

# Argumentation Theory in Health Care

Luca Longo, Bridget Kane, Lucy Hederman

Center for Health Informatics - School of Computer Science and Statistics - Trinity College Dublin

{longol, kanebt, hederman}@scss.tcd.ie

## Abstract

*Argumentation theory (AT) has been gaining momentum in the health care arena thanks to its intuitive and modular way of aggregating clinical evidence and taking rational decisions. The basic principles of argumentation theory are described and demonstrated in the breast cancer recurrence problem. It is shown how to represent available clinical evidence in 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 the clinician decision-making process. A encouraging predictive capacity is compared against the accuracy rate of well-established machine learning techniques confirming the potential of argumentation theory in health care.*

## 1. Introduction

The amount of clinical evidence produced in health care has been rapidly increasing thanks to the advent of Electronic Health Records and other technologies for assisting clinicians, in their activities. Although this shift is good for the advance of science and knowledge, it introduces difficulties for health practitioners 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. Intuitively, there is a tangible need to aggregate pieces of evidence in a way that reduces uncertainty and complexity, minimizing inconsistency and incompleteness. Decision-making processes call for a methodology capable of aggregating clinical knowledge intuitively, accurately and able to cope with inconsistency and conflicts among pieces of evidence.

To demonstrate *Argumentation Theory (AT)* in practice, we use a public available dataset concerned with breast cancer prognosis and/or recurrence [1]. Recurrence is a phenomenon that defines the return of a cancer after treatments 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 minimizing unnecessary therapies. A possible reasoning process an clinician can perform to predict recurrence in a woman who had a breast cancer surgery, is described in the following. 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. Similarly, from the clinician's previous experience, if the degree of malignancy of the cancer is low, there is a reason to believe recurrence is not likely but, if the number of involved nodes in the cancer area is high, the risk of recurrence is high. Arguments in such a discussion are knowledge-based arguments: they are built according to the knowledge, experience and expertise of the clinician. The natural question following such a discussion is: *What is the final outcome of the reasoning process?* Let us assume the outcome is recurrence: the cancer is likely to appear again. The problem now is that the clinician will have to justify the decision. The previous argument-based analysis 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 has been gaining momentum thanks to its ability to reason about unstructured clinical situations, in uncertain domains where available information is partial and sometimes contradictory. It can be modeled via *AT*, a paradigm that studies how arguments can be expressed, sustained and discarded, as well as the validity of the conclusion reached. It deals with formal and computable models, inspired by human-like reasoning. Its goal is to apply formal semantics [2] among arguments to compute their justification status.

This paper investigates the role of argumentation theory in health care, reviewing existing work in section 2 and introducing the building blocks of the theory in section 3. Section 4 describes how the theory can be used in a illustrative domain: the breast cancer recurrence prediction. A critical discussion of experimental results follows in Section 5, highlighting the advantages and potential of argumentation theory for enhancing and qualitatively explaining complex decision-making processes. Section 6 summaries our contribution and delineates future work.

## 2 Related work

Argumentation theory 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 Artificial Intelligence (AI). 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 [3], commonly used by humans, as well as explanatory reasoning, not available in standard non-monotonic logics such as default logic. Non-monotonic reasoning [4] 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 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. The Expected Utility Theory was firstly considered, a classical quantitative formal approach to decision-making aimed at arriving at rational decisions [9]. The theory advises to compute the expected utility of different possible choices and choose the one with the highest value. Despite its power, the difficulty with the Expected Utility Theory is that it requires lots of data to estimate the probabilities. Furthermore, the probabilistic and utility-based inferences are unnatural to clinicians and difficult to understand for patients. In the early 1990s, J. Fox and his group at Cancer Research UK proposed an approach to decision-making under uncertainty [9] 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 [10, 13]. These models have been inspired by the important contribution of Dung [8] who proposed a very abstract methodology to deal with a set of arguments that attack each other along with seman-

tics 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 [2]. An example of such a study is the work of Williams and Williamson aimed at combining probability with logic for breast cancer prognosis [13]. 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 variable and to perform the prognosis. Argumentation theory has been used to aggregate clinical evidence as well. The work of Hunter and Williams [10] was focused on aggregating clinical knowledge in an intuitive way and on reducing the inconsistency, complexity, volume and incompleteness of the available evidence.

## 3 Formal Argumentation Theory

The underlying idea behind argumentation theory is that, given a set of 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. Suppose a clinician is examining a 40 year old adult patient already in post-menopausal status, who had a breast cancer operation, with a small tumor removed. He may reason with 3 arguments, in respect to predicting recurrence: **A**: 'patient's medium age is a reason to believe *no recurrence*' **B**: 'patient's post-menopausal status is a reason to believe *recurrence*' **C**: 'the size of removed tumor is low, *no recurrence* is believed'



Figure 1. Arguments and reinstatement (AF1)

We say that an argument **B** *defeats* argument **A** if, and only if (iff), **B** is a reason against **A**. In our example, argument **B** defeats argument **A**, in fact, **B** supports *recurrence* and **A** supports *no recurrence*. If the internal structure of arguments, as well as the reasons why they defeat each other, are not detailed, the structure is called *abstract argumentation framework* [8]. An *argumentation framework* is a set of (abstract) arguments and binary defeat (attack) relations among these arguments. It specifies a directed graph in which arguments are presented as nodes and the defeat relations with arrows. Formally, 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  $(A, B) \in def$ . Now the question is which arguments should ultimately be accepted. In fig. 1 **A** is defeated by **B**, and apparently **A** should not be accepted since it has a coun-

terargument. 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 arguments can be accepted, is needed. This criterion is referred to as *semantic*, and given an argumentation framework it specifies zero or more sets of acceptable arguments, called *extensions*. Different argument based semantics have been proposed [2], but here we focus on complete, grounded and preferred semantics [8]. The issue of argument semantics can be clarified using the labeling approach proposed by Wu and Caminada [15] (we refer the reader to their works for formal definitions). Each argument is either *in*, *out*, *undec* according to the following conditions: I) an argument is labelled *in* iff all its defeaters are labelled *out*, and II) an argument is labelled *out* iff it has at least one defeater that is 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. In Figure 1, for argument 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 that A, C are *in* and B is *out*.

- A set of argument is called *conflict-free* iff it does not contain any argument A and B such that A defeats B.
- A set of arguments is said to *defend* an argument C iff each defeater of C is defeated by an argument in Args.
- A function  $F$  yields the arguments defended by a given set of arguments. It specifies the set of argument that are acceptable, in the line of Dung [8].
- Given an argumentation framework, and  $Args$  a conflict-free set of Arguments, Args is said to be a *complete extension* iff  $Args = F(Args)$ .

In the argumentation framework  $AF1$  (Fig. 1) there is just one complete extension,  $\{A, C\}$ , conflict-free that defends exactly itself. Note  $\{A, B, C\}$  is also a fixpoint of  $F$ , but is not a complete extension since it is not conflict-free. Complete semantics have a important property: more than one complete extension may exist. However, in some situation it is advantageous to apply a semantic that is guaranteed to yield exactly one extension: this is called *Grounded semantic*. The idea is to select the complete extension in which the set of *in*-labelled arguments is minimal.

The *grounded extension* coincides with the complete extension in which *in* is minimised, *out* is minimised and *undec* is maximised. In Figure 1 the grounded extension is  $\{A, C\}$ . 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 semantic* deals with this limitation. The idea is that preferred semantics, instead of maximizing *undec* arguments, it maximises *in* arguments (and also *out*). It are based on the concept of admissibility.

- A set of arguments is *admissible* iff it is conflict-free and defends at least itself.
- Given an argumentation framework  $(Ar, def)$ , and  $Args \subseteq Ar$ , Args is said to be a *preferred extension* iff Args is a maximal admissible set.

In Figure 1 there exist only one preferred extension which is  $A, C$ . For any argumentation framework, there exist at least one preferred extension. A preferred extension coincides with those labellings in which *in* is maximal and *out* is maximal. It is important to note that every preferred extension is a complete extension and every grounded extension is a complete extension. The idea of complete extension is that, a complete labeling can be seen as a subjective but reasonable point of view that an agent, in our example the clinician, 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 [15].

## 4 Study case

In this study, we have used the Ljubljana breast cancer dataset<sup>1</sup>, repeatedly used in many machine learning studies (1986 up to 2011. Example: [13]). 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 (TD) or after operation (AO), as shown in Table 1. For 81 patients, the illness reappeared after 5 years, while 196 did not have recurrence: data was verified after collection.

Let us assume an expert's knowledge-base is only based on the attributes considered in the dataset (Table 1), and it includes some relationship between attributes and the phenomenon of recurrence. We interviewed an expert in this field, and asked her to provide such relationships with tangible and valid literature references, according to her past experience, intuitions, presumptions and interpretations. This preliminary set of evidence is depicted in Table 2. From this evidence, an argumentation framework can now be built, thus a set of arguments and a set of defeat relations among arguments is needed.

<sup>1</sup>provided by the University Medical Center, Institute of Oncology, Ljubljana, ex-Yugoslavia, by M. Zwitter and M. Soklic (physicians)

Evidence	Attribute	Description	Reference(s)
1	Age	The strongest risk factor for breast cancer is age: the older the woman, the higher the risk of cancer (and presumably recurrence).	[11]
2	Menopausal	Pre-menopausal status is a reason to believe recurrence is not likely.	Presumption
3	Tumor size	In general, the greatest diameter (in mm) of the excised tumor, the greater the chance of recurrence.	Interpretation
4	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.	Interpretation
5	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.	Interpretation & Intuition
6	Degree of malignancy	The histological grade of the tumor affect recurrence. If it is 1, tumors consist of cells that, while neoplastic, retain many of their usual characteristics, then 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.	Interpretation & Intuition
7	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.	[12]
8	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.	[6]
9	Irradiation	Radiotherapy for breast cancer reduces recurrence	[7]
10	Age + Menopausal status	Late menopause increases the risk of breast cancer (and presumably recurrence). Women who have undergone the menopause have a lower risk of breast cancer (and presumably recurrence) than pre-menopausal women of the same age.	[5]

**Table 2. A possible expert's knowledge-base**

Label	Type	Time	Values range
Age	N	TD	10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90-99
Menopause	C	TD	lt40, ge40, premeno
Tumor-size	N	AO	0-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59
Inv-nodes	N	AO	0-2, 3-5, 6-8, 9-11, 12-14, 15-17, 18-20, 21-23, 24-26, 27-29, 30-32, 33-35, 36-39
Node-caps	C	AO	yes, no
Deg-malign	C	AO	1, 2, 3
Breast	C	AO	left, right
Breast-quad	C	AO	left-up, left-low, right-up, right-low, central
Irradiation	C	AO	yes, no
Output class	C	TD	no-recurrence-events, recurrence-events

**Table 1. Breast cancer dataset attributes**

- An *argument* is a set of premises whose union supports a claim:  $Arg : premise + \dots + premise \rightarrow claim$
- A *premise* is modeled via membership function.
- A *claim* is what reasonably follows the *premise(s)*.
- (Membership function) For any set  $X$ , a membership function on  $X$  is any function  $f : X \rightarrow [0, 1] \in \mathfrak{R}$ . 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  $X$ . (0 indicates non membership while 1 fully membership; intermediate values refers to fuzzy members, which partially belong to  $X$ ). For each fuzzy set  $X$ , more than one membership function can be defined.

The rationale behind adopting membership is that commonly people tend to perceive an entity as being 'low' or 'high', 'small' or 'big' and so forth. Intermediate clusters such as 'medium' might be used, but generally, people avoid more complex scales. For instance, an age of 10 is commonly perceived being a 'low' age, with 90 being 'high'. However, 50 can be considered 'high' by an expert, and 'medium' by another. This subjectivity can be influenced by factors such as past experience, knowledge or the context. Therefore, membership functions are useful

for modeling vaguely defined sets and dealing with human reasoning that is approximate rather than fixed and exact.

At this stage, the pool of arguments, needed for the breast cancer prediction problem, can be built considering the expert's knowledge base (Table 2) and her definition of membership functions over attributes (Figure 2). The output of this process is shown in Table 3. An example of an argument is:  $A : LOW\ Age \rightarrow 'no\ recurrence'$  where the *attribute* considered in the premise is 'Age', modeled by membership function  $f_{Age}^{LOW}()$  and 'no recurrence' is the claim that reasonably follows the premise. The argument can be read: 'a low age is a reason to believe no recurrence'.

Attribute: Age, Evidence: 1			Attribute: Menopause, Evidence: 2		
Arg	M. f()	Claim	Arg	M. f()	Claim
A	low	no recurrence	D	pre	no recurrence
B	medium	no recurrence	E	post-lt40	no recurrence
C	high	recurrence	F	post-ge40	recurrence
Attribute: Tumor size, Evidence: 3			Attribute: Node invas., Evidence: 4		
Arg	M. f()	Claim	Arg	M. f()	Claim
G	low	no recurrence	I	low	no recurrence
H	high	recurrence	J	high	recurrence
Attribute: Node Caps, Evidence: 5			Attribute: Deg. Malig., Evidence: 6		
Arg	M. f()	Claim	Arg	M. f()	Claim
K	true	recurrence	M	low	no recurrence
L	false	no recurrence	N	high	recurrence
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	M. f()	Claim
			O	lower	no recurrence
			P	upper	recurrence
Attribute: Irradiation, Evidence: 9			Attribute: Age+Menop., Evidence: 10		
Arg	M. f()	Claim	Arg	M. f()	Claim
Q	true	no recurrence	R	high + pre	recurrence

**Table 3. Pool of arguments**

Membership degrees that produce values greater than zero will activate the corresponding argument, defining a new set of arguments. For instance, let us consider a record of the dataset (as defined in Table 1), related to a patient:

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.

the activated arguments are (of table 3) B, D, H, I, L, N, O.

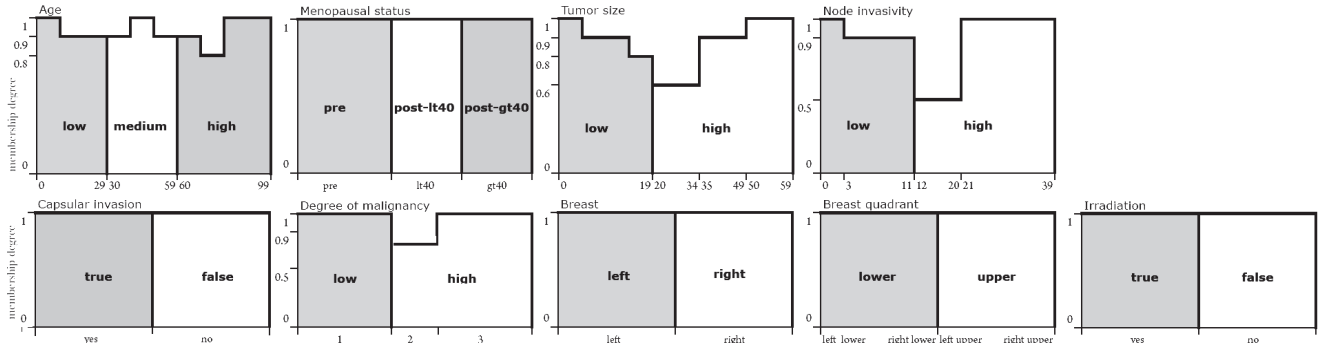


Figure 2. A possible expert's Membership functions definition

Once the activated arguments set is built, the set of defeat relations among arguments need to be formed, to complete the argumentation framework. Logically, two arguments supporting different claims attack each other. In our case study, all the arguments supporting 'recurrence' attack all the arguments supporting 'no recurrence'. For the considered record, the framework built is depicted in fig. 3.

Grounded and preferred argumentation semantics can now be run on the argumentation framework built, for each patient, to decide which arguments should be accepted.

$$AF = (\{B, D, H, I, L, N, O\}, \{(H, B), (H, D), (H, I), (H, L), (H, O), (B, H), (D, H), (I, H), (L, H), (O, H), (N, B), (N, D), (N, I), (N, L), (N, O), (B, N), (D, N), (I, N), (L, N), (O, N)\})$$

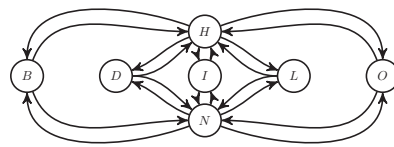


Figure 3. Argumentation framework example

## 5 Evaluation and analysis of results

The grounded semantic is a skeptical way of computing acceptable arguments: it returns just one grounded extension. 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 remaining case, *AT* fails in predicting the recurrence status: further evidence is needed.

The preferred semantic is a credulous way of computing acceptable arguments: it returns multiple coherent extensions, each composed by arguments sharing the same claim. For each record of the 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'. Two

different coherent points of view are available to support the clinician decision-making process in respect to predicting the recurrence status. One possible automatic strategy, the one tested in this study, is the selection of the biggest preferred extension. Out of 286 patients, 192 recurrence status were successfully predicted. In 6 cases, the two preferred extensions had the same size, therefore a further strategy was required to let an extension emerge. In this study we selected the extension with less uncertainty. In other words, each argument in one preferred extension was built upon a certain attribute, activating one membership function and producing a membership degree. The average of membership degrees was computed, for each preferred extension, and the extension with the highest average was chosen (less uncertain). In 3 cases this strategy was successful while in other 3 cases the recurrence status was not predicted. However, this is just an illustrative strategy that can be made available to the clinician for supporting the final decision.

Intuition suggests that, when possible, a skeptical approach should be adopted, therefore the common claim supported by the arguments in the grounded extension should be used in the decision-making process. If not possible, a more credulous approach should be applied and a strategy for selecting the winning preferred extension should be designed. An interesting advantage of *AT* is that, the 9 incomplete records, part of the original dataset, can still be considered using our approach. The only difference is that, as one or more objective values are missing, fewer arguments will be activated, but still extensions and justification status can be computed. For these 9 records, the related grounded extension was empty and our strategy positively predicted 3 cases, by analysing the corresponding preferred extensions.

Results are encouraging. An overall accuracy rate of 71.6% was obtained using *AT*:  $(7 + 192 + 3 + 3) = 205$  out of  $277 + 9 = 286$ . This rate was compared with the accuracy of some machine learning classifiers on the same dataset, with the Weka software [14], with a 10-fold cross-validation. The classification via logistic function produced an accuracy of 68.8%, the alternating decision tree (ADTree) classifier predicted the input in 73.7% cases, a Bayes network was successful 72% of the times and a multilayer perceptron

classifier had a precision of 64.7%. These rates suggest that the preliminary 71.6% obtained with *AT* is promising and further investigation can be carried out. Although there is a initial cost for embedding an expert knowledge's base into formal arguments, *AT* does not require subsequent training of data. However, *AT* has advantages and properties that most machine learning techniques do not have.

- *Inconsistency and Incompleteness*: *AT* easily deals with incomplete and inconsistent data; missing data is simply discarded and even if an argument cannot be expressed, the argumentative process can be still executed with available arguments. This property is powerful when a full dataset is incomplete or not available;
- *Expertise and Uncertainty*: *AT* captures expertise in a organised fashion, subject to uncertainty;
- *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 formal argument, input of the argumentation framework; therefore *AT* provides an intuitive approach for reasoning on available evidence, even if partial and inconsistent;
- *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/interpreted;
- *Extensibility and Updatability*: *AT* is an an extensible paradigm that allows to retract a decision in the light of new evidence: an argumentation framework can be updated with new arguments.

## 6. Conclusions and future work

Argumentation theory is an emerging paradigm adopted in health care for aggregating clinical evidence intuitively and modularly. It is not a statistical tool and it is close to the way humans reason under uncertainty with incomplete and inconsistent evidence. In this study we have presented the basic building blocks of argumentation theory, and applied its principles in a real-world health scenario: the breast cancer recurrence prediction. We have shown how to translate an expert knowledge-base in arguments and how to explicitise defeat relations among them in an organised argumentation framework. Subsequently, we have demonstrated how to interpret the outcomes of the argumentative process, and how to enhance an expert decision-making process. We compared the predictive capacity of our argumentation theory-based model against well known machine learning tools. Results are extremely promising, suggesting that further research can be carried out to provide clinicians with an intuitive argumentation framework for supporting and justifying decisions.

Future work will focus on investigating the practicability of the approach and the identification of those areas that would most benefit from the theory. An interface will allow clinicians to access the argument structure and the extensions, for a given case.

## References

- [1] Breast cancer: an environmental disease - the case for primary prevention. Technical report, The UK Working Group on the Primary Prevention of Breast Cancer., 2005.
- [2] P. Baroni, M. Caminada, and M. Giacomin. An introduction to argumentation semantics. *The Knowledge Engineering Review*, 26(4):365–410, November 2011.
- [3] P. Baroni, G. Guida, and S. Mussi. Full non-monotonicity: a new perspective in defeasible reasoning. *European Symposium on Intelligent Techniques*, pages 58–62, 1997.
- [4] C. F. Chang, A. Miller, and A. Ghose. 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, 2010.
- [5] C. Chen, N. Weiss, P. Newcomb, W. Barlow, and E. White. Hormone replacement therapy in relation to breast cancer. *Jama*, 287(6):734–41, February 2002.
- [6] P. Darbre. Recorded quadrant incidence of female breast cancer in great britain suggests a disproportionate increase in the upper outer quadrant of the breast. *Anticancer Res.*, 3c(25):254302550, May 2005.
- [7] A. B. de Gonzalez, R. Curtis, E. Gilbert, C. Berg, S. Smith, M. Stovall, and E. Ron. Second solid cancers after radiotherapy for breast cancer in seer cancer registries. *British Journal of Cancer*, 102(1):220–226, November 2010.
- [8] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.
- [9] J. Fox, V. Patkar, and R. Thomson. Decision support for healthcare: the proforma evidence base. *Informatics in Primary Care*, 14(1):49–54, 2006.
- [10] A. Hunter and M. Williams. Argumentation for aggregating clinical evidence. *22nd International Conference on Tools with Artificial Intelligence*, 1:361–368, 2010.
- [11] H. Ma, L. Bernstein, M. Pike, and G. Ursin. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. *Breast Cancer Res*, 4(8):R43, July 2006.
- [12] H. Weiss, S. Devesa, and L. Brinton. Laterality of breast cancer in the united states. *Cancer Causes Control.*, 5(7):539–543, September 1996.
- [13] M. Williams and J. Williamson. Combining argumentation and bayesian nets for breast cancer prognosis. *Journal of Logic, Language and Information*, 15(1-2):155–178, 2006.
- [14] I. H. Witten, E. Frank, and M. A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier, 2011.
- [15] Y. Wu, M. Caminada, and M. Podlaszewski. A labelling based justification status of arguments. *13th International Workshop on Non-Monotonic Reasoning, Studies in Logic*, 3(4):12–29, 2010.