

## Combining Fuzzy Logic and Formal Argumentation for Legal Interpretation

Célia da Costa Pereira, Andrea G. B. Tettamanzi, Beishui Liao, Alessandra  
Malerba, Antonino Rotolo, Leendert van der Torre

► **To cite this version:**

Célia da Costa Pereira, Andrea G. B. Tettamanzi, Beishui Liao, Alessandra Malerba, Antonino Rotolo, et al.. Combining Fuzzy Logic and Formal Argumentation for Legal Interpretation. ICAIL 2017 - 16th International Conference on Artificial Intelligence and Law, Jun 2017, London, United Kingdom. pp.49-58, 10.1145/3086512.3086532 . hal-01672343

**HAL Id: hal-01672343**

**<https://hal.inria.fr/hal-01672343>**

Submitted on 24 Dec 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Combining Fuzzy Logic and Formal Argumentation for Legal Interpretation

Célia da Costa Pereira  
Université Côte d’Azur, CNRS, I3S,  
France  
celia.pereira@unice.fr

Andrea G. B. Tettamanzi  
Université Côte d’Azur, CNRS, I3S,  
France  
andrea.tettamanzi@unice.fr

Beishui Liao  
Zhejiang University, China  
baiseliao@zju.edu.cn

Alessandra Malerba  
University of Bologna, Italy  
alessandramalerba87@gmail.com

Antonino Rotolo  
University of Bologna, Italy  
antonino.rotolo@unibo.it

Leendert van der Torre  
University of Luxembourg,  
Luxembourg  
leon.vandertorre@uni.lu

## ABSTRACT

The interpretation of a norm is often uncertain and conflicting. In this paper we propose a model for arguing about legal interpretation, which considers the problems of vagueness. After motivating our adoption of graded categories as a tool to tackle the problem of open texture in legal interpretation, we introduce a model based on fuzzy logic and argumentation. Then, we conduct a case study by using an example from medically assisted reproduction.

## CCS CONCEPTS

•Applied computing → Law; •Theory of computation → Logic;

## KEYWORDS

Legal Interpretation, Fuzzy Logic, Argumentation

### ACM Reference format:

Célia da Costa Pereira, Andrea G. B. Tettamanzi, Beishui Liao, Alessandra Malerba, Antonino Rotolo, and Leendert van der Torre. 2017. Combining Fuzzy Logic and Formal Argumentation for Legal Interpretation. In *Proceedings of ICAIL ’17, London, United Kingdom, June 12-16, 2017*, 10 pages. DOI: 10.1145/3086512.3086532

## 1 INTRODUCTION

Norms are like plans which aim at achieving the social goals the members of a society decided to share [6, 7]. The legislator tries to specify all the circumstances which a norm applies to and all the exceptional contexts where it does not apply, but, as well known in the planning community of AI, universal plans rarely are a practicable strategy. An agent should rather produce a partial plan and revise it when part of it becomes unfeasible. In the same way as replanning allows an agent to revise its plans while keeping fixed its original goals, whenever *legal interpretation* associates purposes to norms it allows them to be adapted after their creation to the unforeseen situations in order to achieve the social goals they have been planned for. After all, not only the world changes, giving rise to circumstances unexpected to the legislator who introduced the norm, but even concepts can change with respect to the one constructed by the law to describe the applicability conditions of norms

(see, e.g., all the problems concerning the application of existing laws to privacy, intellectual property or technological innovations in healthcare). This adaptation can be made only at the moment of evaluating whether a given behavior in a particular situation should be considered as a violation, i.e., by judges in courts.

It is well-known in legal theory that the interpretation of legal rules is often uncertain: legal language is vague, the concepts used to describe a legal rule are not always precise, and the purpose of the rule may be differently perceived (see, e.g., [9, 13, 18]). Indeed, ordinary and legal languages are inherently general and abstract, vague, and open-textured, meaning with this last concept that it is not possible to anticipate all potential occurrences falling within the application scope of any legal norm [12, 19]: in other words, there persists the ineliminable potential that a definition of an empirical concept bounded in all now-foreseeable dimensions can break down in the face of unforeseen and unforeseeable events.

It is not obvious that all legal norms work as plans: consider, e.g., norms stating human rights. Although some influential philosophers, such as Scott Shapiro [31], argued that the law has an inherent teleological nature and that norms are plans, we do not commit here to this philosophical claim. As we will see, we rather move from the doctrinal and judicial practice (covering many norm types in most legal systems) according to which the use of normative purposes is useful to interpret legal provisions when the literal meaning of them is not clear and, thus, teleological interpretation comes to be one important option for judges<sup>1</sup>.

In the context of Normative MAS and AI&Law Boella *et al.* [6, 7] introduced a model of legal interpretation based on the purpose of a norm, and [37] offered a different logical framework for handling the role of purposes in legal interpretation. Both works do not incorporate quantitative methods for uncertainty. Rotolo *et al.* [27] built a qualitative logical machinery for reasoning about interpretive canons. Making use of deontic defeasible logics, they showed the distinguishing structure of interpretive argumentation: “the claim that a legal text ought or may be interpreted in a certain way can be supported or attacked by arguments, whose conflicts may have to be assessed according to further arguments.” Other important and more recent works in AI&Law that contributed at different extent to this topic are [1, 2, 2, 20].

ICAIL ’17, London, United Kingdom

© 2017 ACM. This is the author’s version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *Proceedings of ICAIL ’17, June 12-16, 2017*, <http://dx.doi.org/10.1145/3086512.3086532>.

<sup>1</sup>In this sense, we also see that our paper does not claim to cover all aspects of legal imprecision.

In this paper we study legal interpretation by considering the purpose of norms and incorporating uncertainty measures. In particular, we address the following problem:

**Research question** Given a norm, a specific case, and different interpretations of legal concepts of the norm (w.r.t. the case):

- 1) how to measure the uncertainty of each legal concept?
- 2) how to resolve the conflicts between various interpretations?

The general idea is to use fuzzy logic to measure the uncertainty of legal concepts (i.e., the fact that there are several categories an instance can be associated to and we are not able to choose the “real” one), and argumentation to handle the conflicts between different interpretations of norms. More precisely, we combine a fuzzy argumentation system [33] to represent the interpretations, and fuzzy labeling to evaluate the status of fuzzy arguments [8]. As done in many logical analyses of legal reasoning, our model is not purely descriptive and it is rather meant to offer a rational reconstruction for explaining and checking the robustness of interpretive arguments. Our point is thus that a formal model for legal impreciseness must be cognitively sound, i.e., that it works on reliable cognitive assumptions.

The layout of this paper is as follows. First, we motivate our adoption of graded categories as a tool to tackle the problem of open texture in legal interpretation. Second, we introduce a model of fuzzy argumentation and fuzzy labeling. Third, we conduct a case study by using an example from medically assisted reproduction. Finally, we offer some conclusions.

## 2 FROM OPEN TEXTURE TO GRADED CATEGORIES

### 2.1 Flexible Legal Interpretation Based on Graded Categories

Legal systems are the product of human mind and are then written in a natural language. This implies two facts:

- the basic processes of human cognition have to be taken into account when interpreting norms;
- as natural languages are inherently vague and imprecise so are norms.

The application of laws to a new situation is a metaphorical process: the new situation is mapped on to a situation in which applying law is obvious, by analogy. Here, by metaphor we mean using a well understood, prototypical situation to represent and reason about a less understood, novel situation. Metaphors are one of the basic building blocks of human cognition [17].

Norms are written with references to categories. As pointed out by Lakoff [16], “Categorisation is not a matter to be taken lightly. There is nothing more basic than categorization to our thought, perception, action, and speech”. “The classical view that categories are defined by shared properties is not entirely wrong [...] but [it] is only a small part of the story”. It is now clear that categories may be based on prototypes. Some categories are vague or imprecise; some do not have gradation of membership, while others do. The category “US Senator” is well defined, but categories like “rich person” or “tall man” are graded, simply because there are different

degrees of richness and tallness. However, it is important to notice that these degrees of membership depend both on the the context in which the norm will be applied and on the goal associated to the norm. To be considered tall in the Netherlands is not the same as to be considered tall in Portugal, for example. We have thus first to consider the context and then the goal associated to the norm.

We propose fuzzy logic as a suitable technical tool to capture the imprecision related to categories. More precisely, a category may be represented as a fuzzy set: the membership of an element to a category is a graded notion.

As a result, we get that a norm may apply to a given situation only to a certain extent and different norms may apply to different extents to the same situation.

**2.1.1 Fuzzy Logic.** Fuzzy logic was initiated by Lotfi Zadeh with his seminal work on fuzzy sets [36]. Fuzzy set theory provides a mathematical framework for representing and treating vagueness, imprecision, lack of information, and partial truth. Fuzzy logic is based on the notion of fuzzy set, a generalization of classical sets obtained by replacing the characteristic function of a set  $A$ ,  $\chi_A$  which takes up values in  $\{0, 1\}$  ( $\chi_A(x) = 1$  iff  $x \in A$ ,  $\chi_A(x) = 0$  otherwise) with a *membership function*  $\mu_A$ , which can take up any value in  $[0, 1]$ . The value  $\mu_A(x)$  is the membership degree of element  $x$  in  $A$ , i.e., the degree with which  $x$  belongs to  $A$ . A fuzzy set is completely defined by its membership function. In fact, we can say that a fuzzy set *is* its membership function.

**Operation on Fuzzy Sets.** The usual set-theoretic operations of union, intersection, and complement can be defined as a generalization of their counterparts on classical sets by introducing two families of operators, called triangular norms and triangular conorms [22, 29, 30]. A triangular norm (or t-norm) is a binary operation  $T : [0, 1] \times [0, 1] \rightarrow [0, 1]$  satisfying the following conditions for  $x, y, z \in [0, 1]$ :

- (1)  $T(x, y) = T(y, x)$  (commutativity);
- (2)  $T(x, T(y, z)) = T(T(x, y), z)$  (associativity);
- (3)  $y \leq z \Rightarrow T(x, y) \leq T(x, z)$  (monotonicity);
- (4)  $T(x, 1) = x$  (neutral element 1).

A well-known property about t-norms is:

$$T(x, y) \leq \min(x, y). \quad (1)$$

A triangular conorm (or t-conorm or s-norm), dual to a triangular norm, is a binary operation  $S : [0, 1] \times [0, 1] \rightarrow [0, 1]$ , whose neutral element is 0 instead of 1, with all other conditions identical to those of a t-norm:

- (1)  $S(x, y) = S(y, x)$  (commutativity);
- (2)  $S(x, S(y, z)) = S(S(x, y), z)$  (associativity);
- (3)  $y \leq z \Rightarrow S(x, y) \leq S(x, z)$  (monotonicity);
- (4)  $S(x, 0) = x$  (neutral element 0).

A well-known property about t-conorms is:

$$S(x, y) \geq \max(x, y). \quad (2)$$

If  $T$  is a t-norm, then  $S(x, y) \equiv 1 - T(1 - x, 1 - y)$  is a t-conorm and *vice versa*:  $T$  and  $S$  in this case form a *dual pair* of a t-norm and a t-conorm. Noteworthy examples of such dual pairs are:

- $T_M(x, y) = \min\{x, y\}$ ,  $S_M(x, y) = \max\{x, y\}$  (minimum t-norm and maximum t-conorm or Gödel t-norm and t-conorm);
- $T_P(x, y) = xy$ ,  $S_P(x, y) = x + y - xy$  (product t-norm and t-conorm or probabilistic product and sum);
- $T_L(x, y) = \max\{x + y - 1, 0\}$ ,  $S_L(x, y) = \min\{x + y, 1\}$  (Łukasiewicz t-norm and t-conorm or bounded sum).

For a given choice of a dual pair of a t-norm and a t-conorm  $(T, S)$ , given two fuzzy sets  $A$  and  $B$  and an element  $x$ , the set-theoretic operations of union, intersection, and complement are thus defined as follows:

$$\mu_{A \cup B}(x) = S(\mu_A(x), \mu_B(x)); \quad (3)$$

$$\mu_{A \cap B}(x) = T(\mu_A(x), \mu_B(x)); \quad (4)$$

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x). \quad (5)$$

## 2.2 The Case for Fuzzy Categories

The choice of explicitly representing the vagueness of norms, as determined by their reference to graded concepts, using fuzzy set theory is by no means obvious nor generally accepted in the legal domain. After all, judicial decisions must always be crisp and the very objective of legal interpretation is the elimination of vagueness and ambiguity. Therefore, a brief critical discussion and justification of our choice is needed.

From a legal standpoint, a fundamental objection to our proposal might be that concepts or institutional facts in norms are not really vague, but *open-textured*: “a concept is said to be open-textured if its extension is not determined for all cases in advance of its application” [5], but it has a recognized procedure for deciding whether it applies in a particular case. In law, this procedure is the judicial decision making through which the court rules the case. As a consequence of the judicial decision making, no doubt remains. Therefore, there would be no need to make use of degrees of membership.

To respond to this objection, we begin by observing that graded categories are a reality. They have been studied empirically and their existence may be inferred from linguistic phenomena. In some cases (for instance, the Kay-McDaniel theory of color terms [15]), the researchers have even been able to precisely pinpoint their dependency on the neurophysiology of human perception. Now, vague terms (i.e., words that refer to graded categories) abound in everyday language and, as we have seen, legal interpretation relies primarily on ordinary language. What we want to model by taking gradedness of categories into account is the procedure for deciding whether (or to which degree) a concept applies in a particular case. That is to say that, if the machinery we are proposing works correctly, as a result of its application we should obtain exactly what some authors call “open texture”: starting from a norm expressed by means of vague concepts (represented by graded categories) and from arguments about its purposes, we should be able to compute to which degree it applies to a particular case.

Related to this same objection is the observation that in the legal domain one needs to eventually reach a yes-or-no decision. However, this is the case also in most applications of fuzzy logic: for example, in control one needs to determine a precise numerical set point for an actuator. Several defuzzification techniques have been

proposed and are used to extract a precise numerical value from a fuzzy set. The reason why fuzzy logic is found useful by practitioners is that intermediate calculations starting from input data which are imprecise down to the defuzzification step can be specified intuitively and naturally based on the expert’s understanding of the problem and their translation in terms of mathematical calculations does not overlook the inherent imprecision of the inputs and carries it forward in order to give a faithful account of the imprecision of the results.

Dismissing the use of fuzzy logic on the grounds that a precise result is required would be like pretending that all the input data are known precisely (when they are not) in order to be able to carry out precise calculations. In other words, we do not deny that legal reasoning must eventually come up with a yes-or-no decision; what we argue is that there is much to gain by explicitly modeling the vagueness and imprecision of legal concepts and taking it into full account throughout the reasoning process that leads to the final binary decision! Furthermore, even when an all-or-nothing decision is required, having a clear picture of the amount of imprecision or uncertainty associated with it may be highly relevant: think of the “beyond any reasonable doubt” principle in criminal law.

Other, quite usual critiques of the use of fuzzy logic in the legal domain are those voiced, for instance, by Bench-Capon and Sergot [5]:

- (1) Difficulty of attaching any meaning to statements such as “it is 0.342 true that it is the case that John is tall”.
- (2) Difficulty to determine the membership function of a fuzzy set (linked to the previous one).
- (3) The arbitrariness of the rules of combination (i.e., the dual pair of a t-norm and a t-conorm chosen to define intersection and union).

One possible reply to Critique #1 is that the actual values attributed to the objects are not important *per se*. What is important is that they let us get a ranking among the objects we are considering and make comparisons among them. They are qualitative, not quantitative degrees. An alternative reply is to define and justify precise operational rules to compute degrees of membership of elements in fuzzy concepts, which was done, e.g., in [14, 24] and it is what we are going to do in Section 2.3 below.

As for Critique #3, the choice of a particular dual pair is hardly arbitrary, being usually justified by the specific interpretation of membership degrees. For instance, one popular choice, if the membership degrees are interpreted as qualitative degrees, is to adopt the  $(T_M, S_M) = (\min, \max)$  dual pair, as  $\min$  and  $\max$  do not introduce any arbitrary value; if, instead, the membership degrees are interpreted as probabilities, the adoption of the  $(T_P, S_P)$  is natural.

As for Critique #2, this substantially concerns the methods to determine a membership function for a fuzzy concept. This problem is well known in the field of fuzzy logic and many solutions to it have been proposed. Within this legal-interpretation context, we think that an appropriate solution would be to apply one of the methods proposed in the literature for obtaining membership functions from consensus. An alternative would be to apply one of the methods described in [10] for the construction of possibility distributions which are fuzzy sets. A third possibility, which is the

one we propose below, is to define the membership function of a category based on a similarity measure.

Be it what it may, our proposal is independent of the particular method employed to gather information about the definition of graded categories. It just assumes one such method is available.

### 2.3 Representing Norms

A norm  $r$  may be represented as a rule  $b_1, \dots, b_n \Rightarrow l$  such that  $l$  is the legal effect of  $r$  (such as an obligation linked to the norm) (cf. [28]). A norm has then a conditional structure (if  $b_1, \dots, b_n$  hold, then  $l$  ought to be the case). An agent is compliant with respect to this norm if  $l$  is obtained whenever  $b_1, \dots, b_n$  is derived. Often, logical models of legal reasoning assume that conditions of norms give a complete description of their applicability (for a discussion of the literature, see [28]).

However, this assumption is too strong, due to the complexity and dynamics of the world. Norms cannot take into account all the possible conditions where they should or should not be applied, giving rise to the so called “penumbra”: a core of cases which can clearly be classified as belonging to the concept. By a penumbra of hard cases, membership of the concept can be disputed. Moreover, not only the world changes as also pointed out in [18], giving rise to circumstances unexpected for the legislator who introduced the norm, but even the ontology of reality can change with respect to the one constructed by the law to describe the applicability conditions of norms. See, e.g., the problems concerning the application of existing laws to privacy, intellectual property or technological innovations in healthcare. To cope with unforeseen circumstances, the judicial system, at the moment in which a case concerning a violation is discussed in court, is empowered to interpret, i.e., to adjust the scope of norms, under some restrictions not to go beyond the purpose from which the norms stem.

The clauses of a norm often refer to imprecise concepts (open texture), which can take up different meanings depending on the purpose of the norm. As we have said, those imprecise concepts are a product of the human mind and, more precisely, of a categorization process. According to prototype theory, which is one of the most prominent and influential accounts of the cognitive processes of categorization, each category is defined by one or, in some variants [35], more prototypes, which are typical exemplars of it. A prototype may be regarded as being represented by a property list which has salient properties of the objects that are classified into the concept.

We may formalize these notions in a way that is compatible with an underlying knowledge representation standard and technical infrastructure like the ones provided by the W3C for the Semantic Web, i.e.: OWL (based on description logics) for the terminological part and RDF for the assertional part. This would allow a practical implementation of our proposal using state-of-the-art knowledge engineering technologies. Nevertheless, we will keep our formalization abstract for the sake of clarity.

*Definition 2.1 (Language).* Given a knowledge base  $K$ , an *atom* is a unary or binary predicate of the form  $C(s)$ ,  $R(s_1, s_2)$ , where the predicate symbol  $C$  is a concept name in  $K$  and  $R$  is a role name in  $K$ ,  $s, s_1, s_2$  are *terms*. A *term* is either a variable (denoted by  $x, y, z$ )

or a constant (denoted by  $a, b, c$ ) standing for an individual name or data value.

According to this formalisation, an individual object  $o$  is described by all the facts of the form  $C(o)$ ,  $R(o, y)$  and  $R(y, o)$  such that  $K \models C(o)$ ,  $K \models R(o, y)$  and  $K \models R(y, o)$ . We will call these facts the *properties* of  $o$ .

*Definition 2.2 (Graded Category).* A *graded category*  $\tilde{C}$  is described by a non-empty set of prototypes  $\text{Prot}(\tilde{C}) = \{o_1, o_2, \dots, o_n\}$ , where each  $o_i \in \text{Prot}(\tilde{C})$  is an individual name in  $K$ .

We can consider that the choice of the actual (more plausible) category with respect to a prototype may be seen as if the prototype represented a kind of generalisation, which applied deductively, will allow to “classify” (categorise) new “problems” (instances) [3].

The membership of an instance to a category depends on its similarity to (one of) its prototype(s). Using a similarity measure with values in  $[0, 1]$  allows us to represent graded categories as fuzzy sets. A similarity measure of that kind may be defined. Here, we adapt the contrast model of similarity proposed by Tversky [34]. In such a model, an object is represented by means of a set of features and the similarity between two objects is defined as an increasing function of the features in common to the two objects (which we may call *common features*), and as a decreasing function of the features that are present in one object but not in the other (which we may call *distinctive features*).

*Definition 2.3 (Number of Common Features).* Given two objects (i.e., individuals)  $a, b$  in  $K$ , the number of their common features  $c(a, b)$  is defined as

$$\begin{aligned} c(a, b) &= \|\{C : K \models C(a) \wedge K \models C(b)\}\| \\ &+ \|\{(R, c) : K \models R(a, c) \wedge K \models R(b, c)\}\| \\ &+ \|\{(c, R) : K \models R(c, a) \wedge K \models R(c, b)\}\|, \end{aligned}$$

where  $\wedge$  represents the *and* operator.

*Definition 2.4 (Number of Distinctive Features).* Given two objects (i.e., individuals)  $a, b$  in  $K$ , the number of their distinctive features  $\text{dis}(a, b)$  is defined as

$$\begin{aligned} \text{dis}(a, b) &= \|\{C : K \models C(a) \oplus C(b)\}\| \\ &+ \|\{(R, c) : K \models R(a, c) \oplus R(b, c)\}\| \\ &+ \|\{(c, R) : K \models R(c, a) \oplus R(c, b)\}\|, \end{aligned}$$

where  $\oplus$  represents the *exclusive or*.

It might be the case, in a given application, that some features be more important than others. This might be taken into account by defining different weights for each feature, depending on the application.

*Definition 2.5 (Object Similarity).* Given two objects (i.e., individuals)  $a, b$  in  $K$ , their *similarity* is defined as

$$s(a, b) = \frac{c(a, b)}{c(a, b) + \text{dis}(a, b)}.$$

This similarity function satisfies a number of desirable properties. For all individuals  $a, b$ ,

- (1)  $0 \leq s(a, b) \leq 1$ ;
- (2)  $s(a, b) = 1$  if and only if  $a = b$ ;
- (3)  $s(a, b) = s(b, a)$ .

We may now define the notion of membership degree of an object  $o$  in a graded category.

*Definition 2.6.* Given a graded category  $\tilde{C}$  and an arbitrary individual name  $o$ , the *degree of membership* of  $o$  in  $\tilde{C}$  is given by

$$\mu_{\tilde{C}}(o) = \bigvee_{p \in \text{Prot}(\tilde{C})} s(o, p).$$

Since the category of an item in the left-hand-side of a rule may be vague or imprecise, the degrees of truth of such an item with respect to the actual (known, believed) situation may be partial. This implies that a rule can be partially activated, i.e., the state of affairs to be reached thanks to the compliance to that rule can be uncertain.

Let us consider the following rule  $r: b_1, \dots, b_n \Rightarrow l$ , where the clauses  $b_i$  have the form “ $o_i$  is  $\tilde{C}_i$ ” and let  $\tilde{C}_1, \dots, \tilde{C}_n$  be the categories of  $b_1, \dots, b_n$ , respectively. A clause  $b_i$  of a norm involving a graded category may thus be true only to a degree. The premise of the norm will then be partially true and a norm may thus apply only to some extent.

If the membership of an instance in a category depends on its similarity to the prototype of the category and also on the purpose of the norm, then we must conclude that both the prototype of a category and the similarity measure used to compute the membership might vary as a function of the purpose. While it may be hard to see how the similarity measure could change as a function of purpose, it is reasonable to assume that the legislators may have different prototypes in mind for a category with the same name when they write norms for different purposes.

This amounts to assuming that, given a graded category  $\tilde{C}$ , its set of prototypes may vary as a function of the purpose or goal  $G$  of the norm. We will write  $\text{Prot}(\tilde{C} \mid G)$  to denote the set of the prototypes of category  $\tilde{C}$  when the purpose of a norm is  $G$ .

The degree of truth  $\alpha_{iG}$  of clause  $b_i = “o_i$  is  $\tilde{C}_i”$ , given that the purpose of the norm is  $G$ , may be computed as

$$\alpha_{iG} = \mu_{\tilde{C}_i}(o_i \mid G) = \bigvee_{p \in \text{Prot}(\tilde{C}_i \mid G)} s(o_i, p). \quad (6)$$

*Definition 2.7.* The *degree to which the premise*  $b_1, \dots, b_n$  of rule of the form  $b_1, \dots, b_n \Rightarrow l$  is satisfied, given that the purpose of  $r$  is  $G$ , is given by

$$\text{Deg}(b_1, \dots, b_n \Rightarrow l \mid G) = \bigwedge_{i=1, \dots, n} \alpha_{iG}.$$

The state of affairs which will be reached thanks to the compliance of  $r$  will be associated with the truth degree of  $\text{Deg}(r \mid G)$  – this is also the degree associated to  $l$  after the activation of  $r$ .

### 3 FUZZY ARGUMENTATION AND FUZZY LABELING

In recent years, several research efforts have been made to combine formal argumentation and fuzzy logic such that: the uncertainty of arguments can be measured by their fuzzy degrees, while the conflicts between arguments be properly handled by Dung’s argumentation semantics. Among them, [33] proposed a quantitative preference based argumentation system, called F-ASPIC. Based on ASPIC and the fuzzy set theory, it can be used to model structured argumentation with fuzzy concepts. But, it is not clear how the status of a fuzzy argument is evaluated. Meanwhile, in [8], the

authors introduced a labeling-based approach to evaluate the status of fuzzy arguments. So, in this paper, we combine these two approaches, to lay a foundation for legal interpretation.

#### 3.1 Fuzzy Argumentation System

We propose to use a fuzzy argumentation system which is based on Tamani and Croitoru’s F-ASPIC [33], with some adaptations to make it fit our framework, and with the addition of the fuzzy labeling algorithm proposed by [8].

In our framework, we do not need to represent rules with different degrees of importance, as Tamani and Croitoru do. In our framework, unlike in Tamani and Croitoru’s F-ASPIC argumentation system, the antecedent of a rule may be partially satisfied, if it involves graded categories. As a consequence, the consequent of that rule will have a partial truth degree and an argument depending on that rule will have a partial membership in the set  $\mathcal{A}$  of “active” arguments [8]. Although from a semantical point of view these gradual notions of partial truth or satisfaction are quite different from Tamani and Croitoru’s notion of *importance* and *strength*, they lead to a mathematical treatment which is formally identical. Our main adaptation of F-ASPIC is therefore to replace, in the wording and in the formalism, these notions. In this paper, for the sake of simplicity, we assume that every element of the language and every rule are fallible. Hence, we do not differentiate between strict rules and defeasible rules, as the ASPIC+ does, but we assume that we only have defeasible rules. Technically, this does not make things simpler (partial truth is basically preserved via strict rules, since they encode indisputable inferences).

*Definition 3.1 (Fuzzy argumentation system).* A fuzzy argumentation system, denoted as *FAS*, is a tuple  $(\mathcal{L}, cf, \mathcal{R}, n, \text{Deg})$  where

- $\mathcal{L}$  is a logical language,
- $cf$  is a contrariness function (in this paper, we only consider the classical negation  $\neg$ ),
- $\mathcal{R}$  is the set of (defeasible) inference rules of the form  $\phi_1, \dots, \phi_m \Rightarrow \phi$  (where  $\phi_i, \phi \in \mathcal{L}$ ),
- $n: \mathcal{R} \mapsto \mathcal{L}$  is a naming convention for rules,
- $\text{Deg}: \mathcal{R} \rightarrow [0, 1]$  is a function returning the degree of activation of a rule, given a grounding of the formulas occurring in it. Intuitively,  $\text{Deg}(r)$  represents the degree of truth of the antecedent of  $r$ .

In the original F-ASPIC system, fuzzy arguments are then constructed with respect to a fuzzy knowledge base  $\mathcal{K}$ , assigning a degree of importance  $\mu_{\mathcal{K}}(p)$  to each proposition  $p \in \mathcal{L}$ . In our framework, however, we do not attach a degree of importance to propositions of formulas *per se*, but we need to evaluate a degree of truth of their grounding with respect to graded categories. To be more precise, the atomic propositions that are liable to have a partial degree of truth are those of the form “ $x$  is  $C$ ”, where  $C$  is a graded category. Given a substitution of variable  $x$  with an individual object  $o$ , the truth value of the grounding “ $o$  is  $C$ ” will be given, as suggested in the previous section, by the similarity measure  $s(o, p)$  of  $o$  to one of the prototypes  $p$  of  $C$  (i.e., one  $p$  in the set  $\text{Prot}(C)$ ). To this aim, we keep the same symbol  $\mathcal{K}$ , but we regard it as a fuzzy valuation function.

*Definition 3.2 (Fuzzy Valuation Function).* A fuzzy valuation function in a FAS  $(\mathcal{L}, cf, \mathcal{R}, n, \text{Deg})$  is a fuzzy set  $\mathcal{K} : \mathcal{L}_{\text{ground}} \rightarrow [0, 1]$  such that:

- if  $\phi \in \mathcal{L}_{\text{ground}}$  is a ground atomic proposition of the form “ $o$  is  $C$ ”, with  $C$  a graded category,

$$\mathcal{K}(o \text{ is } C) = \bigvee_{p \in \text{Prot}(C)} s(o, p); \quad (7)$$

- if  $\phi \in \mathcal{L}_{\text{ground}}$  is a ground atomic proposition not involving graded categories,  $\mathcal{K}(\phi) \in \{0, 1\}$ ;
- if  $\phi, \psi \in \mathcal{L}_{\text{ground}}$ ,

$$\begin{aligned} \mathcal{K}(\neg\phi) &= 1 - \mathcal{K}(\phi) \\ \mathcal{K}(\phi \wedge \psi) &= T(\mathcal{K}(\phi), \mathcal{K}(\psi)) \\ \mathcal{K}(\phi \vee \psi) &= S(\mathcal{K}(\phi), \mathcal{K}(\psi)) \end{aligned}$$

where  $T$  represents a triangular norm and  $S$  an associated triangular co-norm.

**REMARK 1.** Let  $r : b_1, \dots, b_n \Rightarrow l$  a rule. In a very simple case, the degree of activation  $\text{Deg}$  of  $r$  simply corresponds to the value returned by the Fuzzy Valuation Function  $\mathcal{K}(\bigwedge_{1 \leq k \leq n} b_k)$ .

*Definition 3.3 (Fuzzy argument).* A fuzzy argument  $A$  on the basis of an argumentation theory with fuzzy valuation function  $\mathcal{K}$  and a fuzzy argumentation system is

- (1)  $\phi$  if  $\phi \in \mathcal{L}$  with:  $\text{Prem}(A) = \{\phi\}$ ,  $\text{Conc}(A) = \phi$ ,  $\text{Sub}(A) = \{A\}$ ,  $\text{Rules}(A) = \emptyset$ .
- (2)  $A_1, \dots, A_m \Rightarrow \phi$  if  $A_1, \dots, A_m$  are arguments such that there exists a rule  $\text{Conc}(A_1), \dots, \text{Conc}(A_m) \Rightarrow \psi$  in  $\mathcal{R}$ . In this case,  $\text{Prem}(A) = \text{Prem}(A_1) \cup \dots \cup \text{Prem}(A_m)$ ,  $\text{Conc}(A) = \psi$ ,  $\text{Sub}(A) = \text{Sub}(A_1) \cup \dots \cup \text{Sub}(A_m) \cup \{A\}$ ,  $\text{Rules}(A) = \text{Rules}(A_1) \cup \dots \cup \text{Rules}(A_m) \cup \{\text{Conc}(A_1), \dots, \text{Conc}(A_m) \Rightarrow \psi\}$ .

Given an argument  $A$ ,  $\text{Conc}(A)$  denotes the conclusion of  $A$ ,  $\text{Prem}(A)$  the set of the premises of  $A$ ,  $\text{Sub}(A)$  the set of the sub-arguments of  $A$  (including  $A$  itself), and  $\text{Rules}(A)$  the set of rules involved in  $A$ .

Then, the degree of activation of each argument is measured by a fuzzy degree, called *strength of argument* in F-ASPIC, which can also be interpreted as a degree of membership in the set of active arguments, defined as follows.

*Definition 3.4 (Strength of argument).* Given a fuzzy argument  $A$ , its strength, denoted  $\mathcal{A}(A)$ , is defined as follows:

- if  $A$  is of the form  $\phi$ , then  $\mathcal{A}(A) = \mathcal{K}(\phi)$ ;
- otherwise,

$$\mathcal{A}(A) = \bigvee_{r \in \text{Rules}(A)} T \left( \text{Deg}(r), \bigwedge_{\phi \in \text{Prem}(A)} \mathcal{K}(\phi) \right). \quad (8)$$

Then, with respect to the notions of rebut, undercut and defeat in ASPIC, the counterparts in the setting of fuzzy argumentation are defined as follows.

Unlike F-ASPIC, our framework does not require the definition of a fuzzy counterpart of the rebut, undercut, and defeat relation. We rely on the usual crisp relations, defined as follows.

*Definition 3.5 (Attacks).*  $A$  attacks  $B$  iff  $A$  undercuts, rebuts or undermines  $B$ , where<sup>2</sup>:

- $A$  undercuts  $B$  (on  $B'$ ) iff  $\text{Conc}(A) = \neg n(r)$  for some  $B' \in \text{Sub}(B)$ .
- $A$  rebuts  $B$  (on  $B'$ ) iff  $\text{Conc}(A) = \neg\phi$  for some  $\exists B' \in \text{Sub}(B)$  of the form  $B'_1, \dots, B'_m \Rightarrow \phi$ .
- $A$  undermines  $B$  (on  $B'$ ) iff  $\text{Conc}(A) = \neg\phi$  for some  $B' = \phi$ ,  $\phi \in \text{Prem}(B)$ .

*Definition 3.6 (Defeat).*  $A$  defeats  $B$  iff  $A$  undercuts  $B$  on  $B'$ , or  $A$  rebuts (undermines)  $B$  on  $B'$  and  $\mathcal{A}(A) \not\leq \mathcal{A}(B')$ .

We use  $\mathcal{A}$  and  $\mathcal{D}$  to denote, respectively, the fuzzy set of active arguments (whose membership is their strength) and the defeat relation between them. Then, a fuzzy argumentation framework is represented as  $\mathcal{F} = (\mathcal{A}, \mathcal{D})$ .

This fuzzification of  $\mathcal{A}$  provides a natural way of associating strengths to arguments, and suggests rethinking the labeling of an argumentation framework in terms of fuzzy degrees of argument acceptability [8]. The status of arguments can thus be evaluated by means of Fuzzy AF-labeling.

*Definition 3.7 (Fuzzy AF-labeling).* Let  $(\mathcal{A}, \mathcal{D})$  be a fuzzy argumentation framework. A fuzzy AF-labeling is a total function  $\alpha : \mathcal{A} \mapsto [0, 1]$ .

*Definition 3.8 (Fuzzy Reinstatement labeling).* Let  $(\mathcal{A}, \mathcal{D})$  be a fuzzy argumentation framework, and  $\alpha$  be a fuzzy AF-labeling. We say that  $\alpha$  is a fuzzy reinstatement labeling iff, for each argument  $A$ ,

$$\alpha(A) = \min\{\mathcal{A}(A), 1 - \max_{B:(B,A) \in \mathcal{D}} \alpha(B)\}. \quad (9)$$

In [8], it has been made clear that given a fuzzy argumentation framework, its fuzzy reinstatement labeling may be computed by solving a system of  $n$  non-linear equations, where  $n = \|\text{supp}(\mathcal{A})\|$ , i.e., the number of arguments belonging to some non-zero degree in the fuzzy argumentation framework, of the same form as Equation 9, in  $n$  unknown variables, namely, the labels  $\alpha(A)$  for all  $A \in \text{supp}(\mathcal{A})$ .

This can be done quite efficiently using an iterative method as follows: we start with an all-in labeling (a labeling in which every argument is labeled with the degree it belongs to  $\mathcal{A}$ ). We denote by  $\alpha_0 = \mathcal{A}$  this initial labeling, and by  $\alpha_t$  the labeling obtained after the  $t^{\text{th}}$  iteration of the labeling algorithm.

*Definition 3.9.* Let  $\alpha_t$  be a fuzzy labeling. An iteration in  $\alpha_t$  is carried out by computing a new labeling  $\alpha_{t+1}$  for all arguments  $A$  as follows:

$$\alpha_{t+1}(A) = \frac{1}{2} \alpha_t(A) + \frac{1}{2} \min\{\mathcal{A}(A), 1 - \max_{B:(B,A) \in \mathcal{D}} \alpha_t(B)\}. \quad (10)$$

Note that Equation 10 guarantees that  $\alpha_t(A) \leq \mathcal{A}(A)$  for all arguments  $A$  and for each step of the algorithm.

The above definition actually defines a sequence  $\{\alpha_t\}_{t=0,1,\dots}$  of labelings, whose convergence has been proven [8]. We may now define the fuzzy labeling of a fuzzy argumentation framework as the limit of  $\{\alpha_t\}_{t=0,1,\dots}$ .

<sup>2</sup>The function  $n$  is a naming convention for rules, which maps each rule to a well-formed formula in  $\mathcal{L}$ : see [21].

*Definition 3.10.* Let  $\langle \mathcal{A}, \mathcal{D} \rangle$  be a fuzzy argumentation framework. A fuzzy reinstatement labeling for such argumentation framework is, for all arguments  $A$ ,

$$\alpha(A) = \lim_{t \rightarrow \infty} \alpha_t(A). \quad (11)$$

Once this fuzzy reinstatement labeling has been computed,  $\alpha(A)$  will give the degree to which each argument  $A$  in the framework is accepted; this degree may be used to compute the corresponding degree to which the purpose of a norm is  $G$ :

$$\alpha(G) = \max_{A: Conc(A)=G} \alpha(A). \quad (12)$$

As it is clear from the above definitions, an argument may be accepted partially and thus the purpose of a norm may be uncertain. Now, different strategies may be used to deal with such an uncertainty. One possibility is to consider the purpose  $G$  for which  $\alpha(G)$  is maximal. Another is to evaluate the norm with respect to all purposes such that  $\alpha(G) > 0$  and then combine the results weighted by their corresponding  $\alpha(G)$ .

#### 4 INTERPRETING A NORM WITH FLEXIBILITY

As we noticed, in addition to taking graded categories into account, any norm is always associated with a purpose (or set of purposes): that is what is called the *purpose* or *goal* of the norm. The idea is then to capture the fact that, when a legislator states a norm, she has in mind a state of affairs to be reached through the compliance with that norm. With that in mind, the degree to which a concept in the rule belongs to a category would also depend on the purpose associated with the rule. In other words, given a norm like  $b_1, \dots, b_n \Rightarrow l$ , the degree associated to  $l$  depends on the degrees of truth of conditions  $b_i$ . These degrees depend in turn on the purpose associated to the norm: for example, the more is the extent in which the prohibition to smoke in public spaces promotes the purpose *public health*, the greater is the degree of applicability of a rule like  $Public\_Space \Rightarrow No\_Smoking$  assuming the fuzziness of the concept  $Public\_Space$ . However, the actual purpose of the legislator can be controversial [18]: for example, not enough evidence or factual information might be available which could help discovering what the legislator was intending when writing a norm. Note that the historical purpose could be obsolete due to social, economic or political change, and the legislator has not reacted in a timely manner or at all. Here, as done in legal theory (cf. [23, 28]), we adopt an objective (teleological) approach to interpretation, which means that the purpose of a norm is the one that any rational interpreter would assign to it. Hence, we use an argumentative (logical) system which will determine which purpose, with respect to the current knowledge, is the most plausible (best argued for) purpose of a norm.

Sterility, as a gradual concept, makes a good example of the point we want to make: the definition of its membership function may depend on the purpose of the norm.

For instance, if the purpose of the Italian Legislative Act n. 40/2004 on “Medically Assisted Reproduction” (on which we base our Case Study; see below) is to avoid abuses like character selection (“I want a boy, not a girl”) and so to demote eugenic manipulations, which is probably the correct historical interpretation,

given the public debates in Italy at the time that act was passed, then the definition of *sterile* will have to be rather strict. For instance, the membership might be defined as

$$\mu_{sterile}(x) = 1 - \frac{k}{n(x)}, \quad (13)$$

where  $0 < k \leq 1$  a constant and  $n(x)$  is the number of ovulations after which repeated attempts at obtaining pregnancy succeed for couple  $x$ , independently of whatever happens next. With  $k = 1$ , this would yield zero if the woman gets pregnant after the first ovulation, 0.5 after the second ovulation, 0.75 after the third, etc. If one considers that an embryo is a human being since its conception, that’s a reasonable definition. If one considers, as Italian laws on abortion implicitly do, that an embryo “becomes” human being, as it were, after a certain number of weeks of pregnancy, then one could modify the above definition by defining the “success” of an attempt as when the woman gets pregnant and she does not have a spontaneous abortion before her embryo becomes human.

If, at the other extreme, the purpose is to maximize the chances of propagation of the couple’s genes, then the definition will have to take into account the capability of the offspring to reach maturity and beget offspring at their turn. According to such a definition, a couple with a severe genetic anomaly like thalassemia major would have a very high membership in the fuzzy set of sterile couples (perhaps even a full membership if the probability of having an embryo not affected by the disease is zero). Couples with other, less severe, genetic anomalies might have a sterility degree comprised between 0 and 1, whereas a couple begetting only perfectly healthy children would have a sterility degree of zero. (See, e.g., the discussion in *Ordinanza Corte Costituzionale* n. 369, 9 November 2006.)

One might come up with other definitions reflecting other, perhaps less extreme, possible purposes of the norm, but this should be enough to convince the reader of the fact that (i) vague concepts like “sterile” can be understood as graded (= fuzzy) concepts (ii) whose definition may depend on the purpose of the norm.

##### 4.1 Applying the Framework: A Case Study

The case study in our paper is the application of the Italian Legislative Act 40/2004 on “Medically Assisted Reproduction.” Before the declaration of unconstitutionality ruled by the Constitutional Court (opinion 96/2015), the statute included section 4, par. 1: “The recourse to medically assisted reproduction techniques is allowed only [...] in the cases of sterility or infertility [...]” Purpose of the discussion is to see whether this provision can be interpreted so that non-sterile or fertile couples, in which one or both spouses are immune carriers of a serious genetic anomaly, could access those techniques.

These couples are able to conceive and bear a child, though the probability that the baby will contract the disease is high. These diseases are normally severely disabling, provoke physical dysfunctions, often prevent the full psychological development of the baby, and can cause premature death. The mentioned medical techniques can detect the illness in advance and consequently let the parents take aware decisions about the pregnancy.

The legislative act does not explicitly define ‘sterility’ and ‘infertility’: On the basis of art. 7, l. 40/2004, every three years, the Ministry of Health is required to promulgate a decree containing

the updated guidelines for the application of the law. According to these guidelines, the terms ‘sterility’ and ‘infertility’ are considered synonyms and refer to the lack of conception, in addition to those cases of certified pathology, after 12/24 months of regular sexual relations in a heterosexual couple.

In civil law systems, when it comes to statutory interpretation, one option is teleological interpretation, according to which, when interpreting a provision, judges often take into account what explicit or implicit purposes can be ascribed to the norm [18, 23].

As for the purposes, law 40/2004 states as follows:

**Art. 1, on “Purposes”.** Par.1: In order to favour the solution of reproduction problems caused by human sterility or infertility, it is allowed the recourse to medically assisted reproduction techniques, according to the conditions and the modalities provided for by the present law, which guarantees the respect of the rights of all the subjects involved, included the conceived baby.

Let us also consider the following norm from art. 4 of L. 40/2004:

The use of techniques of medically assisted procreation is [...] confined to the cases issue of infertility or [...] sterility certified by a medical procedure.

Besides, law n. 40/2004 is connected to other statutes of the legal system. In particular, the Italian Legislative Act 194/1978 on “Social Protection of Maternity and Abortion” provides also for the possibility of a therapeutic abortion if, during pregnancy, it is medically ascertained a pathological condition, among which those relating to significant anomalies or malformations of the baby, that puts at risk the physical or psychic health of the woman.” Severe genetic diseases are thus included. Moreover, along law 194/1978, the chance of a serious danger for the life of the woman is seen as a reason to proceed to abortion. This second legislative act is thus meant to promote the right to health both of the mother and the child.

In light of the previous remarks, we can outline a list of interpretive arguments supporting different interpretations. Main target is to see what interpretation better promotes the purposes that can be ascribed to the norm, if a purpose can be considered prominent, and what attacks can occur.

In what follows we present a plausible set of rules representing norms and interpretive legal arguments about such norms [27]. In both cases, fuzzy argumentation is related to the promotion of legal purposes.

In particular, the following (defeasible) rules can identify the basic the interpretive arguments  $arg_1, arg_2, arg_3$ , respectively, at stake:

- $r_1 : \neg Ste(x), Rsn\_Exp\_Life(x) \Rightarrow \neg Med\_Rpr(x)$
- $r_2 : Med\_Rpr(x), Genetic\_Dis(x), Well\_Being(x) \Rightarrow Sol\_Rep\_Prob(x)$
- $r_3 : \neg Sol\_Rep\_Prob(x), Genetic\_Dis(x) \Rightarrow \neg Rsn\_Exp\_Life(x)$
- $r_4 : Gener\_Child(x) \Rightarrow \neg Ste(x)$

where

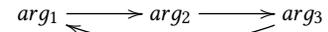
- $Ste(x) = “x$  is sterile”,

- $Med\_Rpr(x) = “x$  can access to medically assisted reproduction techniques”,
- $Rsn\_Exp\_Life(x) = “x$  grants a reasonably expected life”,
- $Genetic\_Dis(x) = “x$  is affected by a serious genetic disease”,
- $Well\_Being(x) = “x$  enjoys psychological well-being”,
- $Sol\_Rep\_Prob(x) = “legally solved for  $x$  the reproduction problems”,$
- $Gener\_Child(x) = “x$  can generate children”.

Consider the case mentioned above: a couple is actually able to conceive and generate children ( $Gener\_Child(CP)$ ), but they are both carriers of a serious genetic disease ( $Genetic\_Dis(CP)$ ), which does not allow children to live for more than a few years. Then according to the above rules, we have the following arguments:

- $arg_1 = \neg Sol\_Rep\_Prob(CP), Genetic\_Dis(CP) \Rightarrow \neg Rsn\_Exp\_Life(CP)$
- $arg_2 = Gener\_Child(CP) \Rightarrow \neg Ste(CP) \Rightarrow Rsn\_Exp\_Life(CP), \neg Ste(CP) \Rightarrow \neg Med\_Rpr(CP)$
- $arg_3 = Med\_Rpr(CP), Genetic\_Dis(CP), Well\_Being(CP) \Rightarrow Sol\_Rep\_Prob(CP)$ .

The attack relation between arguments are:  $arg_1$  attacks  $arg_2$ ,  $arg_2$  attacks  $arg_3$ , and  $arg_3$  attacks  $arg_1$ . Then, we have the following argumentation framework:



**Figure 1: An argumentation framework**

Let us consider these purposes:

- **Hlth\_Of\_MnC** = “purpose: the right to health both of the mother and the child”; this purpose is associated to rule  $r_2$ , i.e., we assume that  $r_2$  promotes purpose **Hlth\_Of\_MnC**;
- **No\_Eugenic** = “purpose: no eugenic selection”; this purpose is associated to rules  $r_1$  and  $r_4$ , i.e., we assume that  $r_1$  and  $r_4$  promote purpose **No\_Eugenic**.

For the sake of illustration, let us also assume that only two concepts are fuzzy:  $Gener\_Child$  and  $Well\_Being$ . Hence, if we consider, for example,  $r_4$ , this means that fuzziness depends only on the fact that rule  $r_4$  makes the degree of  $\neg Ste(CP)$  as dependent on the degree of capability of generating children by CP. No other source of vagueness are considered for  $r_4$ . Analogous considerations apply to rule  $r_2$  in regard to  $Well\_Being$ .

Given these purposes, we can measure the *degrees to which the premise* of rules  $r_2$  and  $r_4$  are satisfied by CP.

- **Rule  $r_4$** : Let us assume that only one prototype  $p_1$  is associated to  $Gener\_Child$  and **No\_Eugenic** (for example, a standard fertile couple statistically identified in the population of couples in which, among others, the expected life of children is greater than 50 years and the incidence of genetic diseases is less than 20%). Clearly, these are distinctive features that differentiates  $p_1$  with respect to CP: suppose that the overall distinctive features are  $d_1, \dots, d_6$ , while the common features are  $c_1, \dots, c_4$ .

If we apply Definition 2.5, then  $s(\text{CP}, p_1) = \frac{4}{4+6} = \frac{4}{10} = 0.4$ . Since  $p$  is the unique prototype for  $\text{Gener\_Child}$  with respect to  $\text{No\_Eugenic}$  and that  $G$  for  $r_4$  is  $\{\text{No\_Eugenic}\}$ , then it is easy to check that (see, in particular, Definitions 2.6 and 3.2)

$$\begin{aligned} \mu_{\text{Gener\_Child}}(\text{CP}) &= \text{Deg}(r_4 | G) = \\ &= \mathcal{K}(\text{Gener\_Child}(\text{CP})) = 0.4. \end{aligned}$$

- **Rule  $r_2$ :** Let us assume that only one prototype  $p_2$  is associated to  $\text{Well\_Being}$  and  $\text{Hlth\_Of\_MnC}$ , and that the overall distinctive features are  $d'_1, \dots, d'_{16}$ , while the common features are  $c'_1, \dots, c'_4$ . For the same reason<sup>3</sup>,

$$\begin{aligned} s(\text{CP}, p_2) &= \mu_{\text{Well\_Being}}(\text{CP}) = \text{Deg}(r_2 | G') = \\ &= \mathcal{K}(A(r_2)) = 0.2. \end{aligned}$$

Given these degrees of activation of rules, the following table illustrates how to apply the machinery of fuzzy labeling to this scenario, given the above degrees of activation of the rules that determine the strength of arguments. As we noticed, we defined the fuzzy labeling of a fuzzy argumentation framework as the limit of  $\{\alpha_t\}_{t=0,1,\dots}$ . The convergence is obtained quickly: a small number of iterations is enough to get so close to the limit.

t	$\alpha_t(\text{arg1})$	$\alpha_t(\text{arg2})$	$\alpha_t(\text{arg3})$
0	1	0.4	0.2
1	0.9	0.2	0.2
2	0.85	0.15	0.2
3	0.825	0.15	0.2
4	0.8125	0.1625	<b>0.2</b>
5	<b>0.8</b>	0.175	↓
6	↓	<b>0.2</b>	

**Table 1: Fuzzy labeling**

Therefore,  $\text{arg}_1$  is accepted to degree 0.8 while  $\text{arg}_2$  and  $\text{arg}_3$  are given a much lower acceptance degree, namely 0.2. In other words,  $\text{arg}_1$  is much more acceptable than  $\text{arg}_2$  and  $\text{arg}_3$ . It is important to observe that these degrees just represent an order of plausibility, as if saying that  $\text{arg}_1$  is four times as plausible as  $\text{arg}_2$  or  $\text{arg}_3$ .

## 5 CONCLUSION

We have proposed a framework for legal interpretation capable of taking graded, purpose-dependent institutional facts into account, which uses argumentation to handle conflicts between different interpretations of legal concepts. The idea of using argumentation to solve conflicts in legal reasoning is not new (see, e.g. [26]), but our proposal is original in that we use argumentation to identify the most likely purpose of a norm, which in turn circumscribes the interpretation of the categories (institutional facts, legal concepts) referred to by the norm. Open problems, which will be the object of future investigation, include the identification of candidate purposes from the legal rules and the extraction of arguments from the text of the norms. The definition of the membership function

<sup>3</sup>Note that  $A(r_2)$  stands for  $\text{Med\_Rpr}(\text{CP}) \wedge \text{Genetic\_Dis}(\text{CP}) \wedge \text{Well\_Being}(\text{CP})$ .

of the fuzzy sets representing graded categories might appear arbitrary; a set of guidelines would be welcome to somehow restrict the discretionary power related to their definition.

Several works in the literature of AI & Law have considered the role of purposes in the legal interpretation. Indeed, this idea is standard in legal theory and the purpose of legal rules are recognised by jurists as decisive in clarifying the scope of the legal concepts that qualify the applicability conditions for those rules [4, 11, 25, 32]. [4, 25] use purposes/goals and values in frameworks of case based reasoning for modelling precedents mainly in a common law context. [32] analyses a number of legal arguments even in statutory law, which include cases close to the ones discussed here. [11] addresses, among others, the problem of reconstructing extensive and restrictive interpretation. This is done in Reason-Based Logic, a logical formalism that can deal with rules and reasons: the idea is that the satisfaction of rules' applicability conditions is usually a reason for application of these rules, but there can also be other (and possibly competing) reasons, among which we have the purposes that led the legislator to make the rules. More recently, works such as [6, 7, 37] proposed formal models for teleological interpretation in statutory law. All these approaches in AI & Law highlight the importance of rule purposes/goals. However, it seems that no work so far has attempted to couple this view with fuzzy logic and argumentation. In this perspective, we believe that this paper may contribute to fill a gap in the literature.

## ACKNOWLEDGMENTS

The authors have received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 690974 for the project MIREL: *MI*ning and *RE*asoning with *LE*gal texts.

## REFERENCES

- [1] Michal Araszkiwicz and Tomasz Zurek. Comprehensive framework embracing the complexity of statutory interpretation. In *Legal Knowledge and Information Systems - JURIX 2015: The Twenty-Eighth Annual Conference, Braga, Portugal, December 10-11, 2015*, pages 145–148, 2015.
- [2] Michal Araszkiwicz and Tomasz Zurek. Interpreting agents. In *Legal Knowledge and Information Systems - JURIX 2016: The Twenty-Ninth Annual Conference*, pages 13–22, 2016.
- [3] Kevin D. Ashley. Reasoning with cases and hypotheticals in HYPO. *International Journal of Man-Machine Studies*, 34(6):753–796, 1991.
- [4] T.J.M. Bench-Capon. The missing link revisited: The role of teleology in representing legal argument. *Artif. Intell. Law*, 10(1-3):79–94, 2002.
- [5] Trevor Bench-Capon and Marek Sergot. Toward a rule-based representation of open texture in law. In Charles Walter, editor, *Computer Power and Legal Language: The Use of Computational Linguistics, Artificial Intelligence, and Expert Systems in the Law*, pages 39–60. Praeger, 1988.
- [6] G. Boella, G. Governatori, A. Rotolo, and L. van der Torre. *Lex Minus Dixit Quam Voluit, Lex Magis Dixit Quam Voluit: A formal study on legal compliance and interpretation*. In P. Casanovas, U. Pagallo, G. Sartor, and G. Ajani, editors, *AICOL-1/IVR-XXIV and AICOL-II/JURIX 2009 Revised Selected Papers*, pages 162–183, 2009.
- [7] G. Boella, G. Governatori, A. Rotolo, and L. van der Torre. A logical understanding of legal interpretation. In F. Lin, U. Sattler, and M. Truszczynski, editors, *KR 2010*, 2010.
- [8] Célia da Costa Pereira, Andrea G. B. Tettamanzi, and Serena Villata. Changing one's mind: Erase or rewind? possibilistic belief revision with fuzzy argumentation based on trust. In Toby Walsh, editor, *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI'11), Barcelona, Catalonia, Spain, July 16–22, 2011*, pages 164–171. AAAI, 2011.
- [9] Anthony D'Amato. Legal uncertainty. *California Law Review*, 71(1):1–55, 1983.
- [10] Didier Dubois and Henri Prade. Practical methods for constructing possibility distributions. *Int. J. Intell. Syst.*, 31(3):215–239, 2016.

- [11] J. Hage. *Reasoning with Rules: An Essay on Legal Reasoning and Its Underlying Logic*. Kluwer, 1997.
- [12] H. L. A. Hart. *The Concept of Law*. Clarendon Press, Oxford, 1994.
- [13] Ph. Heck. *Begriffsbildung und Interessensjurisprudenz*. Mohr Siebeck, Tübingen, 1932.
- [14] Juergen Hollatz. Analogy making in legal reasoning with neural networks and fuzzy logic. *Artificial Intelligence and Law*, 7:289–301, 1999.
- [15] Paul Kay and Chad K. McDaniel. The linguistic significance of the meanings of basic color terms. *Language*, 54(3):610–646, September 1978.
- [16] G. Lakoff. *Women, Fire, and Dangerous Things*. University of Chicago Press, Chicago, 1987.
- [17] G. Lakoff and M. Johnson. *Metaphors We Live By*. University of Chicago Press, Chicago, 1980.
- [18] D. Liebwald. Law's capacity for vagueness. *Int J Semiot Law*, 26:391–423, 2013.
- [19] D.N. MacCormick and R.S. Summers, editors. *Interpreting Statutes: A Comparative Study*. Ashgate, 1991.
- [20] Alessandra Malerba, Antonino Rotolo, and Guido Governatori. Interpretation across legal systems. In *Legal Knowledge and Information Systems - JURIX 2016: The Twenty-Ninth Annual Conference*, pages 83–92, 2016.
- [21] Sanjay Modgil and Henry Prakken. The ASPIC<sup>+</sup> framework for structured argumentation: a tutorial. *Argument & Computation*, 5(1):31–62, 2014.
- [22] M. Navara. Triangular norms and conorms. *Scholarpedia*, 2(3):2398, 2007. revision #137537.
- [23] Aleksander Peczenik. *On Law and Reason*. Kluwer, 1989.
- [24] Lothar Philipps. Vague legal concepts and fuzzy logic: An attempt to determine the required period of waiting after traffic accidents. *Informativa e Diritto*, II:37–51, 1993.
- [25] H. Prakken. An exercise in formalising teleological case-based reasoning. *Artif. Intell. Law*, 10:113–133, 2002.
- [26] Henry Prakken and Giovanni Sartor. A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law*, 4:331–368, 1996.
- [27] Antonino Rotolo, Guido Governatori, and Giovanni Sartor. Deontic defeasible reasoning in legal interpretation: two options for modelling interpretive arguments. In Ted Sichelman and Katie Atkinson, editors, *ICAIL 2015*, pages 99–108, 2015.
- [28] G. Sartor. *Legal reasoning: A cognitive approach to the law*, volume 5 of *A Treatise of Legal Philosophy and General Jurisprudence*. Springer, Berlin, 2005.
- [29] B. Schweizer and A. Sklar. Statistical metric spaces. *Pacific J. Math.*, 10(1):313–334, 1960.
- [30] B. Schweizer and A. Sklar. *Probabilistic metric spaces*. North Holland series in probability and applied mathematics. North Holland, 1983.
- [31] S.J. Shapiro. *Legality*. Harvard University Press, 2011.
- [32] D.B. Skalak and E.L. Rissland. Arguments and cases: An inevitable intertwining. *Artif. Intell. Law*, 1:3–44, 1992.
- [33] N. Tamani and M. Croitoru. A quantitative preference-based structured argumentation system for decision support. In *Proceedings of the 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1408–1415, 2014.
- [34] Amos Tversky. Features of similarity. *Psychological Review*, 84(4):327–352, 1977.
- [35] W. Vanpaemel, G. Storms, and B. Ons. A varying abstraction model for categorization. In B. Bara, L. Barsalou, and M. Bucciarelli, editors, *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, pages 2277–2282, Mahwah, NJ, 2005. Lawrence Erlbaum.
- [36] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
- [37] Tomasz Zurek and Michal Araszkievicz. Modeling teleological interpretation. In *International Conference on Artificial Intelligence and Law, ICAIL '13, Rome, Italy, June 10-14, 2013*, pages 160–168, 2013.