 and Model 6	Percentage (%) difference between Midrange and Average output of Model 6	Percentage (%) difference between Midrange and Average output of Model 1
Name	Min	Max	Midrange			
Class A	0.00	15.00	7.50	21.97%	7.11%	28.97%
Class B	25.00	45.00	35.00	0.16%	0.09%	0.07%
Class C	55.00	75.00	65.00	0.08%	0.08%	0.00%
Class D	85.00	100.00	92.50	2.14%	0.54%	2.69%
P3 NOT Class A and NOT Class C (with P3 neither fully MF_1 nor fully MF_2)	7.50	65.00	36.25	25.54%	59.97%	35.79%
P3 NOT Class B and NOT Class C (with P3 neither fully MF_1 nor fully MF_2)	35.00	65.00	50.00	0.42%	28.60%	28.20%
P3 NOT Class C and NOT Class D (with P3 neither fully MF_2 nor fully MF_3)	65.00	92.50	78.75	4.41%	7.87%	3.46%
P4 NOT Class A and NOT Class B (with P4 neither fully MF_1 nor fully MF_2)	7.50	35.00	21.25	12.05%	3.38%	8.68%
P4 NOT Class C and NOT Class D (with P4 neither fully MF_2 nor fully MF_3)	65.00	92.50	78.75	3.23%	13.62%	16.83%

In general, the selection of defuzzification method is context and problem dependent (Ross 2010). As stated in section 3.6.2, Hellendoorn and Thomas (1993) (as cited in Ross 2010) have defined five criteria against which to measure defuzzification methods. These are *continuity, disambiguity, plausibility, computational simplicity,* and *weighting method.*

In this generic research prototype with the specific not real evaluation data, classes, fuzzy models, etc., it is shown that all models have outputs within the range for classes A, B, C and D. For other classes (Not Class A, B, C or D), in total, Model 2 and 3 have more output results within the expected range (Table 6-6). However, as discussed before, although Model 1 and Model 6 have some outputs results as "NOK", they are very close to expected MIN (difference is considered as marginal) and hence they could be considered within the range. In addition, the output of Models 1 and 6 (for classes A, B, C or D) is not constant (see description of *disambiguity* in section 3.6.2) like the behaviour mentioned above for Models 2, 3, 4 and 5 and the output is adjusted depending on the input data.

Finally, the results and estimation of outputs performed based on the defined fuzzy models of this generic research prototype and with the constraints mentioned in section 6.1. Any change or tuning of the fuzzy models (e.g. membership functions, rules, etc.) or data could also change some output results. Therefore, some of the results indicated as "NOK" might be due to the need of fuzzy model tuning.

According to Ross (2010, p. 111) *"as with many issues in fuzzy logic, the method of defuzzification should be assessed in terms of the goodness of the answer in the context of the data available."*

In general, Mamdani-type fuzzy models enable to express better human knowledge by also expressing the output with fuzzy sets. Hence, it is easier for the experts to express knowledge and this is the reason of the widely acceptance of this method for decision-making applications using fuzzy logic (Hamam and Georganas 2008; Sivanandam et al. 2007). Sugeno-type fuzzy models are considered more computationally efficient using weighted average (Jassbi et al. 2006; Sivanandam et al. 2007). In the fuzzy inference presented in this chapter, no performance issues identified. However, it is noted that the data vector is small. Possibly, for bigger data vectors and for cases that are more complex this computation efficiency might be concerned. Finally, Sugeno model can be used by adaptive techniques for optimisation of fuzzy model, which best models the data (e.g. ANFIS) (Hamam and Georganas 2008; Kaur and Kaur 2012; Sivanandam et al. 2007).

6.7 Summary

This chapter examined fuzzy modelling and reasoning. This performed with literature research and the development of generic research prototype. The research prototype consists of six (6) MISO fuzzy inference systems as a prototype. Five fuzzy models are of Mamdani type and one is of Sugeno type. In addition, variations of the five fuzzy inference systems of Mamdani type examined with different defuzzification methods. Information about decisions and constraints of this research prototype is described in section 6.1. During the assessment, various aspects examined such as Linguistic fuzzy modelling and interpretability as presented by Gacto et al. (2011). In addition, an analysis of results performed based on the outputs of various models. Mamdani is considered easier for the experts to express knowledge with linguistic terms and fuzzy sets. This is also supported by the bibliography since it is a widely accepted method for decision-making applications using fuzzy logic (Hamam and Georganas 2008; Sivanandam et al. 2007). The output of Mamdani models can be expressed with fuzzy sets and using linguistic terms. On the contrary, the membership functions of output of Sugeno fuzzy model are defined as function of input variables. In this prototype, the output of Sugeno fuzzy model is constant (zero-order Sugeno model) and the membership functions defined based on the corresponding membership functions of output of the Mamdani fuzzy inference system (using the centroid). The output of Sugeno fuzzy model is a function of the input so it is considered that it not easy to estimate it. On the other hand, Sugeno fuzzy models are more computational efficient according to the literature research (Hamam and Georganas 2008; Jassbi et al. 2006; Sivanandam et al. 2007). Nevertheless, no performance issues identified for Mamdani fuzzy modes of this particular research prototype. However, it is noted that the data vector is small. Possibly, for bigger data vectors and for cases that are more complex this computation efficiency might be concerned. Finally, adaptive techniques can be used for constructing or optimising Sugeno-type fuzzy models, which best models the data (Sivanandam et al. 2007). ANFIS technique could be used for optimising Sugeno model. This technique is discussed in section 3.7 but it has not been examined in the context of this prototype.

Chapter 7. Conclusions

Previous chapters presented how this research conducted and what activities performed. The purpose of this chapter is to provide some conclusions from this research. In addition, it examines whether this research achieved to accomplish the research objectives. Finally, it provides some ideas for future work.

7.1 General Conclusions

An approach presented in this thesis for fuzzy knowledge-based approach to risk analysis and in particular for the analysis and detection of the risk of a *physical entity* by utilising fuzzy reasoning and semantic modelling. Particularly, this was examined for the Customs domain. As mentioned, importance of human knowledge and modelling of knowledge and semantics in the context of risk analysis for this domain are some of the motivations for this thesis. This risk analysis would support decision-making for further treatment and actions accordingly.

A bibliographic research performed under this thesis showing that there are several researches and works conducted for fraud detection systems and risk analysis in various areas. Several techniques have been used including fuzzy logic. Some of the works refers to fraud detection and risk analysis in Customs domain. Those works mainly examine the application of Neural Networks, Fuzzy Logic, Data Mining, Outlier detection and Statistical methods as detection techniques.

As presented also from the literature review, the *risk* is closely related to *uncertainty*. Fuzzy logic and fuzzy sets can be used for modelling *uncertainty* related to *imprecise information* and *vagueness*.

As extensively discussed in previous chapters, fuzzy modelling and fuzzy reasoning can consider imprecise knowledge and vagueness. Fuzzy logic is a technique, which has been applied in various fields and is used with success for decision-making and inference purposes. Fuzzy models and fuzzy inference systems could be used for expressing human knowledge using linguistic terms. As mentioned before, human knowledge and expertise are very important for that process. Therefore, it is considered that fuzzy logic could be used for supporting risk analysis by using linguistic terms. Bibliographic research shows that fuzzy logic has been examined for supporting decision-making for risk analysis and detection in customs (Singh and Sahu 2004; Singh et al. 2003). In the work of Singh and Sahu (2004) it is stated among others that the proposed system using fuzzy logic considers the human intelligence and also suggestions of the systems are more closely to decision-making ability of customs officers.

Another important element of the fuzzy knowledge-based approach presented in this thesis is the use of ontologies for semantic modelling with purpose to improve the communication, understanding, and interoperability in the context of risk analysis. As mentioned at the beginning of the thesis, when decision-making is complex including assessment of many parameters, deep knowledge is required. Therefore, semantic knowledge should be acquired, which needs deep understanding of various concepts and their relationships in the domain. Consequently, it is considered that semantic modelling can enable the unambiguous definition of concepts and the modelling of complex relationships. The role of ontologies in this research is the one mentioned previously, i.e. for representing semantics and modelling complex relationships related to risk analysis. In addition, ontologies are integrated with fuzzy rule-based reasoning. This would improve the communication and understanding. As described previously, semantic modelling is one of the elements of the presented approach. The following could be summarised for the presented work in this area:

- ✓ an architecture of ontologies assists to define ontologies with concepts at various levels and enables modularity, maintainability, re-usability and extensibility, especially for complex domains such as Customs.

- ✓ Ontologies that are more generic can model high-level or abstract concepts. Ontologies that are more specific have particular purpose. For instance, “Physical Entity Ontology” models the semantics and attributes of the particular *physical entity* enabling better understanding of the *physical entity’s* attributes. On the other hand, the “Fuzzy Risk Model Ontology” defines specific concepts of *fuzzy risk model* for risk analysis with fuzzy logic technique.
- ✓ decomposition of Ontologies is considered as a matter of decision for organising the ontologies. However, it is considered that the “Generic Customs Ontology” presented above could be further decomposed or integrated with other upper Ontologies and/or middle-level Ontologies for re-using existing concepts from existing published ontologies.
- ✓ representation of concepts at various levels illustrated with examples (e.g. domain concepts or concepts related to “Fuzzy Risk Model Ontology” based on the principles of the fuzzy logic). As a summary, this enables the clear definition of concepts:
  - hierarchical structure of concepts;
  - modelling of equality between concepts using equivalent classes (modelling same concepts (synonymous) with different names);
  - defining enumeration in data properties and restricting the allowed data (e.g. fuzzy operator data property);
  - using object properties for visualizing the various complex relationships between concepts;
  - defining property restrictions (e.g. cardinality restrictions) through object properties in order to enrich the definition of relationships;

  - using value partition (e.g. example of Weigh Risk defined as the union of “High Risk”, “Medium Risk” and “Low Risk”).

- ✓ loose integration of “Fuzzy Risk Model Ontology” with fuzzy rule-based reasoning.

- ✓ Finally, ontology evaluation is very important. Ontology evaluation is also discussed in section 5.4. As mentioned, the ontologies of this work have not been validated from any domain expert, official body or any organisation. The development of ontologies is for research purposes and for exploring their benefits for communication, common understanding, and interoperability in complex domains as mentioned previously. Besides, the Ontologies must always be enriched which implies that the continuous evaluation and formal validation of the Ontologies is required. However, it is considered that similar OWL components and approaches would be used to model other concepts and their relationships.

The second main element of the presented concept is the use of fuzzy modelling and fuzzy reasoning. The fuzzy modelling and fuzzy reasoning are examined with bibliographic research on the domain and on the application of fuzzy logic. In addition, they examined with a research prototype. Information about decisions and constraints of this research prototype is described in Chapter 6. Following fuzzy modelling and fuzzy reasoning investigation, some main points are highlighted in the next paragraphs:

- ✓ It has been extensively discussed that HHFC of HFS techniques can be used in complex domains. In such case, there are many variables and hence the number of required rules is increased exponentially. Therefore, this depends on the number of variables of the particular risk analysis of the *physical entity*. More modular *fuzzy risk models* can be defined with HHFC/HFS approach. However, their relationships and execution as well as other parameters should be defined as part of the risk analysis definition of a particular *physical entity*. This is also discussed in section 4.1.2.

- ✓ Principles and techniques for defining a fuzzy model have been discussed and should be considered in fuzzy modelling activity. Linguistic fuzzy modelling (LFM) and interpretability (Gacto et al. 2011) discussed in Chapter 6. This examined in relation to complexity at the level of fuzzy partitions, semantics interpretability at fuzzy partition and complexity at the level of rule base. Therefore, it is believed that the interpretability of linguistic fuzzy rule-based systems as discussed in the literature should also be considered in the definition of fuzzy models apart from accuracy,
- ✓ Following the assessment of six MISO fuzzy inference systems developed in the context of prototype (Chapter 6), it is considered that a Human can express easier the fuzzy rules or define linguistic terms for both input and output with Mamdani type. Human knowledge is important in the risk analysis activity as described in the motivations of this thesis. On the contrary, the output (constant) membership functions of Sugeno fuzzy model is defined in this research prototype based on the corresponding membership functions of output of the Mamdani fuzzy inference system (using the centroid). This happened because relationship between input and output should be known as a function for defining Sugeno fuzzy model outputs.
- ✓ On the other hand, Sugeno fuzzy models are more computational efficient according to the literature research described in section 3.6.2. Nevertheless, no performance issues identified for fuzzy models of this particular research prototype. However, it is noted that the data vector is small. Possibly, for bigger data vectors and for cases that are more complex this computation efficiency might be concerned.

- ✓ Changing environments require adaptive techniques (Torra 2001). This is also presented in the approach of Chapter 4 with the Assistance/Optimisation. ANFIS technique could be used for optimising Sugeno model. This technique is also discussed in section 3.7 but it has not been examined in the context of this prototype. Nevertheless, this technique could be examined and applied by selecting the appropriate parameters and by considering issues such as the *curse of dimensionality* (Wei et al. 2007).

- ✓ From the assessment of six MISO fuzzy inference systems developed in the context of this research prototype (Chapter 6), it shown that all models estimate the output within the range for classes A, B, C and D. For other classes (Not Class A, B, C or D), Model 2 and 3 have more output results within the expected range (Table 6-6). However, as stated during the analysis, the output results of Model 1 and Model 6 indicated as "NOK" are very close to expected MIN (difference is considered as marginal). Therefore, they could be considered within the range. Another observation is that the output of Models 1 and 6 (for classes A, B, C or D) is not constant (see description of *disambiguity* in section 3.6.2) and hence there is no ambiguity in the output value and it is adjusted depending on the input data. This is in contrast to the behaviour mentioned above for Models 2, 3, 4 and 5. Therefore, it could be considered that Mamdani with Centroid defuzzification method (Model 1) and Sugeno (Model 6) have better outputs (more output results within the expected range) and disambiguity in the results. This includes the classes other than A, B, C or D, which it is considered that have some fuzziness.

- ✓ Considering the constraints of this research discussed in Chapter 6, it is deemed that the application of fuzzy modelling and reasoning on real scenarios need further analysis and investigation.

Finally, it is considered that Risk management is broader, multi-dimensional process involving a number of task, activities, and practises. The presented approach is focused on the analysis and detection of the risk for a *physical entity* based on the outputs of the risk management process. Therefore, the discussion focuses on this and it might be possible this approach to be combined or complemented with other approaches or techniques if necessary and following assessment.

7.2 Future Work

The Optimisation/Assistance has been discussed in the approach presented in Chapter 4. The ANFIS technique is mentioned in previous sections as a technique for constructing or optimising Sugeno-type fuzzy inference systems. Hence, it will be interesting to examine deeper the use of this technique in a future research for the purpose of risk analysis.

Due to the constraints mentioned in Chapter 6, it is considered that fuzzy modelling and fuzzy reasoning needs further analysis and investigation with possibly real scenarios in order to be able to have more results for analysis and evaluation.

References

Acampora, G. and Loia, V. 2005. Using FML and Fuzzy Technology in Adaptive Ambient Intelligence Environments. *International Journal of Computational Intelligence Research* 1(2), pp. 171–182.

ADONIS. 2014.*ADONIS Community Edition 3.0* [Online]. Place: BOC Group. Available at: http://www.adonis-community.com/ [Accessed: 3/2/2014]

Alavala, C. R. 2008. *Fuzzy Logic and Neural Networks: Basic Concepts and Applications*. New Age International Publishers.

Albarrak, K. M. and Sibley, E. H. eds. 2011. *A survey of methods that transform data models into ontology models*. Information Reuse and Integration (IRI), 2011 IEEE International Conference on. 3-5 Aug. 2011.

Antoniou, G. and Harmelen, F. v. 2004. *A Semantic Web Primer*. London, UK: MIT press.

Awad, M. E. 1996. *Building Expert Systems: Principles, Procedures and Applications*. United States of America: West Publishing Company.

Bai, Y. and Wang, D. 2006. Fundamentals of Fuzzy Logic Control — Fuzzy Sets, Fuzzy Rules and Defuzzifications. In: Bai, Y., Zhuang, H. and Wang, D. eds. *Advanced Fuzzy Logic Technologies in Industrial Applications*. Springer London, pp. 17-36.

Bobillo, F. and Straccia, U. eds. 2008. *fuzzyDL: An expressive fuzzy description logic reasoner*. Fuzzy Systems, 2008. FUZZ-IEEE 2008. (IEEE World Congress on Computational Intelligence). IEEE International Conference on. 1-6 June 2008.

Bobillo, F. and Straccia, U. 2009. An OWL Ontology for Fuzzy OWL 2. In: Rauch, J., Ras, Z., Berka, P. and Elomaa, T. eds. *Foundations of Intelligent Systems*. Vol. 5722. Springer Berlin / Heidelberg, pp. 151-160.

Bobillo, F. and Straccia, U. 2011. Fuzzy ontology representation using OWL 2. *International Journal of Approximate Reasoning* 52(7), pp. 1073-1094.

Bolton, R. J. and Hand, D. J. 2002. Statistical Fraud Detection: A Review. *Statistical Science* 17(3), pp. 235-255.

Bragaglia, S., Chesani, F., Ciampolini, A., Mello, P., Montali, M. and Sottara, D. 2010. An Hybrid Architecture Integrating Forward Rules with Fuzzy Ontological Reasoning. In: Graña Romay, M., Corchado, E. and Garcia Sebastian, M.T. eds. *Hybrid Artificial Intelligence Systems*. Vol. 6076. Springer Berlin Heidelberg, pp. 438-445.

Braglia, M., Frosolini, M. and Montanari, R. 2003. Fuzzy criticality assessment model for failure modes and effect analysis. *International Journal of Quality & Reliability Management* 20(4), pp. 503-524.

Breiman, L. 1994. *Bagging Predictors*. University of California.

Breiman, L. 1996. Bagging Predictors. *Machine Learning* 24(2), pp. 123-140.

Cameron, E. and Peloso, G. F. 2005. *Risk Management and the Precautionary Principle: A fuzzy logic model.* Society for Risk Analysis. pp. 901-911.

Chiaberge, M., Di Bene, G., Di Pascoli, S., Lazzerini, B., Maggiore, A. and Reyneri, L. M. eds. 1995. *Mixing fuzzy, neural and genetic algorithms in an integrated design environment for intelligent controllers*. Systems, Man and Cybernetics, 1995. Intelligent Systems for the 21st Century., IEEE International Conference on. 22-25 Oct 1995.

Dahal, K., Hussain, Z. and Hossain, M. A. eds. 2005. *Loan risk analyzer based on fuzzy logic*. e-Technology, e-Commerce and e-Service, 2005. EEE '05. Proceedings. The 2005 IEEE International Conference on. 29 March-1 April 2005.

Davies, J., Fensel, D. and Harmelen, F. v. eds. 2003. *Towards the Semantic Web - Ontology-driven Knowledge Management.* West Sussex, UK: John Wiley & Sons Inc., p. 288.

de Ru, W. G. and Eloff, H. P. 1996. Risk analysis modelling with the use of fuzzy logic. *Computers & Security* 15(3), pp. 239-248.

Deshmukh, A. and Talluru, T. L. N. eds. 1997. *A rule based fuzzy reasoning system for assessing the risk of management fraud*. Systems, Man, and Cybernetics, 1997. 'Computational Cybernetics and Simulation'., 1997 IEEE International Conference on. 12-15 Oct 1997.

DGTAXUD. 2004. Standarised Framework for Risk Management in the Customs Administrations of the EU. *EUROPA- Taxation and Customs Union* [Online]. Available at: http://ec.europa.eu/taxation_customs/resources/documents/framework_doc.pdf [Accessed: 16/05/2013].

DGTAXUD. 2010.*Customs Glossary* [Online]. Place: Available at: http://ec.europa.eu/taxation_customs/common/glossary/customs/index_en.htm [Accessed: 30/08/2013]

DGTAXUD. 2013.*Introduction to Risk Management* [Online]. Place: Available at: http://ec.europa.eu/taxation_customs/customs/customs_controls/risk_management/index_en.htm [Accessed: 08/06/2013]

DGTAXUD. 2014a.*Customs Controls - General* [Online]. Place: Available at: http://ec.europa.eu/taxation_customs/customs/customs_controls/general/index_en.htm [Accessed: 10/04/2014]

DGTAXUD. 2014b.*Customs declaration* [Online]. Place: Available at: http://ec.europa.eu/taxation_customs/customs/procedural_aspects/general/declaration/index_en.htm [Accessed: 10/04/2014]

DGTAXUD. 2014c.*Customs legislation* [Online]. Place: Available at: http://ec.europa.eu/taxation_customs/common/legislation/legislation/customs/index_en.htm [Accessed: 09/04/2014]

DGTAXUD. 2014d.*Electronic customs* [Online]. Place: Available at: http://ec.europa.eu/taxation_customs/customs/policy_issues/electronic_customs_initiative/index_en.htm [Accessed: 09/04/2014]

DGTAXUD. 2014e.*EU Customs strategy* [Online]. Place: Available at: http://ec.europa.eu/taxation_customs/customs/policy_issues/customs_strategy/index_en.htm [Accessed: 7/4/2014]

DGTAXUD. 2014f.*EU Customs Union – unique in the world* [Online]. Place: Available at: http://ec.europa.eu/taxation_customs/customs/policy_issues/facts_and_figures/eu_customs_union_unique_en.htm [Accessed: 09/04/2014]

Digiampietri, L. A., Roman, N. T., Meira, L. A. A., Filho, J. J., Ferreira, C. D., Kondo, A. A., Constantino, E. R., Rezende, R. C., Brandao, B. C., Ribeiro, H. S., Carolino, P. K., Lanna, A., Wainer, J. and Goldenstein, S. 2008. Uses of artificial intelligence in the Brazilian customs fraud detection system. In: *Proceedings of the 2008 international conference on Digital government research.* Montreal, Canada. 1367864: Digital Government Society of North America, pp. 181-187.

Dimakopoulos, T. and Kassis, K. 2008. D6. Domain specific ontology - Risk Assessment for Customs in Western Balkans. RACWeB - IST PROJECT 045101.

Dorronsoro, J. R., Ginel, F., Sgnchez, C. and Cruz, C. S. 1997. Neural fraud detection in credit card operations. *Neural Networks, IEEE Transactions on* 8(4), pp. 827-834.

Dubois, D. and Prade, H. 2001. Possibility Theory, Probability Theory and Multiple-Valued Logics: A Clarification. *Annals of Mathematics and Artificial Intelligence* 32(1-4), pp. 35-66.

EEC 1992. Council Regulation (EEC) No 2913/92 of 12 October 1992 establishing the Community Customs Code. Official Journal. p. 88.

EEC 1993. Commission Regulation (EEC) No 2454/93 of 2 July 1993 laying down provisions for the implementation of Council Regulation (EEC) No 2913/92 establishing the Community Customs Code. p. 779.

EEC 2008a. Commission Regulation (EC) No 1192/2008 of 17 November 2008 amending Regulation (EEC) No 2454/93 laying down provisions for the implementation of Council Regulation (EEC) No 2913/92 establishing the Community Customs Code. Official Journal. p. 51.

EEC 2008b. Regulation (EC) No 450/2008 of the European Parliament and of the Council of 23 April 2008 laying down the Community Customs Code (Modernised Customs Code). Official Journal. p. 64.

EUROSTAT 2005. Geonomenclature. p. 108.

Falconer, S. 2010.*OntoGraf Protege Plugin* [Online]. Place: Available at: http://protegewiki.stanford.edu/wiki/OntoGraf [Accessed: 21/03/2014]

Farvaresh, H. and Sepehri, M. M. 2011. A data mining framework for detecting subscription fraud in telecommunication. *Engineering Applications of Artificial Intelligence* 24(1), pp. 182-194.

Fawcett, T. and Provost, F. 1997. Adaptive Fraud Detection. *Data Min. Knowl. Discov.* 1(3), pp. 291-316.

Feng, Y., Gengui, Z. and Jinqiu, L. eds. 2007. *A BP Neural Network Approach on Risk Evaluation of Ventures in China Customs*. Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007. International Conference on. 21-25 Sept. 2007.

Friedlob, G. T. and Schleifer, L. L. F. 1999. Fuzzy logic: application for audit risk and uncertainty. *Managerial Auditing Journal*, pp. 127-135.

Frosdick, S. 1997. The techniques of risk analysis are insufficient in themselves. *Disaster Prevention and Management* 6(3), pp. 165 - 177.

Güneri, A. F., Ertay, T. and Yücel, A. 2011. An approach based on ANFIS input selection and modeling for supplier selection problem. *Expert Systems with Applications* 38(12), pp. 14907-14917.

Gacto, M. J., Alcalá, R. and Herrera, F. 2011. Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures. *Information Sciences* 181(20), pp. 4340-4360.

Garcia, P. A. A., Schirru, R., Frutuoso, E. and Melo, P. F. 2005. A fuzzy data envelopment analysis approach for FMEA. *Progress in Nuclear Energy* 46(3–4), pp. 359-373.

Gargama, H. and Chaturvedi, S. K. 2011. Criticality Assessment Models for Failure Mode Effects and Criticality Analysis Using Fuzzy Logic. *Reliability, IEEE Transactions on* 60(1), pp. 102-110.

Geourjon, A.-M., Graziosi, G. R. and Laporte, B. 2010. How to modernize risk analysis and the selectivity of Customs controls in developing countries? *WCO News* [Online] (62). Available at: http://wcoomdpublications.org/downloadable/download/sample/sample_id/91/ [Accessed: 02/01/2012].

Geourjon, A.-M. and Laporte, B. 2005. Risk management for targeting customs controls in developing countries: A risky venture for revenue performance? *Public Administration and Development* 25(2), pp. 105-113.

Gomez-Perez, A. 1995. Some ideas and examples to evaluate ontologies. In: *Proceedings of the 11th Conference on Artificial Intelligence for Applications.* IEEE Computer Society,

Gruber, T. R. 1993. A translation approach to portable ontology specifications. *Knowledge Acquisition* 5(2), pp. 199-220.

Gruber, T. R. 1995. Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies* 43(5–6), pp. 907-928.

Guillaume, S. and Charnomordic, B. 2012. Fuzzy inference systems: An integrated modeling environment for collaboration between expert knowledge and data using FisPro. *Expert Systems with Applications* 39(10), pp. 8744-8755.

Hamam, A. and Georganas, N. D. eds. 2008. *A comparison of Mamdani and Sugeno fuzzy inference systems for evaluating the quality of experience of Hapto-Audio-Visual applications*. Haptic Audio visual Environments and Games, 2008. HAVE 2008. IEEE International Workshop on. 18-19 Oct. 2008.

Hellendoorn, H. and Thomas, C. 1993. Defuzzification in fuzzy controllers. *Journal of Intelligent and Fuzzy Systems*.

Hilas, C. S. 2009. Designing an expert system for fraud detection in private telecommunications networks. *Expert Systems with Applications* 36(9), pp. 11559-11569.

Hodge, V. and Austin, J. 2004. A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review* 22(2), pp. 85-126.

Horácek, P. and Binder, Z. 1997. Hierarchical fuzzy controllers. *Annual Reviews in Control* 21, pp. 93-101.

Horrocks, I. and Patel-Schneider, P. 2004. Reducing OWL entailment to description logic satisfiability. *Web Semantics: Science, Services and Agents on the World Wide Web* 1(4), pp. 345-357.

Huang, H.-D., Acampora, G., Loia, V., Lee, C.-S. and Kao, H.-Y. eds. 2011. *Applying FML and Fuzzy Ontologies to Malware Behavioural Analysis*. Fuzzy Systems (FUZZ), 2011 IEEE International Conference on. 27-30 June 2011.

Iancu, I. 2012. A Mamdani Type Fuzzy Logic Controller. In: Dadios, E.P. ed. *FUZZY LOGIC – CONTROLS, CONCEPTS, THEORIES AND APPLICATIONS*. InTech.

IEEE 2001. IEEE Standard for Software Life Cycle Processes - Risk Management.*IEEE Std 1540-2001*. pp. i-24.

Iordache, E. and Voiculet, A. V. 2007. Customs Risk Management in the European Union. *The Romanian Economic Journal* [Online] 25. Available at: http://www.rejournal.eu/Portals/0/Arhiva/JE%2025/JE%2025%20Iordache%20Voiculet.pdf.

Ishibuchi, H., Murata, T. and Gen, M. 1998. Performance evaluation of fuzzy rule-based classification systems obtained by multi-objective genetic algorithms. *Computers & Industrial Engineering* 35(3–4), pp. 575-578.

ISO. 2006.*ISO 3166-1-alpha-2 code* [Online]. Place: Available at: http://www.iso.org/iso/english_country_names_and_code_elements [Accessed: 28/12/2008]

ISO 2009. ISO 31010:2009 - Risk management -- Risk assessment techniques. International Organisation for Standardisation (ISO).

Jang, J.-S. R. 1993. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *Systems, Man and Cybernetics, IEEE Transactions on* 23(3), pp. 665-685.

Jassbi, J. J., Serra, P. J. A., Ribeiro, R. A. and Donati, A. eds. 2006. *A Comparison of Mandani and Sugeno Inference Systems for a Space Fault Detection Application*. Automation Congress, 2006. WAC '06. World. 24-26 July 2006.

Jianhong, L. and Dezhao, C. eds. 2008. *Mapping Rules Based Data Mining for Effective Decision Support Application*. Business and Information Management, 2008. ISBIM '08. International Seminar on. 19-19 Dec. 2008.

Jinting, Y., Yinghui, L. and Zilai, S. eds. 2009. *Research and Control of Audit Risk Based on Fuzzy Comprehensive Evaluation*. Management and Service Science, 2009. MASS '09. International Conference on. 20-22 Sept. 2009.

Jupp, S., Moulton, G., Rector, A., Stevens, R. and Wroe, C. 2007. A Practical Guide To Building OWL Ontologies Using Protege 4 and CO-ODE Tools. The University Of Manchester.

Jurgutis, A. and Simutis, R. eds. 2011. *An Investor Risk Profiling Using Fuzzy Logic-based Approach in Multi-Agents Decision Support System*. 17th International Conference on Information and Software Technologies - IT 2011. Kaunas, Lithuania. Kaunas University of Technology.

Kai Meng, T. ed. 2009. *On Fuzzy Inference System Based Failure Mode and Effect Analysis (FMEA) Methodology*. Soft Computing and Pattern Recognition, 2009. SOCPAR '09. International Conference of. 4-7 Dec. 2009.

Kaur, A. and Kaur, A. 2012. Comparison of Mamdani-Type and Sugeno-Type Fuzzy Inference Systems for Air Conditioning System. *International Journal of Soft Computing and Engineering (IJSCE)* 2(2), pp. 323-325.

Khanmohammadi, S. and Jassbi, J. 2012. A Fuzzy Approach for Risk Analysis with Application in Project Management. In: Azeem, M.F. ed. *FUZZY INFERENCE SYSTEM – THEORY AND APPLICATIONS*. InTech.

Klir, G. J. and Yuan, B. 1995. *Fuzzy Sets and Fuzzy Logic - Theory and Application*. Prentice Hall PTR.

Kosko, B. 1992. *Neural Networks and Fuzzy Systems - A dynamical systems approach to machine intelligence*. Prentice Hall.

Laleh, N. and Azgomi, M. A. 2009. A Taxonomy of Frauds and Fraud Detection Techniques. In: Prasad, S.K., Routray, S., Khurana, R. and Sahni, S. eds. *Information Systems, Technology and Management*. Vol. 31. Springer Berlin Heidelberg, pp. 256-267.

Laporte, B. 2011. Risk management systems: using data mining in developing countries' customs administrations. *WCO Customs Journal* [Online] 5(1). Available at: http://www.worldcustomsjournal.org/media/wcj/-2011/1/Laporte.pdf [Accessed: 06/01/2012].

Larose, D. T. 2005. *Discovering knowledge in data: An Introduction to Data Mining*. John Wiley & Sons, Inc.

Lee, C.-S., Wang, M.-H., Acampora, G., Loia, V. and Hsu, C.-Y. eds. 2009. *Ontology-based intelligent fuzzy agent for diabetes application*. Intelligent Agents, 2009. IA '09. IEEE Symposium on. March 30 2009-April 2 2009.

Liu, J., Tan, Y.-H. and Hulstijn, J. 2009. IT Enabled Risk Management for Taxation and Customs: The Case of AEO Assessment in the Netherlands. In: *Proceedings of the 8th International Conference on Electronic Government.* Linz, Austria. 1617628: Springer-Verlag, pp. 376-387.

Liu, K. F.-R. and Yu, C.-W. 2009. Integrating case-based and fuzzy reasoning to qualitatively predict risk in an environmental impact assessment review. *Environ. Model. Softw.* 24(10), pp. 1241-1251.

Loia, V. 2011. Fuzzy Ontologies and Fuzzy Markup Language: A Novel Vision in Web Intelligence. In: Mugellini, E., Szczepaniak, P., Pettenati, M. and Sokhn, M. eds. *Advances in Intelligent Web Mastering – 3*. Vol. 86. Springer Berlin Heidelberg, pp. 3-10.

Lukasiewicz, T. and Straccia, U. 2008. Managing uncertainty and vagueness in description logics for the Semantic Web. *Web Semant.* 6(4), pp. 291-308.

Marhavilas, P. K., Koulouriotis, D. and Gemeni, V. 2011. Risk analysis and assessment methodologies in the work sites: On a review, classification and comparative study of the scientific literature of the period 2000–2009. *Journal of Loss Prevention in the Process Industries* 24(5), pp. 477-523.

Miller, M. ed. 2004. *Probabilistic Risk Analysis and the Concept of Bayesian Networks*. Third Workshop - An Event of German Chapter of System Safety Society. Bielefeld. Center of Interdiciplinary Research (ZiF), University of Bielefeld.

Naaz, S., Alam, A. and Biswas, R. 2011. Effect of different defuzzification methods in a fuzzy based load balancing application. *IJCSI International Journal of Computer Science Issues* 8(5), pp. 261-267.

Negoita, C. V. 1985. *Expert Systems and Fuzzy Systems*. The Benjamin/Cummings Publishing Company, Inc.

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y. and Sun, X. 2011. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decis. Support Syst.* 50(3), pp. 559-569.

Nota, G. ed. 2011. *Risk Management Trends.* InTech.

Obrst, L. 2010. Ontological Architectures - Theory and Applications of Ontology: Computer Applications. In: Poli, R., Healy, M. and Kameas, A. eds. Springer Netherlands, pp. 27-66.

OMG. 2011a.*Business Process Model and Notation (BPMN) v2.0* [Online]. Place: Available at: http://www.omg.org/spec/BPMN/2.0 [Accessed: 04/02/2014]

OMG. 2011b.*Unified Modeling Language™ (UML®) v2.4.1.* [Online]. Place: Available at: http://www.omg.org/spec/UML/2.4.1 [Accessed: 04/02/2014]

OMG. 2013.*Introduction To OMG's Unified Modeling Language™ (UML®)* [Online]. Place: Available at: http://www.omg.org/gettingstarted/what_is_uml.htm [Accessed: 04/02/2014]

Orchard, R. ed. 2001. *Fuzzy Reasoning in Jess: The FuzzyJ Toolkit and FuzzyJess*. Third International Conference on Enterprise Information Systems. Setubal, Portugal, July 7-10.

OWLGrEd. 2013.[Online]. Place: Available at: http://owlgred.lumii.lv/index.html [Accessed: 30/08/2013]

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.

Pejic-Bach, M. ed. 2010. *Invited Paper: Profiling Intelligent Systems Applications in Fraud Detection and Prevention: Survey of Research Articles*. Intelligent Systems, Modelling and Simulation (ISMS), 2010 International Conference on. 27-29 Jan. 2010.

Phua, C., Lee, V., Smith, K. and Gayler, R. 2005. A comprehensive survey of Data Mining-based Fraud Detection Research. *Artificial Intelligence Review*.

Protégé. *Protégé website* [Online]. Place: Available at: http://protege.stanford.edu/ [Accessed: 20/03/2014]

Rao, D. H. and Saraf, S. S. eds. 1996. *Study of defuzzification methods of fuzzy logic controller for speed control of a DC motor*. Power Electronics, Drives and Energy Systems for Industrial Growth, 1996., Proceedings of the 1996 International Conference on. 8-11 Jan 1996.

Roman, N. T., Ferreira, C. D., Meira, L. A. A., Rezende, R., Digiampietri, L. A. and Filho, J. J. 2009. Attribute-value specification in customs fraud detection: a human-aided approach. In: *Proceedings of the 10th Annual International Conference on Digital Government Research: Social Networks: Making Connections between Citizens, Data and Government.* 1556224: Digital Government Society of North America, pp. 264-271.

Ross, T. J. 2010. *Fuzzy Logic with Engineering Applications*. Third ed. Wiley.

Shanks, G., Tansley, E. and Weber, R. 2003. Using ontology to validate conceptual models. *Commun. ACM* 46(10), pp. 85-89.

Shao, H., Zhao, H. and Chang, G.-R. eds. 2002. *Applying Data Mining to detect fraud behavior in Customs declaration*. Proceedings of the First International Conference on Machine Learning and Cybernetics. Beijing, 4-5 November 2002.

Shapiro, A. F. 2004. Fuzzy logic in insurance. *Insurance: Mathematics and Economics* 35(2), pp. 399-424.

Singh, A. K. and Sahu, R. eds. 2004. *Decision Support System for Customs Examination*. Second IEEE International Conference on Intelligent Systems.

Singh, A. K., Sahu, R. and Ujjwal, K. eds. 2003. *Decision Support System in Customs Assessment to Detect Valuation Frauds*. Engineering Management Conference,

2003. IEMC '03. Managing Technologically Driven Organizations: The Human Side of Innovation and Change. 2-4 Nov. 2003.

Sivanandam, S. N., S.Sumathi and Deepa, S. N. 2007. *Introduction to Fuzzy Logic using MATLAB*. Springer-Verlag.

Smith, M. K., Welty, C. and McGuinness, D. L. eds. 10 February 2004. *OWL Web Ontology Language Guide.* W3C Recommendation.

Tay, K. M. and Lim, C. P. 2006. Application of Fuzzy Inference Techniques to FMEA Applied Soft Computing Technologies: The Challenge of Complexity. In: Abraham, A., de Baets, B., Köppen, M. and Nickolay, B. eds., Vol. 34. Springer Berlin / Heidelberg, pp. 161-171.

Tchankova, L. 2002. *Risk Identification - Basic Stage in risk management.* Environmental Management and Health. pp. 290-297.

Torra, V. 2001. Fuzzy Knowledge Based Systems and Chance Discovery. In: Terano, T., Ohsawa, Y., Nishida, T., Namatame, A., Tsumoto, S. and Washio, T. eds. *New Frontiers in Artificial Intelligence*. Vol. 2253. Springer Berlin Heidelberg, pp. 491-495.

Truel, C. 2010. *A Short Guide to Customs Risk*. Gower Publishing Limited.

Uschold, M. and Gruninger, M. 1996. Ontologies: Principles, methods and applications. *Knowledge Engineering Review* 11, pp. 93--136.

Valdez, J. M. G., Sandoval, G. L., Garza, A. A. and Castillo, O. 2007. *Object Oriented Design and Implementation of an Inference Engine for Fuzzy Systems.* Engineering Letters. August 15.

Virtanen, P. and Helander, N. eds. 2010. *Knowledge Management.* In-teh.

W3C. 2001.*XSLT Requirements 2.0* [Online]. Place: Available at: http://www.w3.org/TR/xslt20req [Accessed: 07/03/2014]

W3C. 2004.*SWRL: A Semantic Web Rule Language Combining OWL and RuleML* [Online]. Place: Available at: http://www.w3.org/Submission/SWRL/ [Accessed: 05/03/2014]

W3C. 2012.*OWL 2 Web Ontology Language - Structural Specification and Functional-Style Syntax (Second Edition)* [Online]. Place: Available at: http://www.w3.org/TR/2012/REC-owl2-syntax-20121211/ [Accessed: 18/03/2014]

Wang, Y.-M., Chin, K.-S., Poon, G. K. K. and Yang, J.-B. 2009. Risk evaluation in failure mode and effects analysis using fuzzy weighted geometric mean. *Expert Systems with Applications* 36(2, Part 1), pp. 1195-1207.

WCO. 2010. Background Paper on Risk Management. [Online]. Available at: http://www.wcoomd.org/en/events/event-history/2010/~/media/2BA0488B399D4651BF8F0B6009C8040E.ashx [Accessed: 08/06/2013].

WCO. 2011. WCO Customs Risk Management Compendium. [Online]. Available at: http://www.wcoomd.org/en/topics/enforcement-and-compliance/instruments-and-tools/rmc.aspx [Accessed: 04/09/2013].

Wei, M., Bai, B., Sung, A. H., Liu, Q., Wang, J. and Cather, M. E. 2007. Predicting injection profiles using ANFIS. *Information Sciences* 177(20), pp. 4445-4461.

Wei, X., Ye, P., Jian, M., Shou-Yang, W., Gang, H., Shuo, Z. and Yu-Hua, Q. eds. 2008. *Fraud detection in telecommunication: A rough fuzzy set based approach*. Machine Learning and Cybernetics, 2008 International Conference on. 12-15 July 2008.

Witten, I. H. and Frank, E. 2005. *Data Mining: Practical Attribute Ranking Machine Learning Tools and Techniques*. Morgan Kaufmann Publishers.

Wlodarczyk, T. W., Rong, C., O'Connor, M. and Musen, M. 2011. SWRL-F: a fuzzy logic extension of the semantic web rule language. In: *Proceedings of the International*

Conference on Web Intelligence, Mining and Semantics. Sogndal, Norway. 1988735: ACM, pp. 1-9.

Wolpert, D. H. 1992. Stacked generalization. *Neural Networks* 5, pp. 241--259.

Yaguinuma, C. A., Santos, M. T. P., Camargo, H. A. and Reformat, M. eds. 2013. *A FML-based hybrid reasoner combining fuzzy ontology and Mamdani inference*. Fuzzy Systems (FUZZ), 2013 IEEE International Conference on. Hyderabad, 7-10 July 2013.

Ye, F., Zhou, G. and Lu, J. 2007. The Risk-Evaluation Model in Customs Based on BP Neural Networks. In: *Proceedings of the Third International Conference on Natural Computation - Volume 03.* 1305716: IEEE Computer Society, pp. 377-380.

Yufeng, K., Chang-Tien, L., Sirwongwattana, S. and Yo-Ping, H. eds. 2004. *Survey of fraud detection techniques*. Networking, Sensing and Control, 2004 IEEE International Conference on. 2004.

Zang, B., Li, Y., Xie, W., Chen, Z., Tsai, C.-F. and Laing, C. 2008. An ontological engineering approach for automating inspection and quarantine at airports. *J. Comput. Syst. Sci.* 74(2), pp. 196-210.

Zhu, X. 2008. *Semi-Supervised Learning Literature Survey*. Computer Sciences, University of Wisconsin-Madison.