

# Qualifying and Quantifying Uncertainty in Digital Humanities: A Fuzzy-Logic Approach

Patricia Martin-Rodilla\*

Centro Singular de Investigación en  
Tecnoloxías Intelixentes (CiTIUS)  
Universidade de Santiago de Compostela  
Jenaro de la Fuente Domínguez, s/n  
15782 Santiago de Compostela, Spain  
patricia.martin.rodilla@usc.es

Martín Pereira-Fariña

Institute of Heritage Sciences (Incipit)  
Spanish National Research Council  
(CSIC)  
Avda. Vigo, s/n  
15705 Santiago de Compostela, Spain  
martin.pereira-farina@incipit.csic.es

Cesar Gonzalez-Perez

Institute of Heritage Sciences (Incipit)  
Spanish National Research Council  
(CSIC)  
Avda. Vigo, s/n  
15705 Santiago de Compostela, Spain  
cesar.gonzalez-perez@incipit.csic.es

## ABSTRACT

Research in digital humanities involves the need for conscious and explicit handling of data uncertainty. Recently, some initiatives have highlighted the importance of considering this uncertainty from the conceptual model to the final phases of implementation of software tools. Although the conceptual proposals for handling data uncertainty in the humanities have proliferated successfully, there is still a gap in bringing these proposals to actual implementations of software systems, especially due to the need to quantify this uncertainty and adopt more analytical paradigms based on margins of error, away from the conceptualization of data in the humanities. Trying to close this gap and avoid these paradigms, this paper presents a framework based on fuzzy logic that implements aspects of epistemic uncertainty (uncertainty related to our knowledge about the object of study) in digital humanities, providing algebraic and computational support to implementation. The framework proposes a solution to implement software systems that manage epistemic uncertainty, allowing comparisons, aggregations or reasoning with various levels of uncertainty in humanistic data. The solution is implemented in a real digital humanities project, illustrating its possibilities.

## CCS CONCEPTS

• **Information systems** → **Uncertainty**; • **Applied computing** → **Arts and humanities**; • **Software and its engineering** → **System description languages**; • **Theory of computation** → **Data modeling**; • **Theory of computation** → **Incomplete, inconsistent, and uncertain databases**.

## KEYWORDS

Uncertainty, Fuzziness, Imprecision, Digital Humanities, Fuzzy Logic, Knowledge representation, Conceptual modelling, ConML.

## 1 INTRODUCTION

Research practices in humanities necessarily deal with a high level of vague information, entailing the implicit coexistence with diverse degrees of imprecision or uncertainty, either due to the ignorance of certain aspects of the information by the researchers

—called epistemic vagueness— or due to the nature of the entities studied are intrinsically vague —called ontological vagueness— [1]. In the first case, we can talk about uncertainty and it is the result of situations where our knowledge about something is unclear and/or incomplete, for example when we are not sure about the exact location of a certain ancient city or about the belonging of an object or document to a certain cultural group or historical period. In the second case, the denoted objects have ill-defined borders, they have a grey area, such as the boundaries of a mountain which are impossible to determine clearly even with the best measurement methods.

The need to explicitly manage vague information, especially the epistemically vague form, is given by the nature of humanistic practice, in which the entities, problems or hypotheses handled contain high degrees of uncertainty when studying realities that are difficult to assess, such as eminently human aspects (such as social behaviors and actions, the language, artistic aspects, etc.) or to observe (for example in fields of study such as history, archeology or similar where the object of study is situated in the past, with a high level of uncertainty in our knowledge about it).

Recently, this epistemic vagueness has been addressed from an innovative perspective, 1) avoiding approximations based on the concepts of margin of error or similar, which come from analytical disciplines and assume that uncertainty is necessarily related to *error*, and 2) representing uncertainty explicitly as a valuable element from the point of view of humanistic research, capturing it from the early stages of work [2, 3].

This approach proposes the creation of conceptual models [4, 5] as a first step for their implementation in databases or similar software processing devices [3], while supporting the explicit management of uncertainty throughout. However, most of the works reviewed show that there is still a gap between conceptual proposals for humanities information systems and their subsequent implementation in the form of software systems. This gap is due to the inherent need of software systems to quantify uncertainty in order to ensure that it can be treated as a valuable dimension of the humanistic information that is expressed. In other words, we must offer a complete proposal of quantitative implementation for uncertainty if we want to be able to implement information systems

that can handle uncertain information and convey this uncertainty to researchers and other users.

Addressing this gap following the same vein proposed by these recent works, this paper develops and validates a framework based on fuzzy logic, a type of logic specifically created for the management of vague predicates or sentences. This approach allows us to implement uncertainty-aware conceptual models in through the qualification and quantification of epistemic uncertainty. The resulting artefacts are software systems that can process and expresses several degrees of uncertainty in a linguistic way that is intuitive to their users.

The paper is organized as follows: section 2 reviews the current approaches on computational uncertainty management and implementation, including the ConML conceptual modelling language [6], which includes uncertainty support. Section 3 presents our fuzzy-logic framework, taking ConML as reference. Section 3 describes the implementation of a fuzzy-logic system for a real-world database within a digital humanities project. Finally, section 4 discusses the results and future implications of this work.

## 2 FROM UNCERTAINTY MODELLING TO ITS COMPUTATIONAL TREATMENT

Previous works have studied uncertainty as a component of information from various disciplines, such as philosophy [7], mathematics [8, 9], linguistics [10, 11] or engineering [12-14]. Specifically, [15] proposes a classification of approaches in modeling uncertain information in two large blocks. On the one hand, we can find contributions outside the humanities, mostly statistical, mathematical and engineering approaches. Most of these approaches come from fields outside the humanities, and employ a paradigm based on margins of error. Thus, their application to research problems in the humanities is complex, as they require adaptations that we will detail later in our solution. Also, the strong foundation on error theory hinders their applicability even more, as error is one source of uncertainty but not the only one and, often, not even the most important. On the other hand, some recent approaches from the digital humanities seem promising too, especially those related to metadata frameworks (but without explicit uncertainty support) [16-19] or entity-based information modelling with some degree of representation of certainty for specific cases, such as CIDOC-CRM [20] or the *ad hoc* creation of folksonomies or similar mechanisms to solve particular uncertain cases. Recently, a more general solution to this problem is attempted by the ConML modelling language [6], based on the object-oriented paradigm and with explicit support for uncertainty (as well as other “soft” issues such as temporality or subjectivity) in conceptual models for humanities [21].

These approaches contribute alternatives for the initial conceptualization of uncertain information. However, only the first group (those based on a margin of error foundation) have been developed in the form of implementation proposals for software systems (databases, repositories, etc.) that allow us to quantify that uncertainty and perform computations about it. Existing implementation solutions are therefore based on 1) analytical disciplines whose informational vagueness approach differs

substantially from that in the humanities and 2) implementations fundamentally based on the concept of margin of error or similar, which (incorrectly) assume that epistemic uncertainty is always due to the deviation of our knowledge in relation to a “true” reference value, and try to compute that deviation. These approaches appear, for example, in uncertainty models for weather forecasting or in product recommendation systems for e-commerce, where uncertainty is not taken as a valuable asset for researchers, but as a quantifiable but not desirable informational characteristic. Having said this, there are some attempts to apply these approaches to the humanities, in the fields of geography or archaeology [22, 23], in GIS software [24-26] or some *ad hoc* fuzzy systems for small sets of variables that avoid the concept of margin of error, such as fuzzy logic models used in linguistic or archaeological studies [27]. It is also possible to find some applications whose *ad hoc* relational and non-relational databases handle implicit degrees of vagueness, although these implementations are always linked to specific projects or domains [3, 28] and are therefore difficult to generalize and apply elsewhere.

In summary, the gap between the conceptual support of uncertainty achieved in recent works for humanities domains and the existing proposals for explicit implementation is clear. To the best of our knowledge, no work covers this gap. The work presented here tackles this issue, taking the ConML modeling language as a basis to supports the explicit modelling of epistemic vagueness, and constructing an implementation framework based on fuzzy logic for information having epistemic uncertainty in the humanities.

The following sections briefly introduce the conceptual mechanisms that allow to express in ConML the epistemic uncertainty and how fuzzy logic allows its quantification and implementation in software systems.

### 2.1 A Theoretical Framework for Uncertainty Modelling

ConML is a general-purpose and simple conceptual modelling language designed to be affordable to users with no previous experience in information technologies. Through a graphical notation and textual annotations, it allows users to express basic linguistic statements such as existence, identity, predication, classification and subsumption, [1] (Chapter 2) through conventional object-oriented constructs such as classes, attributes, associations, objects, values and links. ConML superficially resembles UML [29, 30] but it is much simpler, and geared towards conceptual modelling rather than the specification of software systems. In addition, ConML supports the modelling of “soft” issues that are especially relevant to the humanities and social sciences, such as temporality, subjectivity and vagueness [21, 31]. A comprehensive description of ConML is out of scope in this paper, but it can be found in [1, 6]

In this paper we only focus on uncertainty or epistemic vagueness, which is one of the two kinds of vagueness handled by ConML (the other one being ontological vagueness). Uncertainty is mainly expressed through statements, in the sense that a particular subject can evaluate his/her statement by means of a

degree of truth. For example, if I state that “Alexander the Great had five children”, I may want to qualify it by adding that “it is likely”.

In this manner, any kind of statement (be it about existence, identity, predication, classification or subsumption) can in principle be moderated by a certainty qualifier. In practice, ConML supports certainty qualifiers for existence and predication, and support for other kinds of statements (identity, classification and subsumption) is ongoing work. In this manner, with ConML we can represent statements such as “it is probable that there is a person who such and such” (existence) or “it is very uncertain whether the temple was over 25 metres high” (predication). The ConML notation to express this consists of specific symbols that are added at the end of a statement in parenthesis. For example:

- $p: \text{Person} (+)$
- $b.\text{Use} = \text{Trading} (*)$

The first example states that it is probable that there is an entity  $p$  of type *Person*. The second example states that entity  $b$  (a building) is certainly used for trading purposes. The symbols and meanings employed by ConML to express certainty are as follows:

- *Certain*, indicated by an asterisk “\*”
- *Probable*, indicated by a plus sign “+”
- *Possible*, indicated by a swung dash “~”
- *Improbable*, indicated by a dash “-”
- *Impossible*, indicated by an exclamation mark “!”

This conceptualisation describes certainty (or its absence) in a 5-point scale by using specific linguistic labels and matching symbols. Of course, other scales and labels would be possible. In this paper, we present a proposal that, to some extent, questions the existing ones, so that better alternatives can be considered.

In any case, expressing levels of certainty through linguistic labels alone may be very useful for humans (qualifying the certainty presented in any model), but is not optimal for computers, as comparison, aggregation, reasoning and other automated computations become very difficult. For this reason, in this paper we present an approach to add a quantification layer on top of this conceptualisation of uncertainty. By adding quantitative values to each of the labels above by means of fuzzy sets, we will be able to make automated approximate inferences from data sets having embedded certainty qualifiers.

## 2.2 Fuzzy edges for levels of uncertainty

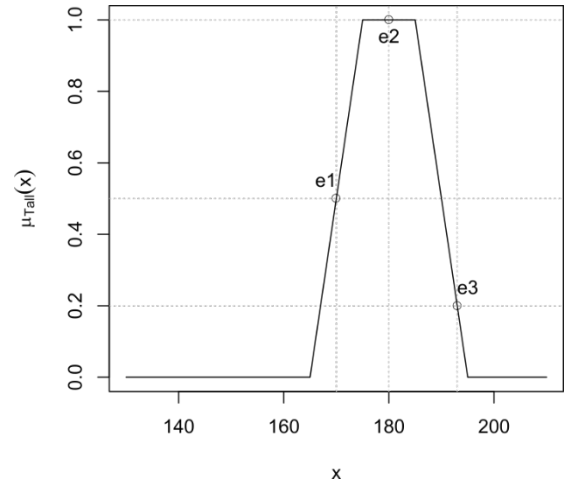
Fuzzy logic proposes the concept of linguistic variable [32], a formal framework for handling variables which values are words or sentences in a natural or artificial language. For instance, the predicate “height” is a linguistic variable when it takes linguistic values such as “tall”, “very tall”, “not so tall”, “short”, etc.

Formally, a linguistic variable is a quintuple  $(X, T(X), U, G, M)$  in which  $X$  is the name of the variable;  $T(X)$  is the term-set of  $X$ ,  $U$  is a universe of discourse;  $G$  is a syntactic rule which generates the

terms in  $T(X)$ ; and  $M$  is a semantic rule which defines the meaning of each value of  $X$ ,  $M(X)$ , where  $M(X)$  is a fuzzy subset of  $U$ .

A fuzzy set is defined as a class of objects characterised by a membership function which assigns to each object a grade of membership in the interval  $[0,1]$  [33]. Fuzzy sets are essentially devoted to the computational representation of vague predicates, such as “tall”, since it is impossible to determine sharply the border between “tall” and “not-tall”, if we accept that someone which height is 1.80cm is tall, is 1.79cm not tall at all? Let us consider the an universe  $U$  with the height in cm, and three individuals with the following height:  $e^1=170\text{cm}$ ,  $e^2=180\text{cm}$ ,  $e^3=193\text{cm}$ , and the fuzzy set tall defined as a trapezoidal fuzzy number:  $\mu_{\text{Tall}}(x) = \{165,175,185,195\}$ , where the Support =  $\{165,195\}$  and the Core =  $\{175,185\}$  (see **Figure 1**). The evaluation of the three individuals gives us the following result:

$$\begin{aligned}\mu_{\text{Tall}}(e^2) &= 0.5 \\ \mu_{\text{Tall}}(e^2) &= 1 \\ \mu_{\text{Tall}}(e^3) &= 0.2\end{aligned}$$



**Figure 1. Fuzzy set modelling the predicate "Tall".**

One of the main strengths of fuzzy logic is its context-dependent definition. The height 1.80cm should be considered “extremely tall” if our universe of reference are pygmy peoples (adult men are on average less than 150 cm) while it only should be considered as “tall” if our reference are Swedish people. With the linguistic variable, we define different fuzzy sets for the same set of values according to the context; being this relationship characterised by a compatibility function, which establishes the meaning of the linguistic variable according to the context.

The linguistic variable can be also applied to different types of concepts, such as truth or probability [32]. Usually, each linguistic variable takes between five and nine values [34] and compatibility function is result of expert knowledge and contextual information.

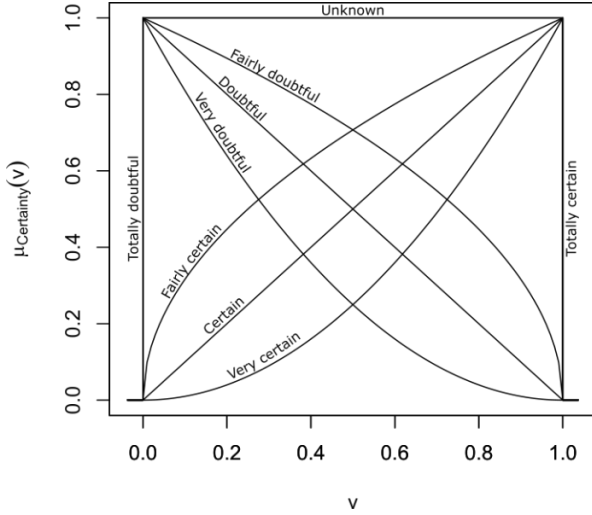
## 3 SOLUTION

In this paper, we define the notion of certainty defined in Section 2.1 as a linguistic variable in order to address the degree of confidence or belief that a subject has on a particular statement. We

define nine possible linguistic values: Thus, the linguistic variable Certainty (C), is composed by the following terms (see **Figure 2**):

- Totally certain:  $\mu_{TotallyCertain}(v) = 1$  if  $v = 1$ ;  $v < 1$   $\mu_{TotallyCertain} = 0$
- Very certain:  $\mu_{VeryCertain}(v) = v^2$
- Fairly certain:  $\mu_{FairlyCertain}(v) = \sqrt{v}$
- Certain:  $F(v) = v$
- Doubtful:  $F(v) = 1 - v$
- Very doubtful:  $\mu_{VeryDoubtful}(v) = (1 - v)^2$
- Fairly doubtful:  $\mu_{FairlyDoubtful}(v) = \sqrt{1 - v}$
- Totally doubtful:  $\mu_{TotallyDoubtful}(v) = 1$  if  $v = 0$ ;  $v > 0$   $\mu_{TotallyDoubtful} = 0$
- Unknown:  $\mu_{Unknown}(v) = 1$

The values for Certainty are the result of a combination of the strong negation (i.e. the values for Doubtful are obtained as  $1-v$ ) with the use of the linguistic hedges “very” and “fairly” [35], which modify the meaning of the corresponding label according to the definition showed above.



**Figure 2. Linguistic variable of Certainty.**

Our proposal defines Certainty (C) as a linguistic variable composed by nine possible fuzzy values. Applying the Certainty variable to ConML implies that, instead of the certain qualifiers defined on 2.1 section, we can now go deeper into the semantics and quantification of the uncertainty. Thus, we can assign one of the valid values for C to instances of features to express how certain we are of the associated predication. In other words, any feature that we add to a ConML model can be qualified with one of this nine possibilities, expressing with this label the degree of epistemic uncertainty that is necessary to manage on each case. Going a step further, the possible linguistic values and its corresponding formulas allow us to quantify the uncertainty on each feature value on software implementation.

### 3.1 Application in a Real Scenario

The proposed fuzzy logic framework has been applied to a real scenario in Digital Humanities within the DICTOMAGRED research project [36], carried out at the Institute for Medieval and Renaissance Studies and Digital Humanities (IEMYRhd) [37], University of Salamanca, Spain. DICTOMAGRED includes a software tool for humanities specialists to “retrieve information about the location of toponyms in North Africa as they appear in historical sources of medieval and modern times” [36, 38]. DICTOMAGRED has been previously studied [15] as a project with demanding needs for informational uncertainty management, especially regarding epistemic uncertainty, due to the lack of knowledge about some of the main topics studied (toponyms, their origins, historical sources, associated geographical settlements, etc.). Thus, ConML models previously created for DICTOMAGRED included certainty qualifiers to express epistemic uncertainty, but they lacked the necessary algebraic support to quantify them, thus hindering operationalization possibilities. More information about ConML models for DICTOMAGRED can be found here [3]

Let us illustrate how the proposed fuzzy logic framework works on DICTOMAGRED. **Figure 3** shows a fragment of the ConML class diagram representing DICTOMAGRED historical sources, what toponyms appear, and their associated geographical areas, as well as some information about political periods and rulers than appear in the text sources. In addition, we can find two object diagrams, representing the two specific cases of the toponyms *Ashir* and *Biskra*. Diagrams in **Figure 3** follow the ConML notation, including the certainty qualifiers described in section 2.1. Now, in order to apply the proposed fuzzy logic framework, values must be specified for these certainty qualifiers, drawing from the nine possibilities defined at the top of this section.

In the first case, the *Ashir* toponym appears on textual historical sources as a founded city by the Ziries dynasty on 936 B.C., so we qualified UsedIn=936 B.C. as Certain. However, we are not sure about the current use of the toponym, because we do not have data and we did not carry out interviews or other methods to obtain information on the current use of *Ashir*. Thus, we qualified UsedIn=2019 A.D. (*Anno Domini*) as Very Doubtful. The current location for the old city is not clear, so we qualified the *XCoord* and *YCoord* values for the geographic locations as Very Doubtful.

In the case of *Biskra*, the toponym appears on historical sources referred to 3000 B.C., so we qualified UsedIn=3000 B.C. as Certain. We also can assign precise current coordinates for locating *Biskra*, due to *Biskra* is an existing city today. We qualified the *XCoord* and *YCoord* values for the geographic locations as Totally Certain. For the same reason, we also assigned Totally Certain to UsedIn=2019 A.D. (we certainly know that the toponym is currently in use). As we can see on this scenario, it is necessary to treat each uncertainty expression instance one by one, qualifying the situation with a linguistic label.

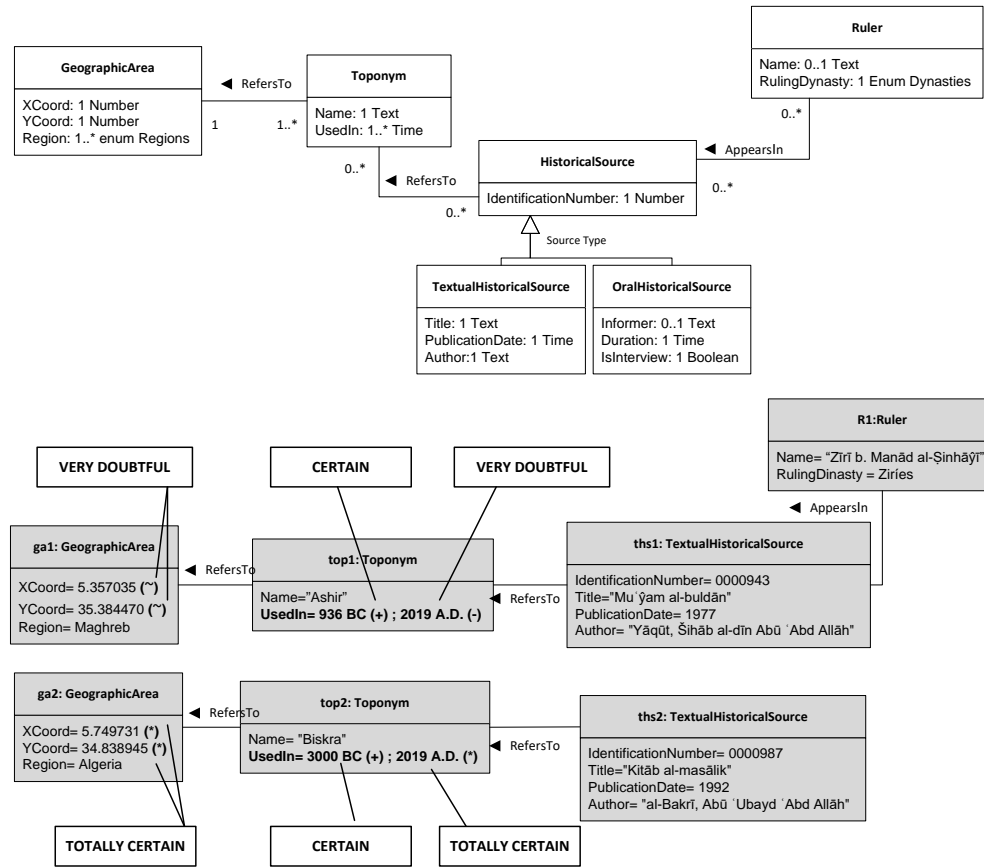


Figure 3. ConML class model fragment and object models for toponyms *Ashir* and *Biskra* in the DICTOMAGRED project

Once we have labelled each value, it is possible to quantify each uncertainty level, thanks to the formulas provided by the fuzzy logic framework. The definition of the fuzzy sets associated to each linguistic label is the result of an agreement from expert knowledge, and they can be changed and tuned to each specific project, domain or application. The advantage of the linguistic variable is that the same set of labels can be defined by means of different membership functions. In addition, fuzzy linguistic variables can also be updated according to the feedback received from users, making this approach robust and extensible. The quantification allows us, for instance, to perform searches that rely on the informational certainty implying aggregations (e.g. retrieve all toponyms for which we are sure, with C = Totally certain, that are in use today), ponderations (e.g. “>” or “<” operators, such as retrieve all toponyms for which UsedIn value are upper “>” Doubtful), or sorts (e.g. retrieve toponyms sorted by their use in time). With the framework proposed it is also possible to compose different fuzzy labels or draw some approximate conclusions. For instance, let us define the reliability of certain textual source or geographic area as the result of the aggregation of the linguistic values that label them. For instance, in Figure 3, ga1 has the

*XCoord* and *YCoord* attributes qualified as “very doubtful” and these can be aggregated by means of the T-norm product (a type of fuzzy conjunction), obtaining the result shown in Figure 4, which can be phrased as “ga1 is extremely doubtful” or “ga1 is very very doubtful”.

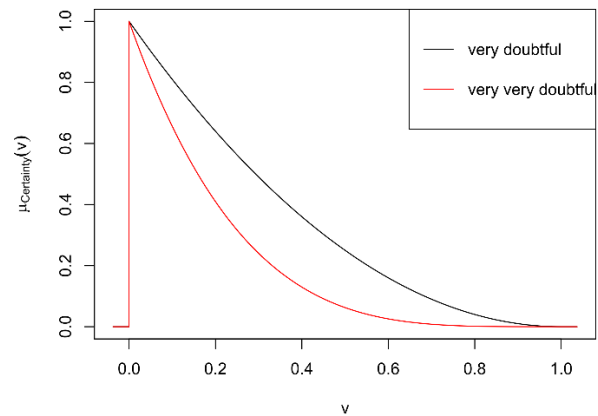


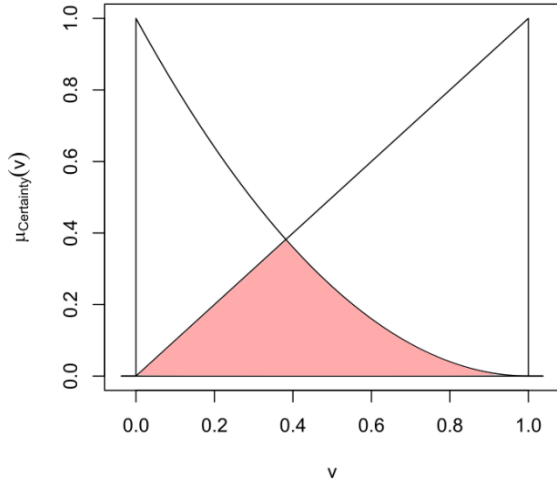
Figure 4. The result of the T-norm product between “very doubtful” and “very very doubtful”.

Following the same approach, “ths1” source refers to “top1” toponym that refers to the “gal” geographic area. If the conceptual model implements these associations, the reliability of “ths1” can be defined by another T-norm. In this case, we select the minimum (T-norm min), because it is less restrictive than the product:

$Min(toponym\$UsedIn(geographicArea\$XCoord\&YCoord))$   
in which:

- the  $\$$  symbol refers to the uncertainty value associated to the different attribute values.
- $XCoord$  &  $YCoord$  refers to the aggregated uncertainty value between the two coordinates of the toponym top1.
- $\$UsedIn$  refers to the uncertainty value given to the use of the toponym today.
- $Min$  denotes the set of the minimum values of the conjunction of the labels “very doubtful” and “certain”.

**Figure 5** shows the reliability area for “ths1”, which maximum degree of confidence is “certain = 0.4”, which can be phrased as “ths1 is general very doubtful and its degree of certainty is not higher than 0.4”. In addition, this approach serves as a basis for exploiting information in additional ways, especially in relation to visualization methods that take into account the expressed uncertainty [39]. For example, uncertainty values could be used to visualize a toponym map using colors or other visual cues to indicate different levels of reliability of software systems.



**Figure 5. Application of the T-norm min for the aggregation of “very doubtful”, “very doubtful” and “certain”**

## 4 CONCLUSIONS

Many works advocate the importance of treating the uncertainty of information in software systems for humanistic disciplines. Taking ConML as a starting point, we have proposed in this paper a framework based on fuzzy logic to go a step further in the implementation of software systems like these, introducing a complete algebraic system that allows us to perform computations on information labelled with uncertainty values. Thus, we can now compare, aggregate and reason about the uncertainty contained in information, as illustrated with the case of the DICTOMAGRED project.

The presented framework is totally general in terms of the nature of the information that is modeled and subsequently implemented and can be applied to any domain. In this manner, it covers the gap between conceptual modelling and implementation of software systems. However, the proposal still lacks support for some relevant aspects, which constitutes our ongoing research. Firstly, and in connection to the fact that the fuzzy set associated to each linguistic label is the result of an agreement from expert knowledge, the presented approach lends itself to research to find out what the most usual or agreed-upon fuzzy sets are for each field of study or family of settings. Having pre-obtained values for this would make applications of this approach faster and more straightforward. Secondly, the presented framework does not support uncertainty due to changes of opinion, that is, situations in which researchers change their mind about facts as a response to new or more information becoming available. In order to express these changes of opinion, we are evaluating the possibility to model recursive operators that mimic the linguistic mechanism of reported speech, such as [39, 40].

Finally, the application of the proposed fuzzy logic framework to a great variety of cases will help us detect improvement areas and potential adjustments to be made depending on the knowledge base taken for each fuzzy implementation, as well as the evaluation of possible implications for the framework related with the choice of different technologies for its implementation, such as programming languages or storage paradigms.

However, the proposal still lacks support for some relevant aspects, which constitutes our ongoing research. Firstly, and in connection to the fact that the fuzzy set associated to each linguistic label is the result of an agreement from expert knowledge, the presented approach lends itself to research to find out what the most usual or agreed-upon fuzzy sets are for each field of study or family of settings. Having pre-obtained values for this would make applications of this approach faster and more straightforward. Secondly, the presented framework does not support uncertainty due to changes of opinion, that is, situations in which researchers change their mind about facts as a response to new or more information becoming available. In order to express these changes of opinion, we are evaluating the possibility to model recursive operators that mimic the linguistic mechanism of reported speech, such as [40, 41].

Finally, the application of the proposed fuzzy logic framework to a great variety of cases will help us detect improvement areas and potential adjustments to be made depending on the knowledge base taken for each fuzzy implementation, as well as the evaluation of possible implications for the framework related with the choice of different technologies for its implementation, such as programming languages or storage paradigms.

## ACKNOWLEDGMENTS

This work was partially supported by Spanish Ministry of Economy, Industry and Competitiveness under its Competitive Juan de la Cierva Postdoctoral Research Programme (FJCI-2016-28032 and IJCI-2017-31477).

## REFERENCES

- [1] Gonzalez-Perez, C. *Information Modelling for Archaeology and Anthropology: Software Engineering Principles for Cultural Heritage*. Springer, 2018.
- [2] PROVIDEH CHIST-ERA Call 2016 - VADMU Topic - PROVIDEDH 2018.2018  
<http://www.chistera.eu/projects/providedh>
- [3] Martin-Rodilla, P. and Gonzalez-Perez, C. *Conceptualization and Non-Relational Implementation of Ontological and Epistemic Vagueness of Information in Digital Humanities*. Multidisciplinary Digital Publishing Institute, 2019.
- [4] Gonzalez-Perez, C. *A conceptual modelling language for the humanities and social sciences* Xplore, 2012.
- [5] Malta, M. C., González-Blanco, E., Cantón, C. M. and Del Rio, G. A Common Conceptual Model for the Study of Poetry in the Digital Humanities.
- [6] Incipit ConML Technical Specification. ConML 1.4.4.2015  
[http://www.conml.org/Resources\\_TechSpec.aspx](http://www.conml.org/Resources_TechSpec.aspx)
- [7] Lakoff, G. *Hedges: A study in meaning criteria and the logic of fuzzy concepts*. Springer, 1975.
- [8] Bouchon-Meunier, B. *Strengths of Fuzzy Techniques in Data Science*. 2018.
- [9] Lacerda, M. J. and Crespo, L. G. *Interval predictor models for data with measurement uncertainty*. 2017.
- [10] Zhou, H., Wang, J.-q. and Zhang, H.-y. Multi-criteria decision-making approaches based on distance measures for linguistic hesitant fuzzy sets. *Journal of the operational research Society*, 69, 5 (2018), 661-675.
- [11] Zadeh, L. A. Fuzzy logic= computing with words. *IEEE transactions on fuzzy systems*, 4, 2 (1996), 103-111.
- [12] Sarma, A. D., Benjelloun, O., Halevy, A., Nabar, S. and Widom, J. Representing uncertain data: models, properties, and algorithms. *The VLDB Journal—The International Journal on Very Large Data Bases*, 18, 5 (2009), 989-1019.
- [13] Ottomanelli, M. and Wong, C. K. *Modelling uncertainty in traffic and transportation systems*. Taylor & Francis, 2011.
- [14] Xiao, J., Pinel, P., Pi, L., Aranega, V. and Baron, C. *Modeling uncertain and imprecise information in process modeling with uml*. 2008.
- [15] Martin-Rodilla, P. and Gonzalez-Perez, C. *Representing Imprecise and Uncertain Knowledge in Digital Humanities: A Theoretical Framework and ConML Implementation with a Real Case Study*. ACM, 2018.
- [16] Consortium, T. Text Encoding Initiative (TEI).2016  
<http://www.tei-c.org/index.xml>
- [17] Isaksen, L., Simon, R., Barker, E. T. E., Ca, P. d. S., #241 and amares. Pelagios and the emerging graph of ancient world data. In *Proceedings of the Proceedings of the 2014 ACM conference on Web science* (Bloomington, Indiana, USA, 2014). ACM, [insert City of Publication],[insert 2014 of Publication].
- [18] Commons, P. Pelagios Commons WebSite (Pelagios 6 project).2018 <http://commons.pelagios.org/>
- [19] Community, L. D. Linked Data WebSite.2018  
<http://linkeddata.org/>
- [20] Van Ruymbeke, M., Hallot, P. and Billen, R. Enhancing CIDOC-CRM and compatible models with the concept of multiple interpretation. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4(2017), 287.
- [21] Gonzalez-Perez, C. *Modelling temporality and subjectivity in ConML: Short paper*. IEEE, 2013.
- [22] Ramos-Soto, A., Alonso, J. M., Reiter, E., van Deemter, K. and Gatt, A. *An empirical approach for modeling fuzzy geographical descriptors*. IEEE, 2017.
- [23] Barceló, J. A. and Bogdanovic, I. *Mathematics and Archaeology*. CRC Press, 2015.
- [24] Mink, P. B., Ripy, J., Bailey, K. and Grossardt, T. H. Predictive archaeological modeling using gis-based fuzzy set estimation: A case study in Woodford County, Kentucky(2009).
- [25] Lieskovský, T., Ďuračiová, R. and Karell, L. Selected mathematical principles of archaeological predictive models creation and validation in the GIS environment. *Interdisciplinaria archaeologica. Natural Sciences in Archaeology*, 4, 2 (2013), 33-46.
- [26] De Runz, C., Desjardin, E., Piantoni, F. and Herbin, M. *Using fuzzy logic to manage uncertain multi-modal data in an archaeological GIS*. 2007.
- [27] Taheri, S. M., Ghadim, F. I. and Kabirian, M. *Application of Fuzzy Inference Systems in Archaeology*. 2019.
- [28] Niccolucci, F. and Hermon, S. A fuzzy logic approach to reliability in archaeological virtual reconstruction. *Proceedings of CAA2004 Budapest, Archaeolingua*(2004).
- [29] OMG UML 2.4.1 Superstructure Specification. August, 2012.  
<http://www.omg.org/>
- [30] ISO/IEC *Information technology -- Object Management Group Unified Modeling Language (OMG UML) Part 1: Infrastructure. ISO/IEC 19505-1:2012.*, 2012.
- [31] Gonzalez-Perez, C., Martín-Rodilla, P. and Blanco-Rotea, R. Expressing Temporal and Subjective Information about Archaeological Entities. *Across Space and Time*(2013), 326.
- [32] Zadeh, L. A. The concept of a linguistic variable and its application to approximate reasoning—I. *Information sciences*, 8, 3 (1975), 199-249.
- [33] Zadeh, L. A. Fuzzy sets. *Information and control*, 8, 3 (1965), 338-353.
- [34] Miller, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 2 (1956), 81-97.
- [35] Lakoff, G. Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. *Journal of Philosophical Logic*, 2, 4 (1973), 458-508.
- [36] Dictomagred DICTOMAGRED: Diccionario de Toponimia Magrebí.2018 <https://dictomagred.usal.es/>
- [37] IEMYR Instituto de Estudios Medievales y Renacentistas y de Humanidades Digitales IEMYRhd 2018 <http://iemyr.usal.es/>
- [38] Manzano Rodríguez, M. A. *El Magreb desde Oriente: un acercamiento al Libro de los países de Ahmad al-Yaqubi*. Universidad de Sevilla, 2017.
- [39] Therón, R., Losada, A. G., Benito, A. and Santamaría, R. *Toward supporting decision-making under uncertainty in digital humanities with progressive visualization*. ACM, 2018.
- [40] Mitra, P. and Wiederhold, G. *An Ontology-Composition Algebra*. Springer Berlin Heidelberg, 2004.
- [41] Yen, J. and Langari, R. *Fuzzy logic: intelligence, control, and information*. Prentice-Hall, Inc., 1999.