Exact Reasoning over Imprecise Ontologies

Mustapha Bourahla(✉)

Computer Science Department, University of M'sila,
Laboratory of Pure and Applied Mathematics (LMPA),
BP 166 Ichebilia, M'sila 28000, Algeria
mbourahla@hotmail.com

Abstract. A real world of objects (individuals) is represented by a set of assertions written with respect to defined syntax and semantics of description logic (formal language). These assertions should be consistent with the ontology axioms described as terminology of knowledge. The axioms and the assertions represent ontology about a particular domain. A real world is a possible world if all the assertions and the axioms over its set of individuals, are consistent. It is possible then to query the possible world by specific assertions (as instance checking) to determine if they are consistent with it or not. However, ontology can contain vague concepts which means the knowledge about them is imprecise and then query answering will not possible due to the open world assumption if the necessary information is incomplete (it is currently absent). A concept description can be very exact (crisp concept) or exact (fuzzy concept) if its knowledge is complete, otherwise it is inexact (vague concept) if its knowledge is incomplete. In this paper we propose a vagueness theory based on the definition of truth gaps as ontology assertions to express the vague concepts in Ontology Web Language (OWL2) (which is based on the description logic SROIQ(D)) and an extension of the Tableau algorithm for reasoning over imprecise ontologies.

Keywords: Vagueness · Ontology · OWL · Description logics · Automatic reasoning

1 Introduction

Formalisms for dealing with vagueness have started to play an important role in research related to the Web and the Semantic Web [7,13]. Ontologies are the definition of domain concepts (extensions) and the relations between them. Formal ontologies are expressed in well-defined formal languages (for example, OWL2) [3,6] that are based on expressive description logics (for example, SROIQ(D)) [1,14,4]. We say ontology is vague if it has at least a vague definition of a concept. A concept (an extension) is vague if it defines a meaning gap with which we cannot decide the membership of certain objects (vague intension).

We state the problem with the following example. Assume an ontology defining a concept called *Expensive* in a domain about cars. The meaning of the concept is vague. This vagueness is pervasive in natural language, but until now

A. Amine et al. (Eds.): CIIA 2015, IFIP AICT 456, pp. 355–366, 2015.
DOI: 10.1007/978-3-319-19578-0_29

is avoided in ontologies definitions. For the concept *Expensive*, we can define three sub-extensions, definitely expensive extension (there are some car prices that we regard as definitely expensive), definitely cheap extension (others we regard as definitely cheap cars) and a vagueness extension, average car prices are neither expensive nor cheap. The source of this indecision is the imprecise definition of concepts that is caused by lack of rigorous knowledge.

Related Works: Almost all concepts we are using in natural language are vague (imprecise). Therefore common sense reasoning based on natural language must be based on vague concepts and not on classical logic. The rising popularity of description logics and their use, and the need to deal with vagueness, especially in the Semantic Web, is increasingly attracting the attention of many researchers and practitioners towards description logics able to cope with vagueness. There are many works in literature for dealing with vagueness and most of them express it as a concept property as those based on fuzzy logics.

The notion of a fuzzy set proposed by Lotfi Zadeh [15] is the first very successful approach to vagueness. Fuzzy description logics (FDLs) are the logics underlying modes of reasoning which are approximate rather than exact, assertions are true to some degree [13,2,8,12]. In this case, any concept instance will have a degree of membership that is determined by a defined fuzzy function. The vagueness under fuzzy theory is treated by extended fuzzy description logics that are supported by fuzzy semantics and fuzzy reasoning. The fuzzy description logics are applied in many domains. The fuzzy knowledge base is interpreted as a collection of constraints on assertions. Thus, the inference is viewed as a process of propagation of these constraints.

Assertions in fuzzy description logics, rather being satisfied (true) or unsatisfied (false) in an interpretation, are associated with a degree of truth using semantic operators, where the membership of an individual to the union and intersection of concepts is uniquely determined by its membership to constituent concepts. This is a very nice property and allows very simple operations on fuzzy concepts. In addition to the standard problems of deciding the satisfiability of fuzzy ontologies and logical consequences of fuzzy assertions from fuzzy ontologies, two other important reasoning problems are the best truth value bound problem and the best satisfiability bound problem.

In our work, the concepts are treated as having a fixed meaning (not a balanced meaning), shared by all users of the ontology; we propose instead that the meaning of a vague concept evolves during the ontology evolution, from more vague meaning to less vague meaning until it reaches if possible, a situation where it becomes non-vague concept. This meaning instability is the base of our vagueness theory that is used for reasoning over vague ontologies. Both theories represent two different approaches to vagueness. Fuzzy theory addresses gradualness of knowledge, expressed by the fuzzy membership, whereas truth gap theory addresses granularity of knowledge, expressed by the indiscernibility relation. The result of reasoning over vague ontology using truth gap theory is the posterior description that represents a revision of the prior description on the light of the evidence provided by acquired information. This property can be used to draw conclusions from prior knowledge and its revision if new evidence is available.

The other closest work to ours is the work in [10] which presents a framework for adjusting numerical restrictions defining vague concepts. An inconsistency problem can happen when aligning the original ontology to another source of ontological information or when ontology evolves by adding learned axioms. This adjustment is used to repair the original ontology for avoiding the inconsistency problem by modifying restrictions parameters called adaptors specified as concept annotations. The idea of this work is close to ours in the sense that we reduce the truth gaps when adding new assertions as learned knowledge to the ontology to guide the reasoning process which will play the same role as adjusting the vague concept restrictions. However, this work differs from our approach by the repair (modification) process applied on the original ontology to avoid introduced inconsistency. In our approach, we define the vague concepts as super concepts over restriction definitions. So, we don't have the problem of inconsistency to repair the ontology.

This paper is organized as follows. We begin in Section 2, by presenting Ontology Web Language (OWL2) and its correspondent description logic (SROIQ(D)). In Section 3, a vagueness theory is proposed to show how to express vague concepts and to describe the characteristics of vague ontologies. Section 4 presents the extended version of Tableau algorithm, to reason over imprecise ontologies. At the end, we conclude this paper by conclusions and perspectives.

2 Description Logics and Ontology Web Language

Ontologies are definitions of concepts and the relationships between them. They can be represented formally using formal languages. These formal description languages are based on well-defined Description Logics (DLs) [1], a family of knowledge representation formalisms. OWL2 DL is a variant of SROIQ(D) [4], which consists of an alphabet composed of three sets of names. The set $\mathcal{C}$ of atomic concepts corresponding to classes interpreted as sets of objects, the set $\mathcal{R}$ of atomic roles corresponding to relationships interpreted as binary relations on objects and the set $\mathcal{I}$ of individuals (objects). It consists also of a set of constructors used to build complex concepts and complex roles from the atomic ones. The roles (object or concrete) are called properties; if their range values are individuals (relation between individuals) then they are called object (abstract) properties. If their range values are concrete data (relation between individual and a concrete data) then they are called data (concrete) properties. The set of SROIQ(D) complex concepts can be expressed using the following grammar:

$$C ::= \top \mid \bot \mid A \mid \{a\} \mid \neg C \mid C \sqcap D \mid C \sqcup D \mid \exists o.Self \mid \forall o.Self \mid \\ \exists o.C \mid \forall o.C \mid \exists c.P \mid \forall c.P \mid \geq n\ s.C \mid \leq n\ s.C \tag{1}$$

Where $\top$ is the universal concept, $\bot$ is the empty concept, A is an atomic concept, a an individual, C and D are concepts, o an object role, c a concrete role, s a simple role w.r.t. $\mathcal{R}$, and n a non-negative integer. P is a predicate over a concrete domain that can have the form

$$P ::= DataType\,[\sim value] \mid P \sqcap P \mid P \sqcup P, \quad \sim \in \{<, \leq, >, \geq\} \tag{2}$$

The data type can be any recognized data type as integer, real, etc. This syntax allows expressing concepts and roles with a complex structure. However, in order to represent real world domains, one needs the ability to assert properties of concepts and relationships between them. The assertion of properties is done in DLs by means of an ontology (or knowledge base). A SROIQ(D) ontology is a pair $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, where $\mathcal{T}$ is called a terminological box and $\mathcal{A}$ is called an assertional box. The terminological box consists of a finite set of axioms on concepts and roles. There are inclusion axioms on concepts, object and concrete roles to define a hierarchy (taxonomy) on the names of concepts and roles, (we write $C \sqsubseteq D$ to denote inclusion axioms on concepts, where C and D are concepts, $C \equiv D$ as an abbreviation for $C \sqsubseteq D \wedge D \sqsubseteq C$ and $r_1 \sqsubseteq r_2$ for role inclusion, where r_1 and r_2 are object (concrete) roles, the same equivalence abbreviation can be applied on roles). The assertional box consists of a finite set of assertions on individuals. There are membership assertions for concepts ($C(a)$ means the object (individual) a is member of C), membership assertions for roles ($o(a, b)$ means the objects a and b are related by the object property o and $c(a, d)$ means the object a has the data property (concrete role) c with a value equals d).

Thus, the assertional box $\mathcal{A}$ of a knowledge base, provides a description of a world. It introduces individuals by specifying their names, the concepts to which they belong, and their relations with other individuals. The semantics of the language uses either the closed world assumption or the open world assumption. With the closed world assumption, we consider that the world is limited to what is stated. It is this assumption that is normally adopted in databases. In description logics, it is rather the assumption of the open world which prevails. This open world assumption has an impact in the way of making inferences in description logics. The inference is more complex with the assumption of the open world; it is often called to consider several alternative situations for the proof. Another important aspect of description logic is that it does not presuppose the uniqueness of names (the standard names). That is, two different names do not necessarily mean that there is case to two separate entities in the described world. To be sure that two different entities a and b are represented, should be added the assertion $a \neq b$ to the assertional box $\mathcal{A}$.

3 Vagueness Theory for Imprecise Ontologies

We define a concept C as vague if it has a deficiency of meaning. Thus, the source of vagueness is the capability of meaning (it has borderline cases). For example, the concept *Expensive* is extensionally vague and it remains intentionally vague in a world of expensive and non-expensive cars. This means that there are truth-value gaps where a vague concept is extensionally (intensionally) definitely true (*tt*), definitely false (*ff*) and true or false (*tf*). Let us consider the following ontology.

$$\mathcal{O} = \left\langle \begin{array}{l} \mathcal{T} = \left\{ \begin{array}{l} Dom(price) \equiv \top, Rge(price) \equiv (int[\geq 0]) \sqcap (int[\leq 100]), \\ Dom(speed) \equiv \top, Rge(speed) \equiv (int[\geq 100]) \sqcap (int[\leq 300]), \\ ExpensiveCar \equiv Car \sqcap \exists price.\left(int\left[\geq 50\right]\right), \\ NonExpensiveCar \equiv Car \sqcap \exists price.\left(int\left[\leq 30\right]\right), \\ SportsCar \equiv Car \sqcap \exists speed.\left(int\left[\geq 200\right]\right), \\ NonSportsCar \equiv Car \sqcap \exists price.\left(int\left[\leq 150\right]\right), \\ ExpensiveCar \sqsubseteq Expensive, \\ NonExpensiveCar \sqsubseteq \neg Expensive, \\ SportsCar \sqsubseteq Sports, \\ NonSportsCar \sqsubseteq \neg Sports, \\ ExpensiveSportsCar \equiv Car \sqcap Expensive \sqcap Sports \end{array} \right\}, \\ \mathcal{A} = \left\{ \begin{array}{l} Car(a), Car(b), Car(c), Car(d), ExpensiveSportsCar(c), \\ price(a, 25), price(b, 55), price(c, 40), price(d, 45), \\ speed(a, 220), speed(b, 250), speed(c, 160), speed(d, 180) \end{array} \right\} \end{array} \right\rangle \quad (3)$$

Where *price* and *speed* are two concrete roles with the universal concept as their domains and their ranges *Rge* are defined by two integer intervals. In this knowledge base (ontology), we assume the price of a definitely expensive car (*ExpensiveCar*) is greater than or equal to fifty units and it is less than or equal to one hundred units, and a definitely no-expensive car (*NonExpensiveCar*) has a price between zero and thirty units. The concept *Expensive* and its complement are subsuming two complex concept expressions (*ExpensiveCar* and *NonExpensiveCar*). Each concept expression contains a sub-expression that is defined as quantified (universal or existential) restriction on a concrete role (for example, the concrete role is *price* and the restricted sub-expressions are $\exists price.\left(int\left[\geq 50\right]\right)$ for the concept *ExpensiveCar* and $\exists price.\left(int\left[\leq 30\right]\right)$ for the concept *NonExpensiveCar*). By the same way, we define the vague concept *Sports* and the concept *ExpensiveSportsCar* as a conjunction of the concepts *Car*, *Expensive* and *Sports*.

We have taken advantage of the open world assumption in description logics to define vague concepts. This ontology satisfies the assertions $Expensive(b)$ and $\neg Expensive(a)$ but the assertions $Expensive(d)$ and $(\neg Expensive)(d)$ are both not satisfied. With this knowledge base (ontology), we will assign *tt* to $Expensive(b)$, *ff* to $Expensive(a)$, and *tf* to $Expensive(d)$. This means, there is a deficiency of meaning (truth value gaps) between *Expensive* and $\neg$*Expensive*. Consequently, the concept *Expensive* is considered vague and the same thing for the vague concept *Sports*. The assertion $ExpensiveSportsCar(c)$ is considered as acquired information to state that c is an expensive sports car in spite of the fact that the terminology does not imply this assertion from information of the object c. We will see how this acquired (learned) information will be used to decide on other instances.

Thus, the satisfaction of a membership assertion to a vague concept depends on the concrete property value and the truth gaps. The vagueness definition of a concept will create one or more truth gaps. These are convex intervals (or ordered sequences) of values from a concrete domain with which the satisfaction of a membership assertion to the vague concept cannot be decided. There are two borderline values for each interval (or sequence). They are the lower (l) and the upper (u) bounds of a truth gap. Thus, we associate with each vague concept C a set of truth gap assertions according to a concrete role r (or to different concrete roles) used by its description.

These truth gaps assertions can be formulated using the description logic SROIQ(D) as a result of ontology description pre-processing. This will augment the ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ by the membership and property assertions to be $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \cup \{C(tt), (\neg C)(ff),\ r(x_i, l_i),\ r(y_i, u_i)\} \rangle$ if C is checked to be a vague concept according to a concrete role r, tt and ff are considered as two additional dummy individuals. The individuals x_i, y_i are either tt or ff with the conditions $x_i \neq y_i \wedge x_{i+1} = y_i$, l_i and u_i are numerical values from the range of the concrete role r with $l_i < u_i < l_{i+1}$ for $1 \leq i < n$, where n is the number of the truth gaps. This description should verify the following vagueness consistency.

Lemma 1. *(vagueness consistency). The truth gaps set defined of any vague concept C associated with a role r (or a set of roles) should verify the condition of acceptability (vagueness consistency), this means $\forall i = 1, \cdots, n-1 : x_i \neq y_i \wedge y_i = x_{i+1} \wedge l_i < u_i < l_{i+1} < u_{i+1}$. This vagueness consistency condition can be formulated using the assertions on the dummy individuals tt and ff as*

$$(\{C(tt), r(tt, d_1), r(tt, d_2), (\neg C)(ff), r(ff, d)\} \subseteq \mathcal{A} \Rightarrow d \notin [d_1, d_2]) \wedge$$
$$(\{(\neg C)(ff), r(ff, d_1), r(ff, d_2), C(tt), r(tt, d)\} \subseteq \mathcal{A} \Rightarrow d \notin [d_1, d_2]) \quad (4)$$

A non-vague (crisp) concept C will have an empty set of truth gaps according to any concrete role r.

The intuition for this vagueness theory is as follows. An ontology is considered the knowledge base of an intelligent agent; if the ontology (knowledge base) $\mathcal{O}$ contains a vague concept C with respect to a concrete role r and one of its truth gaps has the smallest interval $[l, u]$, where the assertions $r(tt, u), r(ff, l)$ are in $\mathcal{O}$. The agent cannot decide if an individual (object) a with r-property value within the interval $[l, u]$ if it belongs to C or to its complement (we say that the knowledge base is incomplete). We assume that at a moment, assertions like $C(a), r(a, d)$ are added to the ontology $\mathcal{O}$, where $l < d < u$. This new information will change the ontology agent beliefs by reducing the truth gap interval to be $[l, d]$. We call that the individual a is similar to the dummy individual tt because they belong to the same concept C. Then, the individual tt will inherit the property of a. Now, if we add the assertions $(\neg C)(b), r(b, d')$ with $u > d' > d$, this will produce a vagueness inconsistency according to this vagueness theory because the agent has already changed its beliefs so that every property assertion of an individual with respect to the concrete role r where its range is greater than d should be member of the concept C. This vagueness theory is used to adjust the truth intervals (or the truth gaps) described in the original ontology by acquired new information.

3.1 Semantics for Vagueness Theory

The formal semantics of DLs is given in terms of interpretations. A SROIQ(D) interpretation is a pair $I = (\Delta^I, (.)^I)$ where Δ^I is a non-empty set called the domain of I, and $(.)^I$ is the interpretation function which assigns for every $A \in \mathcal{C}$ a subset $(A)^I \subseteq \Delta^I$, for every $o \in \mathcal{R}$ a relation $(o)^I \subseteq \Delta^I \times \Delta^I$, called object role, for every $c \in \mathcal{R}$ a relation $(c)^I \subseteq \Delta^I \times \mathcal{D}$, called concrete role ($\mathcal{D}$ is a data type as integer and string) and for every $a \in \mathcal{I}$, an element $(a)^I \in \Delta^I$. We say the interpretation I is a model of a SROIQ(D) ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$, if it satisfies all the assertions in $\mathcal{T}$ and $\mathcal{A}$. In addition, it is a model of any satisfied assertion by the ontology $\mathcal{O}$. If C_r is a vague concept with respect to the concrete role r and $x_i, y_i \in \{tt, ff\}$, then the interpretation function is extended to complex concepts and roles according to their syntactic structure.

$$
\begin{aligned}
(\top)^I &= \Delta^I \\
(\bot)^I &= \emptyset \\
(\{a\})^I &= (a)^I \\
(r)^I &= (r)^I \cup \{(x_i, l_i), (y_i, u_i) \wedge x_i \neq y_i = x_{i+1} \wedge l_i < u_i,\ i = 1, \cdots, n\} \\
(C_r)^I &= (C)^I \cup \{tt\} \\
(\neg C_r)^I &= (\neg C)^I \cup \{ff\} \\
(C_r \sqcap D_r)^I &= (C_r)^I \cap (D_r)^I \\
(C_r \sqcup D_r)^I &= (C_r)^I \cup (D_r)^I \\
(\exists o.Self)^I &= \{a \in \Delta^I | \exists (a,b) \in (o)^I \wedge a = b\} \\
(\forall o.Self)^I &= \{a \in \Delta^I | \forall (a,b) \in (o)^I \Rightarrow a = b\} \\
(\exists o.C_r)^I &= \{a \in \Delta^I | \exists (a,b) \in (o)^I \wedge b \in (C_r)^I\} \\
(\forall o.C_r)^I &= \{a \in \Delta^I | \forall (a,b) \in (o)^I \Rightarrow b \in (C_r)^I\} \\
(\exists c.P)^I &= \{a \in \Delta^I | \exists (a,d) \in (c)^I \wedge P(d)\} \\
(\forall c.P)^I &= \{a \in \Delta^I | \forall (a,d) \in (c)^I \Rightarrow P(d)\} \\
(\geq n\ s.C_r)^I &= \{a \in \Delta^I |\ |\{b | (a,b) \in (s)^I \wedge b \in (C_r)^I\}| \geq n\} \\
(\leq n\ s.C_r)^I &= \{a \in \Delta^I |\ |\{b | (a,b) \in (s)^I \wedge b \in (C_r)^I\}| \leq n\}
\end{aligned}
$$

Where n is the number of truth gaps for the vague concept C, $P(d)$ means the value d verifies the predicate P and $|S|$ is the cardinality of the set S. The predefined concepts like the universal concept $\top$, the empty concept $\bot$, the atomic concepts A and the nominative concepts $\{a_1, a_2, \cdots, a_n\}$ are defined as crisp concepts and then they will not be considered as vague concepts.

Example 1. The new ontology after generation of truth gap assertions on the original ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ described in (3), is $\mathcal{O}^{new} = \langle \mathcal{T}^{new}, \mathcal{A}^{new} \rangle$ where, $\mathcal{T}^{new} = \mathcal{T}$ and using the syntax of SROIQ(D), $\mathcal{A}^{new} = \mathcal{A} \cup \{Expensive(tt)$, $(\neg Expensive)(ff)$, $Sports(tt)$, $(\neg Sports)(ff)$, $price(ff, 0)$, $price(ff, 30)$, $price(tt, 50)$, $price(tt, 100)$, $speed(tt, 200)$, $speed(tt, 300)$, $speed(ff, 100)\}$, $speed(ff, 150)\}$. The new interpretations with these introduced assertions are as follows. The interpretation for the concrete role *price* is $(price)^I = \{(a, 25), (b, 55), (c, 40), (d, 45), (\mathit{ff}, 0), (\mathit{ff}, 30), (\mathit{tt}, 50), (\mathit{tt}, 100)\}$ and for the concrete role *speed*, the interpretation is $(speed)^I = \{(a, 220), (b, 250), (c, 160), (d, 180), (\mathit{ff}, 100), (\mathit{ff}, 150), (\mathit{tt}, 200), (\mathit{tt}, 300)\}$. The interpretation of the vague concept $Expensive_{price}$ is $\{b, c, \mathit{tt}\}$.

This new ontology containing concept truth gaps is considered vague and then it is incomplete for reasoning. An ontology is complete if we can assign only the definite truth values (tt and ff) to assertions. A vague (incomplete) ontology is an ontology that has at least one vague concept and then it is possible to assign the value tf to certain assertions. In addition, a vague ontology should be acceptable (Lemma 1), which means all the truth gap sets should be acceptable. We define a partial order between ontologies that is noted by $\langle \mathfrak{O}, \leq \rangle$, where $\mathfrak{O}$ is a non-empty set of ontologies describing a domain. If $\mathcal{O}_1$ and $\mathcal{O}_2$ are two ontologies from $\mathfrak{O}$ we write $\mathcal{O}_1 \leq \mathcal{O}_2$, if $\mathcal{O}_1$ is less complete than $\mathcal{O}_2$ (we say also that $\mathcal{O}_2$ extends $\mathcal{O}_1$). The relation $\leq$ (we call it also the extension relation) is based on comparison of truth gaps and it is transitive and antisymmetric. By this partial order definition, there is a canonical normal ontology $\mathcal{O}_n$ that is the least complete ontology, which can be extended by other complete ontologies.

The set $\mathfrak{O}$ has a base ontology that corresponds to description of which all other descriptions are extensions. This base ontology is composed of the terminological assertions and eventually some membership assertions. A condition that can be imposed on domain ontology is its completeability. It states that any intermediate ontology can be extended to a complete ontology. We suppose that ontology $\mathcal{O}$ has a vague concept C, with an acceptable set of truth gaps defined by the assertions set $\{C(\mathit{tt}), (\neg C)(\mathit{ff}), r(x_1, l_1), r(y_1, u_1), r(x_2, l_2), r(y_2, u_2), \cdots, r(x_n, l_n), r(y_n, u_n)\}$, then we define the ontology extension ($\oplus$) by the assertions $\{C(a), r(a, d)\}$ as follows.

$$\langle \mathcal{T}, \mathcal{A}\,[C(\mathit{tt}), (\neg C)(\mathit{ff}), \cdots, r(x_i, l_i), r(y_i, u_i), \cdots] \rangle \oplus \langle \mathcal{T}, \{C(a), r(a, d)\} \rangle =$$
$$\left\langle \mathcal{T}, \mathcal{A} \cup \{C(a), r(a, d)\} \cup \left\{ \begin{array}{l} r(x_i, d)\ if\ x_i = \mathit{tt} \wedge l_i < d < u_i \\ r(y_i, d)\ if\ y_i = \mathit{tt} \wedge l_i < d < u_i \end{array} \right\} \right\rangle$$
$$\langle \mathcal{T}, \mathcal{A}\,[C(\mathit{tt}), (\neg C)(\mathit{ff}), \cdots, r(x_i, l_i), r(y_i, u_i), \cdots] \rangle \oplus \langle \mathcal{T}, \{(\neg C)(a), r(a, d)\} \rangle =$$
$$\left\langle \mathcal{T}, \mathcal{A} \cup \{(\neg C(a)), r(a, d)\} \cup \left\{ \begin{array}{l} (x_i, d)\ if\ x_i = \mathit{ff} \wedge l_i < d < u_i \\ r(y_i, d)\ if\ y_i = \mathit{ff} \wedge l_i < d < u_i \end{array} \right\} \right\rangle$$

This extension guarantees the ontology stability if the acquired informations are satisfied by the ontology description to be extended.

Lemma 2. *(stability property of $\langle \mathfrak{O}, \leq \rangle$). Let α be an assertion, we say $\langle \mathfrak{O}, \leq \rangle$ is stable if*

$$\forall \mathcal{O}_1, \mathcal{O}_2 \in \mathfrak{O}, \mathcal{O}_1 \leq \mathcal{O}_2 : \mathcal{O}_1 \models \alpha \Rightarrow \mathcal{O}_2 \models \alpha \ and\ \mathcal{O}_2 \not\models \alpha \Rightarrow \mathcal{O}_1 \not\models \alpha$$

The complete ontology may not be available to remove completely the vagueness, thus it is necessary to work with the most extended ontology. This means, the truth-valuation is based upon the most extended ontology. Ontology can be extended to complete ontology by learned assertions as a process of ontology evolution when using an intelligent agent or inferred assertions. The learned assertions can be imported from other domain ontologies, RDF databases or simply added by the user. In the following, we propose an extension of reasoning that can take into account the proposed vagueness theory.

4 Reasoning over Imprecise Ontologies

An interpretation I is a model of an ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ denoted by $I \models \mathcal{O}$ if I satisfies all the axioms in $\mathcal{T}$ and all the assertions in $\mathcal{A}$. The reasoning is for checking concept and role instances and for query answering over a satisfiable ontology [9,11]. Ontology satisfiability is to verify whether ontology $\mathcal{O}$ admits at least one model where consistency properties should be verified. Concept instance checking is to verify whether an individual a is an instance of a concept C in every model of $\mathcal{O}$, i.e., whether $\mathcal{O} \models C(a)$. Role instance checking is to verify whether a pair (a, b) of individuals is an instance of a role r in every model of $\mathcal{O}$, i.e., whether $\mathcal{O} \models r(a, b)$.

The satisfaction properties will be extended to deal with the vagueness in ontologies. A vague (imprecise) ontology is satisfiable if it generates acceptable truth gaps for all its concepts (note that an empty set of truth gaps is acceptable). For example, if we modify the concept $ExpensiveCar$ in the vague ontology of the previous example to be $Car \sqcap \exists price.\,(int\,[\geq 55]) \sqsubseteq \neg Expensive$, this will change the set of truth gaps assertions associated with the vague concept $Expensive$ to be $\{Expensive(tt)$, $(\neg Expensive)(ff)$, $price(tt, 50)$, $price(tt, 100)$, $price(ff, 0)$, $price(ff, 55)\}$. This set of truth gaps is not acceptable because it is a false set of assertions according to the vagueness consistency stated in (4). Nevertheless, the vague ontology is satisfiable by using the traditional reasoning techniques. However, if we add the assertions $\{Car(d), price(d, 52)\}$ to the assertional box, the vague ontology becomes inconsistent because d is now at the same time expensive car and no-expensive car, although the ontology was initially satisfiable. In the following, we will extend the reasoning Tableau algorithm to cope with the problem of vague ontologies using this proposed vagueness theory.

The principle of this reasoning algorithm is the expansion of a finite configuration $T = \{A_1, \cdots, A_n\}$ of assertions that is represented as a set of subsets, each subset is composed of assertions on individuals, using well defined rules until no rule can be applied on at least one subset (satisfaction) or contradictions (clashes) are observed within all subsets (unsatisfaction). We will have a clash in a subset A_i when a contradiction happens in it. There are three types of contradictions: $\bot(a) \in A_i$, $C(a) \in A_i \wedge (\neg C)(a) \in A_i$, or unacceptable truth gaps assertions. If no expansion rule can be applied in A_i we say that A_i is open. The terminological box should be normalized to apply the expansion rules. It is necessary to begin the inference with formulas that are independent from any terminology. This means elimination of the definitions (equivalence axioms) and subsumptions (inclusion axioms) in the terminological box. If it contains no cycle in the definitions (which will be the case most of the time), it will happen simply by replacing all the terms in the formula by their definitions in the terminology. Obviously, if a term of formula has no definition in terminology, it remains unchanged. We repeat this process until the resulting formula contains no term which has a definition in the terminology.

For reasoning over vague ontologies using the proposed vagueness theory, we have added the following two expansion rules that should be applied after every expansion by a classical Tableau rule (the reader can be referred to [4,5,9] for the classical Tableau rules). We will get a clash (contradiction) if any new set of truth gaps assertions is not acceptable (Lemma 1 and Equation 4). The configuration length depends on ontology description and property being checked. Using the DL syntax of SROIQ(D), these two rules can be formulated as

$$V - Rule^{+}(DL) : \frac{A_i \in T \wedge \{C(a), r(a,d), C(tt)\} \subseteq A_i}{(T \setminus A_i) \cup (A_i \cup \{r(tt,d)\})} r(tt,d) \notin A_i$$

$$V - Rule^{-}(DL) : \frac{A_i \in T \wedge \{(\neg C)(a), r(a,d), (\neg C)(ff)\} \subseteq A_i}{(T \setminus A_i) \cup (A_i \cup \{r(ff,d)\})} r(ff,d) \notin A_i$$

These two rules will augment the assertions subset A_i by the property assertion $r(tt,d)$ if A_i contains the assertion $C(a) \wedge r(a,d) \wedge C(tt)$ (the rule $V - Rule^{+}(DL)$) or by the property assertion $r(ff,d)$ if A_i contains the assertion $(\neg C)(a) \wedge r(a,d) \wedge (\neg C)(ff)$ (the rule $V - Rule^{-}(DL)$). We explain this algorithm extension on a simple example of an instance checking using the ontology described in (3). We want to check the membership of the individual d (instance checking) to the class $ExpensiveSportsCar$ ($\mathcal{O} \models ExpensiveSportsCar(d)$). This means that we want to prove that $(\neg ExpensiveSportsCar)(d)$ is inconsistent with the ontology description. After elimination of terminological axioms and normalization as preliminary steps before applying Tableau Rules, we have:

$$T^0 = \left\{ A_0^0 = \left\{ \begin{array}{l} ((\neg Car) \sqcup (\neg Expensive) \sqcup (\neg Sports))(d), Car(a), Car(b), \\ Car(c), Car(d), price(a,25), price(b,55), price(c,40), price(d,45), \\ speed(a,220), speed(b,120), speed(c,160), speed(d,180), \\ (\neg Expensive)(a), Expensive(b), Expensive(c), Sports(a), \\ (\neg Sports)(b), Sports(c), Expensive(tt), price(tt,50), price(tt,100), \\ (\neg Expensive)(ff), price(ff,0), price(ff,30), Sports(tt), \\ speed(tt,200), speed(tt,300), (\neg Sports)(ff), \\ speed(ff,100), speed(ff,150) \end{array} \right\} \right\}$$

Using the classical expansion rule of the disjunction, we obtain the configuration:

$$T^1 = \left\{ \begin{array}{l} A_0^1 = A_0^0 \cup \{(\neg Car)(d)\}, \\ A_1^1 = A_0^0 \cup \{(\neg Expensive)(d)\}, \\ A_2^1 = A_0^0 \cup \{(\neg Sports)(d)\} \end{array} \right\}$$

We observe a clash in the subset A_0^1 (it contains $Car(d)$ and $(\neg Car)(d)$). By applying the rules $V - Rule^{+}(DL)$ and $V - Rule^{-}(DL)$, we get:

$$T^2 = \left\{ \begin{array}{l} A_0^2 = A_0^1 = \square, \\ A_1^2 = A_1^1 \cup \left\{ \begin{array}{l} price(tt, 55), price(tt, 40), price(ff, 25), price(ff, 45), \\ speed(tt, 220), speed(ff, 120), speed(tt, 160) \end{array} \right\}, \\ A_2^2 = A_2^1 \cup \left\{ \begin{array}{l} price(tt, 55), price(tt, 40), price(ff, 25), \\ speed(tt, 220), speed(ff, 120), speed(tt, 160), speed(ff, 180) \end{array} \right\} \end{array} \right\}$$

It is clear that the subset A_1^2 of assertions contains unacceptable truth gaps assertions (the following implication $\{Expensive(tt),\ price(tt, 40),\ price(tt, 50),\ (\neg Expensive)(ff),\ price(ff, 45)\} \in A_1^2 \Rightarrow 45 \notin [40, 50]$ is false). The same thing for the subset A_2^2, where the implication $\{Sports(tt),\ speed(tt, 160),\ speed(tt, 200),\ (\neg Sports)(ff),\ speed(ff, 180)\} \in A_2^2 \Rightarrow 180 \notin [160, 200]$ is also false. Thus a clash is observed in the two subsets which makes d a member of $ExpensiveSportsCar$. The principle of this approach is as follows. Without this vagueness theory, d which has the price of 45 (greater than 30 and less than 50) and the speed of 180 (greater than 150 and less than 200) cannot be decided by the classical reasoners, as $Expensive$, $Sports$, $\neg Expensive$ and $\neg Sports$ because the definitions of $Expensive$ and $Sports$ are vague. However, the ontology contains an assertion indicating that the price 40 of c is an expensive price ($Expensive(c)$) and its speed 160 makes it a sports car; this information can help the reasoner to decide that the car d of price 45 and of speed 180 is also an expensive sports car.

5 Conclusion

In this paper, we have presented a vagueness theory to deal with the problem of ontologies containing vague concepts. The vague property (characteristic) of a concept is based in general, on certain concept data properties that may generate truth gaps. With the traditional reasoning methods, it is not possible to decide the membership of an individual (object) to a vague concept (class) if its data property is in the truth gap. Ontologies could have extension (evolution), where assertions may be added, intentionally or as result of inferences. This ontology evolution can reduce the truth gaps and then logically it will be possible to infer on previously undecided assertions. This proposed vagueness theory is used to extend the current reasoning method to take into account this vagueness notion. Implementation of this approach is one of our perspectives.

References

1. Baader, F.: What's new in description logics. Informatik-Spektrum 34(5), 434–442 (2011)
2. Bobillo, F., Delgado, M., Gomez-Romero, J., Straccia, U.: Joining gödel and zadeh fuzzy logics in fuzzy description logics. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 20(4), 475–508 (2012)

3. Hitzler, P., Krötzsch, M., Parsia, B., Patel-Schneider, P.F., Rudolph, S. (eds.): OWL 2 Web Ontology Language: Primer. W3C (2009)
4. Horrocks, I., Kutz, O., Sattler, U.: The even more irresistible SROIQ. In: Proc. of the 10th Int. Conf. on Principles of Knowledge Representation and Reasoning, KR 2006, pp. 57–67. AAAI Press (2006)
5. Horrocks, I., Sattler, U.: A tableau decision procedure for SHOIQ. Journal of Automated Reasoning 39(39–3), 249–276 (2007)
6. Krötzsch, M.: OWL 2 profiles: An introduction to lightweight ontology languages. In: Eiter, T., Krennwallner, T. (eds.) Reasoning Web 2012. LNCS, vol. 7487, pp. 112–183. Springer, Heidelberg (2012)
7. Lukasiewicz, T., Straccia, U.: Managing uncertainty and vagueness in description logics for the semantic web. J. Web Sem. 6(4), 291–308 (2007)
8. Lukasiewicz, T., Straccia, U.: Description logic programs under probabilistic uncertainty and fuzzy vagueness. Int. J. Approx. Reasoning 50(6), 837–853 (2009)
9. Lutz, C., Milicic, M.: A tableau algorithm for DLs with concrete domains and GCIs. Journal of Automated Reasoning 38(1–3), 227–259 (2007)
10. Pareti, P., Klein, E.: Learning vague concepts for the semantic web. In: Porc. Joint WS on Knowledge Evolution and Ontology Dynamics. In Conj. with ISWC 2011, vol. 784, CEUR workshop proceedings (2011)
11. Pérez-Urbina, H., Horrocks, I., Motik, B.: Efficient query answering for owl 2. In: Bernstein, A., Karger, D.R., Heath, T., Feigenbaum, L., Maynard, D., Motta, E., Thirunarayan, K. (eds.) ISWC 2009. LNCS, vol. 5823, pp. 489–504. Springer, Heidelberg (2009)
12. Stefan, B., Peñaloza, R.: Consistency reasoning in lattice-based fuzzy description logics. Int. J. Approx. Reason (2013)
13. Straccia, U.: Foundations of Fuzzy Logic and Semantic Web Languages. CRC Studies in Informatics Series. Chapman & Hall (2013)
14. Turhan, A.-Y.: Introductions to description logics – A guided tour. In: Rudolph, S., Gottlob, G., Horrocks, I., van Harmelen, F. (eds.) Reasoning Weg 2013. LNCS, vol. 8067, pp. 150–161. Springer, Heidelberg (2013)
15. Zadeh, L.A.: Knowledge representation in fuzzy logic. IEEE Transactions on Knowledge and Data Engineering 1(1), 89–100 (1989)