

2014

## Defeasible Reasoning and Argument-Based Systems in Medical Fields: an Informal Overview

Luca Longo

*Technological University Dublin*, [luca.longo@tudublin.ie](mailto:luca.longo@tudublin.ie)

Pierpaolo Dondio

*Technological University Dublin*, [pierpaolo.dondio@tudublin.ie](mailto:pierpaolo.dondio@tudublin.ie)

Follow this and additional works at: <https://arrow.tudublin.ie/scschcomcon>



Part of the [Computer Sciences Commons](#)

### Recommended Citation

Longo, L. & Dondio, P. (2014). Defeasible reasoning and argument-based systems in medical fields: an informal overview. *Computer-Based Medical Systems (CBMS): 27th International Symposium, 27-29 May, Mount Sinai, New York*. doi:10.1109/CBMS.2014.126

This Conference Paper is brought to you for free and open access by the School of Computer Sciences at ARROW@TU Dublin. It has been accepted for inclusion in Conference papers by an authorized administrator of ARROW@TU Dublin. For more information, please contact [arrow.admin@tudublin.ie](mailto:arrow.admin@tudublin.ie), [aisling.coyne@tudublin.ie](mailto:aisling.coyne@tudublin.ie).



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 License](#)

# Defeasible reasoning and argument-based systems in medical fields: An informal overview

Luca Longo<sup>1,2</sup>, Pierpaolo Dondio<sup>1</sup>

<sup>1</sup>School of Computing, Dublin Institute of Technology

<sup>2</sup>School of Computer Science and Statistics, Trinity College Dublin

luca.longo@dit.ie, pierpaolo.dondio@dit.ie

**Abstract**—The first aim of this article is to provide readers informally with the basic notions of defeasible and non-monotonic reasoning, logics borrowed from artificial intelligence. It then describes argumentation theory, a paradigm for implementing defeasible reasoning in practice as well as the common multi-layer schema upon which argument-based models are usually built. The second aim is to describe the selection of argument-based applications in the medical and health-care sectors. Finally, the paper will conclude with a summary of the features, which make defeasible reasoning and argumentation theory attractive, that emerge from the applications under review. The target reader is a medical or health-care practitioner, with limited skills in formal knowledge representation and logic, interested in enhancing evidence modelling and aggregation.

## I. INTRODUCTION

The capability of deriving defeasible conclusions with partial information is an important aspect of modern medical systems. In order to achieve such a capability, humans routinely resort to the so-called *default knowledge*, a main feature of which is that it can be used in a reasoning process even if the preconditions for its application are only partially known. These preconditions, whose truth is not explicitly verified, are assumed to hold defeasibly, that means in the absence of explicit information to the contrary. In the event that new information becomes available and the falsity of such preconditions can be deduced, then the conclusions derived from the application of the default knowledge have to be retracted. This type of reasoning is known as *defeasible reasoning* [2]. Default knowledge is represented by using *defaults* that are specific inference rules. These are expressions of the form:  $p(x) : j_1(x), \dots, j_n(x) \longrightarrow c(x)$  where  $p(x)$  is the prerequisite of the default,  $j(x)$  is the justification and  $c(x)$  is the consequent. If  $p(x)$  is known and if  $j(x)$  is consistent with what is known, then  $c(x)$  can be defeasibly deduced. In other words, if it is believed that the prerequisite is true, and each of the  $n$  conditions (justifications) can be assumed since they are consistent with current beliefs, then this leads to believe the truth of the conclusion.

Defeasible reasoning, unlike standard deductive reasoning, is *non-monotonic*. Intuitively this means that adding new premises may lead to removing, rather than adding new conclusions. More specifically, if the conclusion  $p$  follows from a set of premises  $A$  (denoted as  $A \vdash p$ ), in standard monotonic reasoning it also holds that  $A, B \vdash p$  namely  $t$ , if and only if any additional set of premises  $B$  is added to  $A$ , the conclusion  $p$  is still valid. This property is called *monotonicity*:

conclusions are not affected by new evidence hence the set of conclusions monotonically increases. This is not the case in real life, and more particularly, in non-monotonic reasoning where conclusions can be retracted when new evidence is available. Consider the following example: *X has undergone breast cancer surgery and subsequently radiotherapy. Radiotherapy minimises the risk of cancer recurrence*, so it is possible to derive that *X has a low risk of breast cancer recurrence*. If in addition to the fact that *X has undergone cancer surgery and subsequently radiotherapy*, it is found out that *X had a cancer with high degree of lymph node involvement*, then the conclusion that *X has a low risk of cancer recurrence* has to be retracted, as a special exception has been raised. Non-monotonic logic relies on the idea that the pieces of knowledge employed in a reasoning activity such as *X has a low risk of cancer recurrence* may admit exceptions and it is impossible to include a full list of exceptions within the reasoning rules [2]. In these cases, the premise of a certain rule is only partially specified and a conclusion can be derived from the premises, assuming that no exception occurs, that means that all the implicit premises of the rule are satisfied. In the case where an exception subsequently arises then the derived conclusion has to be retracted. The basic idea of non-monotonic inferences is that, when more information is obtained, some previously accepted inference may no longer hold. Defeasible reasoning has increasingly gained attention in the medical sector because it supports reasoning over partial, incomplete and dynamic evidence and knowledge, where several exceptions can arise according to various circumstances. Argumentation theory (AT), an important sub-field of artificial intelligence (AI), provides state-of-the-art computational models of defeasible reasoning (DR).

This paper is structured as follows. Firstly, AT is introduced with an emphasis on its role in defeasible reasoning. This is followed by a detailed description of the multi-layered pattern upon which argument-based systems are usually structured follows. A brief overview of some practical applications of AT in clinical domains is then presented followed by a discussion highlighting the main advantages of DR and AT in decision-making and knowledge representation. A conclusion summarises the paper.

## II. ARGUMENTATION THEORY

Argumentation theory (AT) is a multi-disciplinary research subject ranging from law to philosophy and linguistic, with aspects borrowed from psychology and sociology. AT has

gained interest in artificial intelligence as it provides the basis for computational models inspired by the way humans reason. These models have extended classical reasoning approaches, based on deductive logic, that were proving increasingly inadequate for problems requiring non-monotonic reasoning and explanatory reasoning not available in standard non-monotonic logics [8]. AT focuses on how pieces of evidence, seen as arguments, can be represented, supported or discarded in a defeasible reasoning process, and it investigates formal models to assess the validity of the conclusions achieved [34]. AT differs from many traditional monolithic non-monotonic logics because it envisages a modular and intuitive process, supporting the explanation of each reasoning step, making the reasoning and inference processes more explanatory. Thanks to the above features AT has been employed for tasks like practical reasoning, decision support, dialogue and negotiation [3], [31], [34].

In a nutshell, argumentation deals with the interactions between possibly conflicting arguments, arising when different parties argue for and against some conclusions or when different pieces of evidence are available [23]. Arguments can be regarded as ‘tentative proofs for propositions’ [19] in a logical language whose axioms represent premises in the domain under consideration. In general, the premises are not consistent because they may lead to incompatible conclusions. As already mentioned, these conflicts may arise either during the defeasible reasoning activity of a single agent or in the context of a multi-agent dialogue. These modes are referred to as monological and dialogical argumentation, respectively. Accordingly, monological models [4] focus on the internal structure of an argument, meaning its components (like premises, rules, conclusions) and their relations. Dialogical models focus instead on argument conflicts and their solutions and typically regard arguments as monolithic entities, whose internal structure is abstracted away as far as the conflict resolution process is concerned. Roughly speaking, monological models concern the production and construction of arguments while dialogical models concern management of their conflicts, that means the actual arguing process. A third classification of models, referred to as *rhetorical*, has also been proposed in which neither the monological nor the dialogical structure is considered [4]. Here, the rhetorical nature of arguments is stressed. More specifically, the audience’s perception of arguments and how they can be employed as a means of persuasion is considered [16] [25]. Argument-based systems are generally built upon five layers [30], as it follows.

#### Layer 1: definition of the internal structure of arguments

The internal representation of arguments is addressed by monological models. Often an argument is internally represented with a set of premises, and a conclusion follows from them with the application of some rule. Many argumentation systems do not make any distinction between premises. However, arguments actually used in human reasoning may follow a more articulated structure where different premises play different roles, as in the argument model first introduced by Toulmin [35] composed of six parts (fig. 1).

- *Claim (C)*: an assertion or a claim (conclusion) that has a potentially controversial nature;
- *Data (D)*: statements specifying facts/beliefs previously

established related to a situation in which the claim is made;

- *Warrant (W)*: statement that justifies the derivation of the conclusion from the data;
- *Backing (B)*: a set of information that ensures the trustworthiness of a warrant. It is the grounds underlying the reason. A backing is invoked when the warrant is challenged;
- *Qualifier (Q)*: a statement that expresses the degree of certainty associated with the claim;
- *Rebuttal (R)*: a statement introducing a situation in which the conclusion might be defeated.

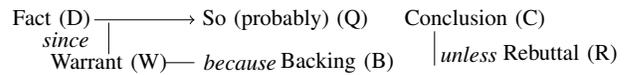


Fig. 1: An illustration of Toulmin’s argument representation

Toulmin’s model plays a significant role in highlighting the elements that might form a natural argument, and provides a useful basis for knowledge representation. Another well-known monological paradigm has been proposed by Reed and Walton to model the notion of *arguments as product* [37], [32]. It is based upon the notion of an *argumentation scheme* and it is useful for identifying and evaluating a variety of argumentation structures in everyday discourse [4]. These argumentation schemes are aimed at capturing common stereotypical patterns of reasoning that are non-monotonic and defeasible in nature. Let us suppose that B and H are discussing chemotherapy, and that H is not in favour of it because H thinks that it has a high emotional impact on the person due to the risk of alopecia phenomenon and should be discouraged. H’s argument is: *Dr. P (expert in psychology) says that chemotherapy affects the emotional state*. It appears that H’s argument is implicitly *an appeal to expert opinion*. In addition, it is evidently an instance of *argument from consequences*. These two schemes can be used by H to build her point of view. H is claiming that negatively affecting the emotional state is a bad consequence of an action. The argument is based upon the assumption that, since the bad outcome is a consequence of chemotherapy, therefore chemotherapy should not be applied. This can be represented by the following argumentation chain:

- Dr. P., an expert psychologist, says that chemotherapy negatively affects emotional state, because he has knowledge of patients emotions
- chemotherapy negatively affects emotional state
- negatively affecting the emotional state is a bad thing
- anything that leads to bad consequences is a bad practice
- chemotherapy is a bad practice.

Walton identified 25 different argumentation schemes, each including a set of *critical questions* such as ‘is the expert E in a position to know about the proposition A?’. Critical questions provide a sort of checklist about the validity conditions for the application of a specific argument scheme. Intuitively, critical questions make the defeasibility of argument schemes explicit and indicate some canonical ways to build the relevant counterarguments. Toulmin’s model as well as Reed and Walton’s approach do not specify the way different argument structures can be aggregated nor how they can interact or conflict in the dynamics of an argumentation process. Their main aim is modelling internal argument structure, specifying the criterion of acceptance for argument. The reader can refer to [4] for a description of other monological models of arguments.

#### Layer 2: definition of the conflicts between arguments

Monological models, aimed at representing the internal structure of arguments are complemented by dialogical models, focused on the relationships between arguments and, in particular, their conflicts. The latter investigates the issue of invalid arguments that appear to be valid (fallacious arguments). *Conflicts* (also called attacks or defeats, sometimes with slightly different meanings) are key notions in AT. In the AT literature several kinds of conflicts have been considered. We recall the classification proposed in [29] encompassing the three classes of undermining, rebutting and undercutting attacks.

- *undermining attack*: an argument can be attacked on one of its premises by another whose conclusion negates that premise. Example: ‘*alcohol consumption is low according to X so X has a low risk of recurrence*’ can be undermined by *the alcohol level emerging from a blood test is high so X has a high consumption of alcohol*

- *rebutting attack* occurs when an argument negates the conclusions of another argument. Example: ‘*radiotherapy minimises this risk*’ so *X has a low risk of breast cancer recurrence* can be rebutted by ‘*X is an old patient, the strongest risk for breast cancer is age, so the risk of recurrence is high*’.

- *undercutting attack* occurs when an argument uses a defeasible inference rule and is attacked by arguing that there is a special case that does not allow the application of the rule itself [27]. Example: ‘*radiotherapy minimises the risk of recurrence so X has a low risk of breast cancer recurrence*’ can be undercut by ‘*paper Z demonstrated that radiotherapy failed several times in curing breast cancer so it is not always an effective method to reduce recurrence*’.

#### Layer 3: evaluation of conflicts and definition of valid attacks

Conflict between arguments, although an important notion, does not embody any approach for evaluating an attack. The determination of the success of an attack, from one argument to its target, is another important aspect of argument-based systems. Generally an attack, sometimes referred to as ‘defeat’, has a form of a binary relation between two arguments. Some authors distinguish a relation in a weak form (attacking another argument and not weaker) or in a strong form (attacking another argument and stronger). The former is generally referred to as ‘defeat’ whereas the latter as ‘strict defeat’ [29]. Defeat relations are determined in various ways, depending on the argumentation system. Often, they are influenced by the domain of application and are usually defeasible. For instance, in those domains where observations are important, defeat relations might depend on the reliability of tests as well as on observers having different expertise. To establish whether an attack can be considered a successful defeat, a trend in AT is devoted to the consideration of the *strength* of arguments. In this respect a key concept is represented by the inequality of the strength of arguments that has to be accounted for in the computation of sets of arguments and counterarguments [9]. Several works have adopted the notion of *preferentiality* among arguments [24]. The information necessary to decide whether an attack between two arguments is successful is often assumed to be pre-specified, and implemented as an ordering of values or a given partial preference. However, according to [24], the information related to preferentiality of arguments might be contradictory, as the preferences may vary depending on the context and on different subjects who can assign different strengths, to different arguments, employing

different criteria. This led the author to propose the concept of *meta-level argument*: a simple node in a graph of nodes where preferentiality is abstractly defined, by creating a new attack relation that comes from a preference argument. Meta-level arguments allow no commitment regarding the definition of the preferences of arguments, rendering the reasoning process simpler. As opposed to the preferentiality approach, another branch of artificial intelligence is devoted to associate weights to attack relations instead of arguments [9], [22]. The classical binary relation of attack has been extended in [18] by employing the notion of *fuzzy relations* borrowed from Fuzzy Logic. This approach allows the representation of the degree to which an argument attacks another one. Similarly in [7] probabilities are assigned to arguments and defeat relations referring to the likelihood of their existence and thus capturing the inherent uncertainties in the argumentation system. In summary, the individuation of the proper valid attacks can be implemented through the notion of strength of argument, preferentiality or strength of attack relations.

#### Layer 4: definition of the dialectical status of arguments

Defeat relations, as defined in layer 3, focus on the relative strength of two individual conflicting arguments and do not tell yet what arguments can be seen as justifiable. The final state of each argument depends on the interaction with the others and a definition of their *dialectical status* is needed. This layer is aimed at determining the outcome of an argumentation system usually by splitting the set of arguments in two classes, those that support a certain decision/action and those that do not. Sometimes a further class can contain those arguments that leave the dispute in an undecided status. Multiple actions/decisions can be accounted for in a defeasible reasoning process, thus the number of classes can increase. Modern implementations for computing the dialectical status of arguments are built upon the theory of Dung [8] which, historically speaking, derives from other more practical and concrete works on argumentation [36] and [27]. Dung’s abstract argumentation frameworks (AF) allow comparisons of different systems by translating them into his abstract format [36]. The underlying idea is that given a set of abstract arguments (the internal structure is not considered) and a set of defeat relations, a decision to determine which arguments can ultimately be accepted has to be taken. AF is a directed graph in which arguments are presented as nodes and the attacks as arrows (figure 2). Solely looking at an argument’s defeaters to decide the status of an argument is not enough: it is also important to investigate whether the defeaters are defeated themselves. Generally, an argument *B* *defeats* an argument *A* if and only if *B* is a reason against *A*.

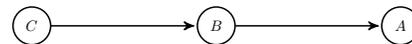


Fig. 2: Argument and reinstatement

Given an AF, the issue is to decide which arguments should ultimately be accepted. In figure 2, *A* is attacked by *B*, and apparently *A* should not be accepted since it has a counterargument. However, *B* is itself attacked by *C* that, in turn, is not attacked 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 scenario it is said that *C* *reinstates* *A* and in order to determine which arguments of an AF can

be accepted, a formal criterion is necessary. This criterion is known as *acceptability semantics*, and given an AF, it specifies zero or more *extensions* (sets of acceptable arguments) [1]. Using a labelling approach by [39], each argument is either *in*, *out* or *undec* according to the following two conditions: 1) an argument is labelled *in* if and only if all its defeaters are labelled *out*, and 2) an argument is labelled *out* if and only if it has at least one defeater labelled *in*. Informally speaking, an argument labelled *in* means it has been accepted, *out* means it has been rejected and *undec* means it can be neither accepted nor rejected.

In the AF of figure 2, for argument  $C$  it holds that all its defeaters are labelled *out* (trivial as  $C$  is not defeated by any argument), thus  $C$  has to be labelled *in*.  $B$  now has a defeater labelled *in* thus it has to be labelled *out*. For  $A$ , it holds that all its defeaters are labelled *out*, so it has to be labelled *in*. As a consequence the resulting status of each argument is:  $Lab(A) = in$ ,  $Lab(C) = in$  and  $Lab(B) = out$ . Thus,  $A$  and  $C$  can be accepted and argument  $B$  has to be rejected. 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  $Args$ . These basic notions drive the proposal of the *complete semantic* aimed at computing *complete extensions* [8]. The idea is that a complete labelling might be viewed as a subjective and reasonable point of view that a designer can consider with respect to which arguments are accepted, rejected or considered undecided. Each point of view can be certainly questioned by someone, but its internal inconsistency cannot be pointed out. The set of complete labellings can be seen as the reasonable positions available to a designer [39].

Complete semantics have an important property: more than one complete extension might exist. However, sometimes it is advantageous to take a skeptical approach, thus a semantic that is guaranteed to generate exactly one extension is the *grounded semantic*. The idea is to select the complete labelling  $Lab$  in which the set of *in*-labelled arguments is minimal. The grounded extension coincides with the complete labelling in which *in* is minimised, *out* is minimised and *undec* is maximised and can be the empty set. In figure 2, the grounded extension is  $\{A, C\}$ . However, this skeptical approach might be replaced by a more credulous one, known as *preferred semantic* [8]. The idea is that, instead of maximising *undec* arguments, it maximises *in* arguments (and also *out* arguments). They are based on the notion of *admissibility*. A set of arguments is admissible if and only if it is conflict-free and defends at least itself. The empty set is admissible in every AF as it is conflict-free and trivially defends itself against each of its defeaters (none). For any AF, there exists at least one preferred extension. Every grounded and every preferred extension is a complete extension. In figure 2, 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 exists:  $\{A, C\}$ . Other semantics exist and the reader is referred to [1].

#### Layer 5: accrual of acceptable arguments

Multiple acceptable extensions of arguments may be computed from the previous layer coinciding with possible consistent points of view that can be considered for describing the

knowledge being modelled and thus employed for decision-making and support. However, sometimes for practical purposes, as in the medical domain, a single decision/action must be taken/performed. Thus a fifth layer is sometimes added to the argumentative schema aimed at extracting the most credible or consistent point of view for taking such a decision/action. This layer includes a strategy for computing, for instance, a credibility degree of each extension that can be used for purposes of comparison. The most credible can be eventually considered and employed to support decision-making.

In summary, the aforementioned five layers give an overall idea of the main components that are usually considered in an argumentative process, and are strictly connected. The first layer deals with monological argumentation while the other layers with dialogical argumentation. Some of these layers can be missing or merged together. For example, when the strength of arguments or attack relations is not considered, layer 3 can be removed. Also, the strength of arguments may be considered in the 5th layer and not only in the 3rd layer.

### III. ARGUMENT-BASED APPLICATIONS IN THE MEDICAL AND HEALTH CARE SECTOR

The literature of defeasible reasoning (DR) is vast in the logic and artificial intelligence communities. Readers can refer to [4] for a taxonomy of argument-based models and to [5] for a review of DR implementations. In this section applications of DR and argument-based systems in the medical and health-care sectors are described.

DR was implemented in the context of the *Aspic* project [12]: a general model for argumentation services. The goal was to develop a theoretical framework for inference, decision-making, dialogue and learning. This may be useful in health-care, with potential applications in the identification of patients treatment options. Examples included a multi-agent scenario where three agents collaborate exchanging pros and cons of alternative interventions and diagnoses towards settling on a justifiable treatment for a patient with chest pain [10]. Argumentation has also been employed as a method for genetic counselling: the goal was to provide an aid both for patients and clinicians in customising, evaluating, visualising and communicating with respect to care plans. Eight cancer genetic counsellors participated in a simulated experiment in which they had to counsel a woman carrying a risk-increasing gene mutation. Information was visually displayed in an organised-fashion, in the form of structured arguments. These helped counsellors in enhancing their discussion and explaining the options available for mitigating the risk of cancer [13]. In general, within the *Aspic* project, arguments are constructed from a knowledge-base of facts, internally modelled with strict and defeasible rules. Both are composed by a set of premises supporting a claim. A tree of these rules can form an argument and a numeric degree of belief is attached to each of them (1 for strict arguments, a partial degree ( $< 1$ ) for defeasible arguments). Various principles to compute this value exist. For example, these might be the ‘weakest link’ (being this the minimum of the strength of its premises and its links) or the ‘last link’ (the maximum strength of its links - no accrual of reasons) [28]. Once arguments are defined, a set of attack relations among them is explicated, always according to the knowledge-base of facts. Dung’s calculus of opposition

[8] is employed to compute a dialectical (justification) status of arguments. Eventually, from the claim of the justified arguments a final inference is drawn, being this usually a decision, a diagnosis or a treatment recommendation.

DR has also been used in the context of medical group-decision support. Here, usually expert clinicians participate in a group discussion to decide on the best treatment for a given patient or case. In [6], for example, a web-prototype to build arguments was presented to a group of oncologists who were supposed to discuss treatment therapy for patients having cancer in the head-to-neck region. These were modelled as natural language propositions related to a particular piece of evidence, acquired from the literature, and linked to a particular treatment choice. Each argument was accompanied by a value indicating the strength of the evidence. A machinery that extended Dung's calculus of opposition [8] was proposed, adding a preference-based accrual of arguments [6]. Further research in the *Aspic* project, in the context of consensus on explanations, focused on understanding how two clinicians, with a disagreement in relation to an anomalous patients response to treatment, exchanged arguments in order to arrive at a consensus [15]. Argumentation theory (AT) has been employed for predicting the recurrence of breast cancer in patients who have undergone a surgery [21] [20]. Here, the knowledge-base of a cancer expert has been translated into arguments with premises supporting either recurrence or non-recurrence of cancer. This monological structure has been subsequently extended adding conflicts among arguments organised in a dialogical structure, always according to the expert's knowledge-base. In turn, they were evaluated with Dung's calculus of opposition using preferred semantics. A strategy was implemented for selecting the most credible preferred extension and thus proposing a justifiable outcome (recurrence or non-recurrence) that was also compared against the outcome of a few machine-learning classifiers. In this context AT performed to the same level as machine learning [20].

Gorogiannis et al. employed argumentation for investigating treatment efficacy and their work [14] was motivated by the fact that, although there was a rapidly-growing dataset of trial results, this dataset was inconsistent, incomplete and required a lot of effort to be sensibly aggregated to enable a single correct decision to be inferred. Therefore, authors proposed an argument-based approach to analyse the available knowledge and present the different possible results to the user. In their frameworks, the monological structure of arguments is a triple  $\langle A, B, C \rangle$  with A the set of evidence from a clinical trial, B an inference rule that links evidence to a claim C. A claim is a comparison between the outcomes of two generic treatments  $t_1$  and  $t_2$  (only two-arm comparisons are treated) that can be either  $t_1 > t_2$  ( $t_1$  is statistically superior to  $t_2$ ),  $t_2 < t_1$  (vice-versa) or  $t_1 \sim t_2$  (no statistical difference). Arguments (clinical tests) might conflict if they entail contradicting claims, thus authors employed the Dung's calculus of opposition, starting with clinical tests and reasons about acceptable and defeated arguments. In [38] this framework is enhanced by allowing the expression of preferences among arguments and by employing descriptive logic to further specify their monological structure. The authors performed a case study on ovarian cancer data showing how the introduction of the dialogical Dung's calculus of opposition can help the selection of relevant/undisputed clinical evidence in a large and fragmented dataset of cases.

[33] describes an application of argumentation to the field of organ transplant called Carrel+. Human-organ is a decision-making process that often illustrates conflicts among medical experts: what may be sufficient for one doctor to discard an organ may not be for another doctor. The framework allows doctors to express their arguments about the viability of an organ and employs argumentation techniques, namely argumentation schemes and critical questions, to combine arguments, to identify inconsistencies and to propose a valid solution considering their relative strength and the available evidence about the organ and the donor. The key difference between [33], [26], [11] and the works [14], [38] is the variability of the monological structure of arguments. In the former works, arguments are hand-crafted and ad-hoc construct built by relying on domain specific expertise and therefore they have a variable monological structure. In the latter works arguments are instead built directly from clinical trial results with a uniform structure that makes the approach less domain-dependent and scalable to large-volume data.

Ultimately, [17] is probably the most complete work applying argumentation to medical decision support. The work is close to the 5-layer schema introduced in previous sections. First, the available evidence, collected from experts or literature, is converted into a monological inductive argument structured as a inference rule; second the medical expert sets preference relations by assigning weight to the arguments (argument A can be preferred to B because, despite having comparable effects, A has fewer side effects than B). Third, meta-arguments can be built about the quality of those created in the first stage (an argument based on a non-randomised small sample is weaker than another based on evidence collected on a large randomised sample). Forth, the dialogical structure is arranged in a Dung style argumentation graph and an argumentation semantic is used for computing their dialectical and acceptability status from which consistent conclusions can be suggested to the decision makers. The study proposes several case studies: diagnosis of glaucoma, treatment of hypertension and treatments of pre-eclampsia.

According to the previous overview, the main benefit provided by defeasible reasoning (DR) and argumentation theory (AT) can be summarised as follows.

- *inconsistency/incompleteness*: DR provides a methodology for reasoning on available evidence, even if this evidence is partial and inconsistent as often happens in medicine;
- *expertise/uncertainty*: AT captures expertise in an organised fashion, employing the notion of arguments, handling uncertainty and the vagueness associated with clinical evidence;
- *intuitiveness*: AT is not based upon statistics or probability and is thus close to the way humans reason. If a designer is inclined to use statistical evidence anyway, this can be modelled as an argument using a monological model. Additionally, vague knowledge-bases can be structured as arguments built with familiar linguistic terms, which is extremely appealing for clinicians;
- *explainability*: AT leads to explanatory reasoning thanks to its incremental, modular way of reasoning with evidence. It provides semantics for computing arguments' justification status, allowing the final decision to be better explained;
- *dataset independency*: AT does not require a complete dataset and it may be useful for emerging knowledge, where quantitative evidence has not yet been gathered;

- *extensibility/updatability*: DR is an extensible paradigm that allows a decision to be retracted in the light of new evidence: an argumentation system can be updated with new arguments;
- *knowledge-bases comparability*: AT allows comparisons of different subjective knowledge-bases. Two clinicians might build their own argumentation frameworks and identify differences in the definition of their beliefs [21] and [20].

#### IV. CONCLUSION

This article has presented an overview of defeasible reasoning (DR) and argumentation theory (AT) with an informal description of the common pattern usually adopted for implementing such reasoning. An overview of current research conducted in the medical and health-care sectors showed where defeasible reasoning has been employed so far. The main drawback concerns the anecdotal evidence that characterises current instances of use. However, any attempt to design and adopt defeasible and argument-based systems is encouraged and supported as are comparisons with other canonical techniques for decision-support, decision-making and knowledge-representation. Finally, a summary of the appealing properties of DR and AT was provided, to guide practitioners in the selection of argument-based applications in medical environments.

#### ACKNOWLEDGEMENT

We would like to thank prof. Pietro Baroni for his feedback.

#### REFERENCES

- [1] P. Baroni, M. Caminada, and M. Giacomin. An introduction to argumentation semantics. *Knowledge Engineering Review*, 26(4):365–410, 2011.
- [2] P. Baroni, G. Guida, and S. Mussi. Full nonmonotonicity: a new perspective in defeasible reasoning. In *ESIT 97, European Symposium on Intelligent Techniques*, pages 58–62, 1997.
- [3] T. J. Bench-Capon and P. E. Dunne. Argumentation in artificial intelligence. *Artificial Intelligence*, 171(10-15):619–641, 2007.
- [4] J. Bentahar, B. Moulin, and M. Blanger. A taxonomy of argumentation models used for knowledge representation. *Art. Intell. Review*, 33(3):211–259, 2010.
- [5] D. Bryant and P. Krause. A review of current defeasible reasoning implementations. *Knowl. Eng. Rev.*, 23(3):227–260, Sept. 2008.
- [6] C. F. Chang, A. Miller, and A. Ghose. Mixed-Initiative Argumentation: Group Decision Support in Medicine. pages 43–50, 2010.
- [7] P. Dondio. Probabilistic argumentation frameworks: Basic properties and computation. In *Highlights on Practical Applications of Agents and Multi-Agent Systems*, pages 263–279, 2013.
- [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(2):321–358, 1995.
- [9] P. E. Dunne, A. Hunter, P. McBurney, S. Parsons, and M. Wooldridge. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175(2):457–486, 2011.
- [10] J. Fox, L. Black, D. Glasspool, S. Modgil, A. Oettinger, V. Patkar, and M. Williams. Towards a general model for argumentation services. In *Proc. AAI Spring Symposium Series 2006 (AAAI, 2006)*.
- [11] J. Fox and S. Das. *Safe and Sound: Artificial Intelligence in Hazardous Applications*. AAAI Press, 1st edition, 2000.
- [12] J. Fox, D. Glasspool, D. Greco, S. Modgil, M. South, and V. Patkar. Argumentation-based inference and decision making—a medical perspective. *IEEE Intelligent Systems*, 22(6):34–41, 2007.
- [13] D. Glasspool, J. Fox, A. Oettinger, and J. Smith-Spark. Argumentation in decision support for medical care planning for patients and clinicians. In *AAAI Spring Symposium: Argumentation for Consumers of Healthcare*, pages 58–63, 2006.
- [14] N. Gorogiannis, A. Hunter, V. Patkar, and M. Williams. Argumentation about treatment efficacy. *Knowledge Representation for Health-Care. Data, Processes and Guidelines*, pages 169–179, 2010.
- [15] M. A. Grando, L. Moss, D. Sleeman, and J. Kinsella. Argumentation-logic for creating and explaining medical hypotheses. *Artif. Intell. Med.*, 58(1):1–13, May 2013.
- [16] F. Grasso. Towards a framework for rhetorical argumentation. In *Proceedings of the 6th workshop on the semantics and pragmatics of dialogue*, pages 53–60, 2002.
- [17] A. Hunter and M. Williams. Argumentation for aggregating clinical evidence. In *ICTAI (1)*, pages 361–368, 2010.
- [18] J. Janssen, M. De Cock, and D. Vermeir. Fuzzy argumentation frameworks. In *Information Processing and Management of Uncertainty in Knowledge-based Systems*, pages 513–520, June 2008.
- [19] P. Krause, S. Ambler, M. Elvang-Gransson, and J. Fox. A logic of argumentation for reasoning under uncertainty. *Comp. Intell.*, 11(1):113–131, 1995.
- [20] L. Longo and L. Hederman. Argumentation theory for decision support in health-care: a comparison with machine learning. In *Brain Informatics*, 2013.
- [21] L. Longo, B. Kane, and L. Hederman. Argumentation theory in health care. In *Int. Symposium on Computer-Based Medical Systems*, 2012.
- [22] D. C. Martinez, A. J. Garcia, and G. R. Simari. An abstract argumentation framework with varied-strength attacks. In *Int. Conf. on Principles of Knowledge Representation and Reasoning*, pages 135–143, 2008.
- [23] P.-A. Matt, M. Morgem, and F. Toni. Combining statistics and arguments to compute trust. In *Int Conf. on Autonomous Agents and Multiagent Systems*, 2010.
- [24] S. Modgil. Reasoning about preferences in argumentation frameworks. *Artificial Intelligence*, 173(9-10):901–934, 2009.
- [25] P. Pasquier, I. Rahwan, F. Dignum, and L. Sonenberg. Argumentation and persuasion in the cognitive coherence theory. In *The 1st int. conf. on computational models of argument*, pages 223–234, 2006.
- [26] V. Patkar and et Al. Evidence-based guidelines and decision support services: A discussion and evaluation in triple assessment of suspected breast cancer. *British Journal of Cancer*, 95(11):1490–1496, 2006.
- [27] J. L. Pollock. Justification and defeat. *Artificial Intelligence*, 67(2):377–407, 1994.
- [28] J. L. Pollock. Defeasible reasoning with variable degrees of justification. *Artificial Intelligence*, 133:233–282, 2001.
- [29] H. Prakken. An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1(2):93–124, 2011.
- [30] H. Prakken and G. Vreeswijk. Logics for defeasible argumentation. In *Handbook of philosophical logic*, volume 4, pages 219–318, 2002.
- [31] I. Rahwan and P. McBurney. Argumentation technology (guest editors). *IEEE Intelligent Systems*, 22(6):21–23, 2007.
- [32] C. Reed and D. Walton. Argumentation schemes in argument-as-process and argument-as-product. In *Proceedings of the conference celebrating informal Logic*, volume 25, 2003.
- [33] P. Tolchinsky and et Al. Increasing human-organ transplant availability: Argumentation-based agent deliberation. *Intelligent Systems, IEEE*, 21(6):30–37, Nov 2006.
- [34] F. Toni. Argumentative agents. In *The Multiconference on Computer Science and Information Technology*, pages 223–229, 2010.
- [35] S. Toulmin. *The use of argument*. Cambridge University Press, 1958.
- [36] G. Vreeswijk. Defeasible dialectics: A controversy-oriented approach towards defeasible argumentation. *Journal of Logic and Computation*, 3:3–27, 1993.
- [37] D. Walton. *Argumentation Schemes for Presumptive Reasoning (Studies in Argumentation Theory)*. Lawrence Erlbaum Associates, Inc., 1996.
- [38] M. Williams and A. Hunter. Harnessing ontologies for argument-based decision-making in breast cancer. In *ICTAI (2)*, pages 254–261, 2007.
- [39] 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.