Defeasible Reasoning and Argumentation Theory for Decision Support in Health Care

Luca Longo

A dissertation submitted to the University of Dublin
in partial fulfillment of the requirements for the degree of
Master of Science in Health Informatics

2012

Declaration

I declare that the work described in this dissertation is, except where otherwise stated, entirely my own work, and has not been submitted as an exercise for a degree at this or any other university.

Signed:

Luca Longo
September, 2012

Permission to lend and/or copy

I agree that the Trinity College Library may lend or copy this dissertation upon request.

Signed:

Luca Longo

September, 2012

Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor Prof. Lucy Hederman for the continuous support of this study and research, for her patience, enthusiasm and motivation. Her guidance helped me in all the time of research and writing of this thesis.

Besides my advisor, I would like to thank Doctor Bridget Kane, for her patience, constructive feedback and her practical vision of this study.

Doctor Pierpaolo Dondio, for his precious insights on defeasible reasoning and our artificial intelligence-based lunches at the Science Gallery in Trinity College Dublin.

Last but not the least, my family for the constant, unconditional and unlimited support, patience and love.

Summary

The amount of evidence produced in clinical environments has been rapidly increasing thanks to the adoption of new technologies, such as Electronic Health Records, for assisting clinicians in their daily activities. Although this shift is good for the advance of science and knowledge, it introduces difficulties for health care practitioners and researchers in terms of degree of efficiency and accuracy in assimilating, acquiring and aggregating clinical evidence. In the *health care* sector, knowledge and new evidence are often heterogeneous and complex, inconsistent and incomplete. These factors play an important role in many clinical decision-making processes, most of the time made under conditions of uncertainty and with partial knowledge and evidence.

Current *clinical decision support systems* are becoming more complex because plausible conclusions need to be extracted from a set of heterogeneous pieces of evidence, sometimes contradictory, and from different points of view and interpretations. They are mainly based on case-base or probability-based reasoning, and they adopt techniques borrowed from Artificial Intelligence such as machine learning or Fuzzy Logic. However, the majority of them requires well structured evidence, not partial and are based on learning from previous data or cases. In addition, the amount of evidence required for the learning process must be high in order to significantly infer recommendations for clinical decisions. These systems manipulate knowledge and evidence in a numerical, usually complex

way, not using familiar terms, thus being not attractive to clinicians. Health care practitioners usually tend to follow a *defeasible reasoning* process for taking plausible decisions. Defeasible reasoning is a kind of analysis and interpretation that is based on reasons that are defeasible: a conclusion can be retracted in the light of new evidence. Indeed decisions are taken by evidence-based knowledge, but the aggregation of pieces of evidence tends to be close to the way humans reason. This kind of reasoning process can be formalised using *Argumentation Theory*, an emerging paradigm, based on arguments, aimed at investigating their consistency and reducing uncertainty.

According to the limitations of current state-of-the-art approaches, clinicians and health practitioners, in general, prefer decision-making support systems that deliver more explanations than numerical aids. In other words, they would adopt qualitative systems rather than quantitative tools. Indeed, numerical outcomes are more accurate than linguistic outcomes, but most of the time they are difficult to be interpreted. Furthermore, the inference process that leads to outcomes, in this context recommendations to decisions, can be hard to be understood by clinical experts, in the case it is only based on a numerical manipulation of evidence and knowledge.

In this dissertation the main issue investigated concerns the *role of defeasible reasoning and argumentation theory for supporting decision-making processes under uncertainty in the health care sector*. The main aim is to support clinicians with a tool for taking plausible and rational medical decisions that can be better justified and explained. The basic principles of argumentation theory are described and demonstrated in a well known health scenario: the breast cancer recurrence problem. It is shown how to represent available clinical evidence in form of arguments, how to define defeat relations among them and how to create a formal argumentation framework. Argumentation semantics are then applied over the built

framework to compute arguments justification status. It is demonstrated how this process can enhance clinician decision-making. In detail, the designed framework is developed according to the knowledge-base of an interviewed expert in cancers and it is used to predict the recurrence of each breast cancer removed from 286 patients. An encouraging predictive capacity of 74% is compared against the accuracy rate of well-established machine learning techniques that performed equally or worse. This results is extremely promising because not only demonstrates how a knowledge-base paradigm, such as the designed model based on Argumentation Theory, can perform equally or better than current learning-based decision support systems, but also it shows a better explanatory capacity and a higher degree of intuitiveness in term of use.

Contents

List of Tables

List of Figures

Acronyms

AAT Abstract Argumentation Theory.

AF Argumentation Framework.

AI Artificial Intelligence.

ANN Artificial Neural Network.

AT Argumentation Theory.

BC Bayesan Classifier.

CBR Case-based Reasoning.

CDSS Clinical Decision Support Systems.

DM Data Mining.

DT Decision Theory.

DTT Decision Trees.

EHRs Electronic Health Records.

ES Expert Systems.

EUT Expected Utility Theory.

FL Fuzzy Logic.

FMF Fuzzy Membership Functions.

FT Fuzzy Theory.

KD Knowledge Discovery.

KNN K-nearest Neighbors Algorithm.

MDTs Multi-disciplinary meetings.

ML Machine Learning.

MLP MultiLayer Perceptron.

NBN Naive Bayesan Network.

Chapter 1

Introduction

The amount of clinical evidence produced in the health care sector has been rapidly increasing thanks to the advent of Electronic Health Records (EHRs) and other technologies for assisting clinicians, in their activities, at different levels. Although this shift is good for the advance of science and knowledge, it introduces difficulties for health practitioners, clinicians and health care researchers in terms of degree of efficiency and accuracy in assimilating and acquiring clinical evidence. In current health care settings, knowledge and new evidence are often heterogeneous and complex, inconsistent and incomplete. These factors play an important role in many clinical decision-making processes, most of the time made under conditions of uncertainty. Uncertainties may concern the diagnosis and accuracies of tests, the background and history of a certain disease, the effects of a specific treatment on a patient or the consequences of an intervention in a group.

1.1 Decision support in medicine

Clinical knowledge is rapidly increasing and as a consequence also the spectrum of options for various treatments and the available evidence-based knowledge associated with them. This evolution is bringing into the health sector the need for inter-disciplinary consultations aimed at providing the highest quality of health care. For this reason, over the last few decades, Multi-disciplinary meetings (MDTs) have appeared as a way to manage multi-disciplinary diseases such as cancers. These meetings have became new standard practices in the UK after the report of Calman-Hine [Calman and Hine., 1995], in 1995, that listed the weaknesses in providing care for cancers along with the lack of consistent standard of care across various areas and regions of the country. Another report [Imperial Cancer Research Fund ICRF, 1995], provided by the Imperial Cancer Research, suggested that 16 thousand lives could be saved, a year, if the standard of health care was consistent and uniform. Let us take a look at an illustrative example, aimed at clarifying the complexity and the uncertainty involved in current clinical complex decision-making processes in group. This example is extracted from a multi-disciplinary team meeting involving different experts discussing a patient with early stage superficial unilateral larynx cancer and presented in [Chang et al., 2009] and [Chang et al., 2010]. The meeting is attended by 3 surgeons (S_1, S_2, S_3) and two radiation oncologists (R_1, R_2). The discussion, in the form of a argumentative process (A is an argument), is aimed at deciding the most appropriate treatment for the patient.

Example 1

- A_1 (S_1): My opinion is to take out the patient's larynx: this has the best cure rate of 99%;
- A_2 (S_2): I agree, taking out the patient's larynx would provide the best cure potential;
- A_3 (S_3): I also agree, taking out the patient's larynx would provide the best cure potential;
- A_4 (R_1): If we take out the patient's larynx, the patient will have no voice;

- A_5 (R_1): However, if you use radiotherapy, there is 97% cure rate from the radiotherapy and about 97% voice quality, which is very good. The 3% who fail radiotherapy can have their larynx removed and most of these will be cured too;

- A_6 (S_2): my opinion is also that the patient should have a hemi-laryngectomy [1]. This will give a cure rate as good as radiation therapy;

- A_7 (S_3): I agree, performing a hemi-laryngetcomy would give a cure rate as good as radiotherapy;

- A_8 (R_1): Yes, I have performed many hemi-laryngectomies, and when I reviewed my case load, the cure rate was 97%, which is as good as that reported internationally for radiotherapy;

- A_9 (R_2): I agree, however you fail to take into account the patient's age. Given the patient is over 75, operating on the patient is not advisable as the patient may not recover from an operation;

- A_{10} (R_1): Yes, however in this case, the patient's performance status is extremely good, the patient will most likely recover from an operation.

- A_{11} (S_2): Reviewing our past case decisions, evidence suggest that we have always performed hemi-laryngectomy, hence my preference is to do the same;

- A_{12} (S_3): I agree, however, there is some new medical literature reporting that the voice quality after a hemi-laryngectomy was only 50% acceptable and the reporting institution was the North American leaders in hemy-laryngectomy, hence we should perform radiotherapy.

---

[1] hemi-laryngectomy is the excision of one side of the larynx.

The discussion above highlights important issues. Firstly, it becomes clear the need for accrual in an argumentative reasoning process. Generally, “accrual” refers to the aggregation and grouping of arguments to support or defeat a particular opinion and it is a delicate issue, as recognised by [Verheij, 1995] [Prakken, 2005] [Lucero et al., 2009]. For instance, argument A_4 is the basis of an attack against the argument A_1: just considering these two arguments in isolation, the decision making process is not trivial. There are evident difficulties in deciding the best treatment for the patient. Similarly between arguments A_1 and A_5. However, when considering the whole set of arguments, it becomes clearer that the more appropriate action to take, that means the final decision, is to perform radiation therapy before removing the larynx of the patient. A second issue is the ability to enforce arguments by repeating them. For example, arguments A_1, A_2 and A_3 enforce the potential decision to remove the larynx, as supported by the three surgeons.

Although in human debate situations it is conceivable that the number of argument is enough to overwhelm any suggestion of the contrary, the three arguments do not add additional information. Therefore, enhancing the strength of a position or option is not always recommendable by simply providing identical multiple arguments. Source expertise, credibility, reputation and background should be taken into account. So the information source is a key point in argumentative and decision making processes. Considering only the accrual of arguments as a consequence of the norms of communities, then removing the patient’s larynx seems to be the most appropriate decision to take. However, if an expert is known to have a particular developed background, with special insight or knowledge, then it may be possible to assign him/her a privileged position, increasing the strength of his/her arguments. Another interesting issue in the previous example is provided by arguments A_8 and A_9. Here the radiologist R_2 does not disagree with the radiologist R_1, rather A_9 defeats the inference rule used by R_1 to construct argument A_8. This phenomenon is know in the literature as "undercut" as the argument A_8 is more specific than the argument A_9. This situation stresses the fact that in real contexts there may exist some exception to general rules, affecting final decisions [Pollock, 1987] [Prakken and Vreeswijk, 2002]. Arguments A_{11} and A_{12} are attacks against user preference and are sensitive to the context. An argumentation system should be dynamic, evolving over time and taking into account past decisions as a way for justifying future decision making.

Let us assume that the physician of the patient decided to perform hemi-laryngetcomy: now he/she needs to justify the decision. If the discussion with other experts had not occurred, the knowledge base would be empty, and the physician would have to present just the arguments supporting the decision to perform hemi-laryngetcomy. However, taking into consideration all the arguments in the illustrative discussion, therefore a defined knowledge base with defeat relations and preferences, then the physician is required to justify the decision along with explaining and addressing all the attacks against his/her decision. This example highlights the uncertainty involved in such a discussion, as well as its complexity and how justifiable medical decisions are difficult to achieve and explain.

1.2 Decision-making in health care

Clinical decisions have been getting more complex because they involve multiple heterogeneous pieces of evidence in a dynamic context, the health care, characterised by a high degree of uncertainty. They can be taken individually or by a group of health practitioners, with different backgrounds, expertise, experience and sometimes with partial and contradictory knowledge-bases. Intuitively, there is a tangible need to aggregate evidence and knowledge in a way that reduces uncertainty and complexity, minimising inconsistency and incompleteness. Decision-making processes call for a method capable of aggregating clinical knowledge intuitively, accurately and able to cope with inconsistency, conflicts among pieces of evidence and to support the justification of a decision.

To clarify the issue behind current clinical decisions, let us take a look at the breast recurrence problem, being cancer one of the commonest non-skin cancers in women. According to a recent report [can, 2005], in Europe every two and a half minutes a woman is diagnosed with breast cancer and every 7 and half minutes, a woman dies from the disease. Consequently, significant research effort is required and important progress has been made, in the last 10 years, both in therapy planning and reduction of the risk of cancer recurrence. Recurrence is a the return of a cancer after treatment and a temporal interval during which its detection is not possible. Predicting recurrence is important for assisting the identification of patients with critical prognosis and minimising unnecessary therapies. A possible reasoning process that a clinician can perform to predict recurrence in a woman who had a breast cancer surgery, is described in the following.

Example 2

Firstly, the clinician may consider the tumor size at the time of the surgery: a low size is a reason to believe recurrence is not likely. However, if the patient is in a post-menopausal state, the risk of recurrence is higher. Secondly, from the clinician's knowledge, if the degree of malignancy of the cancer is low, there is a reason to believe recurrence is not likely but, if this is high, recurrence is more possible. Yet, evidence and knowledge available to the clinician suggest that if the cancer is on the right breast, risk of recurrence is relatively low but if the number of involved nodes in the cancer area is high, the risk of recurrence is extremely high.

Arguments in such a illustrative reasoning process are knowledge-based arguments: they are built according to the knowledge, experience and expertise of the clinician and based on clinical evidence. The natural question following such a process is: *What is the final outcome of the reasoning process?* Usually it is the patient's physician who is in charge of deciding the expected recurrence. Let us assume that the outcome is recurrence: the cancer is likely to appear again. The problem now is that the physician has to justify the decision.

From the example above, it becomes clear that the interest is on the supporting of medical decision making. Potential solutions are not aimed at removing the decision-making power from clinical experts, instead at providing them with tools able to enhance the capacity to make better decisions. In other words, decisions have to be supported, not dictated [Eddy, 1990].

1.3 Defeasible Reasoning

In the previous illustrative scenarios, the breast cancer recurrence prediction and the multidisciplinary meetings on larynx cancer, the discussions follow an argument-based reasoning. This typology of reasoning is a form of *defeasible reasoning*, a kind of analysis and interpretation that is based on reasons that are defeasible: a conclusion can be retracted in the light of new evidence. Defeasible reasoning is also known as common-sense reasoning, as being close to the way humans reason under uncertainty. Defeasible reasoning has been gaining momentum in the health care arena [Chang et al., 2009] thanks to its ability to reason about unstructured clinical situations, characterised by a significant degree of uncertainty, where available information is partial and sometimes contradictory. Defeasible reasoning, in turn, is a form of *non-monotonic reasoning*, a paradigm that deals with the problem of deriving, from one or more knowledge bases, plausible conclusions that are not infallible [Baroni et al., 1997]. Since conclusions are under uncertainty, it can be possible to retract some of them once new information becames available, showing that they are wrong or no longer plausible.

Defeasible reasoning can be practically implemented and modelled via Argumentation Theory (AT), a theoretical research field that studies how arguments can be expressed, sustained and discarded, as well as the validity of the conclusion reached in a reasoning process. It deals with formal and computable models, inspired by human-like reasoning. Its goal is to apply formal semantics among arguments to compute their justification status [Baroni et al., 2011] . Although the topic has proven to be useful in decision-making in fields such as law, logic and Artificial Intelligence (AI), its application in the health care sector is relatively new.

Defeasible reasoning and argumentation theory appear to be valid candidates for building decision making support systems in the context of health care and medicine. They seem to have all the characteristics for handling uncertainty in complex clinical domains, and the capacity to deal with incomplete, partial evidence and recommend decisions from unstructured knowledge bases.

1.4 Investigated issue

The reasoning involved in decision-making processes in the health care sector, as described in previous sections, is uncertain, dealing with partial knowledge and often with contradictory pieces of information. Clinical decisions are becoming more complex due to the fact that plausible conclusions need to be extracted from a set of heterogeneous pieces of evidence, sometimes contradictory, and from different points of view and interpretations. Health care practitioners, most of the time, follow a defeasible reasoning process for taking plausible decisions. Indeed decisions are taken by evidence-based knowledge, but the aggregation of pieces of evidence tends to be close to the way humans reason. This kind of reasoning process can be formalised using AT, an emerging paradigm, based on arguments, aimed at investigating their consistency and reducing uncertainty. The main aim is to support clinicians with a tool for taking plausible and rational medical decisions that can be better justified and explained. In this dissertation the main issue investigated concerns the *role of defeasible reasoning and argumentation theory for supporting decision-making processes under uncertainty in the health care sector*.

1.5 Route map to dissertation

This dissertation investigates the role of defeasible reasoning and argumentation theory in decision-making in clinical and health care environments. In this introduction chapter, a brief summary has been proposed underlying the growing complexity of medical decisions and the degree of uncertainty that characterises them.

Chapter 2 presents a literature review aimed at describing the main paradigms and approaches used in decision-making within health care domains. For each approach, advantages and disadvantages are highlighted and discussed.

Chapter 3 formalises the research question listing the associated objectives. Furthermore, based on the gaps and limitations found in the literature review, the need for a more qualitative approach to reason under uncertainty in complex decision-making processes in the health care sector is highlighted.

Chapter 4 introduces formal Abstract Argumentation Theory (AAT), along with skeptical and credulous approaches to aggregate pieces of evidence. This abstract theory is then extended to produce a formal model for decision-making under uncertainty suitable for clinical environments. Formalisms are clarified with an illustrative example about breast cancer recurrence prediction using a well-know dataset of evidence provided by the University Medical Center, Institute of Oncology, Ljubljana, ex-Yugoslavia. It is shown how experts' knowledge-bases can be translated into arguments, using fuzzy logic, how interactions among them can be explicated and how arguments-based semantics can be run to investigate their consistency and to minimize their uncertainty.

Chapter 5 is aimed at defining an evaluation strategy and discussing the results. It is shown how to interpret outcomes of a defeasible reasoning process, implemented formally with AT. Results are then compared with the ones obtained by running a selection of machine learning classifiers on the same dataset. Advantages and disadvantages, strengths and limitations of the use of AT in the health care sector are then discussed.

Chapter 6 summaries the thesis, underlying its main contribution and highlighting the main findings. Possible areas of applications and future work are eventually discussed.

Chapter 2

State-of-the-art of Decision Support Systems

2.1 Introduction

In this chapter we describe some of the main work on decision support systems, with emphasis on health-care applications. The primary objective of a clinical decision support system is to inform clinicians, and not to automate or dictate decisions. The system should support, explain and justify a potential or a set of decisions, and not infer the final decision. It should provide health-care practitioners with an explanatory tool, highlighting differences of each potential decision, in the presence of uncertainty. Here we describe the state-of-the-art on decision support tools, discussing each of them in terms of theoretical background, use, advantages and disadvantages, strengths and limitations.

2.2 Fuzzy logic

The explosion of the amount of information and evidence available to clinicians is increasing the level of uncertainty in many health domains. Handling uncertainty is a key issue in health care. Fuzzy Logic (FL) is an alternative paradigm of Bayes classifiers, (as described later) perhaps more suitable for handling different types of uncertainty that usually arise in clinical decision and reasoning. For instance, in clinical guidelines, there are many sources of uncertainty [Warren et al., 2000]. FL can represent the vagueness in the data and it is useful for dealing with uncertainty. Although it is based on numerical manipulation, it is more attractive to clinicians than Bayes Networks, due to the fact that it gives them the ability to manipulate knowledge and evidence with more familiar and natural language terms. According to [Warren et al., 2000], six typologies of uncertainty can occur in managing clinical data and, for instance, in formalising and computerizing clinical guidelines:

- *lack of information*: information can be incomplete, partial and sometimes totally missing;

- *non-specificity*: the conversion of the natural language adopted in many clinical guidelines and other health contexts, into standard logic, is not a trivial task; “a high degree of malignancy” of a tumor, or “a low lymph nodes involvement” are not easy terms to represent;

- *probabilistic nature of data and outcomes*: outcomes of clinical tests are often associated with error bounds;

- *vagueness in the formulation of recommendations:* recommendations in general and guidelines recommendations in particular, are performed using vague sentences that do not properly embed the numerical associated strength;

- *discordance:* often, data, knowledge and pieces of evidence are contradictory and discordant;

- *fuzziness in the determination of clinical signs:* many attributes and characteristics of patients are often the outcomes of the interpretation of data by clinicians. These interpretations are often performed subjectively, thus a source of uncertainty.

Fuzzy logic is a paradigm capable of representing most of the types of uncertainties in a more structured and formal way. FL is built upon Fuzzy Theory (FT), a paradigm aimed at dealing and formalising vagueness of data, knowledge and the associated uncertainty [Zimmermann, 2001]. The theory started with the seminal work of Prof. Lofty Zadeh and the major characteristic it that the theory deals well with reasoning that is vague and approximated rather than exact and fixed [Zadeh, 1965] [Zadeh, 1966]. Fuzzy logic is based on fuzzy sets, particular sets modeled by membership functions. Given a value, for a certain attribute, the membership function, defined for that attribute, gives the membership degree of that value to the fuzzy set modeled, in the continuous range 0 to 1. In other words, the degree of membership is indicative of how well a value or element is represented by the set. Membership functions can be discrete or continuous. Let us consider the attribute 'age' of a patient and let us focus on the set A of people over 70. In standard formal way it can be modeled as:

$$age > 70 \in A$$

This formalisation is a binary distinction between a value lower or greater than 70, and does not tell anything else. In fact, a patient 69 years and 364 days old does not belong to the set A, while a patient 70 years and 1 day old is in the set A. Clearly, this formalism leads to undesirable situations because the two patients may be treated in a different way. Fuzzy logic copes and resolves this situation via membership functions. A fuzzy set for people over 70 years could be represented and formalised with a membership function that returns low values for patient younger than 70 and high values for patients over 70. A simple function can be (figure 2.1(a)):

$$f_A(x) = \begin{cases} 1 & if \quad x \geq 70 \\ \frac{x}{70} & if \quad x \leq 70 \end{cases}$$

A more complex definition can be the following sigmoid function (figure 2.1(b)):

$$f_A(x) = \frac{e^{x-70}}{1.1 + e^{x-70}}$$

Although fuzzy representation deals better with vagueness and it is a good approach for handling uncertainty, there is the problem of using the results generated by the logic. Defuzzification is a method to convert the numerical outcomes into something more interpretable, such as a linguistic variable [Warren et al., 2000]. This can produce natural language outcomes to describe the strength of the output.

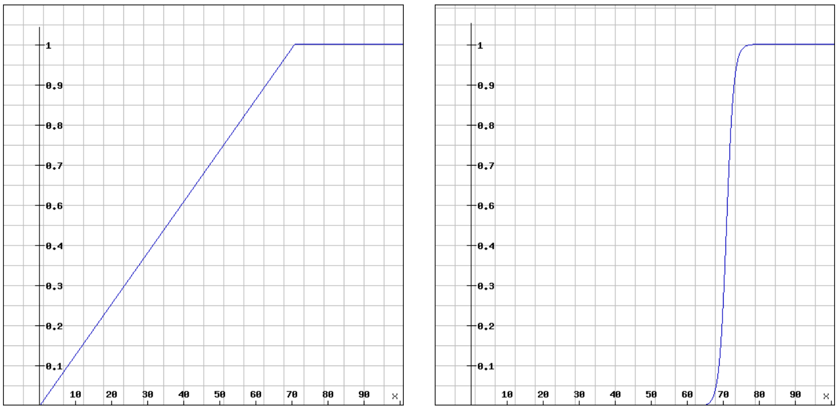


Figure 2.1. Membership function definitions for the attribute 'age': simple and complex definition

In many health care scenarios, the attributes used for describing an entity, for instance a patient, can be multiple. Each of them can be represented by one or more fuzzy sets, but when used in a decision making process, they need to be aggregated in order to recommend a decision. Furthermore, the same membership function of a fuzzy set may be defined subjectively, and arbitrarily by two different designers. Another issue of FL, as with probabilistic-based representations, is the sensitivity of the approach to changes in the membership functions definitions. The time required by an expert to design and define fuzzy sets represents a disadvantage as well as the clinicians' subjective perception of what fuzzy values represent. Yet, the mathematical manipulation of fuzzy sets and how they are actually aggregated can be obscure and hard to understand by clinicians.

2.3 Decision theory

Decision Theory (DT) is a well known approach in clinical environments, with a background at the intersection between mathematics, logic and philosophy [Parmigiani and Inoue, 2009]. DT can be divided into two typologies of theory: descriptive and normative. Normative theory, or prescriptive, is aimed at determining and identifying the best decision and solution to take, in terms of optimality, ideality and rationality. In practice, the main aim of this prescriptive approach is to help people make better decisions. In health care, practical normative techniques are most of the time in the form of computer-aided decision support tools, and are often known and referred to as Clinical Decision Support Systems (CDSS). In practical decision making processes, people do not tend to behave consistently with axiomatic rules. As a consequence, often their own reasoning lead to violation of optimality. For this reason, the descriptive theory, or positive, aims to describe what people actually do in reality.

Normative and descriptive theories, although having two different objectives, are linked and semantically connected. In fact, normative theory creates optimal decisions, generating hypothesis for testing agains actual behaviour. Normative theory can produce different behavioural predictions and options, favoring tests against practical decisions effectively taken. In decision theory, a set of candidates or options are considered towards a decision. These options, often represented numerically, or ordered by their importance, need to be evaluated for recommending the best alternative. If this exists, and it is unique, under DT this should be the solution actually taken.

One of the classical approaches adopted under decision theory is the Expected Utility Theory (EUT). This approach is built upon the utilitarian moral theory: during a decision making process, the decision maker should consider the utility of the outcomes. Generally, ensuring a level of utility that is satisfactory is enough in many cases. However, in health care, this is not always sufficient. Medical decision making should consider the best option that maximises the quality of care of patients as well [Rabin, 1999]. Under EUT, for each option available to the decision maker, a utility score is defined along with the probability that the outcome for that option will occur. Formally:

$$EU^C = \sum_{i=1}^{n} P_i \cdot V_i$$

where n are the available outcomes for a candidate C, P_i represents the probability that has been assigned to the outcome i, and V_i is the value for the outcome i (it can be either positive or negative). EU^C is therefore the expected utility value for a candidate C. Each option considered by the decision maker may be seen as a risky option, with a set of associated outcomes and probabilities of recurrence. Furthermore, since the outcomes of that option, for instance, a medical treatment, cannot be known with absolute certainty, the expected utility theory represents a powerful paradigm to predict choices under uncertainty [Hellinger, 1989]. The main difficulty with EUT is the estimation of the probabilities of each outcome for each option. A reasonable approximation can be achieved if probabilities are derived from a large enough dataset of empirical evidence for a certain specific case. On the other hand, if the amount of the available evidence is small, then the expected utility theory loses power and predictive capacity. Another issue is the difficulty in finding the predictions of the involved probabilities. In fact, generally, physicians tend to over-estimate the probabilities of the success of their plans for treatments. For this reason, it is commonly and generally accepted that, only when values and objective estimates are available, should the expected utility theory be adopted as a decision making support tool.

In the health care sector, decision making processes are characterised by uncertainty both in utility and probability. The numerical values associated with both utilities and probabilities are inherently uncertain and can be calculated by retrospective frequencies and previous evidence. However, the predictive capacity and efficacy of the EUT is unproven at present. As mentioned in [Fox and Das., 2000] clinicians and health care experts are not inclined to adopt a decision support tool that uses numbers automatically calculated but not explained. Decision makers seem to prefer clinical decision support systems that are more qualitative rather than quantitative.

2.4 Expert systems

Expert Systems (ES) are computer-based approaches aimed at emulating the ability of human experts at making decisions. This type of system is part of AI and it is designed to reason about knowledge and solve complex problems like a human expert, refuting to follow standard procedures of a developer [Jackson, 1998]. A typical expert system consists of two parts:

- *the inference engine*: a fixed part that is independent of the expert system;
- *the knowledge base*: a variable part, usually expressed with natural language rules.

A typical knowledge base, or rule base is composed by a set of rules of the form:

IF... THEN...

A few examples are:

- IF 'high lymph nodes involvement' THEN 'high risk of cancer recurrence'
- IF 'low age' THEN 'low chances of cancer recurrence'
- IF 'high degree of malignancy' AND 'high age' THEN 'very high risk of cancer recurrence'

The use of natural languages terms made expert systems very appealing for decision making support and they represented the first real form of success of AI-based software. Rules are the translation of knowledge that is exploited by the expert system. Despite their extensive use in many domains, the main difficulty is the collection of the knowledge from experts. Despite the fact that several approaches have been proposed to translate knowledge bases into formulae, these approaches are most of the time usable only by computer experts.

Inference engines can be executed in two ways:

- batch
- conversational

In batch mode, a system internally embeds all the necessary information and data to process from the beginning, as a classical software program. Although results are provided immediately, the reasoning process is invisible. The conversational method is necessary when the developer knows in advance that not all the necessary data can be provided by the user from the beginning and cannot be subject of questions. In this case, the expert system needs to infer a solution automatically, refining it through gradually requesting the missing information from the users. This process seems to be a dialogue led by an expert aimed at approaching the goal as quickly as possible.

Expert systems are based on everyday language and they can be deployed by experts and general users much faster than conventional programs, bypassing professional software developers and without the need for explaining the subject. Another advantage is that rules can be added, modified or removed thus increasing the scalability of the system. On the other hand, knowledge collection and translation into formal rules is a major issue. In the majority of the cases, developer do not have a methodology to perform such a translation and they are forced to work manually, with a clear increase of potential errors. Furthermore, rule-based systems, despite their extensive usage, are not well understood and rules can be contradictory, poorly written therefore unusable in practice. Yet, many implementations of such systems are built upon a logic that operates on facts that are variable, that means facts whose value is dynamic, changing during the reasoning process. This property is advantageous for developers, but it generates systems less clear to users, less accurate and reliable and they are not capable of producing explanations nor detecting contradictions [Leondes, 2002]. In the health care arena, it is essential that an expert system is accurate, as humans life is involved: a tiny error can cause death, and the system cannot go back and remedy the situation.

2.5 Case-based reasoning

Clinicians and, in general, health-care practitioners, are comfortable with evidence-based medicine. In fact, from the evidence gathered in medicine, most of the current clinical guidelines are derived. Unfortunately, not all the guidelines can be generally applicable, thus case-based system are increasingly taken into consideration as useful decision support tools.

A Case-based Reasoning (CBR) system is not built upon general domain knowledge and does not rely on general guidelines. Rather, it is aimed at extracting specific knowledge from previous clinical cases and evidence. The success of this typology of systems derives from the fact that they approximate human cognitive memory recall, which significantly influence human decision making processes [Watson and Marir, 1994].

An evident advantage of case-based reasoning is represented by its capacity for facilitating sustained learning throughout an application life cycle. In fact, each decision taken from a given case, contributes to decisions for subsequent cases, improving their quality. A clinician, for instance, might select the same or similar option of treatment, for a certain patient, because a previous patient, in a similar situation and with similar attributes, has been recently treated thus. Despite the fact that this way of reasoning seems to be very natural and appropriate for decision making, in reality it does not follow good practices of evidence-based medicine. In fact, this technique does not consider outcomes of the treatments, for previous patients. Another disadvantages of CBR is that, when the number of previous cases is limited, the approach is essentially based on anecdotal evidence. With limited previous evidence, actual knowledge is limited as well, and practical decision making cannot be fully supported. Therefore a useful case-based reasoning support tool requires a large case-base.

A general reasoning process, case-based, as depicted in figure 2.2, can be divided into four main steps [Aamodt and Plaza, 1994]:

- *retrieving* as many similar cases as possible;
- *reusing* the evidence, knowledge and information of similar cases to investigate and solve the current case;

- *revising* the proposed solution for the current case;
- *rataining* the most useful current experience for likely supporting future case-based decision making processes.

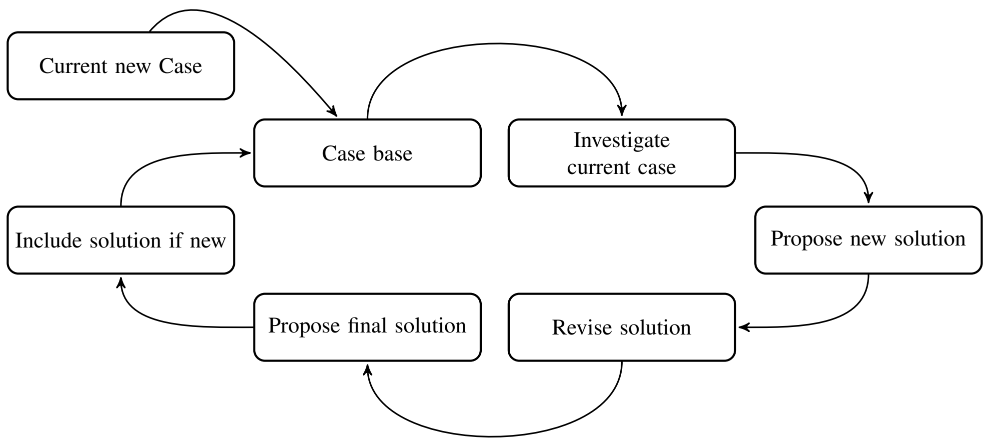


Figure 2.2. Case-based reasoning process loop

For a current new case, similar previous cases have to be retrieved, from the case-base (previous evidence and knowledge). At this stage, the current new case is investigated and a new solution, considering previous solutions, is proposed. Subsequently, the proposed solution may need to be revised in order to account for any difference between new and previous cases. The final solution is then decided on and the new current case is included in the case-base. This reasoning loop is aimed at providing an automated and structured methodology for learning and maintaining the system.

Case-based reasoning is widely accepted and used because it provides many benefits compared to knowledge-based approaches. The attention given by health-care practitioners to case-based systems is mainly explained by the fact that decision and solutions are taken considering previous evidence and data, avoiding the difficult process of eliciting knowledge by experts. Furthermore, a knowledge base requires maintenance, which can be difficult as well. A case-based reasoning decision support system facilitates the maintenance of knowledge, because each new case extends, enhances and updates the case base. According to [Austin, 2008], the updates, in a case-based reasoning system, represent machine learning as an inherent, intrinsic characteristic. Furthermore, they can be easily interpreted by domain experts. On the other hand, implementing CBR tools is not a trivial problem. The increasing speed in producing medical knowledge, collecting and processing data from previous patients and cases, represents a difficulty. As a consequence, new data and evidence may be not always relevant and immediately applicable to new cases and patients. In addition, new treatments and drugs evolve continuously, therefore a case base, for instance, dealing with breast cancers, would have to be generated considering recent cases and findings, which is a significant task. Another disadvantage of CBR is that it cannot directly take into account national guidelines and recommendations, because it relies on previous cases. Yet, even if guidelines are partially considered, they treat information with emphasis on details, difficult to represent and to embody in case-based reasoning systems.

2.6 Bayesian network classifier

Bayesian decision theory, in the field of AI, is believed to be the best approach to decision making under uncertainty. According to the theory a single agent acting in an environment, characterised by uncertainty, and provided with an unlimited amount of computational power and resources, should behave consistently with the theory [Buntine, 1997].

The simplest form of a Bayesian classifier is known as Naive Bayesan Network (NBN). This approach consists of representing each option available, that means each decision, as a class, and computing the probability that a given patient, with certain attributes, falls within the class. Given an option, the Bayesian classifier has to learn the conditional probabilities of each attribute. Subsequently, in order to classify a particular instance, represented by a set of attributes, the Bayes' formal rule (or theorem) is used to calculate the probability of a class. Eventually, a decision-making support tool can select and recommend the class, that means the option, with the highest posterior probability [Friedman et al., 1997].

Formally, the Bayes' rule, or theorem is:

$$P(C|A_i...A_n) = \frac{P(C) \cdot P(A_i...A_n|C)}{P(A_i...A_n)}$$

where P stands for probability, C is a class, A is an attribute. The Naive Bayes classifier assumes that all the attributesA_i are conditionally independent for a give class C. The formula above can be rewritten and expressed in natural language as:

$$Posterior = \frac{prior \times likelihood}{evidence}$$

In practice, the evidence is available and it is constant. Therefore, the denominator of the formula can be not considered, as always constant and independent of the class C. The numerator is equivalent to the joint probability model and $P(A_i...A_n)$ can be decomposed into the product:

$$P(A_i...A_n) = P(A_1 = v_{1k}|C)...P(A_n = v_{nk}|C)$$

with v_{nk} the specific actual values of the attributes. The Bayes' theorem can thus be written as:

$$P(C) \cdot \prod_{j=1}^{n} P(A_j = V_{jk}|C)$$

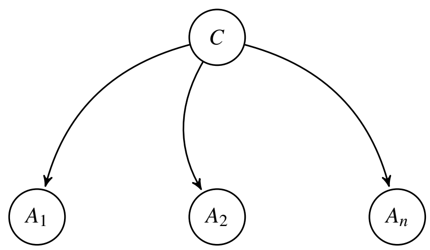


Figure 2.3. Bayes simple naive network

For each class, or option, the theorem is applied and the classification can be made by selecting the Maximum a Posteriori criterion and thus, extracting the corresponding class.

The NBN diagram depicted in figure 2.3 is a simple naive case. However, in complex health care scenarios, more complicated networks can be built, able to handle interdependencies among attributes as well. These typologies of graphs represent a well structured way of dealing with uncertainty in the data. Despite this, the main difficulty is that, in case of missing values, further algorithms are needed to fill gaps. In most health care contexts, the chances of missing data are high, therefore the amount of evidence required to build a robust Bayes classifier must be high. Furthermore, realistically speaking, an expert cannot be expected to translate his/her knowledge base into a numerical network such as the Bayes one.

2.7 Machine learning

Machine Learning (ML) is a major branch of AI that has developed rapidly in the last few decades. For any intelligent behaviour, as agreed by many researchers [Michalski et al., 1997] [Mitchell, 1997], the basic requirement is represented by learning. Modern clinical environments are very often equipped with data collections tools and devices that have been continuously contributing to the increment of the amount of evidence available to clinicians. ML is a paradigm that provides several approaches for analysis of data and it is useful for dealing with large amounts of information. In healthcare it has been mainly used in medical diagnosis in contextual specialised problems. Machine learning requires that data, for instance related to a patient, usually in the form of features, is provided, along with the known correct outcome, for instance the known correct diagnosis. Subsequently a learning algorithm is executed on a, preferably, big dataset in order to extract association rules or models among input features with correct outcomes. This developed network of rules can then be used as a classifier, for a new unknown record, in order to predict the likely correct outcome. This output can be used for enhancing and supporting decision making. Clinicians are provided with a tool for assisting their decisions, improving, for instance, diagnostic speed, reliability and accuracy [Kononenko, 2001].

Current machine learning algorithms have proved useful for revealing interesting properties and interrelationships of heterogeneous data. Among these various types of algorithms, the main approaches are Artificial Neural Network (ANN), Decision Trees (DTT), Bayesan Classifier (BC) and K-nearest Neighbors Algorithm (KNN). We do not provide technical explanations of these classifiers, because it is out of the scope of the dissertation, but we refer the reader to [Michalski et al., 1997] [Mitchell, 1997]. What needs to be known is that they can be divided into:

- supervised
- unsupervised

In supervised learning , given a set of examples (usually provided by experts), representative of all the variations, for instance, of a disease, along with the actual outcome, for each example, a machine learning algorithm is able to discover associations among the examples' attributes in order to approximate the given outcome. Figure 2.4 is an example of an architecture [Eom et al., 2008] of a three-layered MultiLayer Perceptron (MLP) network, a special implementation of ANN. The model is aimed at classifying a cardiovascular disease. For a given patient sample to be diagnosed, a input array of values, representing selected proteins, is provided to the input layer of the network. The output layer is composed to a set of output class, in this case classes of diagnosis. The internal layer is trained and weights are modified according to the inputs arrays provided in the learning process. The classification can be done by giving to the input layer a new array of proteins, and the network will select the class with the maximum value. This value can then be used to support a decision making process [Eom et al., 2008]. In unsupervised learning, the actual outcome is not be provided to the network that is limited to discover clusters of similar examples. This approach is often referred to as Data Mining (DM) and Knowledge Discovery (KD).

Approaches to supervised machine learning are very accurate in connecting inputs to outputs, so attributes to potential decisions. The main advantage is that there is no need to provide and build a specific algorithm for a specific health domain, in order to identify a certain disease. However, the examples, input of a learning algorithm, have to be selected carefully in order to achieve a good prediction capacity and reliability of the system. Also identifying predictive features can be an issue. Approaches to unsupervised machine learning extends, in somehow, case-based approaches, providing an automatic way of classifying and analysing complex dataset. Newer medical approaches to machine learning have started considering and combining the decisions of multiple classifiers in order to increase the reliability of the whole system on a particular problem [Matja et al., 1996]. According to [Kononenko, 2001], physicians in general have found that this combination of classifiers was the best way of increasing the reliability of systems along with improving the comprehensibility of diagnosis.

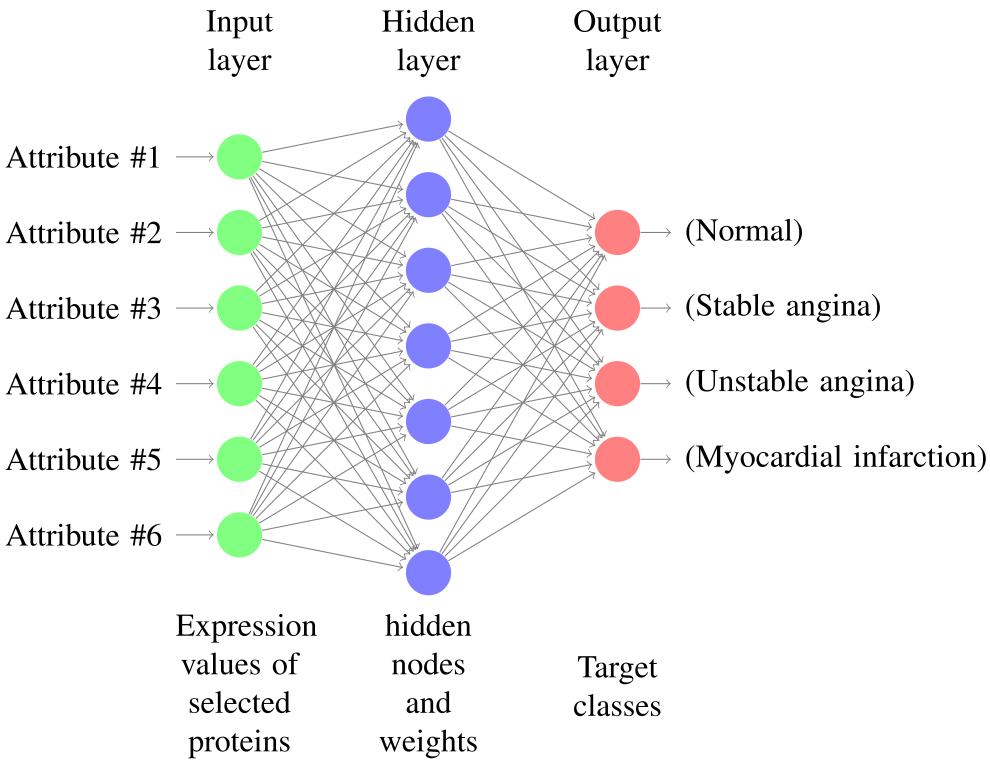


Figure 2.4. Artificial neural network example for cardiovascular disease classification

Despite this advantage, machine learning-based techniques have not been fully accepted in clinical environments, due to their limited explanatory capacity. In practice, although accurate, the inference process of the algorithm is complex, and most of the time unknown by health practitioners. Furthermore, they rely on big datasets of evidence and knowledge, and perform poorly on small datasets. The number of attributes, representing a single example input, is static, that means it cannot increase or decrease during the learning process. In addition, an explicit technique needs to be implemented in the case where the value of a single attribute of an example is missing. This technique needs to fill the gap, introducing a form of error and uncertainty. If a new attribute, for instance related to a patient, needs to be considered and added to describe an entity (example), then all the records (all the examples) within the same dataset need to be adjusted with the new attribute: this task may require time. The network is therefore static, and if new pieces of evidence about an entity need to be considered, the network will need to be built again and trained.

2.8 Argumentation theory and defeasible reasoning

Argumentation theory AT has evolved from its original primary context as a sub-discipline in philosophical logic, to emerge, in the last decade, as an important area of logic-based AI. Within AI, research in argumentation theory has produced significant contributions to the modeling and analysis of defeasible reasoning and the development of formal methods for negotiation and dialogue processes in multi-agent systems. Within health care, argumentation methods have been used for decision making, in areas such as prescribing, risk assessment and therapy selection. The theory has become an important topic in computer science, resulting from a multi-disciplinary approach at the intersection of philosophy and law, and with elements drawn from psychology and sociology [Baroni et al., 2011].

Argumentation theory systematically studies how arguments can be expressed, built, sustained or discarded in a debate, as well as the validity of the conclusions reached. The theory gained importance with the introduction of formal and computable models, inspired by human-like reasoning. These models extended the classical reasoning models based on deductive logic that appeared increasingly inadequate for problems requiring non-monotonic reasoning [Baroni et al., 1997], commonly used by humans, as well as explanatory reasoning, not available in standard non-monotonic logics such as default logic. Non-monotonic reasoning [Chang et al., 2010] differs from standard deductive reasoning because in the former a conclusion can be retracted in the light of new pieces of evidence, whereas in the latter the set of conclusions always grows. Argumentation lends itself to explanatory reasoning because the argumentative reasoning is composed of modular and intuitive steps, thus avoiding the monolithic approach of many traditional logics for non-monotonic reasoning.

The reasoning required in many decision-making processes in health care are both non-monotonic and explanatory. Argumentation theory is also suitable when the available information may be uncertain and incomplete, as in health care, where there may be relevant but partially conflicting information. In health care, the first application area of argumentation-based methodologies is decision-making. In the early 1990s, J. Fox and his group at Cancer Research UK proposed an approach to decision-making under uncertainty [Fox et al., 2006] in which logical methods are used to develop arguments for and against competing clinical hypotheses such as diagnoses, or actions such as therapies. Despite its simplicity, the approach has proved to be surprisingly effective for constructing practical decision support systems, and it has high acceptability to clinicians.

Recently, models of argumentation have become increasingly concerned not only with the formulation of individual arguments for beliefs or actions but also how arguments may interact, particularly how certain kinds of argument can attack and defeat other arguments [Hunter and Williams, 2010, Williams and Williamson, 2006] [Longo et al., 2012]. These models have been inspired by the important contribution of Dung [Dung, 1995] who proposed a very abstract methodology (Abstract Argumentation Theory AAT) to deal with a set of arguments that attack each other along with semantics to compute extensions of arguments that can be seen as justifiable. Over the years, Dung's grounded and stable semantics have been applied in several studies and further semantics have been proposed [Baroni et al., 2011]. An example of such a study is the work of

Williams and Williamson aimed at combining probability with logic for breast cancer prognosis [Williams and Williamson, 2006]. The authors proposed to represent background knowledge about breast cancer with logical formal arguments, the qualitative explanation of the prognosis, while they propose a Bayesian network to capture the probabilistic relationships among the variables and to perform the prognosis. Argumentation theory has been used to aggregate clinical evidence as well. The work of Hunter and Williams [Hunter and Williams, 2010] was focused on aggregating clinical knowledge in an intuitive way and on reducing the inconsistency, complexity, volume and incompleteness of the available evidence.

2.9 Discussion

Health care environments are becoming more structured, thanks to technologies and methodologies that help clinicians and health care practitioners in their daily activities. This transition, however, increases the amount of knowledge available to them introducing complexity and uncertainty. More evidence is indeed useful for supporting decision making of clinicians, but the drawback is that this new additional knowledge needs to be properly managed and aggregated to remain a valid support. CDSS are aimed at providing clinicians with a tool for enhancing the quality of the decision taken, and are not aimed at substituting them. Clinicians have anyway the final word of a decision, but a good decision support tool can recommend to them why an option is more consistent and accurate than another one. In other words, supporting tools need to have explanatory capabilities, they should not infer decisions, but provide consistent alternatives and options. Several reasoning approaches have been proposed in CDSS, however, they have their own advantages and disadvantages.

Fuzzy logic is known to be a very useful approach for handling and representing different typologies of uncertainties. It is built upon Fuzzy theory, a paradigm capable of representing the vagueness of entities and attributes involved in health domains with fuzzy sets. A further advantage is represented by its capacity of dealing with human reasoning that is vague, approximated rather than exact and fixed. However, there is the problem of practically using the numerical results because they need to be translated in something more interpretable by clinicians. Moreover, the time required by an expert to design and define fuzzy sets, is a disadvantage as well as the clinicians subjective perception of what fuzzy values represent. Eventually, the manipulation of fuzzy sets and their aggregation can be difficult for clinicians to understand.

Decision theory and the main classical adopted approach of Expected Utility Theory have both a long history. They are based on probabilities and are aimed not only at proposing the optimal option, but also at recommending the option that maximises the utility for patients. Although intuitive in many cases, the main difficulty is to estimate the probabilities of each option, usually overestimated by clinicians. Furthermore, if the amount of evidence available is not sufficient, then the theory loses power.

Expert systems are computer-based tools focused on emulating the ability of human of making decision by using simple 'If ... Then...' rules derived by knowledge bases. Although they use natural language terms, the main disadvantage is the collection of the knowledge bases from experts and their translation into formal rules. Moreover the fact that they need to be practically implemented by computer experts is a major drawback. Yet, expert systems mostly fail to detect contradictory rules as well as the dynamism of changing information and knowledge bases.

Case-base reasoning is an intuitive tool for clinicians because they are comfortable with evidence-based medicine when guidelines cannot be applied. The aim is to extract specific knowledge from previous clinical cases and evidence, approximating human cognitive memory recall. This evidence is maintained and updated for supporting future clinical decisions. Although it seems to be very natural and appropriate for decision making, practically it does not follow good practices, failing to consider the outcomes of the decisions taken. In addition, when the previous evidence is limited, the approach is essentially based on anecdotal evidence, giving weak support to decisions. On the other hand, the problem of formally translating knowledge-bases is avoided by domain experts and their outcomes can be easily interpreted because they are built upon the same language.

Bayesian network classifiers are believed to be the best approaches for supporting decisions under uncertainty. This paradigm consists of representing each available alternative (a decision) as a class, and computing the probability that a given patient, with certain attributes, falls within that class. Given a class, the Bayesian classifier has to learn the conditional probabilities of each attribute. The option with the highest posterior probability is usually recommended as the best option. Although these classifiers are built upon numerical manipulation, they are more attractive to clinicians because they give them the ability to manage knowledge and evidence, in familiar terms. However, the main drawback is that, in the case of missing data, (high chances in many clinical domains) the approach requires a specific strategy to fill the gaps. Furthermore, realistically speaking, an expert cannot be expected to translate his/her knowledge into a numerical network such as the Bayes network.

Machine learning is a paradigm that provides several approaches for analyzing and dealing with large amounts of data along with its capacity to discover interrelationships among variables. It provides tools for learning from examples, and classifying new inputs into output classes, usually decisions. The main advantage is that there is no need to provide and build a specific algorithm for a specific health domain, in order to identify a certain disease. Although this is appealing to clinicians, the main disadvantage is that the examples, input of a learning algorithm, must be carefully selected to achieve good classification thus accurate recommendations for decisions. It relies on big datasets, whose records are composed by a fixed number of attributes, therefore not suitable for changing information. Moreover it lacks explanatory power, limiting its appeal.

Argumentation theory is an emerging approach for dealing with unstructured evidence, characterised by uncertainty, in dynamic domains. Knowledge bases are expressed in form of arguments, and the theory systematically studies how they can be expressed, built, sustained or discarded in a decision making process, as well as the validity of the outcomes. It implements defeasible reasoning, a form of reasoning commonly used by humans where conclusions, derived by available evidence, can be retracted in the light of new pieces of evidence. It is suitable when information is uncertain, incomplete, and sometimes contradicting. Furthermore, argumentation theory lends to explanatory reasoning because the argumentative reasoning is more modular and intuitive. Although these advantages, the theory is mainly applied in field such as Law and Philosophy, with very tentative applications in clinical decision making support in healthcare.

2.10 Rationale for a qualitative approach to decision support

According to the limitations of current state-of-the-art approaches, as described in chapter 2, clinicians and health practitioners, in general, prefer decision making support systems that deliver more explanations than numerical aids. In other words, they would adopt qualitative systems rather than quantitative tools. Indeed, numerical outcomes are more accurate than linguistic outcomes, but most of the time they are difficult to interpret. Furthermore, the inference process that leads to outcomes, in this case recommendation to decisions, can be hard for clinical experts to understand, in the case it is only based on a numerical manipulation of evidence and knowledge. The main rationale behind this research is to produce a decision support tool that follows a more qualitative approach, capable of delivering self-explanatory outcomes, easier to interpret and manipulate by health care practitioners. Our approach does not exclude the manipulation of numerical-based pieces of evidence, but the fact that these can be embedded in more linguistic-based and understandable arguments, facilitate the inference process towards a defeasible decision recommendation. The main strengths of adopting a more qualitative research approach, as also mentioned by [Miles and Huberman, 1994], are:

- acknowledging the complexity of real-world health care contexts;
- the capacity of recognising differences in the way clinicians and experts manipulate their knowledge, understand and make sense of the same situation;
- a focus of description of the available evidence and interpretation of knowledge and decisions;
- the use of a dynamic and flexible process that accounts for emergent and changing clinical information;
- the possibility of generating new knowledge and insight about a concept or particular situation.

Given these strengths and advantages, we believe that the qualitative research approach is the most suitable for the domain of decision support systems in health care and medicine, as also confirmed in a recent conference in health care [Longo et al., 2012]. Treating clinical decision making as defeasible reasoning processes seems to be a valid hypothesis that needs to be validated through analysis and experimental studies along with their formal representation via argumentation theory.

In the next chapters it is shown how Argumentation Theory could be successfully applied in healthcare for enhancing decision-making support systems. The goal is to provide clinicians with a framework for reasoning with unstructured pieces of evidence, characterised by uncertainty, often contradictory, and to support their decisions. The basic building blocks of Argumentation Theory are firstly introduced and a computational model, based on arguments, is subsequently designed. Formal descriptions are clarified by describing a well-known healthcare problem: the prediction of breast cancers after surgery. The rationale for selecting such a problem has been the availability of a data-set, public on the Web, containing real-world data of patients who went through a breast cancer operation. This data not only can be used for testing the designed computational model, but also as the input of many machine learning algorithms, useful for comparisons.

Chapter 3

Design and Methodology

3.1 Introduction

It appears that argumentation theory is a potential good candidate for supporting decision making in health care, because it has the capacity to adapt to domains characterised by uncertainty and dynamism in terms of changing information. The theory seems to be suitable for translating the knowledge bases of health experts into something more formal and structured because it considers the vagueness and subjectiveness of this process. Another key point is that it is capable of dealing with evidence that can be fragmented, partial and sometimes contradictory. In addition, the theory is gaining momentum because not only it is concerned with the formulation of individual arguments for beliefs or actions, but because it also considers how arguments interact, attack and defeat each others, thus handling their consistency.

3.2 Research question

According to the gaps and limitations of the current state-of-the-art approaches in decision making under uncertainty in health care domains, it seems that Argumentation Theory is a favourite candidate for supporting decision making reasoning processes in emerging clinical environments. The research question investigated in this thesis can be formalised as in the following:

> *To what extent can argumentation theory support decision making in emerging clinical and complex health care domains?*

3.3 Research objectives

The research question can be addressed by completing the following research objectives:

1. definition of the basic building blocks of argumentation theory AT;
2. design of an extended model of AT for decision making support under uncertainty (main contribution);
3. design of an evaluation strategy for testing the deployed model;
4. identification of those clinical environments that would benefit the most from the adoption of the designed model .

The defined research objectives can be split into sub-objectives (figure 3.1) as in the following.

1. a) definition of Abstract Argumentation Theory AAT;
   b) definition of an argumentation framework (arguments and defeat relationships);
   c) definition of argument-based semantics for assigning justification status to arguments;
2. a) definition of a logic for representing arguments and defeat relations into a organised framework;

    b) definition of a logic for instantiating a sub-framework of arguments and defeat relations for real-world (objective) data available;
    c) definition of a strategy for aggregating arguments, checking their consistency under uncertainty and providing a recommendation to decision makers;

3. a) definition of a comparative evaluation of the designed model against other approaches used in decision making;
    b) comparisons of the predictive capacity and explanatory power of our approach against other approaches;
    c) identification of the strengths and limitations of the designed model;

4. a) identification of health contexts and domains that would likely benefit most from the use of AT
    b) identification of the potential advantages of AT in the identified health domains.

Objective 1 is aimed at identifying and describing the basic building blocks of AAT (a) that considers abstract arguments and their mutual relationships (b). In other words, an argument is represented as a abstract node in a graph of linked nodes, without considering its internal composition. Furthermore, the objective aims to describe current skeptical and credulous approaches for assigning justification statuses to abstract arguments (c). This objective is addressed in chapter 4

Objective 2 concerns the identification of a logic for extending abstract arguments to structured arguments. The goal is to define a logic that allows the translation of clinical knowledge bases into a structured framework. This process will include the representation of arguments as well as the representation of defeat relations among them (a). This process is aimed at providing clinicians with a tool for translating their knowledge-base in a set of more expressive arguments built with familiar terms. Once a contextual argumentation framework has been designed, a logic for instantiating (activating) a sub-framework with real-world data is needed (b). Subsequently, the focus it to propose argument-based semantics capable of investigating the relationships among arguments and rationally aggregating them into internally coherent and consistent sets. This step also concerns the definition of a strategy to weight each argument is terms of uncertainty carried. Eventually, the goal is to define a way of selecting the set of coherent arguments that minimise uncertainty as a recommendation for a rational decision (c). This objective is addressed in chapter 4.

Objective 3 deals with the evaluation of the designed model. In particular a comparative study needs to be defined for comparing the outcomes of the designed approach against the outcomes of other approaches for decision making support under uncertainty (currently adopted in clinical environments) (a). The evaluation strategy will measure the predictive capacity and the explanatory power of the designed model compared to other approaches (b). Eventually, the advantages, strengths and disadvantages, limitations of the proposed decision making support tool need to be identified (c). This objective is addressed in chapter 5.

Objective 4 is aimed at identifying potential health care application areas for the designed solution (a). In particular, the advantages derived from adopting the argument-based model along with tangible support to decision making need to be identified and clarified. Similarly, potential limitations, weaknesses and issues have to be addressed for the success of the proposed solution (b). Eventually the focus is to create a practical tool with attention on intuitiveness and explanatoriness. The system must be intuitive and easy to use and must provide clinicians with more qualitative explanations and not only with numerical outcomes (c). This objective is addressed in chapter 6.

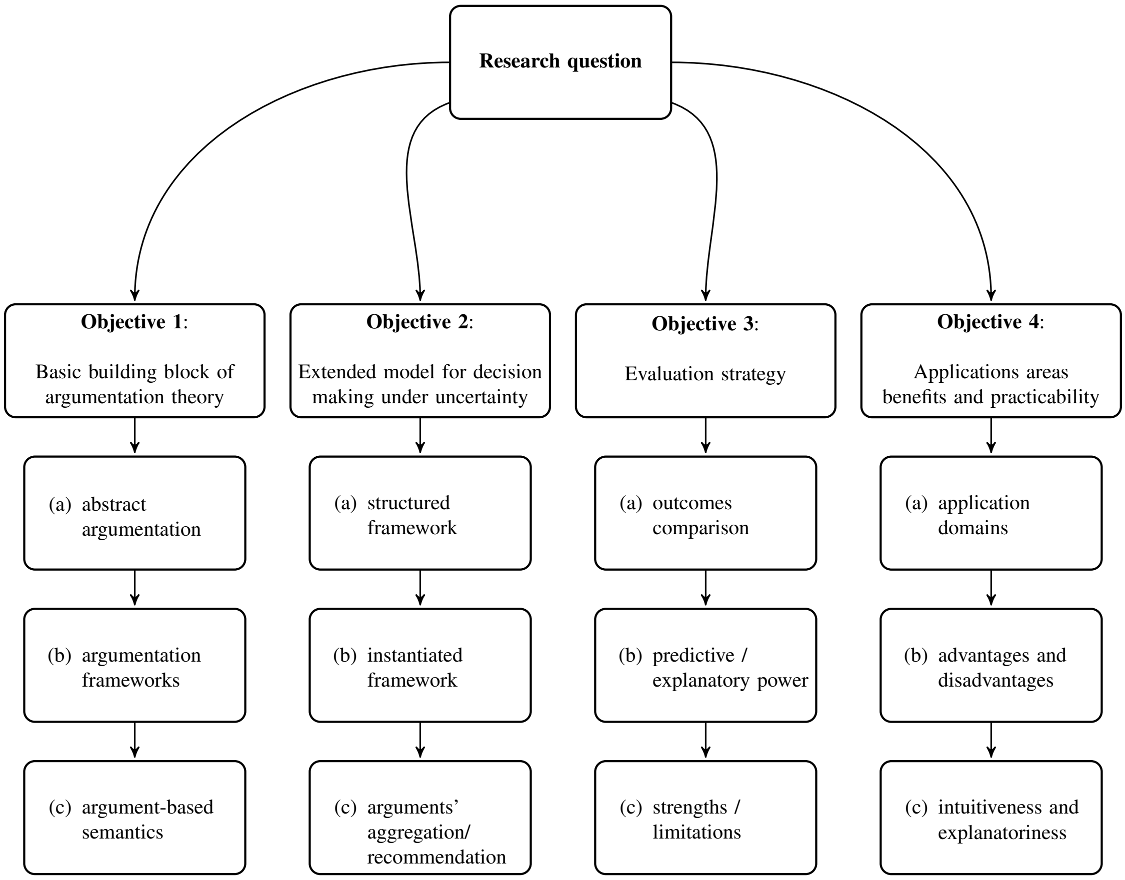


Figure 3.1. Research question and research objectives

Chapter 4

Implementation

4.1 Abstract Argumentation Theory

The first objective of this thesis is to introduce the basic building blocks of argumentation theory, as stated in objective 1. In particular, for addressing the objective 1 (a), we describe the work of Dung on Abstract Argumentation Theory (AAT) [Dung, 1995], upon which modern and current implementations of arguments-based systems are built. His work, historically speaking, derives from other more practical and concrete works on argumentation and defeasible reasoning, such as [Vreeswijk, 1993] and [Pollock, 1987] [Pollock, 1994]. Dung's implementation of abstract argumentation was and is still a success due to the fact that it provides a way, applicable to all type of systems that instantiate his framework, for assigning justification statuses to arguments. It is useful to mention that Dung-style argumentation approaches, contrary to, for instance, standard logic approaches, are not based on the notion of *truth*. These approaches formalise reasoning processes that are defeasible in nature, and are not concerned with truth of propositions, rather they focus on accepting a proposition as true. Dung's frameworks allow comparisons among different systems by translating them into his abstract format [Vreeswijk, 1993]. This property was a breakthrough because it showed how several logics fro non-monotonic reasoning can be translated into his abstract framework. The underlying idea behind argumentation theory is that, given a set of abstract arguments, where some of them defeat (attack) others, a decision is to be taken to determine which arguments can ultimately be accepted. Merely looking at an argument's defeaters to determine the acceptability status of an argument is not enough: it is also important to determine whether the defeaters are defeated themselves. We say that an argument B *defeats* an argument A if and only if B is a reason against A. If the internal structure of arguments, as well as the reasons why they defeat each other are not considered, what is remaining is called an *Argumentation Framework (AF)* [Dung, 1995]. To complete the research objective 1 (b) we now provide a more precise definition of such a framework.

4.1.1 Argumentation framework

An *argumentation framework* is a set of (abstract) arguments and binary attack (defeat) relations among these arguments. It essentially is a directed graph in which arguments are presented as nodes and the attacks as arrows (figure 4.1).

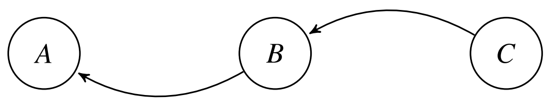


Figure 4.1. Argument and reinstatement

Definition 1 (Argumentation Framework)
An argumentation framework is a pair $AF = (Ar, def)$ where Ar is a set of arguments and $def \subseteq Ar \times Ar$. We say that A defeats B iff (A, B) $\in def$.

The question is which arguments should ultimately be accepted. In figure 4.1, A is defeated by B, and apparently A should not be accepted since it has a counterargument. However, B is itself defeated by C that is not defeated by anything, thus C should be accepted. But if C is accepted, then B is ultimately rejected and does not form a reason against A anymore. Therefore A should be accepted as well. In this situation we say that C *reinstates* A. Due to this issue of *reinstatement*, a formal criterion that determines which of the arguments of an AF can be accepted, is needed, as stated in the objective 1 (c). In the literature of argumentation theory, this criterion is referred to as *semantic*, and given an AF, it specifies zero or more sets of acceptable arguments, called *extensions*.

4.1.2 Argument-based semantics

Various argument-based semantics have been proposed [Baroni et al., 2011], but here we focus on complete, grounded and preferred semantics as proposed in [Dung, 1995]. We start clarifying the issue of argument semantics using the labeling approach by Wu and Caminada [Wu et al., 2010].

Each Argument is either *in*, *out* or *undec* according to the following conditions:

- an argument is labelled *in* if and only if all its defeaters are labelled *out*, and
- an argument is labelled *out* if and only if it has at least one defeater labelled *in*.

Informally, an argument can have the label *in* if one has accepted the argument, *out* if rejected and *undec* if one abstains from taking a position on either accepting or rejecting it.

Definition 2 (Complete labeling)
Let (Ar, def) *be an argumentation framework and* $Lab : Ar \rightarrow \{in, out, undec\}$ *be a total function. We say that Lab is a complete labeling iff it holds:*

- *if* $Lab(A) = in$, $\forall B \in Ar : (B\ def A \supset Lab(B) = out)$

- *if* $Lab(A) = out$, $\exists B \in Ar : (B\ def A \wedge Lab(B) = in)$

- *if* $Lab(A) = undec$, $\neg\forall B \in Ar : (B\ def A \supset Lab(B) = out)$ *&*
  $\neg\exists B \in Ar : (B\ def A \wedge lab(B) = in)$.

Example 3

A concrete version of the argumentation framework of figure 4.1 could concern a reasoning process to predict recurrence of a removed breast cancer from a patient:

(A) the size of the exported breast cancer was high, therefore recurrence is believed to be high;

(B) however, the patient is in pre-menopausal status, so recurrence is low;

(C) but the lymph nodes involvement was high, so the risk of recurrence is high as well.

This example can be interpreted as follows. For C it holds that all its defeaters are labelled *out* (trivial as C does not have any defeater), therefore C must be labelled *in*. B now has a defeater labelled *in* therefore it must be labelled *out*. For A, it holds that all its defeaters are labelled *out*, so it must be labelled *in*. Thus the result is $Lab(A) = Lab(C) = in$ and $Lab(B) = out$. In other words, arguments A and C can be accepted while argument B is rejected. This means that the recurrence of the breast cancer is high.

Definition 3 (Abbreviations)
Let (Ar, def) *be an argumentation framework and let* $A \in Ar$ *and* $Args \subseteq Ar$*.*
For abbreviation:

- A^+ *as* $\{B | A\ def B\}$

- $Args^+$ *as* $\{B|A\ def B\ for\ some\ A \in Args\}$

- A^- *as* $\{B|B\ def A\}$

- $Args^-$ *as* $\{B|B\ def A\ for\ some\ A \in Args\}$

A^+ indicates the arguments defeated by A, A^- indicates the arguments that defeat A. $Args^+$ refers to the arguments that are defeated by the set of arguments $Args^+$ while $Args^-$ refers to the arguments that defeat the set of arguments $Args^-$.

A set of arguments is called *conflict-free* if and only if it does not contain any argument A and B such that A defeats B. A set of arguments *Args* is said to *defend* an argument C if and only if each defeater of C is defeated by an argument in *Args*.

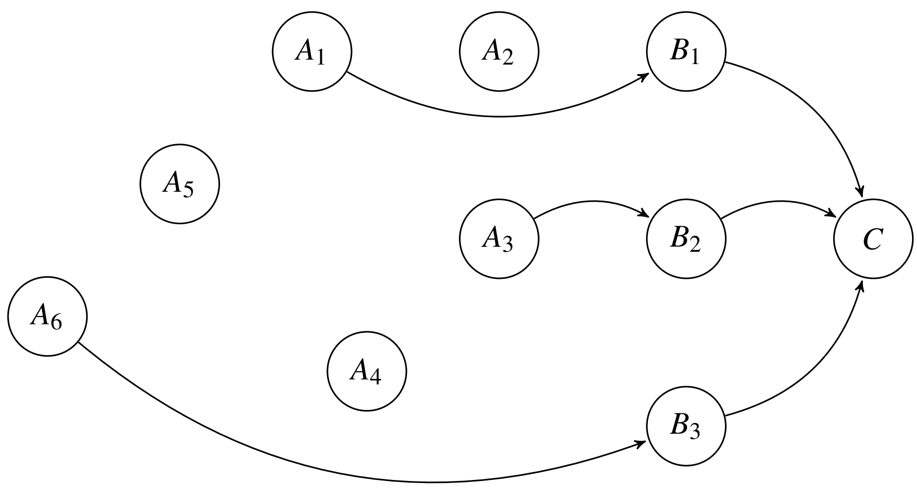


Figure 4.2. The set of arguments $Args = [A_1, ..., A_6]$ defends argument C

Definition 4 (Conflict-freeness)
Let (Ar, def) be an argumentation framework and let $Args \subseteq Ar$. Args is said to be conflict-free iff $Args \cap Args^+ = \emptyset$.

Definition 5 (Defence)
Let (Ar, def) *be an argumentation framework and let Args* $\subseteq Ar$ *and* $B \in Ar$*. Args is said to defend B iff* $B^- \subseteq Args^+$.

Definition 6 (Defence of arguments)
Let (Ar, def) *be an argumentation framework and let Args* $\subseteq Ar$ *and* $B \in Ar$*. We introduce a function* $F : 2^{Ar} \rightarrow 2^{Ar}$ *such that F(Args) =* $\{A | A \textit{ is defended by Args}\}$.

F yields the arguments defended by a given set of arguments. It specifies the set of arguments that are acceptable, in line with Dung [Dung, 1995].

Having introduced the preliminary notions behind AAT we are now able to introduce the first extension-based semantic, according to objective 1 (c). This is the well-known *complete semantic* proposed by Dung [Dung, 1995].

Definition 7 (Complete extension)
Let (Ar, def) *be an argumentation framework and let Args be a conflict-free set of Arguments. Args is said to be a* complete extension *iff* $Args = F(Args)$.

Example 4

In the AF of figure 4.1 there is just one complete extension, $\{A, C\}$, which is conflict-free and defends exactly itself. Note $\{A, B, C\}$ is also a fixpoint of F, but it is not a complete extension since it is not conflict-free.

The idea of complete extensions is that a complete labeling can be seen as a subjective but reasonable point of view that an agent can take with respect to which arguments are *in*, *out* and *undec*. Each point of view is internally coherent and if questioned, the agent can use its own position to defend itself. Someone can disagree with that position, but can not point out an internal inconsistency.

Eventually, the set of all the complete labelings represent all the possible and reasonable positions an agent can take [Wu et al., 2010].

4.1.3 Skeptical approach

Complete semantics have a important property: more than one complete extension may exist. However, sometimes it is advantageous to apply a semantic that is guaranteed to yield exactly one extension: the grounded semantic. The idea is to select the complete labeling *Lab* in which the set of *in*-labelled arguments is minimal.

Definition 8 (Grounded Extension)
Let (Ar, def) be an argumentation framework. The grounded extension is the minimal fixpoint of F.

Note 1
The grounded extension coincides with the complete labelling in which *in* is minimised, *out* is minimised and *undec* is maximised.

Example 5

In the AF of figure 4.1, the grounded extension is $\{A, C\}$.

4.1.4 Credulous approach

Grounded semantics have the advantage that there always exists exactly one grounded extension (it can be the empty set), however this approach is a *skeptical* approach. Preferred semantics deals with this limitation. The idea is that preferred semantics, instead of maximizing *undec* arguments, maximises *in* arguments (and also *out*). They are based on the concept of admissibility. A set of arguments is admissible if and only if it is conflict-free and defends at least itself.

Definition 9 (Admissibility)
Let (Ar, def) be an argumentation framework and let $Args \subseteq Ar$. $Args$ is said to be admissible iff $Args$ is conflict-free and $Args \subseteq F(Args)$.

We are now able to define the concept of preferred extension as follows.

Definition 10 (Preferred extension)
Let (Ar, def) be an argumentation framework and $Args \subseteq Ar$. $Args$ is said to be a preferred extension iff $Args$ is a maximal admissible set.

Note 2
The empty set is admissible in every AF as it is conflict-free and trivially defends itself against each of its (none) defeaters. For any AF, there exist at least one preferred extension. Every grounded and every preferred extension is a complete extension.

Example 6

The admissible sets are $\{C\}$, $\{A,C\}$. $\{B\}$ and $\{A\}$ are not admissible as they do not defend themselves respectively against C and B. Only one preferred extension exits: $\{A,C\}$.

4.2 An extended model for decision-making support

Abstract argumentation theory is an abstract way of representing mutual relationships among arguments and studying their consistency by applying argument-based semantics. The paradigm is really useful for investigating which arguments can be seen as justifiable in a defeasible reasoning process. Although really powerful, it considers abstract arguments and abstract defeat relationships, failing to fully model real-world scenarios, characterised by tangible information and evidence. As stated in objective 2, the next goal concerns the extension of an abstract model of arguments to a more structured model capable of representing real-world arguments and defeat relationships. The aim is to produce a tool for supporting real-world decision making processes in health contexts.

In the next sections, the focus is to internally define arguments and to describe different typologies of attack relationships among them. Subsequently, reusing the argument-based semantic proposed by abstract argumentation, the goal is to design a strategy for selecting the extension of arguments that minimizes the overall degree of uncertainty. In particular, our model handles two typologies of uncertainty:

- *vagueness* in defining a concept of the knowledge-base;
- *uncertainty* in defining arguments as a form of defeasible inference rules;

During these steps, we clarify formal definitions with real-world evidence considering a well-known health care problem: the breast cancer recurrence prognonis/prediction. As mentioned in the introduction chapter, recurrence is a phenomenon that defines the return of a cancer after treatments and a temporal interval during which its detection was not possible. Predicting recurrence is important for assisting the identification of patients with critical prognosis and minimizing unnecessary therapies. This example does not mean to be fully exhaustive and totally correct, rather it is a illustrative way of elucidating formal definitions. Moreover, the rationale behind considering this problem is that a dataset of tangible evidence exists in the literature, with real-world data of patients that had breast cancer surgery. However, the features considered in the dataset are outdated. The availability of this dataset will serve as a basis for comparisons between the argument-based model designed in this thesis and some of the approaches described in chapter 2.

In detail, in order to deal with objective 2, the aims of our solution are:

a to translate the knowledge-base of an expert into a set of structured defeasible arguments and defeat relationships among them; this will produce a structured argumentation framework (Objective 2 a);

b to run argumentation semantics for extracting consistent sets of arguments (Objective 2 b);

c to recommend the set of arguments (recommended option/alternative) that minimises uncertainty (Objective 2 c);

Our proposal is to implement step a) adopting fuzzy sets and degrees of truth for arguments, to execute step b) by using grounded and preferred semantics from abstract argumentation theory and to recommend a set c) that minimises the uncertainty of its arguments.

4.2.1 Incremental implementation through an illustrative case study: Breast cancer recurrence

In this study, we have used the Ljubljana breast cancer dataset[1], repeatedly used in many machine learning studies (1986 up to 2011. Examples: [Williams and Williamson, 2006] [Clark and Niblett, 1987] [Tan and Eshelman, 1988] [Cestnik et al., 1987]). It includes 286 instances of real patients who went through a breast cancer operation. 9 were incomplete and the others are described by 9 possible predictive attributes and a binary outcome class (recurrence or no recurrence). The values (numerical [N] or categorical [C]) were recorded at time of diagnosis or after operation, as shown in Table 4.1. For 81 patients, the illness reappeared after 5 years, while 196 did not have recurrence. The evidence within the dataset was verified by clinicians after collection. In the next paragraphs we design our AT model for this dataset

[1] provided by the University Medical Center, Institute of Oncology, Ljubljana, ex-Yugoslavia, by M. Zwitter and M. Soklic (physicians)

Attribute	Dataset Range
Age	10-19, 20-29, .. , 90-99
Menopausal	lt40, ge40, premeno
Tumor size	0-4, 5-9, ... , 55-59
Node involvement	0-2, 3-5, ..., 36-39
Degree of malignancy	1, 2, 3
Breast	left, right
Breast quadrant	left-up, left-low, right-up, right-low, central
Irradiation	yes, no
Outcome (class)	Recurrence (R), no-Recurrence (NR)

Table 4.1. Dataset attributes of the Ljubljana breast cancer dataset

providing the rationale behind each design decision. It is important to note that an arguments-based model is not derived from a dataset, but from expert knowledge. Throughout the text we use the following notations:

- D is the domain/context considered by an agent;
- Σ is the knowledge-base of an agent;
- P is the problem that needs to be solved with a rational decision by an agent in a domain D according to its knowledge-base Σ.
- α is an attribute, that means a concept/characteristic considered in Σ. For disambiguation, we refer to α^{lab} as the name/label of the attribute, and to α^{val} as the objective value for that attribute.
- c is a claim. In other words, it represents an option, a possible alternative or a possible outcome.

Example 7

In the breast cancer prognosis/recurrence domain:

- $D =$ breast cancer

- P = cancer recurrence prediction.
- Σ contains t attributes, evidence and their mutual interactions (for breast cancer recurrence prediction).
- A possible set of attributes considered in an agent's knowledge base are in table 4.1. For instance, $\alpha_1^{lab} = age$ and $\alpha_1^{val} \in [10..19, 20..29, ., 90..99]$.
- the claims are R (recurrence) and NR (no-recurrence)

4.2.2 Fuzzy membership functions and degrees of truth

To describe an attribute we propose to use fuzzy set theory and degrees of truth [Zadeh, 1965] [Zadeh, 1966]. The rationale behind adopting fuzzy set theory (membership functions) is that commonly people tend to perceive and describe an attribute as being 'low' or 'high', 'small' or 'big' and so forth. Sometimes intermediate clusters are used ('medium'), but generally, in informal reasoning, more complex scales are avoided. Furthermore, 10 is commonly perceived being a 'low' age ('young') and 90 being 'high' ('old'). However, 50 can be considered 'high' by a subject, and 'low' by another one, with different degrees of truth. This subjectiveness is influenced by various factors (past experience, knowledge, context). Membership functions are useful for modeling vaguely defined sets and human reasoning that is approximate rather than fixed and exact. Therefore the main rationale in modeling attributes (concepts) with fuzzy set and membership functions is represented by the facility with which humans (domain experts) can describe their knowledge in a more formal and structured way. Membership functions allow the mapping of an attribute's value to the relative set with degrees of truth. For completeness we remind the reader of the definition of membership functions from fuzzy logic as well as some properties.

Attribute	Agent's Knowledge-base - Expert's description
Age	The strongest risk factor for breast cancer is age: the older the woman, the higher the risk of cancer (and presumably recurrence).
Menopausal	Pre-menopausal status is a reason to believe recurrence is not likely.
Tumor size	In general, the greatest diameter (in mm) of the excised tumor, the greater the chance of recurrence.
Node involvement	Since the axillary lymph nodes act as a primary site of drainage for the breast, they represent common site of early metastasis. The more lymph nodes involved are, the more likely recurrence is. This is probably the most influential factor for recurrence.
Node capsular invasion	If the cancer does metastasis to a lymph node, even if outside the original tumor site, it can remain 'contained' by the lymph node's capsule. However, the tumor may replace the lymph node and penetrating the capsule, invading the surrounding tissues. If capsular invasion, it makes sense that recurrence is more likely.
Degree of malignancy	The histological grade of the tumor affect recurrence. If it is 1, that means tumors consist of cells that, while neoplastic, retain many of their usual characteristics, recurrence is less likely. If it is 2 or 3, tumors consists of highly abnormal cells, with marked variation in cell size, or a high index of mitotic activity in the cells, therefore making recurrence more likely.
Breast	Although breast cancer can occur in either breast, there is no difference in incidence between breasts. It seems a slightly higher (but unexplained) risk of breast cancer, on the left side, exists.
Breast quadrant	The breast may be divided in four quadrant, using the nipple as a central point. Breast cancer more often occurs in the upper outer quadrant, and as a consequence this increase the chance of recurrence.
Irradiation	Radiotherapy for breast cancer reduces recurrence
Outcome (class)	Reappearance of cancer after 5 years No Reappearance of cancer after 5 years

Table 4.2. Dataset attributes and a possible expert's knowledge-base vague description

Definition 11 (Membership function)

For any set X, a membership function on X is any function $f : X \rightarrow [0,1] \in \Re$. Membership functions on X represent fuzzy subsets of X. For an element x of X, the value $f(x)$ is called the membership degree of x in the fuzzy set and quantifies the grade of membership of x to the fuzzy set X.

We indicate $MF_X = \{f | f : X \rightarrow [0,1] \in \Re\ \}$ as the set of membership functions defined over X.

Note 3

For each fuzzy set X, zero or more membership functions can be defined and they can partially overlap, sharing some values of X, but not necessarily returning the same degree of truth for the same input x. A membership value of 0 and 1 indicate respectively non-membership and fully membership: intermediate values refers to fuzzy members partially belonging to X.

Example 8

The attributes in table 4.2 or figures 4.4-4.19, represent fuzzy sets. For the attribute 'Age', for instance, the membership degree $f_{Age}^{Low}(x)$ quantifies the grade of membership of x to the fuzzy subset 'Low' of 'Age'. Possible membership functions can be associated with the fuzzy subset 'low' of the fuzzy set 'Age', as in figure 4.3. Here, an age of 39 has three different degrees of truth, according to the three different functions.

Note 4

A fuzzy membership function can be any function, from the classical straight line to the step-function or complex functions such as sigmoidal functions, logarithmic functions or curves in general. This property provides a designer with a flexible tool for modeling an attribute (concept) and relative sub-classes. In

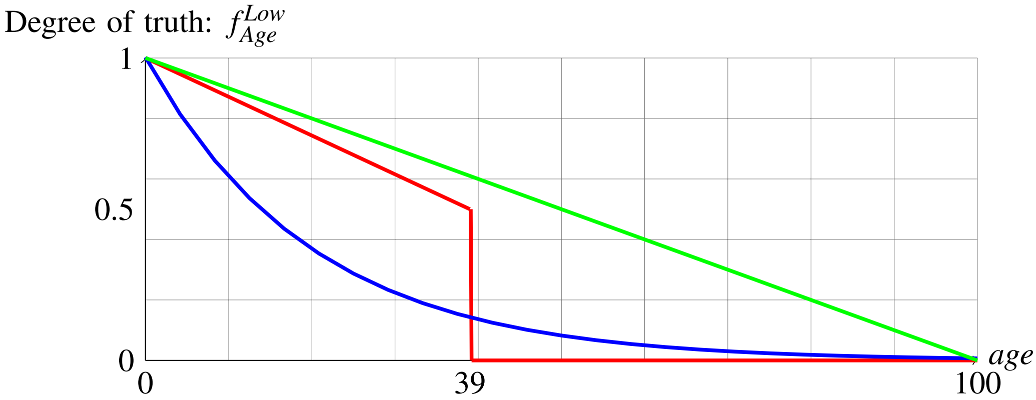


Figure 4.3. Possible membership functions definition for the fuzzy set 'Age' and fuzzy subset 'low'

addition, membership functions are context-aware and user-centered. In other words, they handles the subjectivity of the user in modeling a certain attribute along with the context in which the attribute is taken into consideration. An age of 39 can be considered low with a degree of truth of 0.2 in the breast cancer prognosis context, but considered low with a degree of truth of 0.6 in context of learning (at the age of 39 a user learn in general faster than at the age of 10).

Example 9

In the domain D of breast cancer recurrence, we have interviewed an expert who provided her definition of attributes, via membership functions, as depicted in the following figures:

- Age: low (fig. 4.4), medium (fig. 4.5), high (fig. 4.6);
- Menopause: pre (fig. 4.7), $lt < 40$ (fig. 4.8), $lt > 40$ (fig. 4.9);
- Tumor size: low (fig. 4.10), high (fig. 4.11);

- Node involvement: low (fig. 4.12), high (fig. 4.13);
- Node caps: true (fig. 4.14), false (fig. 4.15);
- Degree of malignancy: low (fig. 4.16), high (fig. 4.17);
- Breast quadrant: lower (fig. 4.18), upper (fig. 4.19);

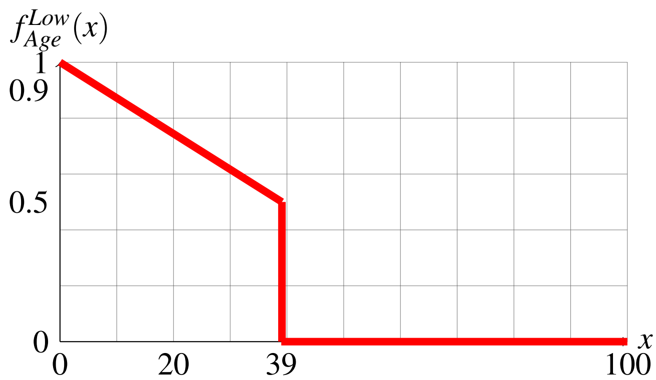


Figure 4.4. Expert definition of membership function 'low' for the attribute 'age'

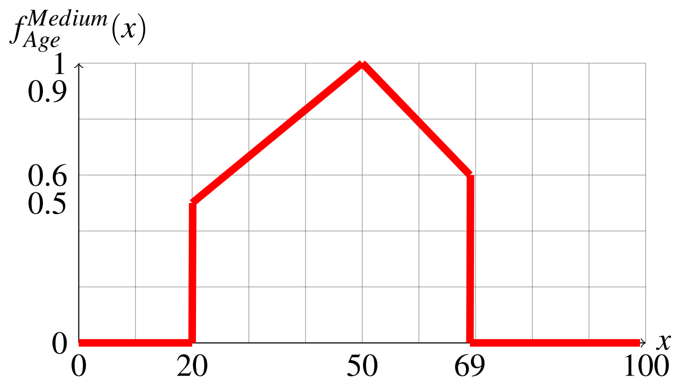


Figure 4.5. Expert definition of membership function 'medium' for the attribute 'age'

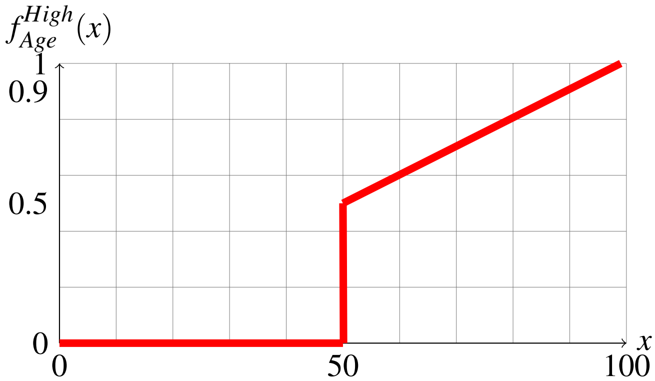


Figure 4.6. Expert definition of membership function 'high' for the attribute 'age'

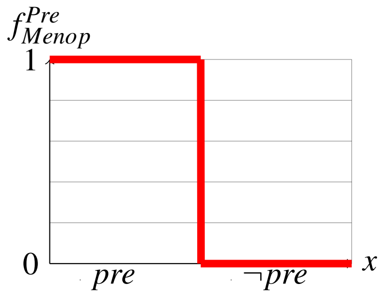


Figure 4.7. Expert definition of membership function 'pre' for the attribute 'menopause'

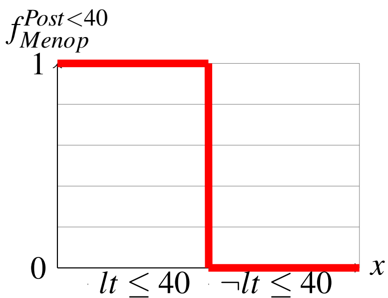


Figure 4.8. Expert definition of membership function '$lt < 40$' for the attribute 'menopause'

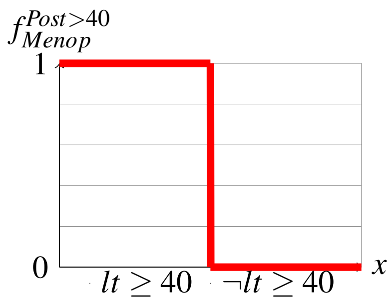


Figure 4.9. Expert definition of membership function '$lt > 40$' for the attribute 'menopause'

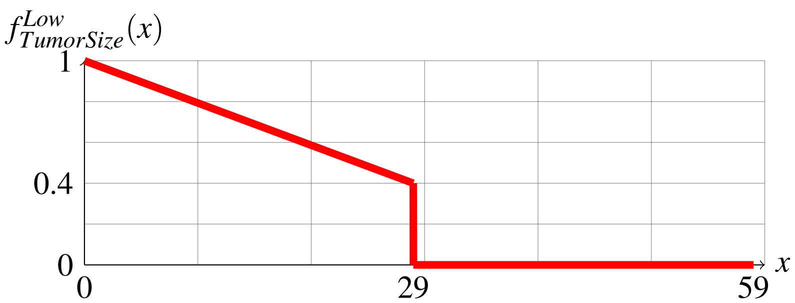


Figure 4.10. Expert definition of membership function 'low' for the attribute 'tumorsize'

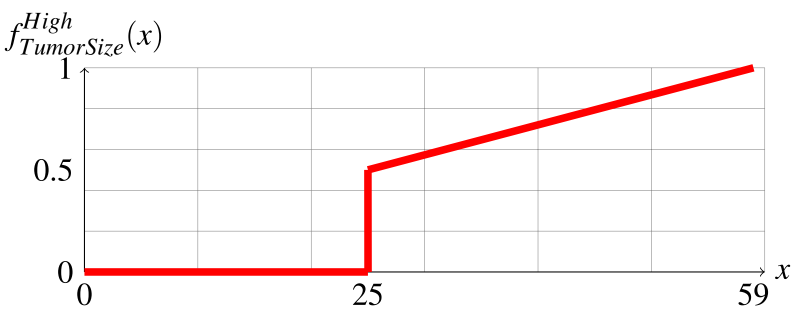


Figure 4.11. Expert definition of membership function 'high' for the attribute 'tumorsize'

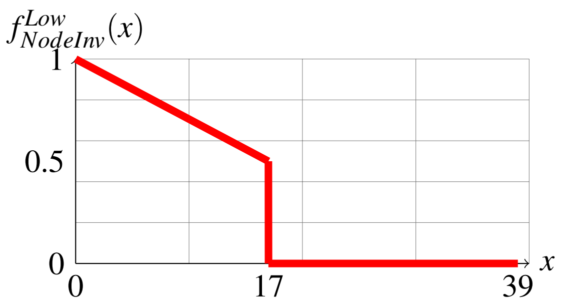


Figure 4.12. Expert definition of membership function ‘low’ for the attribute ‘node involvement’

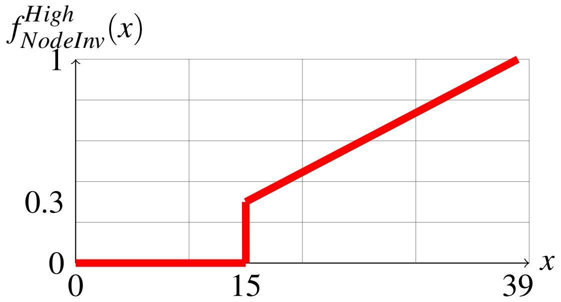


Figure 4.13. Expert definition of membership function ‘high’ for the attribute ‘node involvement’

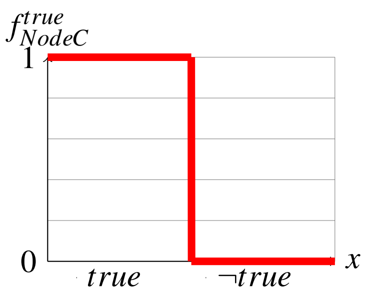


Figure 4.14. Expert definition of membership function ‘true’ for the attribute ‘node caps’

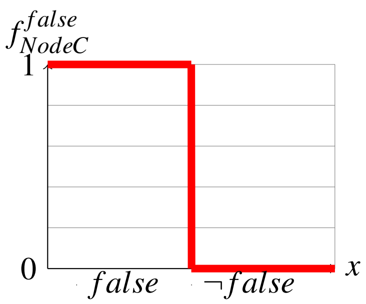


Figure 4.15. Expert definition of membership function 'false' for the attribute 'node caps'

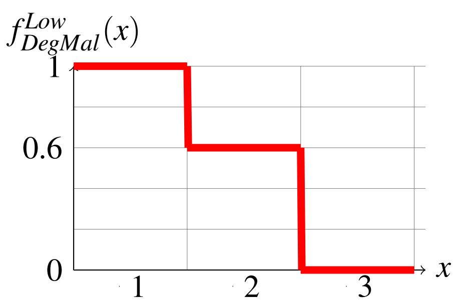


Figure 4.16. Expert definition of membership function 'low' for the attribute 'degree of malignancy'

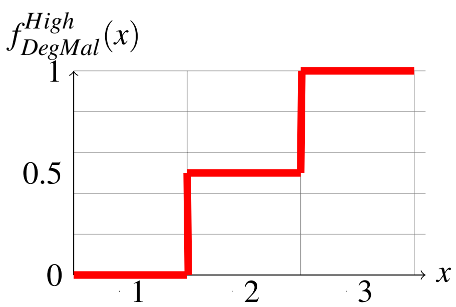


Figure 4.17. Expert definition of membership function 'high' for the attribute 'degree of malignancy'

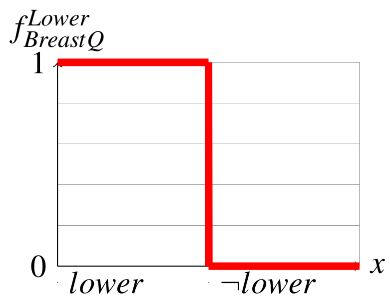


Figure 4.18. Expert definition of membership function 'lower' for the attribute 'breast quadrant'

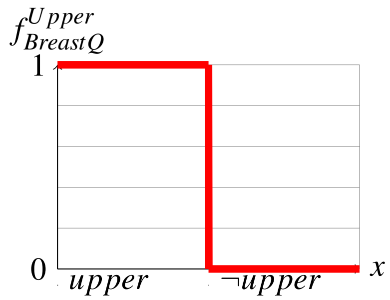


Figure 4.19. Expert definition of membership function 'upper' for the attribute 'breast quadrant'

4.3 From knowledge-base to contextual argumentation framework

4.3.1 Modeling arguments

To translate a knowledge-base into arguments it is necessary to define an argument. Informally, an argument is a defeasible rule (open to defeats) composed by a set of premises and a claim. In other words, from the premises a claim can be inferred.

$$Argument : premises \rightarrow claim$$

This process is intrinsically uncertain, because the inference rules are defeasible in nature, and not strict and totally certain. This is coherent with human reasoning that is uncertain rather than exact. To keep it simple, we propose to use arguments with just one premise and one claim. Here we associate an attribute (for example 'age') as the premise of an argument followed by a claim which is a possible conclusion available to a clinician (Recurrence R or no-Recurrence NR) to support decisions. Therefore a possible internal structure of an argument can be:

$$Low\ Age \rightarrow no - Recurrence(NR)$$

In particular, we propose to model the linguistic variable 'Low Age' as a fuzzy set, exactly as described in the sub-section 4.2.2.

Note 5

In the literature of argumentation theory, an argument can sometime be expressed as a combination of more premises as:

$$ar : (high\ tumor\ size)\ AND\ (low\ age) \rightarrow R$$

The main rationale behind using the first representation is to use just one premise for every argument. The decision is to keep an argument in its basic representation instead of embedding two premises supporting the same claim within the same argument. Furthermore, we adopt this representation because a potential attack can refer to just one premise, and not to both the premises.

According to the literature of argumentation theory, arguments can be considered *forecast* when they are in favor or against a certain claim (but justification is not infallible), and *mitigation arguments*, when defeating (undermining justification for) forecast or other mitigation arguments [Matt et al., 2010].

Definition 12 (Argument - Forecast)
A forecast argument β *is defined over a membership function* f_α *for attribute* α *and a claim c.* $ARG_F : MF_{ATTR} \times C$ *and "*$\beta : f_\alpha \rightarrow c$*" can be read as an argument* β *stating that '*f_α *implies c', or 'there is a reason to believe c from* f_α*' or 'c is what reasonably follows from* f_α*.'*

Example 10

As an example,

$$ar : \text{'}old\ age \rightarrow R'$$

is a forecast argument, with 'Old' being a fuzzy subset of 'age' defined by the membership function f_{Age}^{Old} defined over attribute 'age' and being R ('recurrence') the claim that reasonably follows the premise.

Definition 13 (Argument - Mitigation)
A mitigation argument β *is defined over a membership function* f_α *for the attribute* α *and another argument* δ *(either forecast or mitigation).* $ARG_M : MF_{ATTR} \times ARG_F \cup ARG_M$ *and '*$\beta : f_\alpha \rightarrow \neg\delta$*' can be read as an argument* β *stating that '*f_α *implies* $\neg\delta$*', 'there is a reason to believe* $\neg\delta$ *from* f_α*' or 'the justification of* δ *is undermined by* f_α*.'*

Example 11

As an example:

$$ar : high\ tumor\ size \rightarrow \neg(low\ age \rightarrow NR)$$

is a mitigating argument where 'high' is the membership function defined over 'tumor size', 'low' is the membership function defined over 'age', 'no recurrence' is the claim that follows 'low age' that is undermined by 'high tumor size'. The fact that a low age is a reason to believe no recurrence of a breast cancer is undermined by the fact that the tumor size is high.

At this stage, the pool of arguments for modeling the domain D and investigating the problem P emerges, according to the contextual agent's knowledge-base. In other words, the *pool of arguments* is the set of all the arguments (forecast or mitigation), designed by an agent, according to his knowledge base Σ over the domain D. We refer to this as:

$$ARG_{pool} : \{a \mid a \in ARG_F \cup ARG_M\}.$$

Example 12

Let us assume an agent's knowledge-base is the one shown in table 4.2 (as previously mentioned, we have interviewed an expert in the domain of breast cancer), the membership functions (fuzzy subset) of attributes (fuzzy sets) are figures 4.4-4.19 and the arguments are in table 4.3. Membership functions of figures 4.4-4.19 are two or three-valued functions, simple straight lines or steps-functions. However, other typologies of functions might be adopted. We do not provide references to the evidence provided by the expert.

4.3.2 Modeling defeat relations (attacks)

Once the process of representing arguments has been completed, the set of defeat relations (attacks) among arguments can be explicated (according to an agent's knowledge base). Defeating relations can be *rebuttal* or *undercutting*. The former occurs between two *forecast arguments* contradicting each other because they support opposite mutually exclusive claims (bi-directional attack). The latter occurs when a mitigating argument challenges the claim of a forecast or another mitigating argument (uni-directional attack) or is preferred to it.

Definition 14 (Attack - rebutting)
Given two forecast arguments a, b with $a : f_\alpha \rightarrow c_1$*,* $b : f_\beta \rightarrow c_2$ *we say that a attacks b and we indicate thus by* (a,b)*,* iff $c_1 \neq c_2$ *and* c_1 *and* c_2 *are mutually exclusive.* $ATT_R : ARG_F \times ARG_F$*. As a rebuttal attack is symmetrical it holds that iff* (a,b) *then also* $\exists(b,a)$*.*

Note 6
Generally, arguments supporting different claims do not necessarily attack each other. In fact let us consider the following arguments whose premises lead to a specific action:

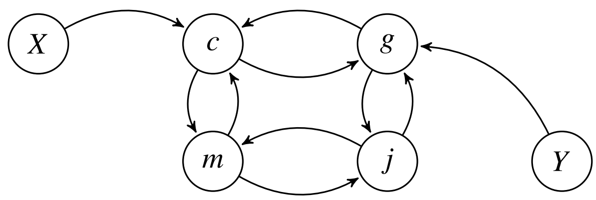
A: 'if sunny' $\rightarrow$ 'go to the beach'
B: 'if windy' $\rightarrow$ 'run in the park'

A and B in this case are not mutually exclusive, and they do not attack each other. In this paper, claims follow a mutual exclusive property. For example in our case study, the claims can be recurrence (R) or (NR), excluding each other. Formally, given two claims c_1 and c_2, with $c_1 \neq c_2$, c_1 logically excludes c_2 and viceversa.

Definition 15 (Attack - undercutting)
We say that a mitigating argument a attacks a forecast or another mitigating argument b, and we indicate thus by (a,b), if in the agent's knowledge base there is evidence suggesting that b is no longer justified because of a. $ATT_U : ARG_M \times (ARG_F \cup ARG_M)$

Example 13

Consider the following subset of arguments from the breast recurrence problem:



- 'c : *high age* $\rightarrow R$'
- 'm : *low degree of malignancy* $\rightarrow NR$'
- 'g : *low tumor size* $\rightarrow NR$'
- 'j : *high node invasivity* $\rightarrow R$'
- 'X : *high node invasivity* $\rightarrow \neg$(*high age* $\rightarrow R$)'
- 'Y : *high tumor size* $\rightarrow \neg$(*low age* $\rightarrow NR$)'

with R for recurrence, NR for no-recurrence. c and j support the same claim (R), so they do not attack each other but both attack and are attacked by m and J, supporting a different claim (NR) (c, j contradict m, J). X is an undercutting argument for c as it undermines its justification. Similarly Y undermines g. The attack relations are thus (c,m), (m,c), (c,g), (g,c), (m,j), (j,m), (g, j), (j,g), (X, c), (Y, g).

Note 7

In (Y, g) we say that Y challenges what is claimed by g, instead in (X, c), X is preferred to c. In fact, in (Y, g), 'high tumor size' invalidates the inference link of g ('g : *low tumor size* $\rightarrow NR$'). In other words, the argument g is not longer justified in the case the tumor size is high. On the other hand, in (X, c), the premise of X ('high node invasivity') invalidates the link of argument c whose claim is 'recurrence'. But having a look at argument j ('j : *high node invasivity* $\rightarrow R$'), the claim is still 'recurrence'. In this case 'high node invasivity' is preferred than 'high age'. The nature of the attack relations, here, remains abstract because their strength is not considered. In other words, if an attack is not symmetrical, the attacked argument will not concur in the final decision as it will not appear in any preferred extensions. Furthermore, in this paper the uncertainty associated with an argument (with the inference rule), that means the strength or agent's belief in the related claim, is not considered, but it represents our future work.

At this stage, the pool of attacks emerges. The pool of the attacks relations is the set of all the rebutting and undercutting attacks built by an agent according to its knowledge base Σ.

$$ATT_{pool} = \{(a,b) \mid (a,b) \in ATT_R \cup ATT_U \text{ and } a,b \in ARG_{pool}.$$

Once the pool of arguments has been defined along with the pool of attack relations a contextual argumentation framework emerges. It represents the knowledge base of an agent contextualised in a formal and organised manner.

Attribute: Age, Evidence: 1			**Attribute: Menopause, Evidence: 2**		
Arg	MF_{Age}	**Claim**	**Arg**	$MF_{Menopause}$	**Claim**
a	low	no rec.	d	pre	no rec.
b	medium	no rec.	e	post-lt40	no rec.
c	high	rec.	f	post-gt40	rec.
Attribute: Tumor size, Evidence: 3			**Attribute: Node invas., Evidence: 4**		
Arg	$MF_{TumorSize}$	**Claim**	**Arg**	$MF_{NodeInv}$	**Claim**
g	low	no rec.	i	low	no rec.
h	high	rec.	j	high	rec.
Attribute: Node Caps, Evidence: 5			**Attribute: Deg. Malig., Evidence: 6**		
Arg	MF_{NodeC}	**Claim**	**Arg**	$MF_{DegMalig}$	**Claim**
k	true	rec.	m	low	no rec.
l	false	no rec.	n	high	rec.
Attribute: Breast, Evidence: 7			**Attribute: Breast quad, Evidence: 8**		
Available evidence suggests that the attribute breast is not influential, thus no argument is built			**Arg**	$MF_{BreastQ}$	**Claim**
			o	lower	no rec.
			p	upper	rec.

Table 4.3. Pool of arguments - illustrative scenario

q: High Age $\rightarrow \neg$ (Menop Pre$\rightarrow$NR)
r: High Age $\rightarrow \neg$ (Lower BreastQ$\rightarrow$NR)
s: High Age $\rightarrow \neg$ (Upper BreastQ$\rightarrow$R)
t: High NodeInv $\rightarrow \neg$ (Low age$\rightarrow$NR)
u: High TumorSize $\rightarrow \neg$ (Low age$\rightarrow$NR)
v: High NodeInv $\rightarrow \neg$ (Lower BreastQ$\rightarrow$NR)
w: High NodeInv $\rightarrow \neg$ (Upper BreastQ$\rightarrow$R)
x: High Tumorsize $\rightarrow \neg$ (Lower BreastQ$\rightarrow$NR)
y: High Age $\rightarrow \neg$ (Upper BreastQ$\rightarrow$R)
z: High TumorSize $\rightarrow \neg$ (Upper BreastQ$\rightarrow$R)

Table 4.4. Undercutting attacks: illustrative scenario

Once arguments, either forecasting or mitigating have been designed along with rebuttal or undercutting defeat relations among them, a contextual framework emerges. Formally, a contextual argumentation framework is a pair:

$$CAF = (ARG_{pool}, ATT_{pool})$$

where ARG_{pool} and ATT_{pool} are respectively the set of arguments and the set of attack relations as in definition 4.3.1 and 4.3.2.

Example 14

Considering the pool of arguments built in table 4.3 and the undercutting attack relations of table 4.4 (provided by the expert), the contextual argumentation frameworks that emerges is:

- $ARG_F = \{a,b,c,d,e,f,g,h,i,j,k,l,m,n,o,p\}$
- $ARG_M = \{q,r,s,t,u,v,w,x,y,z\}$
- $ARG_{pool} = ARG_F \cup ARG_M$
- $ATT_R = \{(\alpha,\beta) \mid (\alpha,\beta)\ \forall\alpha,\beta \in ARG_F$ and '$\alpha : f_\alpha \rightarrow c_1$ ', '$\beta : f_\beta \rightarrow c_2$ ', $c_1 \neq c_2\}$

- $ATT_U = \{(q,d),(r,o),(s,p),(t,a),(u,a),(v,o),(w,p),(x,o),\ (y,p),(z,p)\}$
- $ATT_{pool} = ATT_R \cup ATT_U$
- $CAF = (ARG_{pool}, ATT_{pool})$

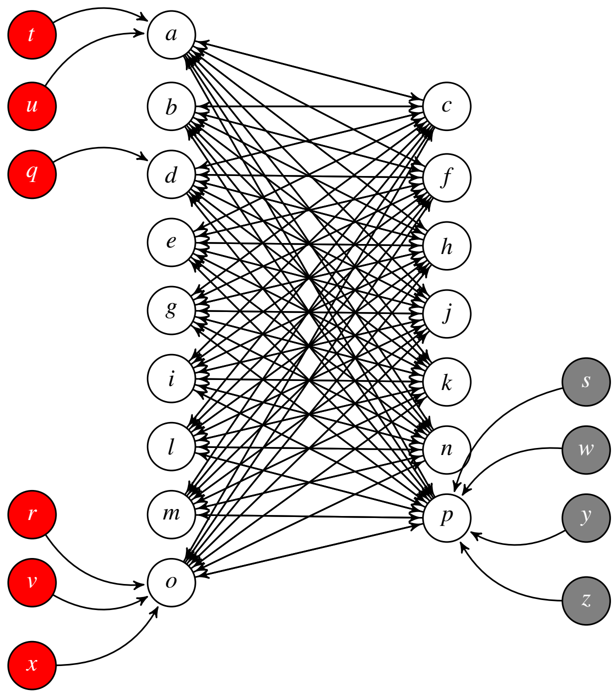


Figure 4.20. Contextual argumentation framework emerged by expert's knowledge-base

Plain white nodes represent forecast arguments, with symmetrical straight black arrows indicating rebuttal attack relationships. Red and gray nodes represent mitigating arguments, with asymmetrical curvy black arrows representing undercutting attack relationships. Red arguments (left side - t, u, q, r, v, x) challenge forecast arguments while gray arguments (right side - s, w, y, z) indicate a preference over forecast arguments (their difference is in note 7).

4.4 Instantiating argumentation frameworks

Once the knowledge base of an agent has been fully modeled through the use of arguments and attacks relations (contextual argumentation framework), it is necessary to individuate which arguments and attacks are objectively activated in a given practical decision-making scenario, as stated in objective 2 (b). A forecast argument is activated if the membership function contained in its premises returns a non-zero value. Similarly, not all the mitigating arguments are activated. A mitigating argument *a* is activated if and only if the membership function contained in its premises returns a non-zero value *and* if and only if the argument defeated by *a* is in the set of activated forecast arguments. In other words, if the mitigating attacker (premise) is not activated, or if the defeated argument is not activated (claim), the existence of the whole mitigating argument does not make sense, thus it is not activated.

4.4.1 Activating arguments

Definition 16 (Activated forecast arguments)
The set of activated forecast arguments is a subset of the pool of arguments: $ARG_F^{Act} \subseteq ARG_{pool}$. For a given argument A defined over the attribute α, a membership function f_α, and a objective value α^{val}, $A \in ARG_F^{Act}\ \ iff\ f_\alpha(\alpha^{val}) > 0$.

Definition 17 (Activated mitigating arguments)
The set of activated mitigating arguments is a subset of the pool of arguments: $ARG_M^{Act} \subseteq ARG_{pool}$. For a given mitigating argument $B : f_\alpha \rightarrow \neg\delta$, defined over the attribute α, a membership function f_α, and a objective value α^{val}, $B \in ARG_M^{Act}\ \ iff\ \ \delta \in ARG_F^{Act}$ and $iff\ f_\alpha(\alpha^{val}) > 0$

4.4.2 Activating defeat relations

The same principle is applied to rebutting and undercutting attacks. A rebutting attack is activated if and only if both the attacker and the attacked are in the set of activated forecast arguments. An undercutting attack (a,b) is activated if and only if the mitigating argument a is in the set of the activated mitigating arguments. Note that from the definition of activated mitigating argument (17) we already know that the attacked argument b is forecast and it is within the set of activated forecast argument. In other words we already know that b exists because if not, the mitigating argument a would not exist. From a mitigating argument we can extract the related undercutting attack.

Definition 18 (Activated rebutting attacks)
The set of activated rebutting attacks is a subset of the pool of attacks: $ATT_R^{ACT} \subseteq ATT_{Pool}$, $(a,b) \in ATT_R^{ACT}\ \ iff\ \ a,b \in ARG_F^{Act}$.

Definition 19 (Activated undercutting attacks)
The set of activated undercutting attacks is a subset of the pool of attacks: $ATT_U^{ACT} \subseteq ATT_{Pool}$ *and* $(a,b) \in ATT_U^{ACT}\ \ iff\ \ a \in ARG_M^{Act}$.

At this stage, the *instantiated argumentation framework* emerges, which is basically a sub-contextual argumentation framework. Formally, an instantiated argumentation framework is a pair:

$$IAF = (ARG_M^{Act} \cup ARG_F^{Act}, ATT_R^{Act} \cup ATT_U^{Act})$$

where ARG_M^{Act} and ARG_F^{Act} are respectively the set of activated forecast and mitigation arguments, ATT_R^{Act}, ATT_U^{Act} are respectively the set of activated rebutting and undercutting attack relations.

Example 15

Let us consider a record of the Ljubljana breast cancer dataset (as in table 4.2), related to a patient, with the following values:

- Age=$40-49$
- Menopause=premeno
- Tumor-size=$30-34$
- Inv-nodes=$0-2$
- Node-caps=no
- Deg-malign=2
- Breast=right
- Breast-quad=right_low
- Irradiation=no

For 'age', 'Tumor-size', 'Inv-nodes', we take respectively the center of each interval, thus 44.5, 32, 1. The membership functions that return degrees of truth greater than zero are:

- f_{age}^{medium}
- f_{Menop}^{Pre}
- $f_{TumorSize}^{High}$
- $f_{NodeInv}^{Low}$
- f_{NodeC}^{False}
- $f_{DegMalig}^{Low}$
- $f_{DegMalig}^{High}$
- $f_{BreastQ}^{Lower}$

As a consequence, the activated arguments and attacks relationships, and the final activated argumentation frameworks are:

- $ARG_F^{Act} = \{b,d,h,i,l,m,n,o\}$
- $ARG_M^{Act} = \{x\}$
- $ATT_R^{Act} = \{(b,h),(h,b),(b,n),(n,b),(d,h),(h,d),(d,n),(n,d),$
  $(i,h),(h,i),(i,n),(n,i),(l,h),(h,l),(l,n),(n,l)(m,h),(h,m),$
  $(m,n),(n,m),(o,h),(h,o),(o,n),(n,o)\}$

- $ATT_U^{Act} = \{(x,o)\}$
- $IAF = (ARG_F^{Act} \cup ARG_M^{Act}, ATT_R^{Act} \cup ATT_U^{Act})$

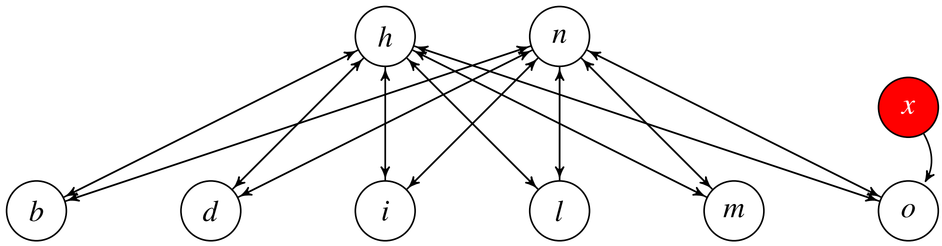


Figure 4.21. Activated argumentation framework for a record of the Ljubljana dataset

---

Note 8

Argument m and n, despite dealing with the same attribute (Degree of malignancy), are both activated, and in this case, with different membership degrees. Argument u, even though the membership function associated with its premise $f_{TumorSize}^{High}$ returns a value greater than zero, the attacked argument $a : LowAge \rightarrow NR$ is not activated (not present in Arf_F), therefore according to definition 17 it is not activated. Similarly argument q is not activated because, even if the attacked argument d is in Arf_F, its premise ($HighAge$) is not satisfied. Instead, the mitigating argument x is activated because definition 17 is satisfied.

4.5 Aggregating and recommending arguments

4.5.1 Running abstract skeptical and credulous semantics

In order to deal with objective 2 (c), we propose to run *argumentation semantics* on the instantiated argumentation framework to obtain *argument-base extensions* and to decide which arguments should ultimately be accepted. In this work we focus on the grounded and the preferred semantics (as described in definitions 8 and 10). The grounded semantic returns always one extension of arguments: if not empty, it contains all the arguments in the instantiated argumentation framework that share the same claim. The preferred semantic instead, may return a set of extensions of arguments hence an heuristic for selecting the *winning extension* and extracting the *winning claim*, is necessary.

We argue that the cardinality of an extension is an important factor to consider. It indicates how many arguments (pieces of evidence) support the same claim (a preferred extension is an admissible, conflict-free set of arguments). However, in the pool of arguments there might be, for instance, 5 arguments supporting a claim c_1 and 10 supporting a claim c_2. Claim c_2 has a higher chance to emerge as winning claim, because the arguments that support it have a higher chance of being activated. In the absence of measures of certainty for arguments, the number of supporting arguments is the best we can do. However, considering just the cardinality of a set of argument is not enough. In order to handle this issue, we propose to consider the degree of truth of an entire extension. Computing the average of the membership degrees of the premise of each argument, in the same extension, is a intuitive solution. However, some membership functions associated with an attribute, (such as $f^{Lower}_{BreastQ}$ or $f^{Upper}_{BreastQ}$ of figure 4.18 and fig. 4.19) are double-valued functions (return just 0 or 1 as membership degree). This lack of shades render their uncertainty null in their activation or exclusion. If 0, the associated argument is not activated at all, otherwise it will be activated with maximum degree of truth. We argue that, a degree of truth of 1 returned by a membership function defined in the continuous in $[0,1] \in \Re$, (multivalued logic) should be considered more contributive that the degree of truth of 1 returned by a double-valued membership function (two-valued logic). However, someone can argue that this is not a strong point because there is no really uncertainty for these attributes. We investigate this issue in future works.

4.5.2 Aggregating degrees of truth and recommending an alternative

For the above reasons we propose to use the fuzzy algebraic product (fuzzy intersection) to aggregate degrees of truth within the same extension. The rationale behind adopting the algebraic product is that we want to follow a pessimistic approach, giving more importance to a single low degree of truth in an extension of arguments rather than high degrees of truth. A chain is as strong as its weakest link.In other words, we want to emphasise the situation in which an attribute activated with a low degree of truth negatively influences the other attributes, within the same extension, even if activated with high degrees of truth. For instance, a degree of truth of 1 is not influential on other degrees of truth within the same set of arguments, but a degree of 0.5 brings down the degrees of truth of the argument in the same set, even if all the arguments are activated with degree of truth of 1. This approach produces a decreasing affect, where the final output is always smaller, or equal, than the smallest contributing membership value. Such a pessimistic approach is then applied to every consisted set of arguments computed by argumentation semantics. Eventually, the extension that maximizes both the number of arguments and the fuzzy algebraic product is declared to be the *winning extension* and therefore the *winning claim* can be extracted, representing the shared claim by those arguments with the highest combined degree of truth.

Definition 20 (Extension degree of truth)
The overall degree of truth d of an argument-base extension E containing n arguments is the product between the percentage of arguments in E of the cardinality of the pool of arguments, and the algebraic product of the degrees of truth of each forecast argument in the extension.

$$E_d = \frac{n}{card(ARG_{pool})} \times \prod_{i=1}^{n} f_{\alpha_i}(x_i)$$

with f_{α_i} the membership function associated to the premise of the forecast argument '$a_i : f_\alpha \rightarrow c$' $\in ARG_F^{Act}$ in the extension E, x_i the input value for the attribute α that activated the argument a_i and $card(ARG_{pool})$ the cardinality of the pool of arguments.

Note 9

Only forecast arguments in the extension are considered in the algebraic product because mitigating arguments do not carry a claim. However, the percentage of arguments within an extension refers to the cardinality of the pool of arguments (forecast + mitigating).

Definition 21 (Winning extension)

The winning extension WE of a set E of n preferred extensions, is the extension with the highest degrees of truth.

$WE = \{A \mid A \in E,\ and\ A_d = max(E_d^1, E_d^2, ..., E_d^n)\}$

Note 10

An extension with a small cardinality may still be proclaimed the winning extension if its algebraic product is higher than other extensions. Similarly an extension with a small algebraic product may still be the winning extension if its cardinality is higher than other extensions.

Definition 22 (Winning claim)

The winning claim c is the claim supported by all the arguments in the winning extension WE. $\forall a \in WE$ with a : '$f_\alpha \rightarrow c$' $\in ARG_F^{Act}$, c is the winning claim.

Example 16

In figure 15, the grounded extension is empty and the two preferred extensions are: $p_1 = \{b, d, i, l, m, x\}$, $p_2 = \{h, n, x\}$. The degrees of truth of each argument of the two computed extensions are:

- b: $f_{age}^{medium} = 0.9$
- d: $f_{Menop}^{Pre} = 1.0$
- i: $f_{NodeInv}^{Low} = 0.97$
- l: $f_{NodeC}^{False} = 1.0$
- m: $f_{DegMalig}^{Low} = 0.6$
- h: $f_{TumorSize}^{High} = 0.6$
- n: $f_{DegMalig}^{High} = 0.5$

Argument x is in both the extensions according to preferred semantics, but as it is a mitigating argument, it does not have an associated claim (recurrence or no-recurrence), and, in this implementation, it does not contribute to the computation of the overall degree of truth of each extension. According to definition 20 we have:

- (b, d, i, l, m) algebraic product: 0.52, % of args. : (5/26=0.19)
  so $E_d = 0.1$
- (h, n) - algebraic product: 0.3, % of args.:(2/26=0.08)
  so $E_d = 0.024$

The winning extension is [b, d, i, l, m] therefore the winning claim is 'no recurrence (NR)'.

4.6 Summary

In this section a formal model of arguments has been designed and implemented, based on argumentation theory AT. The basic building blocks of formal abstract argumentation theory, as firstly proposed by Dung [Dung, 1995], has been described. It has been shown how to create an *argumentation framework* composed of arguments and defeat relations among them. Here arguments are abstract, as well as the nature of defeat relations. In other words, an argumentation framework is a graph of nodes (arguments) and directed arrows among them (defeat relations). The internal structure of argument is not taken into account as well as why an arrow between two arguments exists. On top of an argumentation framework it has been shown how *argumentation semantics* (algorithms) can be executed for the acceptance of arguments. Two semantics have been described: grounded and preferred semantics [Dung, 1995]. The former is a skeptical way of aggregating arguments while the latter is a more credulous way.

The abstract model of arguments has subsequently been extended to handle structured arguments and defeat relations. The new model is built upon Abstract Argumentation Theory (AAT) and Fuzzy Logic (FL). In detail, we have proposed to construct an argument with Fuzzy Membership Functions (FMF) and Fuzzy Theory (FT). Here an argument is composed of a premise, dealing with a certain concept, and modeled with a fuzzy set, followed by a claim which is a possible option or alternative a decision maker could follow in decision-making. Arguments are defeasible in nature and can be considered *forecast* when they are in favor or against a certain claim (but justification is not infallible), and *mitigation* arguments, when defeating (undermining justification for) forecast or other mitigation arguments [Matt et al., 2010]. Subsequently it has been shown how to represent defeat relations among arguments. Two typologies of relations have been defined: rebuttal and undercutting defeat. The former occurs when the claims of two arguments are different, contradicting each other because they are mutually exclusive. The latter occurs when an argument either challenges the validity of a forecast argument or when it is considered more important than another argument (preference).

Once the abstract arguments-based model has been extended to handle structured arguments and defeat relations among them, a designer can translate a knowledge - base into a *contextual argumentation framework*. This step is subjective, it depends on the expertise, past experience and knowledge of the designer. In addition it accounts for the uncertainty in creating arguments and defeat relations. In fact, the premise of each argument is modeled as a fuzzy set, that means with a membership function that returns a specific degree of truth. So an argument can be instantiated only if its premise has a degree of truth greater than zero. All the instantiated arguments form an *instantiated argumentation framework* which is basically a sub-framework of the contextual framework. In other words, not all the designed arguments are activated in a certain context and case, with certain characteristics. As a consequence, not all the defeat relations are activated. Only a defeat relation in which both the defeater and the defeated arguments are instantiated, can be part of the instantiated argumentation framework.

As soon as the instantiated argumentation framework emerges, grounded and preferred abstract semantics can be executed to compute the justification status of the activated arguments, thus computing conflict-free *extensions* of arguments (sets). Each extension is in contradiction with all the other extensions as the arguments in it share a common claim (an option/alternative available to decision maker) which is different than the ones of other extensions. It is then shown how to evaluate the overall degree of truth of an extension, combining all the degrees of truth of each premise of each argument in the same extension. Eventually, the extension with the highest degree of truth is then recommended as the winning extension. From this set, the common claim shared by its arguments can be extracted and recommended to the decision maker.

In this chapter, each formal definition has been clarified with a illustrative scenario: the breast cancer recurrence prognosis/prediction. The main goal was to design a tool for decision-making support based on defeasible reasoning. The entire process starting from a knowledge base of an expert to a potential rational recommendation for a given health-care problem is depicted in figure 4.22. The goal is to provide a designer (expert) with a methodology for representing his/her knowledge base, composed by attributes (concepts), with a set of structured arguments, forming a structured knowledge base (a). Subsequently, the tool allows the designer to formally represent relations among arguments in form of defeats. Structured arguments and defeat (attack) relations form a contextual argumentation framework (b). This represents the entire expertise of the designer about a given health-care problem. The contextual framework might not be fully necessary in a given practical scenario, characterised only by certain attributes and thus structured arguments that activate a instantiated argumentation framework (sub-framework of arguments) (c). On top of this emerging framework, automatic algorithms (argumentation semantics) extract contradicting extensions of arguments (d), from which just one will emerge has the most representative (d) and from which a rational decision/option can be recommended.

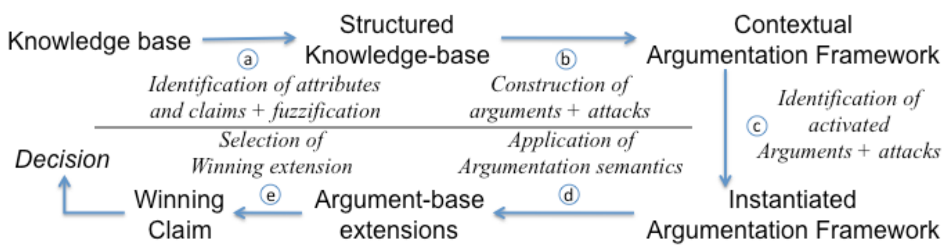


Figure 4.22. A decision-support tool: from knowledge base to a rational recommendation

Chapter 5

Experiments, results and discussion

5.1 Introduction

In this section, the designed arguments-based model, formalised in chapter 4, is evaluated. In particular, as stated in objective 3, the evaluation strategy includes three parts:

a) outcomes comparison against machine learning;

b) evaluation of the predictive and explanatory power;

c) individuation of strengths and limitations.

In order to achieve sub-objective (a), an experiment has been designed. This includes the Ljubljana breast cancer dataset previously described in section 4.2.1 and used in many machine learning studies and experiments [Williams and Williamson, 2006] [Clark and Niblett, 1987] [Tan and Eshelman, 1988] [Cestnik et al., 1987]. The dataset includes 286 instances of real patients who went through a breast cancer operation. Each record of the dataset contains values related to a single patient, about a set of attributes as described in table 4.1. These values were recorded at the time of the operation, that means when the breast cancer was removed. 9 of these instances are incomplete, as the value of some attribute is missing. In 81 cases, that means for 81 patients, the breast cancer reappeared within 5 years of surgery, and for the remaining 205 cases, the cancer did not occur within 5 years. Each value was verified by clinicians making the dataset a good base of real-world evidence.

In order to implement objective 3 (a), we used the dataset with state-of-the-art machine learning supervised techniques, as listed in the following:

- decision tables (Rules-based)
- bayesian network (Bayes-based)
- best-first decision tree (Tree-based)
- classification regression (Meta)
- multilayer perceptron (Function-based)
- logistic (Function-based)
- alternating decision tree (Tree-based)
- naive Bayes (Bayes-based)

The goal was to compare the accuracy of the designed arguments-based model in predicting the right prognosis (recurrence and no-recurrence) against the accuracy of the listed machine learning techniques. We removed the attribute ‘irradiation’ and ‘breast’ from the dataset because they were not used in the expert knowledge-base and thus not used in the arguments-based model, as described in table 4.3. In particular the attribute ‘irradiation’ has been removed also because it is unrelated to the other attributes due to the fact that irradiation occurred after the surgery and not at the time of surgery. We have used WEKA machine learning software [Witten et al., 2011]. We did not implement any machine learning tools, instead we have used robust implementations as provided by the software.

5.2 Machine learning prediction

Six experiments were conducted. In the first, we used a 10-fold cross-validation. This means that the dataset was randomly reordered and then split into 10 folds of equal size. For each iteration, one fold is used for testing the model while the remaining (10-1=9) folds are used for training the classifier. The number of iterations is exactly as the size of the specified folds (10 in this case) which means every fold is used exactly once for testing. The test results are subsequently collected and averaged over all folds. This methodology gives the cross-validation estimate of the accuracy and it is usually adopted when the dataset is limited in size. The folds can be purely random or slightly modified to create the same class distributions in each fold as in the complete dataset. In the second experiment we used a 28-fold cross validation (10 cases out 286 in each fold) while in the third a 40-fold cross validation was used. Subsequently we used the WEKA percentage split functionality. This means that we specified a percentage of split of the records of the dataset both for training and tests. In other words, a 66% split means that 66% of the records of the dataset are used for training the classifier, and the remaining 34% is used to test the model and check its predictive capacity. In the fourth experiment the split was 70% while in the fifth it was 50% and in the sixsth was 30%. These experiments are listed as following and the results are in table 5.1:

- (Exp. 1) 10-fold cross-validation;
- (Exp. 2) 28-fold cross-validation;
- (Exp. 3) 40-fold cross-validation;
- (Exp. 4) 70% split;
- (Exp. 5) 50% split;
- (Exp. 6) 30% split;

Classifier	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	Exp. 6
decision tables	73.42	75.52	73.42	73.25	74.12	74.00
bayesian network	72.37	73.07	73.07	68.60	72.70	73.00
best-first decision tree	66.78	70.62	73.07	62.79	74.12	72.00
regression	70.62	73.07	71.67	66.26	72.72	72.00
multilayer perceptron	65.38	68.88	65.73	58.13	65.03	65.00
logistic	68.88	69.23	70.27	68.60	69.23	63.50
alternating decision tree	74.47	75.17	74.82	65.11	69.93	72.50
naive bayes	72.72	73.07	73.07	68.60	72.02	72.50

Table 5.1. Prediction accuracy percentage of machine learning techniques with the Ljubljana breast cancer dataset.

5.3 Argumentation theory prediction

In order to test the prediction capacity of the argument-based model designed in chapter 4, we compared the winning claim (recurrence or no-reccurence), as per definition 22 against the outcome class (recurrence or no-recurrence) of the Ljubljana dataset, as in table 4.1. We executed this process for all the records of the dataset, that means for all the patients. As described in the formal model of section 4.1.2, we have used grounded and preferred abstract semantics, as firstly proposed by Dung [Dung, 1995].

The grounded semantic is a skeptical way of computing acceptable arguments: it returns just one grounded extension that coincides with the minimal complete extension in which the arguments are labeled *in*, as per definition 8. In our case-study, for 8 records (of 286) a non-empty grounded extension has been computed, that is a conflict-free set of arguments supporting the same claim. A clear coherent position emerges, thus the clinician's point of view about recurrence is well-grounded. In 7 of these 8 cases, the common claim, shared by the arguments in the grounded extension, coincides with the objective output class of the dataset's record (recurrence or not). In the only remaining case, *AT* fails in predicting the recurrence status: available evidence, thus the built arguments are not enough and further evidence is needed.

The preferred semantic is a credulous way of computing acceptable arguments: it returns multiple coherent conflict-free extensions, as per definition 10, each composed by arguments sharing the same claim. For each record of the Ljubljana dataset, according to our way of constructing argumentation frameworks, two preferred extensions exists: one with arguments supporting 'recurrence', and one with arguments supporting 'no recurrence'. Preferred extensions computed coincide with the grounded extension if any. Two different coherent points of view are available to support the clinician decision-making process in respect of predicting the recurrence status. One possible automatic strategy, the one defined in this study, is the selection of the biggest preferred extension along with the aggregation of the degrees of truth of the premises of its internal arguments, as per section 4.5.2. Out of 286 patients, 210 recurrence status were successfully predicted. The claim shared by the arguments of the winning preferred extension can be credulously used for enhancing a clinician's decision-making process. Obtained results are summarised as in the following:

- grounded semantic: 8 computed, 7 successfully predicted;
- preferred semantic: 286 computed, 210 successfully predicted.

Considering the preferred semantic, the prediction rate was $\frac{210}{286} = 73.42\%$

5.4 Argumentation theory vs machine learning

In order to finalise objective 3 (a), a comparison of the accuracy of the machine learning classifiers used against the accuracy of the argument-based model is performed. We refer to the latter as Argumentation Theory AT. Figure 5.1 compares the accuracy of AT against the ones obtained in experiment 1, where a 10-fold cross-validation has been used. Similarly, figure 5.2 and figure 5.3 show respectively the comparisons between AT against experiments 2 and 3 (28-folds and 40-folds cross validation). Figures 5.4, 5.5, 5.6 compare AT against the machine learning classifiers used in experiment 3, 4, 5 (70%, 50% and 30% splits). It is worth noting that the accuracy rate of the argument-based model is always the same: as it is not based on training, n-fold cross-validation, etc, do not apply.

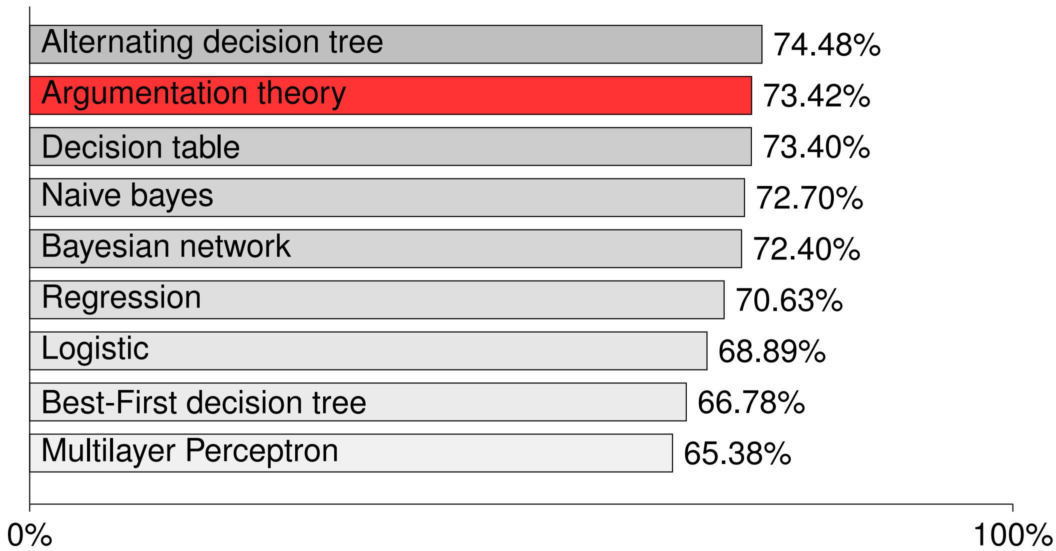


Figure 5.1. Comparison of the accuracy of the arguments-based model against learning-based classifiers (10-folds cross validation)

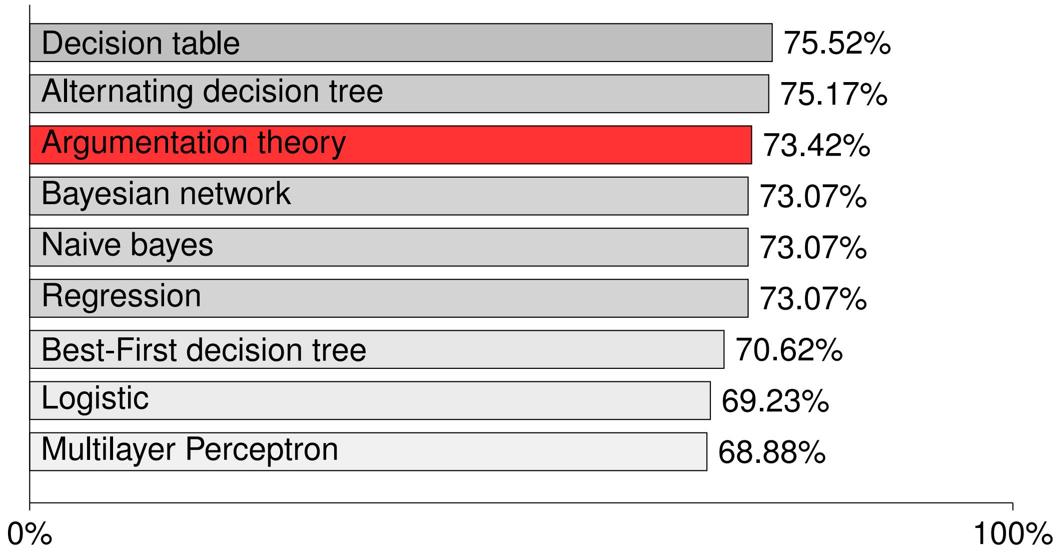


Figure 5.2. Comparison of the accuracy of the arguments-based model against learning-based classifiers (28-folds cross validation)

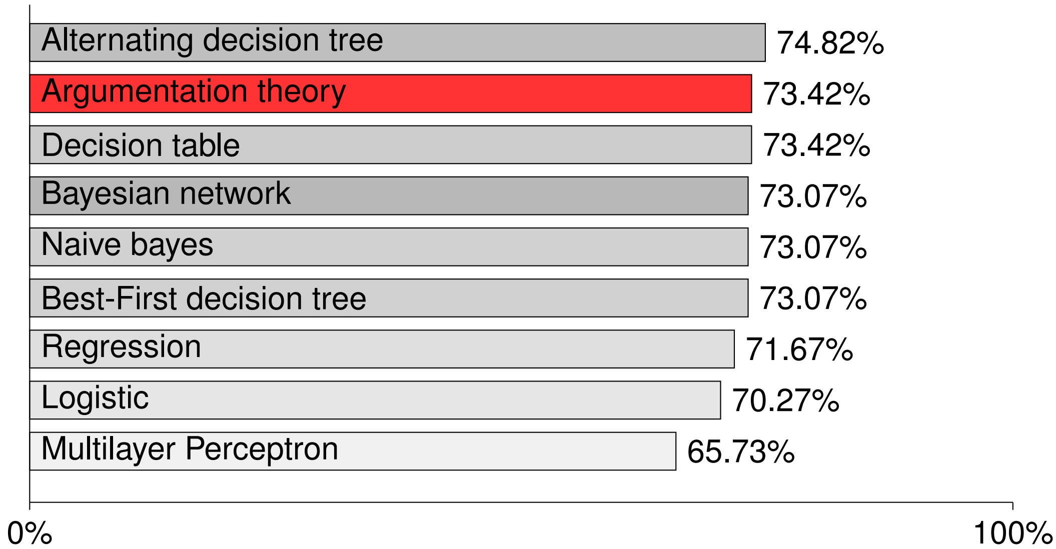


Figure 5.3. Comparison of the accuracy of the arguments-based model against learning-based classifiers (40-folds cross validation)

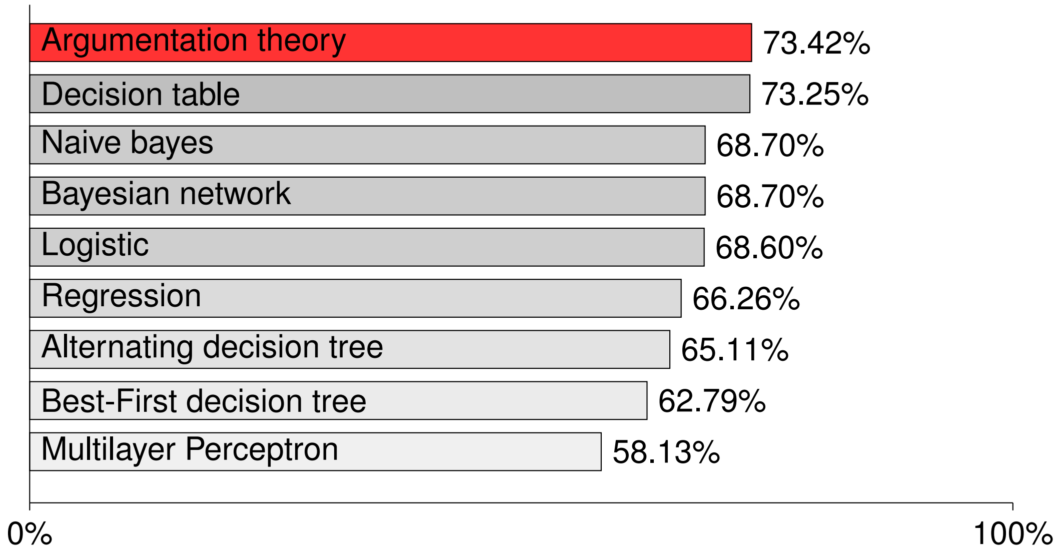


Figure 5.4. Comparison of the accuracy of the arguments-based model against learning-based classifiers (70% split)

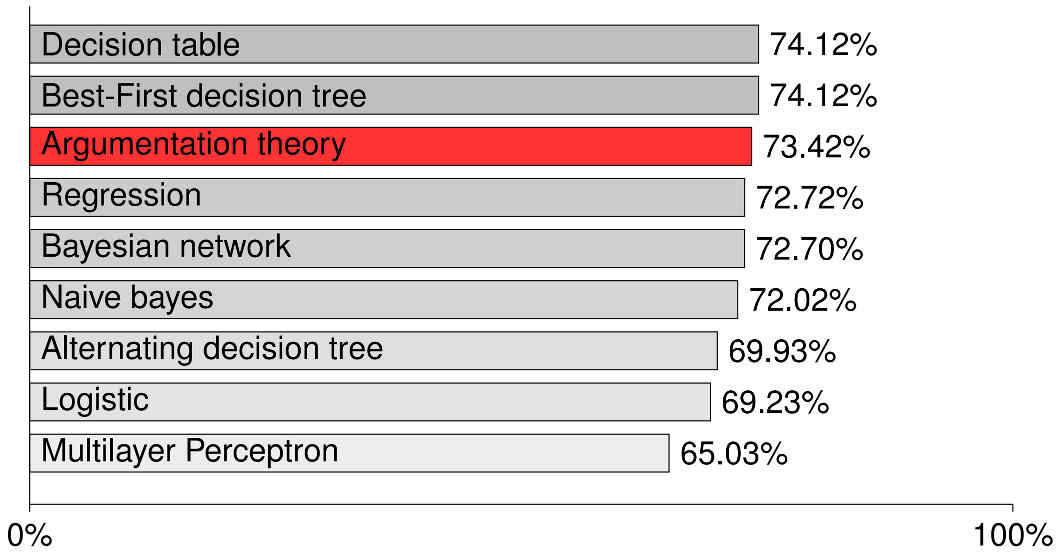


Figure 5.5. Comparison of the accuracy of the arguments-based model against learning-based classifiers (50% split)

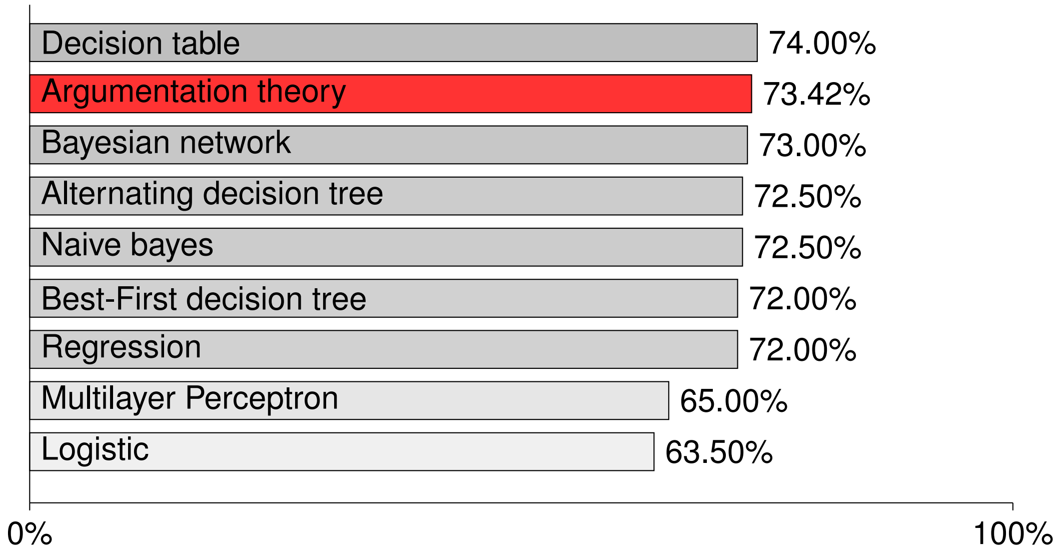


Figure 5.6. Comparison of the accuracy of the arguments-based model against learning-based classifiers (30% split)

5.5 Discussions

At a first glance, by having a quick examination of the results of figures 5.1-5.6, the arguments-based model, built using argumentation theory, produced interesting and encouraging results. In the first 3 figures, corresponding to the first three experiments (varying the folds cross-validation parameter) the argument-based model had an accuracy rate similar to the decision table classifier and the alternating decision tree classifier. It was always superior to the two Bayes-based classifiers (Bayesian network and naive Bayes), even if in a small measure. The best-first decision tree classifier was always inferior, in different measures as well as the function-based classifiers (logistic and multilayer perceptron). The outputs of the classification via regression was almost stable, slightly lower then the argument-based model.

In these three experiments the number of folds for cross-validation has been increased meaning that more records of the dataset in turn have been used for testing and less for training the classifier. With 10-folds cross validation 10% of the records are used for testing and the remaining 90% for training. This process is repeated 10 times varying the fold for test. Similarly, with a 40-folds cross validation, 2.5% of records is used for testing (approximately 7 records) while the rest for training. This process is repeated 40 times. From obtained results, increasing the number of folds, that means having a bigger training set, does not always produce an increment in terms of prediction rate. In general this rate is stable, with minimum variation, except for the best-first decision tree that had a constant increment, as well as the logistic classifier. The Bayesian network and the Naive Bayes classifiers had the same constant behaviour for the three experiments as well as the alternating decision tree classifier. The decision table and the multilayer perceptron classifiers do not follow any constant increment or decrement, but the decision table classifier has always a high accuracy while the multilayer perceptron has always a low accuracy. The arguments-based model, as it is constructed, does not require any training, or special learning processes, thus the output is always the same. We recall that in the experiments we have only evaluated just one ‘expert”s knowledge, which is not trained to fit the data.

In the last three experiments, the percentage of records for training (thus also for testing) varied. This split option is based on the premise that a certain machine learning classifier is evaluated in terms of accuracy in predicting a certain percentage of the data that is held for testing. Ideally, the prediction accuracy obtained by running a classifier would be better when a bigger percentage of training data is used. In these experiments, 70% means that 200 records out of 286 are held for training a model, while the remaining is used for testing. With 50% and 30% the records held out for training are respectively 143 and 86. The records used for the training process are selected randomly. Results are listed in table 5.1 but are not as expected. In fact, decreasing the size of the training set, the accuracy of each classifier increases (except for the multilayer perceptron), which is in contradiction with what ideally hypothesized. This inverted relationship between the size of the training set and the prediction accuracy could be explained by the small size of the dataset. With a bigger training set, more cases and examples are added to the model, making it more robust in prediction and more trustworthy. Vice-versa, a small training set makes prediction less accurate and less trustworthy. So the results of table 5.1, for experiment 4, 5, 6 should not be considered really trustworthy, especially for experiment 5 and 6 where a small training set is used. In fact, considering experiment 4, that can be considered trustworthy enough with a 70% of records used for training, argumentation theory performed better than any other classifier, confirming its potential.

The confusion in analysing outcomes of classifiers is avoided with argumentation theory, where no learning process is adopted. The arguments networks is built according the knowledge base of an expert and it is not based on learning from example or previous cases. So for each case, the built arguments-based model is independent of other cases and the size of the dataset does not really matter. Another interesting property of AT is that, the 9 incomplete records of the Ljubljana dataset, can still be considered using our approach. With machine learning an explicit algorithm is needed to fill in the missing values, so guessing what would have been those values, but with an arguments-based model, this issue is avoided. In fact, an expert knowledge-base is represented in a contextual argumentation framework, as per section 4.3, and if a value for a certain value is missing, it does not activate the related argument, which will not be present in the activated argumentation framework. Although one less argument, the inference process through the use of argument-based semantics can still be run and a rational prediction can be achieved.

5.6 Interpreting outcomes and explaining decisions

One of the main advantages of argumentation theory, compared to machine learning classifiers, is represented by its explanatory power. Firstly, the translation of knowledge-bases into a set of arguments and defeat relations among them is more intuitive, following a modular process that is based upon terms familiar to the clinician and using natural language. Secondly, the outcomes, consequence of the application of the semantic for acceptance of arguments, are conflict-free sets of the same input arguments. In other words, the language and terms do not change; what does change is their aggregation.

Let us consider again the example of figure 4.20, that can be drawn as in in figure 5.7. Here the argumentation framework is indeed complex, but it contains the entire clinician knowledge base in form of arguments, built with familiar terms, contradicting and challenging each other, along with the consideration of their preferences. From this contextual argumentation framework, a subset of arguments and, as a consequence, defeat relations, are activated when considering a real-world scenario. For instance, the activated argumentation framework of the illustrative example of figure 15 can be replotted as in figure 5.8. This activated framework is smaller in terms of number of arguments and defeat relations, representing the current case under examination. On top of this, the application of the preferred semantic, as formalised in section 4.1.2 produces the two preferred extensions, as also computed in the example 16 and depicted in figure 5.9. Argument x is in both the extensions, as it is not defeated by any other mitigating argument, and including it in both the computed preferred extension does not break the property of conflict-freeness of each. Although this, in the implemented arguments-based model, does not participate in the computation of the degree of truth of the overall extension as it does not forecast a claim (recurrence or no-recurrence).

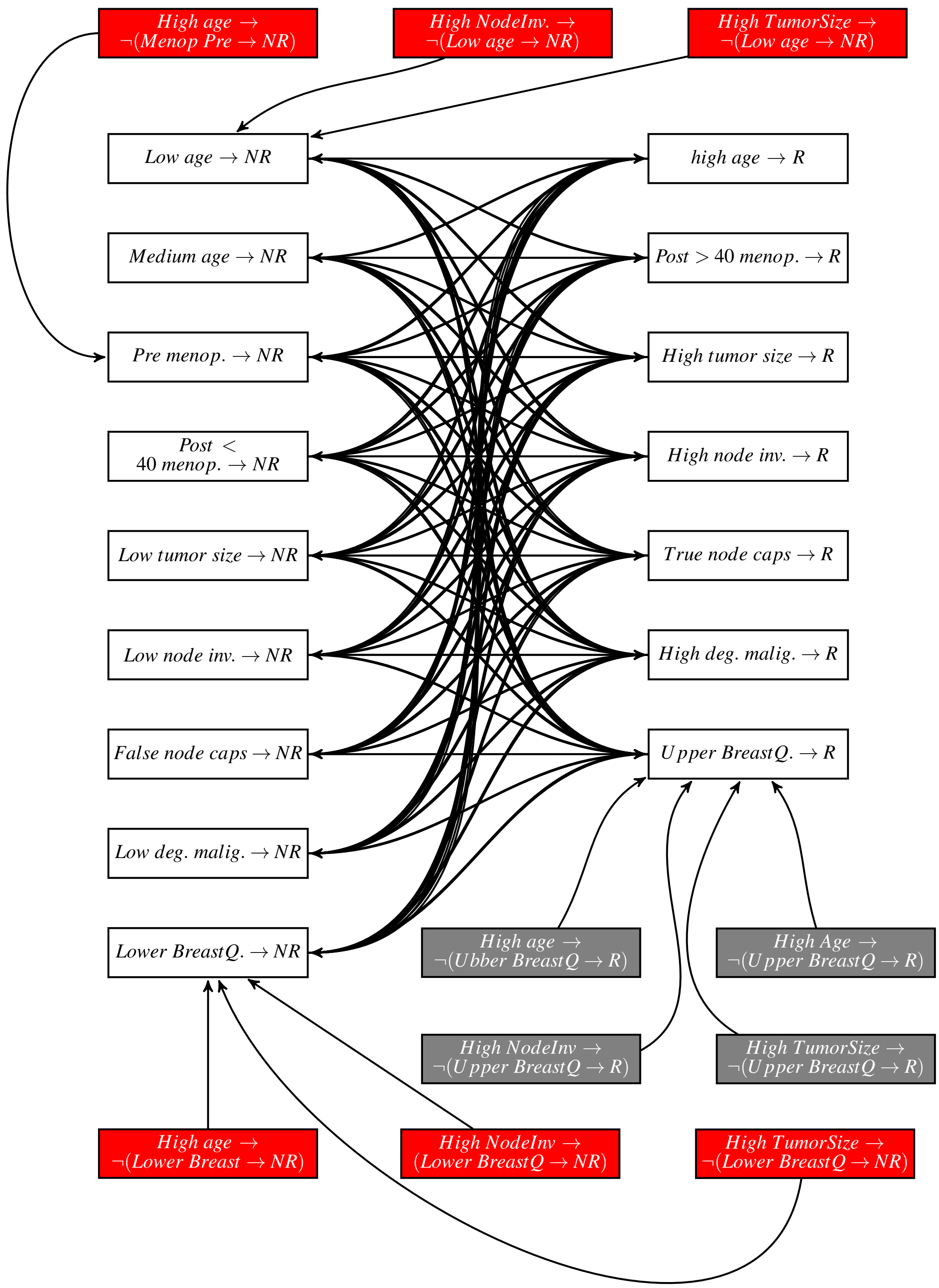


Figure 5.7. Contextual argumentation framework emerged by expert's knowledge-base, built with natural language terms

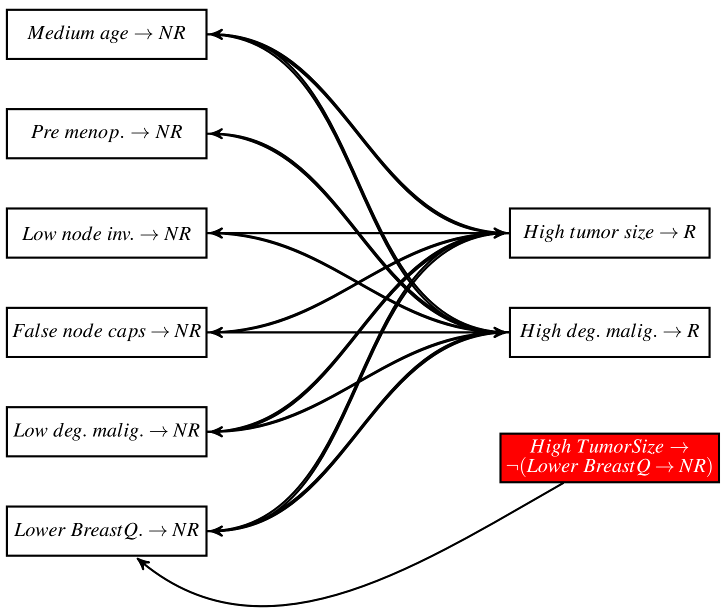


Figure 5.8. Activated argumentation framework emerged by a record of the Ljubljana dataset, built with natural language terms

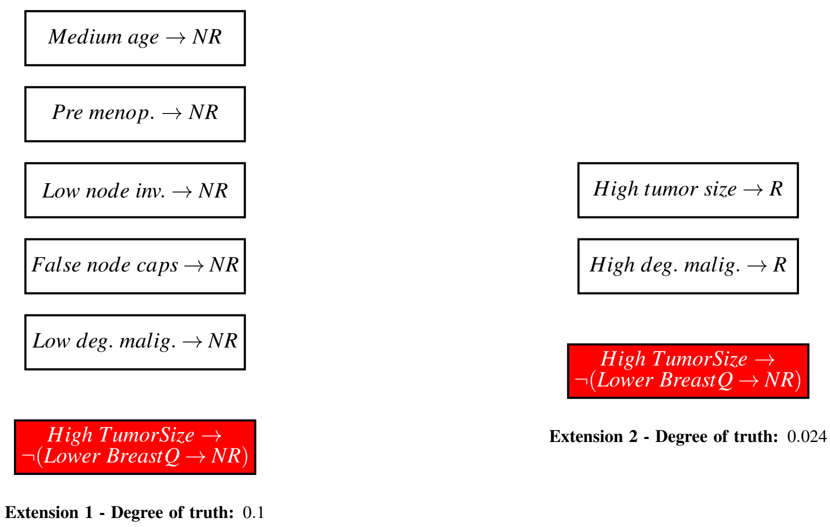


Figure 5.9. Preferred extensions computed on the activated argumentation framework built with natural language terms

It is worth mentioning again that the degree of truth is not in the real range $[0..1]$, so the value of 0.1 of the extension 1 does not mean that the overall degree of truth is 10%. This value is computed according the method proposed in definition 20. The claims of the arguments in each extension are the same, and the recommended claim is no-recurrence (NR) as the recommended (winning) extension is 1. The two extensions can still be easily interpreted, as the arguments in each set are easy to be read. In addition, a clinician can better understand why those arguments are the closest candidates to represent the truth.

5.7 Advantages and limitations of argumentation theory

Argumentation theory AT has several advantages and properties that most machine learning techniques do not have. Indeed, it also has limitations that are worth mentioning. In order to deal with objective 3 (c) we discuss the main strengths and drawbacks of the proposed arguments-based model.

- *Inconsistency and Incompleteness:* *AT* provides a methodology for reasoning on available evidence, even if partial and inconsistent; missing data is simply discarded and even if an argument cannot be expressed, the argumentative process can still be executed with available evidence. This property is powerful when a full dataset is incomplete or not available and when some value within the same record is missing;

- *Expertise and Uncertainty:* *AT* captures expertise in a organised fashion, handling uncertainty; a clinician can translate his/her knowledge-base into an argumentation framework that takes into account the uncertainty of this translation process and the degrees of truth of each argument built;

- *Intuitiveness:* *AT* is not based on statistics or probability. This is close to the way humans reason and if the designer is anyway inclined to use statistical evidence, he/she can embed that information in a structured argument, input of the argumentation framework; knowledge-bases can be formally structured with intuitive arguments built with familiar linguistic terms;

- *Explainability:* *AT* leads to explanatory reasoning thanks to the incremental and modular way of reasoning with available evidence. *AT* provides semantics for computing arguments' justification status thus letting the final decision be better explained and interpreted;
- *Dataset independency*: *AT* does not require a dataset and it may be particularly useful for emerging knowledge where quantity evidence has not yet been gathered;
- *Extensibility and Updatability:* *AT* is an a open and extensible paradigm that allows to retract a decision in the light of new evidence: an argumentation framework can be updated with new arguments and evidence;
- *Knowledge-bases comparability*: *AT* allows comparisons of different subjective knowledge-bases. For instance, two clinicians can subjectively build their own argumentation frameworks according their knowledge-bases. Subsequently they can compare the outcomes of the computed extensions of acceptable arguments, for a given problem, such as the breast cancer recurrence prediction/prognosis. They can identify potential problems in the definition of their arguments, the differences in their knowledge-bases and they can even create an overall argumentation framework built upon their subjective frameworks;

The above characteristics are not shared by machine learning-based tools that are automatic procedures, supervised or not, that try to learn from previous examples or past evidence. Despite the strength of our arguments-based approach, a few weaknesses can be delineated:

- *knowledge-base translation*: the initial process of translating knowledge-bases into structured arguments, even if intuitive, may require effort, above all when several pieces of evidence and knowledge are considered; similarly, indicating defeat relationships among arguments requires an initial cost;

- *lack of learning*: *AT* is not a paradigm based upon learning, so automatic rules or patterns cannot be extracted automatically as in machine learning approaches. However, machine learning approaches rely on big datasets of evidence, and sometimes they requires a big amount of time to complete the learning process;

5.8 Summary

Argumentation theory AT seems to be a fruitful paradigm for building models aimed at supporting decision making process in health care scenarios. In this chapter an experiment has been conducted considering the evidence embedded in the Ljubljana breast cancer dataset. The first aim was to compare the outcome of the designed arguments-based model against few machine learning-based classifiers. AT performed really well, with a predictive capacity of 73.42%, in line with other machine learning tools, and most of the times even superior. Indeed, this is a preliminary result and cannot be generalized, but it represents a interesting finding. However, AT has been shown to have certain properties that are not shared by machine learning-based techniques. Firstly, AT is knowledge-based and not learning-based. This means that with AT knowledge-bases are translated into argumentation frameworks, and with machine-learning, expertise is not taken into account as the goal is to automatically extract knowledge from available data. It is clear that, in the case available data is not enough, machine learning will likely perform poorly, while AT can still be executed with the evidence available. Secondly, AT is a context-aware and user-centered paradigm because it adopts users' language and familiar terms sensitive to the context under consideration. This makes AT a self-explanatory tool where arguments and outcomes can be better interpreted than outputs of machine learning classifiers. Furthermore, the process of translating knowledge-bases into argumentation frameworks is not an impossible process, even if it requires an initial cost. However, the lack of learning from evidence represents the major limitation of argumentation theory, but this could be in part addressed creating arguments that consider the output of machine learning classifiers. In other words, machine learning could be merged with argumentation theory where arguments can be built considering the accuracy of learning-based classifiers, for a certain attribute.

Chapter 6

Conclusions and future work

New technologies are undoubtedly useful for the advance of knowledge, especially in health care and medicine. In these domains, electronic records and other technologies facilitate clinicians' daily activities providing them with a wider range of tools for managing patients' information. However, despite the increasing amount of information available to health care practitioners, clinical decision-making is getting more complex because this new information needs to be considered and managed. In particular, clinical decision support systems need to account for a bigger set of evidence, likely characterised by uncertainty and sometime not coherent. State-of-the-art approaches in decision making under uncertainty have been proved to be indeed useful, but they have their drawbacks.

Fuzzy logic handles various typologies of uncertainties and it is capable of representing the vagueness of attributes. The main advantage is its capacity of dealing with human reasoning that is vague, approximated rather than exact and fixed. The main disadvantage is represented by the time required for translating knowledge-bases in fuzzy sets as well as the subjective perception of what fuzzy values represent. In addition, the manipulation of fuzzy values and their aggregation can be hard to be interpreted by clinical experts.

Decision theory and in particularly the expected utility theory have a long history and are based on probability. The aim is to propose not only the optimal option, but also the one that maximise the utility for patients. Although the theory is intuitive in many cases, the main difficulty is represented by the estimation of the probabilities of each option, usually overestimated by clinicians, thus introducing a form of error. In addition, if the amount of evidence is limited, the approach loses significance in term of support for decision-making.

Expert systems is another approach that facilitates clinicians to take decision by modelling a certain health scenario with 'if...then...' rules according to their knowledge-base. Although they are based on natural language constructs, using clinicians familiar terms, the translation of knowledge-bases into formal rules might require an effort. In addition, this translating process needs to be practically executed by computer experts who need to create the appropriate software. Eventually, these systems often fail to detect contradictory rules as well as not accounting for the dynamism of changing information and knowledge-bases.

Similarly, case-base reasoning is an intuitive approach for clinician because using evidence-based medicine when guidelines cannot be applied. The objective is to extract knowledge from previous clinical cases and evidence in order to take decisions. Current cases are then added to previous cases updating the database of evidence useful as a support for future decisions. Although it is a natural and appropriate approach for decision making, in practice it does not follow good clinical practices because failing to consider the outcomes of the decisions taken. Moreover, when the previous databases of cases is limited, the approach is essentially based on anecdotal evidence, with a limited support for decision-making.

Bayesian networks is a probability-based approach that consists of representing each available alternative or decision as a class, and computing the probability that, for instance, a given patient, with certain attributes, falls within that class. These typologies of classifiers need to learn conditional probabilities of each attribute of a patient. Although they are based on numerical manipulation, they attract clinicians because using natural language familiar terms. However, in the case of missing values for attributes, the approach needs a further strategy to feel in these gaps, introducing a form of error. In addition, the main drawback is the translation of knowledge-bases into a bayes network: clinicians and experts cannot be expected to do it.

Machine learning is a paradigm capable of handling large sets of data and for discovering interrelationships among variables. It is a learning-based approach that learns from previous examples in order to classify future cases. Although a clinician is required to provide just a set of previous cases and evidence, the selection of previous data must be careful in order to be effective for classifying new future cases. In addition, as Bayesian networks and decision theory, it relies on a big dataset that is also composed by a number of fixed attributes, therefore not suitable for changing information. Eventually, it lacks of explanatory power as the predicted cases are just in somehow similar of previous cases, with any additional new findings.

Argumentation theory is an emerging paradigm, recently being considered in health care for aggregating clinical evidence intuitively and modularly. It is a practical implementation of defeasible reasoning, a form of non-monotonic reasoning that is close to human reasoning, where a conclusion can be retracted in the light of new evidence. It is not a statistical tool and it is close to the way humans reason under uncertainty with incomplete and inconsistent evidence. It is not a learning-based approach nor probability-based, but it is a knowledge-based paradigm that can work also when previous evidence is limited and partial. It is built upon the representation of a knowledge-base into a set of arguments and defeat relations among them. It handles contradictions of evidence and it is open to refinements. These properties seem to be appealing for creating decision-support tools that follow a qualitative rather than a quantitative aggregation of evidence. In addition, recommended decisions could be easily interpreted by clinical practitioners at various levels.

6.1 Contributions

The main contribution of this dissertation is the design of a decision-making support framework, based on defeasible reasoning, and implemented by argumentation theory. The basic building blocks of argumentation theory have been firstly introduced, taken from state-of-the art works in artificial intelligence. These mainly refers to the abstract model provided by Dung [?] that, despite its power in handling relationships and extracting consistent and coherent sets of pieces of evidence, it remains an abstract paradigm that does not account for the internal representation of arguments as well as the nature of relationships. In this dissertation arguments are represented by applying concepts from Fuzzy Logic, being this a novel contribution. In addition, in the case the Dung-style semantics, such as the preferred, return multiple consistent set of arguments, a novel technique based, on aggregation of degrees of truth of evidence is proposed.

Argumentation theory basic principles as well as the novel extensions have been formally proposed, informally illustrated and applied in a real-world health scenario: the breast cancer recurrence prediction. It has been shown how to translate an expert knowledge-base into a set of structured arguments, often contradictory, and how to explicate defeat relations among them in an organised argumentation framework. Subsequently, it has been proposed how do aggregate these arguments with skeptical and more credulous approaches, borrowed from abstract argumentation theory. In addition, it has been demonstrated how to interpret the outcomes of the argumentative process, and how to enhance an expert decision-making process.

The predictive capacity of the designed arguments-based model has been tested and compared against well known machine learning tools such as Bayes-based classifiers, neural networks and regressions models. Results are extremely promising: 74% of cases have been successfully predicted, in line with the other classifiers tested. It is worth mentioning that each case has been predicted just with the designed argumentation framework and not at all considering previous cases and evidence or probabilities. Indeed, designed arguments are built according to previous knowledge of an expert, but their aggregation does not follow any learning approach: each case is evaluated autonomously. In other words, the 74% of correctly predicted cases have been achieved without knowing in advance the truth value (recurrence or no-recurrence of breast cancer), contrarily to machine learning tools that manipulate and aggregate different independent attributes to approximate it. Furthermore, machine learning requires a big dataset to accurately approximate truth values. On the other hand, with argumentation theory, each cases is evaluated individually and independently, and the truth values are inferred considering an expert knowledge-base embedded in a formal argumentation framework.

6.2 Future works

Findings suggest that further research can be carried out to provide clinicians with an intuitive framework for supporting and justifying decisions. As stated in objective 4 a, more evaluations in different health care areas are needed before claiming that such an approach has a more general validity. Therefore, future works will be focused on the identification of those areas that would most benefit from the theory as well as the investigation of the main benefits and drawbacks (objective 4 b). In detail, the aim is to apply the proposed framework to further clinical decision problems, testing its effectiveness. Future work will include the investigation of the potential for Argumentation Theory-based decision support tools in multi-disciplinary decision making settings. Not only strengths, limitations and potential areas need to be identified, but the proposed solution must be usable, explanatory and as intuitive as possible, as stated in objective 4 c. For this reasons the author of this dissertation will focus on the design of a digital system that allows clinicians and health care practitioners to intuitively represent their knowledge bases into a set of arguments and defeat relationships among them with a low effort. In detail, the intuitiveness and the usability of the system, as well as the explainability of its outcomes, will be evaluated. The aim is to provide clinicians with qualitative outcomes for explaining recommended decisions.

Bibliography

[can, 2005] (2005). Breast cancer: an environmental disease - the case for primary prevention. Technical report, The UK Working Group on the Primary Prevention of Breast Cancer.

[Aamodt and Plaza, 1994] Aamodt, A. and Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7:39–59.

[Austin, 2008] Austin, M. (2008). *Information Integration and Decision Support for Multidisciplinary Team Meetings on Colorectal Cancer*. Phd Thesis - Wolfson Medical Vision Laboratory, Department of Engineering Science, University of Oxford.

[Baroni et al., 2011] Baroni, P., Caminada, M., and Giacomin, M. (2011). An introduction to argumentation semantics. *The Knowledge Engineering Review*, 26(4):365–410.

[Baroni et al., 1997] Baroni, P., Guida, G., and Mussi, S. (1997). Full nonmonotonicity: a new perspective in defeasible reasoning. *European Symposium on Intelligent Techniques*, pages 58–62.

[Buntine, 1997] Buntine, W. (1997). Bayesian in principle, but not always in practice. *IEEE Expert*, 12(6):24–25.

[Calman and Hine., 1995] Calman, K. and Hine., D. (1995). A policy framework for commissioning cancer services. *The Expert Advisory Group on Cancer to the Chief Medical Officers of England and Wales*.

[Cestnik et al., 1987] Cestnik, G., Konenenko, I., and Bratko, I. (1987). Assistant-86: A knowledge-elicitation tool for sophisticated users. *Progress in Machine Learning*, pages 31–45.

[Chang et al., 2009] Chang, C. F., Ghose, A., and Miller, A. (2009). Mixed-initiative argumentation: A framework for justification management in clinical group decision support. *AAAI Fall Symposium Series*.

[Chang et al., 2010] Chang, C. F., Miller, A., and Ghose, A. (2010). Mixed-initiative argumentation: Group decision support in medicine. *Electronic Healthcare, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering.*, 27(3):43–50.

[Clark and Niblett, 1987] Clark, P. and Niblett, T. (1987). Induction in noisy domains. *In Progress in Machine Learning (from the Proceedings of the 2nd European Working Session on Learning)*, pages 11–30.

[Dung, 1995] Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357.

[Eddy, 1990] Eddy, D. (1990). Clinical decision making: from theory to practice. the challenge. *Jama*, (263):287–290.

[Eom et al., 2008] Eom, J.-H., Kim, S.-C., and Zhang, B.-T. (2008). Aptacdss-e: A classifier ensemble-based clinical decision support system for cardiovascular disease level prediction. *Expert Systems with Applications*, 34(4):2465–2479.

[Fox and Das., 2000] Fox, J. and Das., S. (2000). *Safe and sound: Articial intelligence in hazardous applications*.

[Fox et al., 2006] Fox, J., Patkar, V., and Thomson, R. (2006). Decision support for healthcare: the proforma evidence base. *Informatics in Primary Care*, 14(1):49–54.

[Friedman et al., 1997] Friedman, N., Geiger, D., and Goldszmidt, M. (1997). Bayesian network classiers. *Machine Learning*, 29:131–163.

[Hellinger, 1989] Hellinger, F. J. (1989). Expected utility theory and risky choices with health outcomes. *Med Care*, 27(3):273–279.

[Hunter and Williams, 2010] Hunter, A. and Williams, M. (2010). Argumentation for aggregating clinical evidence. *22nd International Conference on Tools with Artificial Intelligence*, 1:361–368.

[Imperial Cancer Research Fund ICRF, 1995] Imperial Cancer Research Fund ICRF, G. B. (1995). *Our Vision for Cancer*. Imperial Cancer Research Fund.

[Jackson, 1998] Jackson, P. (1998). *Introduction To Expert Systems (3 ed.)*. Addison Wesley.

[Kononenko, 2001] Kononenko, I. (2001). Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in Medicine*, 23:89–109.

[Leondes, 2002] Leondes, C. T. (2002). *Expert systems: the technology of knowledge management and decision making for the 21st century*.

[Longo et al., 2012] Longo, L., Kane, B., and Hederman, L. (2012). Argumentation theory in health care. *25th IEEE International Symposium on Computer-Based Medical Systems. Rome, Italy.*

[Lucero et al., 2009] Lucero, M., Chesnevar, C., and Simari, G. (2009). On the accrual of arguments in defeasible logic programming. *In IJCAI*, pages 804–809.

[Matja et al., 1996] Matja, K., Igor, K., and Toma, S. (1996). Machine learning in prognosis of the femoral neck fracture recovery. *Artifial Intelligence in Medicine*, 8(5):431–451.

[Matt et al., 2010] Matt, P., Morgem, M., and Toni, F. (2010). Combining statistics and arguments to compute trust. *9th International Conference on Autonomous Agents and Multiagent Systems*, 1.

[Michalski et al., 1997] Michalski, R., Bratko, I., and Kubat, M. (1997). *Machine Learning, Data Mining and Knowledge Discovery: Methods and Applications*. John Wiley and Sons.

[Miles and Huberman, 1994] Miles, M. B. and Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Sage Publications, Newbury Park, CA.

[Mitchell, 1997] Mitchell, T. (1997). *Machine Learning*. MCGraw Hill.

[Parmigiani and Inoue, 2009] Parmigiani, G. and Inoue, L. (2009). *Decision theory: principles and approaches*. John Wiley and Sons.

[Pollock, 1987] Pollock, J. (1987). Defeasible reasoning. *Cognitive Science*, 11(4):481–518.

[Pollock, 1994] Pollock, J. L. (1994). Justification and defeat. *Artif. Intell.*, 67(2):377–407.

[Prakken, 2005] Prakken, H. (2005). A study of accrual of arguments, with applications to evidential reasoning. *In Proceedings of the 10th International Conference on Artificial Intelligence and Law. Bologna, Italy*, pages 85–94.

[Prakken and Vreeswijk, 2002] Prakken, H. and Vreeswijk, G. (2002). Logics for defeasibleargumentation. *Handbook of Philosophical Logic*, 4:218–319.

[Rabin, 1999] Rabin, M. (1999). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica*, 68:1281–1292.

[Tan and Eshelman, 1988] Tan, M. and Eshelman, L. (1988). Using weighted networks to represent classification knowledge in noisy domains. *Proceedings of the Fifth International Conference on Machine Learning*, pages 121–134.

[Verheij, 1995] Verheij, B. H. (1995). Accrual of arguments in defeasible argumentation. *In Proceedings of the 2nd Dutch/German Workshop on Nonmonotonic Reasoning.*, pages 217–224.

[Vreeswijk, 1993] Vreeswijk, G. (1993). Defeasible dialectics: A controversy-oriented approach towards defeasible argumentation. *Journal of Logic and Computation*, 3:3–27.

[Warren et al., 2000] Warren, J., Beliakov, G., and Zwaag, B. V. D. (2000). Fuzzy logic in clinical practice decision support systems. In *in 33rd Hawaii International Conference on System Sciences*.

[Watson and Marir, 1994] Watson, I. and Marir, F. (1994). Case-based reasoning: A review. *The Knowledge Engineering Review*, 9:327–361.

[Williams and Williamson, 2006] Williams, M. and Williamson, J. (2006). Combining argumentation and bayesian nets for breast cancer prognosis. *Journal of Logic, Language and Information*, 15(1-2):155–178.

[Witten et al., 2011] Witten, I. H., Frank, E., and Hall, M. A. (2011). *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier.

[Wu et al., 2010] Wu, Y., Caminada, M., and Podlaszewski, M. (2010). A labelling based justification status of arguments. *13th International Workshop on Non-Monotonic Reasoning, Studies in Logic*, 3(4):12–29.

[Zadeh, 1965] Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3):338–353.

[Zadeh, 1966] Zadeh, L. A. (1966). *Fuzzy Sets, Fuzzy Logic, Fuzzy Systems*. World Scientific Press.

[Zimmermann, 2001] Zimmermann, H.-J. (2001). *Fuzzy Set Theory: And Its Applications.* Springer.