

# Handling Uncertainty: An Extension of DL-Lite with Subjective Logic <sup>\*</sup>

Jhonatan Garcia<sup>1</sup>, Jeff Z. Pan<sup>1</sup>, Achille Fokoue<sup>2</sup>, Katia Sycara<sup>3</sup>, Yuqing Tang<sup>3</sup>, and Federico Cerutti<sup>1</sup>

<sup>1</sup> Computing Science, University of Aberdeen, UK

<sup>2</sup> IBM T. J. Watson Research Center, NY, US

<sup>3</sup> Carnegie Melon University, Pittsburgh, US

**Abstract.** Data in real world applications is often subject to some kind of uncertainty, which can be due to incompleteness, unreliability or inconsistency. This poses a great challenge for ontology-based data access (OBDA) applications, which are expected to provide a meaningful answers to queries, even under uncertain domains. Several extensions of classical OBDA systems has been proposed to address this problem, with probabilistic, possibilistic, and fuzzy OBDA being the most relevant ones. However, these extensions present some limitations with respect to their applicability. Probabilistic OBDA deal only with categorical assertions, possibilistic logic is better suited to make a ranking of axioms, and fuzzy OBDA addresses the problem of modelling vagueness, rather than uncertainty. In this paper we propose Subjective DL-Lite (SDL-Lite), an extension of DL-Lite with Subjective Logic. Subjective DL-Lite allows us to model uncertainty in the data through the application of opinions, which encapsulate our degrees of belief, disbelief and uncertainty for each given assertion. We explore the semantics of Subjective DL-Lite, clarify the main differences with respect to its classical DL-Lite counterpart, and construct a canonical model of the ontology by means of a chase that will serve as the foundation for a future construction of an OBDA system supporting opinions.

**Keywords:** Subjective Logic, Query Answering, OBDA, Description Logics

## 1 Introduction

Semantic applications that model real world scenarios often have to deal with uncertainty in the data. This is usually the case when extracting data from web sources,

---

<sup>\*</sup> This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon. This work is partially supported by the EU K-Drive project.

where information might be incomplete or unreliable. Even the method used for extracting the data creates another point of uncertainty, as it is quite common to rely on heuristic algorithms that are prone to errors.

In order for a semantic application to address all these issues, such an application should be able to comply with the following list of requirements:

- Models uncertainty in the data
- Understands the meaning of the underlying data
- Provides answering services over custom user queries
- Determines the reliability of the answers given the available information

In this paper we explore some of the theoretical foundations required to develop a logic capable of supporting ontology-based data access (OBDA) applications that can fulfil these requirements.

Our contributions in this paper are twofold: A) We define the semantics for Subjective DL-Lite in section 4, and B) We present a methodology to build a chase-based canonical interpretation of subjective ontologies in section 5.

## 2 Related Work

Several relevant approaches to model Uncertainty have been proposed in different areas of research. Probabilistic Logic [5] extends axioms in a knowledge base with a probability value that models the degree of trust that we place on the validity of the proposition. Possibilistic Logics [9] offers a similar approach, but its possibility values express the necessity and the validity for a certain proposition, alongside with the plausibility of said proposition to be true. Fuzzy Logic [11], on the other hand, relies on a membership function to establish the degree of membership for a given proposition to a determined value of truth.

These approaches have become popular for the results that they yielded, and many implementations for specific solutions have been produced based on their premises [8, 12, 13]. However, it is our belief that each of these approaches have limitations in their expressivity for modelling Uncertainty. For instance, Probabilistic Logic is by far the most extended approach to handle uncertain information. Yet, every axiom stated in Probabilistic Logic is categorical. That is, let  $A(x) : (p)$  be the axiom assigning a probability of  $p$  to the truth of the statement: "Object  $x$  belongs to concept  $A$ ". Then it is implicitly implied that the probability of  $x$  not being a member of  $A$  is  $1 - p$ . In other words,  $A(x) : (p) \implies \neg A(x) : (1 - p)$ .

The use of Probabilistic Logic does not naturally allow the user to assign some amount of believe to the fact that we might be missing some information about a certain statement. We propose the use of Subjective Logic [7] to overcome this limitation. Subjective Logic was proposed by A. Jøsang as a tool to express structured argument models with an associated degree of truth. It is based on Dempster-Shafer theory of belief, and uses frames of discernment to assign belief masses to given statements. Under Subjective Logic, statements are extended with opinions. An opinion is a triple  $(b, d, u)$ , where  $b$  is a degree of belief on the truth of the statement,  $d$  is the degree of disbelief associated with the statement, and  $u$  is a degree of uncertainty. These three degrees must

sum up to 1 to comply with Kolmogorov’s probabilistic axioms. From an intuitive point of view,  $b$  represents the amount of evidence that support the validity of the axiom,  $d$  represents how much evidence has been collected against the statement, and  $u$  is the amount of evidence that is not available at the moment, but could tilt our confidence either for or against the validity of the axiom.

We will use this approach to model uncertainty in ontologies, extending ABox assertions with subjective opinions. In this manner, we will be able to encapsulate how much information is already known about the validity of a certain axiom, as well as how much information is currently unknown. The application of opinions to axioms will result in some constraints that must hold for the ontology to make sense. These constraints will form the foundation for our reasoning, since they will let us propagate our beliefs through the ontology.

### 3 Preliminaries

#### 3.1 DL-Lite<sub>core</sub>

We will follow the standard syntax and semantics for classical (that is, without uncertainty) description logics, and due to space constraints, will refer the reader to [2] for further details. The subfamily DL-Lite<sub>core</sub> will be used through this paper for simplicity sake, but the results presented in this paper could be extended to other families of description logics.

As usual,  $A$  denotes atomic concept names, and  $r$  denotes atomic role names.  $B$  denotes basic concepts, and  $R$  denotes roles or their inverses. All valid expressions for DL-Lite<sub>core</sub> are built using the following production rules:  $P ::= r \mid r^-$ ,  $B ::= A \mid \exists P \mid \exists P^-$ .

A TBox  $\mathcal{T}$  is a finite set of concept inclusions (CIs)  $B \sqsubseteq B'$ , or  $B \sqsubseteq \neg B'$ . An ABox  $\mathcal{A}$  is a finite set of membership assertions of the form  $A(a)$ ,  $P(a,b)$ . A DL-Lite<sub>core</sub> ontology  $\mathcal{O}$  is a pair  $(\mathcal{T}, \mathcal{A})$ , where  $\mathcal{T}$  is a DL-Lite<sub>core</sub> TBox, and  $\mathcal{A}$  is a DL-Lite<sub>core</sub> ABox.

Following the standard semantics of description logics [2], the semantics of DL-Lite<sub>core</sub> is based on interpretations. An interpretation  $\mathcal{I}$  is a pair  $(\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ , where  $\Delta^{\mathcal{I}}$  is a non-empty set of objects, and  $\cdot^{\mathcal{I}}$  is an interpretation function, which maps every individual  $a \in \mathcal{A}$  to an object  $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$ , every class  $C$  into a subset  $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ , and each role  $R$  to a subset  $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ . An interpretation is a model of a TBox  $\mathcal{T}$  (resp. ABox  $\mathcal{A}$ ) if it satisfies all concept inclusions in  $\mathcal{T}$  (resp. assertions in  $\mathcal{A}$ ). An ABox  $\mathcal{A}$  is consistent with respect to a TBox  $\mathcal{T}$  if  $\mathcal{A}$  and  $\mathcal{T}$  have a common model. We write  $\mathcal{T} \models C \sqsubseteq D$  if for all models  $I$  of  $\mathcal{T}$ ,  $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$  and say that  $C$  is subsumed by  $D$  relative to  $\mathcal{T}$ .

There are a number of common reasoning services that are usually provided when developing an application that deals with ontologies. Among these services, we can name:

- **Instance checking:** Given an individual  $x$  and an concept  $C$ , determine whether or not  $x$  is a member of  $C$ .
- **Instance retrieval:** Given a concept  $C$ , retrieve all the individuals that are members of  $C$ .

- **Consistency checking:** Given a knowledge base  $\mathcal{KB}$ , determine whether or not a model for  $\mathcal{KB}$  exists.
- **Query answering:** Given a knowledge base  $\mathcal{KB}$  and a query  $q$ , determine all the answers for  $q$  that satisfy every model of  $\mathcal{KB}$ .

### 3.2 Subjective ABoxes

A subjective DL-Lite<sub>core</sub> ABox  $\mathcal{SA}$  is an extension of a DL-Lite<sub>core</sub> ABox  $\mathcal{A}$  in which every assertion in  $\mathcal{A}$  is extended with an opinion. An opinion  $\mathbf{w}$  over a statement  $x$  is a triple of positive numbers  $(b, d, u)$  such that  $b + d + u = 1$ ; and in which  $b$  represents the degree of belief assigned to the truth of  $x$ ,  $d$  represents the degree assigned to the falsehood of  $x$ , and  $u$  measures the degree of uncertainty associated with  $x$ . If, during the execution of any reasoning task, an opinion  $\mathbf{w}$  is produced such that  $b + d + u > 1$ , we say that  $\mathbf{w}$  is invalid. We denote with  $b(\mathbf{w})$ ,  $d(\mathbf{w})$ , and  $u(\mathbf{w})$  the degrees of belief, disbelief and uncertainty associated with an opinion  $\mathbf{w}$ , respectively, and with  $\mathcal{W}$  the set of all possible opinions.

**Definition 1.** Let  $w_1 = (b_1, d_1, u_1)$  and  $w_2 = (b_2, d_2, u_2)$  be two opinions about the same assertion  $\alpha$ . We call  $w_1$  a specialisation of  $w_2$  ( $w_1 \preceq w_2$ ) iff  $b_2 \leq b_1$  and  $d_2 \leq d_1$  (implies  $u_1 \leq u_2$ ). Similarly, we call  $w_1$  a generalisation of  $w_2$  ( $w_2 \preceq w_1$ ) iff  $b_1 \leq b_2$  and  $d_1 \leq d_2$  (implies  $u_2 \leq u_1$ ).

### 3.3 Example Scenario

In order to help us illustrate the many properties of the different aspects of Subjective DL-Lite, we present in this subsection a running example set in a medical domain. More specifically, we will consider a medical clinic, in which patients come seeking for a doctor to treat their illnesses. We can have the knowledge domain modelled by an ontology, with relevant relations represented in the TBox, and data for patients and clinical cases instantiated in the ABox. Table 1 illustrates the ontology that we are going to use for our example.

**Table 1.** Scenario knowledge base

$t_1 : \text{GraveDisease} \sqsubseteq \text{Disease}$	$t_2 : \text{MinorDisease} \sqsubseteq \text{Disease}$
$t_3 : \text{GraveDisease} \sqsubseteq \neg \text{MinorDisease}$	$t_4 : \exists \text{hasSymptom} \sqsubseteq \text{SickPatient}$
$t_5 : \text{CriticalPatient} \sqsubseteq \text{Patient} \sqcap \exists \text{hasGraveDisease}$	$t_6 : \text{PandemicDisease} \sqsubseteq \text{GraveDisease}$

$a_1 : \text{presentsPainIn}(\text{patientA}, \text{abdomen}) : (1, 0, 0)$
$a_2 : \text{hasSymptom}(\text{patientA}, \text{nausea}) : (0.4, 0, 0.6)$
$a_3 : \text{hasSymptom}(\text{patientA}, \text{migraine}) : (0.4, 0, 0.6)$
$a_4 : \text{hasFamilyCondition}(\text{patientA}, \text{IBS}) : (0, 0, 1)$
$a_5 : \text{hasPositiveTest}(\text{patientA}, \text{bloodTest}) : (0.9, 0.05, 0.05)$

In our scenario, a patient sees the doctor due to an acute abdominal pain that he is suffering. Being the main reason for the visit, and having no reason to doubt the patient, the doctor proceeds to instantiate with total certainty the fact that the patient suffers an abdominal pain. This is covered by axiom  $a_1$ .

Next, our patient tells the doctor that he also suffers something that he cannot describe very well, midway between a nausea or a migraine. This uncertain claim, stemming from the fact that the patient lacks the expertise to differentiate two distinct symptoms, can be easily modelled in a subjective ontology with axioms  $a_2$  and  $a_3$ . The

rational for this approach is based on the doctor having some reasons to believe that the patient suffers one of the symptoms, but not having enough information that could justify the choice of one over the other. It could be argued that, since either outcome is equally probable, both axioms should be extended with the opinion (0.5,0,0.5) instead. However, this is where the potential of Subjective Logic becomes clear, since such an opinion would reject any other option as the cause of the discomfort. By choosing to have a buffer of 0.1 degree in our uncertainty, we are leaving a door open for any other possible symptom that could be responsible for the malady. Indeed, it could be easily the case that the patient was suffering neither a nausea or a migraine, and instead had an ear infection. This capability of modelling what is unknown to us at the moment is what mainly differentiates Subjective Logic from similar approaches.

Continuing with our example, the doctor wants to know whether the Irritable Bowel Syndrome is present in any member of his relatives. Not even knowing about the disease itself, the patient finds himself unable to confirm, nor discard, any presence of it in his family. This is modelled by axiom  $a_4$ , in which there is no commitment towards the truth nor the falsehood of the claim. The opinion (0,0,1), representing total uncertainty about an axiom, is the most general opinion possible, since any other opinion must necessarily be a specialisation of it. We will call (0,0,1) the default opinion, and assume that any axioms that do not explicitly appear in our ABox are extended with it. With this approach, we reflect the fact that anything not stated in our ontology is unknown, rather than false.

Finally, the doctor decides to run a blood test on the patient, to discard possible diseases that could be responsible for the symptoms. After a couple of days the results arrive, and the test shows that all the values for the patient fall within standard nominal ranges. However, the doctor is aware that these tests have an error margin of five percent, in which either a false positive or negative can be delivered instead of the real result. Knowing this limitation of the tests, the doctor only commits 90% of his confidence to axiom  $a_5$ , reflecting the fact that the test itself is fallible by assigning 5% of his confidence to the disbelief degree of the axiom, and covering for some possible exceptional situations with the use of some uncertainty.

## 4 Semantics

### 4.1 Subjective DL-Lite Semantics

The semantics for a  $SDL\text{-}Lite_{core}$  ABox is given in terms of subjective interpretations. A subjective interpretation  $\mathcal{I}$  is a pair  $(\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ , where the domain  $\Delta^{\mathcal{I}}$  is a non-empty set of objects, and  $\cdot^{\mathcal{I}}$  is a subjective function that maps:

- an individual  $a$  to an element  $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$
- a named class  $A$  to a function  $A^{\mathcal{I}} : \Delta^{\mathcal{I}} \rightarrow \mathcal{W}$
- a named property  $R$  to a function  $R^{\mathcal{I}} : \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \rightarrow \mathcal{W}$

Following the example set in [10], we summarise the semantics for various axiomatic relations in  $SDL\text{-}Lite_{core}$  through Table 2. Top ( $\top$ ) and bottom ( $\perp$ ) are special

**Table 2.** Semantics of Subjective DL-Lite<sub>core</sub>

	Syntax	Semantics
$s_1$ :	$\top$	$\top^{\mathcal{I}}(o) = (1, 0, 0)$
$s_2$ :	$\perp$	$\perp^{\mathcal{I}}(o) = (0, 1, 0)$
$s_3$ :	$\exists R$	$b((\exists R)^{\mathcal{I}}(o_1)) \geq \max_{\cup_{o_2}} \{b(R^{\mathcal{I}}(o_1, o_2))\}$ and $d((\exists R)^{\mathcal{I}}(o_1)) \leq \min_{\cup_{o_2}} \{d(R^{\mathcal{I}}(o_1, o_2))\}$
$s_4$ :	$\neg B$	$(\neg B)^{\mathcal{I}}(o) = \neg B^{\mathcal{I}}(o)$
$s_5$ :	$R^-$	$(R^-)^{\mathcal{I}}(o_2, o_1) = R^{\mathcal{I}}(o_1, o_2)$
$s_6$ :	$B_1 \sqsubseteq B_2$	$\forall o \in \Delta^{\mathcal{I}}, b(B_1^{\mathcal{I}}(o)) \leq b(B_2^{\mathcal{I}}(o))$ and $d(B_2^{\mathcal{I}}(o)) \leq d(B_1^{\mathcal{I}}(o))$
$s_7$ :	$B_1 \sqsubseteq \neg B_2$	$\forall o \in \Delta^{\mathcal{I}}, b(B_1^{\mathcal{I}}(o)) \leq d(B_2^{\mathcal{I}}(o))$ and $b(B_2^{\mathcal{I}}(o)) \leq d(B_1^{\mathcal{I}}(o))$
$s_8$ :	$B(a):w$	$b(w) \leq b(B^{\mathcal{I}}(a^{\mathcal{I}}))$ and $d(w) \leq d(B^{\mathcal{I}}(a^{\mathcal{I}}))$
$s_9$ :	$R(a,b):w$	$b(w) \leq b(R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}}))$ and $d(w) \leq d(R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}}))$

concepts in our ontology. Every object in our domain is a member of top with total certainty. Likewise, we know with total certainty that no object in our domain is a member of bottom. For the rest of the axiomatic rules, the constraints given in Table 2 must hold for every object in the domain. To illustrate how the semantics might be applied, we can have a look at the scenario presented in section 3.3. It is clear that the most trivial interpretation possible is the one that links distinct object to each one of the individual appearing in the ABox, and then assigns to each required axiom the same opinion that it already has in the ABox. We could then proceed to infer new axioms by applying the constraints given by the semantics. For instance, imagine that we agree that the flu is usually a mild sickness, although at least 2% of the population experiment a virulent outcome every year. We can then instantiate the flu for the year 2015 with the following axiom:  $a_6 : \text{MinorDisease}(\text{flu2015}) : (0.9, 0.02, 0.08)$ . This opinion encapsulates our perception that the flu is usually a mild sickness, that this is not the case for 2% of the cases, and that there is a margin for which the actual statistical values of mild cases versus severe cases will fall this year. Now, from table 1, and the semantic rule  $s_6$  from table 2, we can infer the axiom  $a_7 : \text{Disease}(\text{flu2015}) : (0.9, 0, 0.1)$ . Notice how only the belief is propagated from the subclass to the inferred superclass, since any amount committed to the disbelief that *flu2015* is a *MinorDisease* does not justify stating that it is not a *Disease*. Certainly it could be the case that *flu2015* is later declared a *PandemicDisease* instead, thus making it a *GraveDisease*, but a *Disease* nonetheless. Also notice that the semantics for rule  $s_6$  require  $a_7$  to have a belief degree equal or greater than 0.9, but does not specify any exact value. One could argue that any value falling in the range  $[0.9, 1]$  could be chosen as the resulting belief for the inferred axiom, since any of these values comply with the semantics. However, by selecting precisely the lowest possible value in the range, we are maximising the use of available information in our ontology, at the same time that minimise the commitment for the resulting axiom. In other words, this is precisely the value that yields the most general

opinion  $\omega$  that complies with the semantics. Any other opinion of  $a_7$  that complies with the semantics must be a specialisation of  $\omega$ , and vice versa.

One interesting point to remark is that, in subjective environments, positive inclusions can lead to inconsistencies. This is not the case for classical knowledge bases, where inconsistencies were produced by violation of negative inclusions. We can illustrate this property going back to our example scenario. Imagine that the flu for 2015 gets declared a pandemic due to the high rate of spread in the population. Since we apply a series of guidelines and well-defined rules to check the criteria for the declaration of pandemics, we can state that the flu falls within the category of pandemia with total confidence using the following axiom  $a_8 : \text{PandemicDisease}(\text{flu2015}) : (1, 0, 0)$ . It is clear that this axiom introduces an inconsistency in the ontology. Intuitively, it does not make sense to declare  $\text{flu2015}$  to be a minor disease with 90% of confidence at the same time that we state with total certainty that  $\text{flu2015}$  is also a grave disease (as inferred through  $t_6$ ). A direct application of the semantic constraint  $s_7$  lets us spot the inconsistency.

One last note before continuing to the formal definition of inconsistencies for subjective ontologies. In our example, the inconsistency arose due to some dynamic behaviour present in our ontology. That is, there was some initial statement that was refined at a later stage. Although extremely interesting, the task of introducing some dynamic aspect to our ontologies will be left for a future work. We will consider our ABoxes to be static in nature, and the only inconsistencies will arise due to the implicit relations given by axioms at the TBox level.

Let  $\mathcal{K} = (\mathcal{T}, \mathcal{A})$  be a  $\text{SDL-Lite}_{\text{core}}$  knowledge base,  $\alpha$  be an axiom of  $\mathcal{K}$ , and  $\mathcal{I}$  and interpretation of  $\mathcal{K}$ . The following will provide a formal definition of consistency for subjective ontologies:

**Definition 2.**  $\mathcal{I}$  is a model of  $\alpha$ , denoted  $\mathcal{I} \models \alpha$ , if  $\alpha^{\mathcal{I}}$  satisfies all the constraints presented in table 2.

**Definition 3.**  $\mathcal{I}$  is a model of  $\mathcal{K}$ , denoted  $\mathcal{I} \models \mathcal{K}$ , if  $\mathcal{I} \models \alpha$  for each  $\alpha \in \mathcal{K}$

**Definition 4.**  $\mathcal{K}$  is consistent if it has at least one model

**Definition 5.**  $\mathcal{K}$  models  $\alpha$ , denoted  $\mathcal{K} \models \alpha$ , if  $\mathcal{I} \models \alpha$  for every model  $\mathcal{I}$  of  $\mathcal{K}$ .

**Definition 6.**  $\sigma$  is an answer for a query  $q$  over  $\mathcal{K}$ , denoted  $\mathcal{K} \models_q \sigma$ , if  $\sigma$  is an answer of  $q$  for every possible model of  $\mathcal{K}$ .

Finally we need to redefine the meaning of some common reasoning tasks for a subjective ontology  $\mathcal{K}$ :

- **Instance checking:** Given an individual  $x$  and an concept  $C$ , determine the most general opinion  $\omega$  such that  $\mathcal{K} \models C(x) : \omega$ .
- **Instance retrieval:** Given a concept  $C$ , return the set  $\{C(x) : \omega \mid \omega \text{ is the most general opinion such that } \mathcal{K} \models C(x) : \omega\}$ .
- **Query answering:** Given a query  $q$ , return the set  $\{\sigma : \omega \mid \omega \text{ is the most general opinion such that } \mathcal{K} \models_q \sigma : \omega\}$ .

## 5 Canonical Interpretation

Following the example presented in [3], we now will provide a methodology to build a canonical interpretation of a subjective knowledge base  $\mathcal{SK}$ . To achieve this goal, we will follow the notion of chase [1]. In particular, we will adapt the notion of restricted chase adopted by Johnson and Klug in [6]. This restricted chase will be constructed in an iterative manner by applying a series of rules based on TBox axioms. For easiness of exposition, we assume that every assertion  $\alpha$  that does not explicitly appear in the subjective ABox  $\mathcal{A}$  has the vacuous opinion  $(0,0,1)$  associated to it. In a more formal way, our assumption states that we work with the extended subjective ABox  $\mathcal{A}'$  given by  $\mathcal{A}' = \mathcal{A} \cup \{\alpha : (0, 0, 1)\}$ , if  $\alpha : w \notin \mathcal{A}$  for any opinion  $w$ . We will also make use of the function  $ga$ , that takes as input a basic role and two constants, and returns a membership assertion as specified below:

$$ga(R, a, b) = \begin{cases} P(a, b), & \text{if } R = P \\ P(b, a), & \text{if } R = P^- \end{cases} \quad (1)$$

**Definition 7.** Let  $\mathcal{S}$  be a set of DL-Lite<sub>core</sub> membership assertions, and let  $\mathcal{T}_\alpha$  be a set of DL-Lite<sub>core</sub> TBox axioms. Then, an axiom  $\alpha \in \mathcal{T}_\alpha$  is applicable in  $\mathcal{S}$  to a membership assertion  $f \in \mathcal{S}$  if

- (cr1)  $\alpha = A_1 \sqsubseteq A_2, f = A_1(a) : w$ , and  $A_2(a) : w' \in \mathcal{S}$ , with  $b(w) > b(w')$
- (cr2)  $\alpha = A_1 \sqsubseteq A_2, f = A_2(a) : w$ , and  $A_1(a) : w' \in \mathcal{S}$ , with  $d(w) > d(w')$
- (cr3)  $\alpha = A \sqsubseteq \exists R, f = A(a) : w$ , and there does not exist any constant  $b$  such that  $ga(R, a, b) : w' \in \mathcal{S}$ , with  $b(w') > b(w)$
- (cr4)  $\alpha = \exists R \sqsubseteq A, f = ga(R, a, b) : w$ , and  $A(a) : w' \in \mathcal{S}$ , with  $b(w) > b(w')$
- (cr5)  $\alpha = \exists R \sqsubseteq A, f = ga(R, a, b) : w$ , and  $A(a) : w' \in \mathcal{S}$ , with  $d(w') > d(w)$
- (cr6)  $\alpha = A_1 \sqsubseteq \neg A_2, f = A_1(a) : w$ , and  $A_2(a) : w' \in \mathcal{S}$ , with  $b(w') > d(w)$
- (cr7)  $\alpha = A_2 \sqsubseteq \neg A_1, f = A_1(a) : w$ , and  $A_2(a) : w' \in \mathcal{S}$ , with  $b(w') > d(w)$
- (cr8)  $\alpha = A \sqsubseteq \neg \exists R, f = ga(R, a, b) : w$ , and  $A(a) : w' \in \mathcal{S}$ , with  $b(w') > d(w)$
- (cr9)  $\alpha = A \sqsubseteq \neg \exists R, f = A(a) : w$ , and  $ga(R, a, b) : w' \in \mathcal{S}$ , with  $b(w') > d(w)$
- (cr10)  $\alpha = \exists R \sqsubseteq \neg A, f = ga(R, a, b) : w$ , and  $A(a) : w' \in \mathcal{S}$ , with  $b(w) > d(w')$
- (cr11)  $\alpha = \exists R \sqsubseteq \neg A, f = A(a) : w$ , and  $ga(R, a, b) : w' \in \mathcal{S}$ , with  $b(w) > d(w')$

Applicable axioms can be used, i.e., applied, in order to construct the chase of a knowledge base. The chase of a  $\mathcal{SDL}\text{-Lite}_{core}$  KB is a (possibly infinite) set of membership assertions, constructed step-by-step starting from the ABox  $\mathcal{A}$ . At each step of the process, an axiom  $\alpha \in \mathcal{T}$  is applied to a membership assertion  $f \in \mathcal{S}$ . Applying an axiom means refining our opinion about a certain  $f'$ , that might not appear explicitly in  $\mathcal{S}$ . The outcome of the application is a new set  $\mathcal{S}'$  in which  $\alpha$  is no longer applicable to  $f$ .

This construction process heavily depends on the order in which we select both the TBox axiom and the membership assertion in each iteration, as well as what constants we introduce when required. Therefore, we can produce a number of syntactically distinct sets of membership assertions following this process. However, it is possible to



show that the result is unique up to renaming of constants occurring in each such set. In order to achieve this, we select TBox axioms, membership assertions and constant symbols in lexicographic order. We denote with  $\Gamma_A$  the set of all constant symbols occurring in  $\mathcal{A}$ . We assume to have an infinite set  $\Gamma_N$  of constant symbols not occurring in  $\mathcal{A}$ . Finally, the set  $\Gamma_C = \Gamma_A \cup \Gamma_N$  is ordered in lexicographic order.

**Definition 8.** Let  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$  be a *SDL-Lite<sub>core</sub>KB*, let  $\mathcal{T}_\alpha$  be the set of assertions in  $\mathcal{T}$ , let  $n$  be the number of membership assertions in  $\mathcal{A}$ , and let  $\Gamma_N$  be the set of constants defined above. Assume that the membership assertions in  $\mathcal{A}$  are numbered from 1 to  $n$  following their lexicographic order, and consider the following definition

- $\mathcal{S}_0 = \mathcal{A}$
- $\mathcal{S}_{j+1} = \{\mathcal{S}_j \setminus f_{old}\} \cup \{f_{new}\}$

Then, we call *chase* of  $\mathcal{K}$ , denoted  $chase(\mathcal{K})$ , the set of membership assertions obtained as the (possibly infinite) union of all  $\mathcal{S}_j$ , i.e.,

$$chase(\mathcal{K}) = \bigcup_{j \in \mathbb{N}} \mathcal{S}_j \quad (2)$$

The element  $f_{old}$ , presented in definition 8, is the axiom whose opinion is being refined by  $f_{new}$ . The membership assertion  $f_{new}$ , numbered with  $n + j + 1$  in  $\mathcal{S}_{j+1}$ , is obtained as follows:

**Definition 9.** Let  $f$  be the first membership assertion in  $\mathcal{S}_j$  such that there exists a  $\alpha \in \mathcal{T}_\alpha$  applicable in  $\mathcal{S}_j$  to  $f$ , let  $\alpha$  be the lexicographically first TBox axiom applicable in  $\mathcal{S}_j$  to  $f$ , and let  $a_{new}$  be the constant of  $\Gamma_N$  that follows lexicographically all constants occurring in  $\mathcal{S}_j$

*case*  $\alpha, f$  of

- (cr1)  $\alpha = A_1 \sqsubseteq A_2, f = A_1(a) : w, A_2(a) : w' \in \mathcal{S}$   
**then**  $f_{new} = A_2(a) : (b(w), d(w'), 1 - b(w) - d(w'))$
- (cr2)  $\alpha = A_1 \sqsubseteq A_2, f = A_2(a) : w, A_1(a) : w' \in \mathcal{S}$   
**then**  $f_{new} = A_1(a) : (b(w'), d(w), 1 - b(w') - d(w))$
- (cr3)  $\alpha = A \sqsubseteq \exists R, f = A(a) : w, \exists R(a) : w' \in \mathcal{S}$   
**then**  $f_{new} = ga(R, a, a_{new}) : (b(w), d(w'), 1 - b(w) - d(w'))$
- (cr4)  $\alpha = \exists R \sqsubseteq A, f = ga(R, a, b) : w, A(a) : w' \in \mathcal{S}$   
**then**  $f_{new} = A(a) : (b(w), d(w'), 1 - b(w) - d(w'))$
- (cr5)  $\alpha = \exists R \sqsubseteq A, f = ga(R, a, b) : w, \text{ and } A(a) : w' \in \mathcal{S}$   
**then**  $f_{new} = ga(R, a, b) : (b(w), d(w'), 1 - b(w) - d(w'))$
- (cr6)  $\alpha = A_1 \sqsubseteq \neg A_2, f = A_1(a) : w, \text{ and } A_2(a) : w' \in \mathcal{S}$   
**then**  $f_{new} = A_1(a) : (b(w), b(w'), 1 - b(w) - b(w'))$
- (cr7)  $\alpha = A_2 \sqsubseteq \neg A_1, f = A_1(a) : w, \text{ and } A_2(a) : w' \in \mathcal{S}$   
**then**  $f_{new} = A_1(a) : (b(w), b(w'), 1 - d(w) - b(w'))$
- (cr8)  $\alpha = A \sqsubseteq \neg \exists R, f = ga(R, a, b) : w, \text{ and } A(a) : w' \in \mathcal{S}$   
**then**  $f_{new} = ga(R, a, b) : (b(w), b(w'), 1 - b(w) - b(w'))$
- (cr10)  $\alpha = \exists R \sqsubseteq \neg A, f = ga(R, a, b) : w, \text{ and } A(a) : w' \in \mathcal{S}$   
**then**  $f_{new} = ga(R, a, b) : (b(w), b(w'), 1 - b(w) - b(w'))$
- (cr11)  $\alpha = \exists R \sqsubseteq \neg A, f = A(a) : w, \text{ and } ga(R, a, b) : w' \in \mathcal{S}$   
**then**  $f_{new} = A(a) : (b(w), b(w'), 1 - b(w) - b(w'))$

It is worth noting that the application of chase rules can be a source of inconsistencies in the ontology. By increasing the belief degree of an opinion (resp. disbelief), we may put it in conflict with its disbelief degree (resp. belief), rendering the opinion invalid. Having an invalid opinion in our KB means that no interpretation will be able to satisfy it.

With the notion of chase in place we can introduce the notion of canonical interpretation. We define  $can(\mathcal{K})$  as the interpretation  $\langle \Delta^{can(\mathcal{K})}, \cdot^{can(\mathcal{K})} \rangle$ , where:

- $\Delta^{can(\mathcal{K})} = \Gamma_C$
- $a^{can(\mathcal{K})} = a$ , for each constant  $a$  occurring in  $chase(\mathcal{K})$
- $A^{can(\mathcal{K})} : \Gamma_C \rightarrow \mathcal{W}$ , such that  $A(a) : w \in chase(\mathcal{K}) \implies A^{can(\mathcal{K})}(a) = w$
- $P^{can(\mathcal{K})} : \Gamma_C \times \Gamma_C \rightarrow \mathcal{W}$ ,  $P(a_1, a_2) : w \in chase(\mathcal{K}) \implies P^{can(\mathcal{K})}(a, b) = w$

We can also define  $can_i(\mathcal{K}) = \langle \Delta^{can(\mathcal{K})}, \cdot^{can_i(\mathcal{K})} \rangle$  as the interpretation relative to  $chase_i(\mathcal{K})$  instead of  $chase(\mathcal{K})$ .

**Lemma 1.** *Let  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$  be a  $SDL-Lite_{core}$  knowledge base, then  $can(\mathcal{K})$  is a model of  $\mathcal{K}$  iff every opinion  $w$  that appears in  $can(\mathcal{K})$  is valid.  $\square$*

*Proof.(Sketch)*

$\Leftarrow$  If any of the opinions  $w$  that appear in the canonical interpretation is invalid, i.e.,  $b(w) + d(w) > 1$ , then it is obvious that the canonical interpretation is not a model of  $\mathcal{K}$ .

$\Rightarrow$  The fact that  $can(\mathcal{K})$  satisfies all membership assertions in  $\mathcal{A}$  follows from the fact that  $A \subseteq chase(\mathcal{K})$ .  $\square$

**Lemma 2.** *Let  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$  be a  $SDL-Lite_{core}$  knowledge base, then if  $can(\mathcal{K})$  is a model of  $\mathcal{K}$ , every other model of  $\mathcal{K}$  is a specialisation of  $can(\mathcal{K})$ .  $\square$*

*Proof.(Sketch)*

Let  $m$  be a model of  $\mathcal{K}$ , and  $m(\alpha) = \omega$  the opinion assigned by  $m$  to the assertion  $\alpha \in \mathcal{K}$ . Let  $can(\alpha) = \omega_c$  be the opinion assigned by the canonical model to  $\alpha$ , with  $\omega_c \preceq \omega$ . This means that, according to  $m$ ,  $\omega$  is a perfectly valid opinion for  $\alpha$ . However, since  $\omega_c$  is a specialisation of  $\omega$ , we can infer that, while building the chase, there was a semantic constraint applicable to  $\alpha$  that it is not satisfied by  $m$ . Given that there is at least one semantic constraint applicable to  $\alpha$  that is not covered by  $m$ , it is clear that  $m$  is not a model of  $\mathcal{K}$ . We conclude for these reasons that every model of  $\mathcal{K}$  must be at most as general as  $can$ .  $\square$

The implications from lemma 2 are profound and very relevant. Knowing that every model of  $\mathcal{K}$  is a specialisation of the canonical interpretation, we can focus on answering queries over this canonical interpretation. Any answer that is valid for the canonical interpretation will be valid for any other possible interpretation.

Of course, from a practical point of view, we will never construct the chase nor use directly the canonical model, since it might not be feasible to construct the chase for huge collections of data in a reasonable amount of time. Instead, we will apply the chase rules during the rewriting of the query, in such a way that we simulate the propagation

of the beliefs performed during the chase into the final query. Following this approach, we can be sure to obtain a valid answer for our original query, since this will be an answer for the canonical model and, through virtue of lemma 2, an answer for any other interpretation of  $\mathcal{K}$ .

## 6 Conclusions

It is expected that the capability to handle uncertainty in query answering solutions will be a critical requirement for future applications. Precisely to address this problem we propose a subjective extension of DL-Lite, to combine the efficient query answering properties of DL-Lite with the uncertainty modelling of Subjective Logic. Our main contributions come in the form of the theoretical foundation for the justification of the semantics used in Subjective DL-Lite, and the construction of a canonical model through a chase. We have shown that the theory behind this approach is sound, and could be used to develop a query answering application with support for uncertainty. For our future works we still need to demonstrate that every possible interpretation is a specialisation of the canonical model. Thus, any answer given for the canonical model will be an answer for the rest of the interpretation. Finally, in order to develop our query answering application, we need to define the algorithms that will perform inference over the set of axioms of the ontology and collect the answers to the queries. The initial results are promising, and encourages us to continue in this interesting, though challenging, line of research.

## References

- [1] S. Abiteboul, R. Hull, and V. Vianu. *Foundations of Databases*. ADDISON WESLEY Publishing Company Incorporated, 1995.
- [2] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, New York, NY, USA, 2003.
- [3] D. Calvanese, G. Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Tractable reasoning and efficient query answering in description logics: The DL-Lite family. *Journal of Automated Reasoning*, 2007.
- [4] D. Dubois and H. Prade. Possibility theory, probability theory and multiple-valued logics: A clarification. *Annals of Mathematics and Artificial Intelligence*, 32(1-4):35–66, 2001.
- [5] V. Gutiérrez-Basulto, J. C. Jung, C. Lutz, and L. Schröder. A closer look at the probabilistic description logic Prob-EL. In *AAAI*, 2011.
- [6] D. Johnson and A. Klug. Testing containment of conjunctive queries under functional and inclusion dependencies. *Journal of Computer and System Sciences*, 28(1):167 – 189, 1984.
- [7] A. Jøsang. *Subjective Logic*. Book Draft, 2011.
- [8] G. Qi and J. Pan. A tableau algorithm for possibilistic description logic  $\mathcal{ALC}$ . In D. Calvanese and G. Lausen, editors, *Web Reasoning and Rule Systems*, volume 5341 of *Lecture Notes in Computer Science*, pages 238–239. Springer Berlin Heidelberg, 2008.
- [9] G. Qi, J. Z. Pan, and Q. Ji. A possibilistic extension of description logics. In *In Proc. of DL07, 2007*, 2007.

- [10] M. Sensoy, A. Fokoue, J. Z. Pan, T. J. Norman, Y. Tang, N. Oren, and K. Sycara. Reasoning about uncertain information and conflict resolution through trust revision. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems, AAMAS '13*, pages 837–844, Richland, SC, 2013. International Foundation for Autonomous Agents and Multiagent Systems.
- [11] U. Straccia. Reasoning within fuzzy description logics. *Journal of Artificial Intelligence Research*, 14:2001, 2001.
- [12] U. Straccia and I. cnr Via G. Moruzzi. Transforming fuzzy description logics into classical description logics. In *In Proceedings of the 9th European Conference on Logics in Artificial Intelligence (JELIA-04), number 3229 in Lecture Notes in Computer Science*, pages 385–399. Springer Verlag, 2004.
- [13] A. W. Systeme, T. Lukasiewicz, and T. Lukasiewicz. Probabilistic description logics for the semantic web, 2007.