EQuIKa: Epistemic Querying in OWL 2 Ontologies

Anees Mehdi, Sebastian Rudolph, and Jens Wissmann

Institute AIFB, Karlsruhe Institute of Technology, DE
{anees.mehdi,sebastian.rudolph}@kit.edu
Forschungszentrum Informatik Karlsruhe, DE
wissmann@fzi.de

Abstract. Extending ontology querying facilities with the epistemic operator provides practically useful additional functionalities like ontology introspection, integrity constraints checking, etc. In this paper, we present a practical system called EQuIKa capable of epistemic inference on OWL 2 DL Ontologies. It implements our recently developed reduction method of epistemic queries to several subsequent standard reasoning steps. EQuIKa is implemented as a Protégé plugin, featuring a convenient querying interface alike the DL Query tab, enabling epistemic querying directly within the editor. Besides the implementation details, we discuss several optimization issues important for the feasibility of the system in practice. First experiments demonstrate practical feasibility of our system, as the system’s runtime lies in the same order of magnitude as standard reasoning tasks.

1 Introduction

OWL 2 DL is the most expressive yet decidable dialect of the Web Ontology Language (OWL) [?]. It is based on the description logic (DL) $\mathcal{SROIQ}$ [?]. Being a decidable first-order logic fragment, reasoning in $\mathcal{SROIQ}$ is inherently monotonic; addition of information to the ontology never invalidates entailed conclusions. Nevertheless, certain aspects of non-monotonicity are sometimes desired in various semantic applications. In [?], the inadequacy of DLs for the task of matching semantic services is discussed and a solution based on a non-monotonic extension of DLs is proposed. A mild form of non-monotonicity which can still be handled rather conveniently but is sufficient for many purposes is when only the querying language is endowed with non-monotonic constructors. In early 1980s, Hector J. Levesque was the first to present the idea of enriching the query language with the epistemic operator $\mathbf{K}$ [?]. In [?], Raymond Reiter makes a similar arguments and discusses why a language enriched with epistemic operator is desirable for integrity constraint (IC) checking.
Also in the DL community, extensions by epistemic operators have been considered [??.??]. In epistemic extension of DLs, the epistemic operator, also called $K$-operator and paraphrased as “known to be, is allowed in front of a concept (role) to represent the individual (pair of individual) known by an ontology to possess the property described by the concept (role). In other words, a query language with $K$-operators allows for ontology introspection.

The $K$-operator allows for epistemic querying. E.g., in order to formulate queries like “known white wine that is not known to be produced in a French regions we could perform an instance retrieval w.r.t. the concept $K_{WhiteWine} \sqcap \neg \exists K_{locatedIn} \{FrenchRegion\}$. This concept represents all the wines that are explicitly excluded from being French wines but for which there is also no evidence of being French wines either (neither directly nor indirectly via deduction). For an ontology containing

\[
\{ WhiteWine(MountadamRiesling), locatedIn(MountadamRiesling, AustralianRegion) \}
\]

the query would yield $MountadamRiesling$ as a result, since it is known to be a white wine (stated explicitly in the ontology) but not known to be produced in France. Note that querying for the instances of the concept $WhiteWin \sqcap \neg \exists locatedIn \{dlFrenchRegion\}$ yields an empty result. Further, a query language enriched with the $K$-operator allows for IC checking. Such constraint checking capabilities help ontology engineers to determine whether certain properties are satisfied by different elements occurring in the ontology. For example, testing whether the axiom

\[
K_{Wine} \sqsubseteq \exists K_{hasSugar}.\{Dry\} \sqcap \exists K_{hasSugar}.\{OffDry\} \sqcap \exists K_{hasSugar}.\{Sweet\}
\]

is entailed by an ontology allows to check if for every named individual in that ontology that is known to be a wine, it is also known (i.e., it can be logically inferred from the ontology) what degree of sugar it has. Note that this cannot be taken for granted even if $K_{Wine} \sqsubseteq \exists K_{hasSugar}.\{Dry\} \sqcap \exists K_{hasSugar}.\{OffDry\} \sqcap \exists K_{hasSugar}.\{Sweet\}$ is stated in (or can be derived from) the ontology.

In this paper, we present a system called $EQuIKa^1$ which allows for epistemic querying and IC checking. It basically implements and significantly improves the algorithm devised in our early works [??.??]. For the ease of use, we have also implemented a Protégé plugin that allows a convenient deployment of epistemic querying during the ontology design.

---

1. Epistemic Querying Interface Karlsruhe
Table 1. Syntax and semantics of role and concept constructors in $SROIQ$. Thereby $a$ denotes an individual name, $R$ an arbitrary role name and $S$ a simple role name. $C$ and $D$ denote concept expressions.

<table>
<thead>
<tr>
<th>Name</th>
<th>Syntax</th>
<th>Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>inverse role</td>
<td>$R^-$</td>
<td>${(x, y) \in \Delta^I \times \Delta^I \mid (y, x) \in R^I}$</td>
</tr>
<tr>
<td>universal role</td>
<td>$U$</td>
<td>$\Delta^I \times \Delta^I$</td>
</tr>
<tr>
<td>top</td>
<td>$\top$</td>
<td>$\Delta^I$</td>
</tr>
<tr>
<td>bottom</td>
<td>$\bot$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>negation</td>
<td>$\neg C$</td>
<td>$\Delta^I \setminus C^I$</td>
</tr>
<tr>
<td>conjunction</td>
<td>$C \land D$</td>
<td>$C^I \cap D^I$</td>
</tr>
<tr>
<td>disjunction</td>
<td>$C \lor D$</td>
<td>$C^I \cup D^I$</td>
</tr>
<tr>
<td>nominals</td>
<td>${a_1, \ldots, a_n}$</td>
<td>${a^I_1, \ldots, a^I_n}$</td>
</tr>
<tr>
<td>univ. restriction</td>
<td>$\forall R.C$</td>
<td>${x \mid \forall y. (x, y) \in R^I \rightarrow y \in C^I}$</td>
</tr>
<tr>
<td>exist. restriction</td>
<td>$\exists R.C$</td>
<td>${x \mid \exists y. (x, y) \in R^I \land y \in C^I}$</td>
</tr>
<tr>
<td>Self concept</td>
<td>$\exists S:\text{Self}$</td>
<td>${x \mid (x, x) \in S^I}$</td>
</tr>
<tr>
<td>qualified number</td>
<td>$\leq n S.C$</td>
<td>${x \mid #{y \in C^I \mid (x, y) \in S^I} \leq n}$</td>
</tr>
<tr>
<td>restriction</td>
<td>$\geq n S.C$</td>
<td>${x \mid #{y \in C^I \mid (x, y) \in S^I} \geq n}$</td>
</tr>
</tbody>
</table>

phase. Our paper is organized as follows: We first provide some background knowledge as preliminaries in Section ??, where we also discuss a method for epistemic query answering. A major contribution in this work are several optimization rewrite rules in order to ensure the feasibility of EQuIKa in practice. These along with the correctness proof are the matter of Section ???. In Section ???, we discuss several implementation issues. We also provide results of experiments we performed and evaluate EQuIKa with and without our new optimization. Finally, we conclude our work in Section ?? and identify some lines of future work.

2 Preliminaries

We now present an overview of the description logic $SROIQ$ the underlying logic of OWL 2 DL (for details see [?]). For the signature of $SROIQ$ we have finite, disjoint sets $N_I$, $N_C$, and $N_R$ called individual names, concept names and role names respectively, with $N_R$ being partitioned into simple and non-simple roles. These atomic entities can be used to form complex ones as displayed in Table ??.

A $SROIQ$ TBox $\mathcal{T}$ is a finite set of TBox axioms as in Table ???. Similarly, an RBox $\mathcal{R}$ and an ABox $\mathcal{A}$ are finite sets of the corresponding
### Table 2. Syntax and semantics of SROIQ axioms

| Axiom α                | I |= α, if                              |
|------------------------|----------------------------------------|
| $R_1 \circ \cdots \circ R_n \sqsubseteq R$ | $R^I_1 \circ \cdots \circ R^I_n \subseteq R^I$ | RBox axioms |
| $\text{Dis}(S, T)$     | $S^I \cap T^I = \emptyset$             |
| $C \sqsubseteq D$      | $C^I \subseteq D^I$                    | TBox axioms |
| $C(a)$                 | $a^I \in C^I$                          | ABox axioms |
| $R(a, b)$              | $(a^I, b^I) \in R^I$                   |
| $a \equiv b$           | $a^I = b^I$                            |
| $a \neq b$             | $a^I \neq b^I$                         |

axiom types. An ontology $\Sigma$ is a tuple $(\mathcal{T}, \mathcal{R}, \mathcal{A})$ where $\mathcal{T}$ is a SROIQ TBox, $\mathcal{R}$ is a regular SROIQ role hierarchy\(^2\) and $\mathcal{A}$ is a SROIQ ABox.

The semantics of SROIQ is defined via interpretations $I = (\Delta^I, \cdot^I)$ composed of a non-empty set $\Delta^I$ called the domain of $I$ and a function $\cdot^I$ mapping individual names to elements of $\Delta^I$, concept names to subsets of $\Delta^I$ and role names to subsets of $\Delta^I \times \Delta^I$. This mapping is extended to complex role and concept expressions as in Table 2 and finally used to define satisfaction of axioms (see Table 2). We say that $I$ satisfies an ontology $\Sigma = (\mathcal{T}, \mathcal{R}, \mathcal{A})$ (or $I$ is a model of $\Sigma$, written: $I |= \Sigma$) if it satisfies all axioms of $\mathcal{T}$, $\mathcal{R}$, and $\mathcal{A}$. We say that an ontology $\Sigma$ entails an axiom $\alpha$ (written $\Sigma |= \alpha$) if all models of $\Sigma$ are models of $\alpha$.

Next we present the syntax and semantics of SROIQ\(_K\), the extension of SROIQ by $\mathbf{K}$, where we allow $\mathbf{K}$ to appear in front of as well as within concept expressions. Similary we allow $\mathbf{K}$ in front of role expressions. We call a SROIQ\(_K\)-role an epistemic role if $\mathbf{K}$ occurs in it. An epistemic role is simple if it is of the form $\mathbf{K}S$ where $S$ is a simple SROIQ-role. As the traditional semantics leads to unintuitive effects when employed for expressive languages like SROIQ\(_K\), different solutions have been proposed, e.g., in \[ ?\,?\]. We adopt the semantics presented in \[ ?\] for SROIQ\(_K\).

To this end, we present the definition of extended interpretations.

**Definition 1.** An extended SROIQ-interpretation $I$ is a tuple $(\Delta^I, \cdot^I, \varphi^I)$ such that

- $(\Delta^I, \cdot^I)$ is a standard DL interpretation,
- $\varphi^I : N_I \cup \mathbb{N} \rightarrow \Delta^I$ is a surjective function from $N_I \cup \mathbb{N}$ onto $\Delta^I$, such that for all $a \in N_I$ we have that $\varphi^I(a) = a^I$.

\(^2\) We assume the usual regularity assumption for SROIQ, but omit it for space reasons.
Note that the function $\varphi_I$ returns the actual interpretation of an individual, given its (abstract) name, under the interpretation $I$. We lift $\varphi_I$ to sets of names and let $\varphi_I^{-1}$ denote the corresponding inverse. Based on the notion of extended interpretations, we define an extended epistemic interpretation for SROIQK as a pair $(I, W)$, where $I$ is an extended SROIQ-interpretation and $W$ is a set of extended SROIQ-interpretations. The extended interpretation function $I^W$ is then defined as follows:

$$
\begin{align*}
a^{I,W} &= a^I \text{ for } a \in N_I \\
A^{I,W} &= A^I \text{ for } A \in N_C \\
R^{I,W} &= R^I \text{ for } R \in N_R \\
\top^{I,W} &= \Delta^I \text{ (the domain of } I) \\
\bot^{I,W} &= \emptyset \\
(C \cap D)^{I,W} &= C^{I,W} \cap D^{I,W} \\
(C \cup D)^{I,W} &= C^{I,W} \cup D^{I,W} \\
(\neg C)^{I,W} &= \Delta^I \setminus C^{I,W} \\
(\forall R.C)^{I,W} &= \{p_1 \in \Delta^I \mid \forall p_2.(p_1, p_2) \in R^{I,W} \rightarrow p_2 \in C^{I,W}\} \\
(\exists R.C)^{I,W} &= \{p_1 \in \Delta^I \mid \exists p_2.(p_1, p_2) \in R^{I,W} \land p_2 \in C^{I,W}\} \\
(\leq n R.C)^{I,W} &= \{d \mid \#\{e \in C^{I,W} \mid (d, e) \in R^{I,W}\} \leq n\} \\
(\geq n R.C)^{I,W} &= \{d \mid \#\{e \in C^{I,W} \mid (d, e) \in R^{I,W}\} \geq n\} \\
(KC)^{I,W} &= \varphi_I \left(\bigcap_{J \in W} \varphi_J^{-1}(C^{J,W})\right) \\
(KR)^{I,W} &= \varphi_I \left(\bigcap_{J \in W} \varphi_J^{-1}(R^{J,W})\right)
\end{align*}
$$

Again, for an epistemic role $(KR)^-$, we set $[(KR)^-]^J,W := (KR^-)^J,W$.

The semantics of TBox, RBox and ABox axioms follows exactly that for the classical semantics. We now define the notions of modelhood. An extended epistemic model of an ontology $\Sigma = (T, R, A)$ is a maximal non-empty set $\mathcal{M}$ of extended interpretations such that $(I, \mathcal{M})$ satisfies $T$, $R$ and $A$ for each $I \in \mathcal{M}$. An ontology $\Sigma$ is satisfiable (under the extended semantics) if it has an extended epistemic model. Similarly the ontology $\Sigma$ entails an axiom $\alpha$, written $\Sigma \models \alpha$, if for every extended epistemic model $\mathcal{M}$ of $\Sigma$, we have that for every $I \in \mathcal{M}$, the extended epistemic interpretation $(I, \mathcal{M})$ satisfies $\alpha$. Note that a K-free ontology $\Sigma$, one without any occurrence of K, admits a unique extended epistemic model, which is the set of all models of $\Sigma$ extended by all possible surjective mappings that map individuals names and elements of $\mathbb{N}$ to the elements of their domain. We denote this model by $\mathcal{M}(\Sigma)$. A detail discussion on the extended semantics can be found in [?], where a method of translating epistemic concept expressions into equivalent K-free ones is also presented. Since EQuIKa implements this method, we recall the translation function in the following. Note that the translation itself re-
quires to check entailment of (K-free) axioms, hence it is not strictly syntactical and it depends on the underlying knowledge base.

**Definition 2.** Given a SROIQ knowledge base Σ, we define a function \( \tilde{\Phi}_\Sigma \) mapping SROIQ concept expressions to SROIQ concept expressions (where we let \( \emptyset = \bot \)):

\[
\begin{align*}
C \mapsto C & \quad \text{if } C \text{ is an atomic or one-of concept, } \top \text{ or } \bot; \\
\text{KD} & \mapsto \begin{cases} 
\top & \text{if } \Sigma \models \tilde{\Phi}_\Sigma(D) \equiv \top \\
\{a \in N_I \mid \Sigma \models \tilde{\Phi}_\Sigma(D)(a)\} & \text{otherwise}
\end{cases} \\
\exists K.S.\text{Self} & \mapsto \begin{cases} 
\exists S.\text{Self} & \text{if } \Sigma \models \top \subseteq \exists S.\text{Self} \\
\{a \in N_I \mid \Sigma \models S(a, a)\} & \text{otherwise}
\end{cases} \\
C_1 \cap C_2 & \mapsto \tilde{\Phi}_\Sigma(C_1) \cap \tilde{\Phi}_\Sigma(C_2) \\
C_1 \cup C_2 & \mapsto \tilde{\Phi}_\Sigma(C_1) \cup \tilde{\Phi}_\Sigma(C_2) \\
-C & \mapsto -\tilde{\Phi}_\Sigma(C) \\
\exists R.D & \mapsto \exists R.\tilde{\Phi}_\Sigma(D) & \text{for non-epistemic role } R; \\
\forall R.D & \mapsto \forall R.\tilde{\Phi}_\Sigma(D) & \text{for non-epistemic role } R; \\
\forall K.P.D & \mapsto -\tilde{\Phi}_\Sigma(\exists K.P.\neg D) \\
\geq n S.D & \mapsto \geq n S.\tilde{\Phi}_\Sigma(D) & \text{for non-epistemic role } S; \\
\geq n K.S.D & \mapsto \begin{cases} 
\cup \{a \in N_I \mid \Sigma \models S(a, b) \cap \tilde{\Phi}_\Sigma(D)\} & \text{if } \Sigma \models \top \subseteq \exists S.\text{Self} \\
\cup \{\tilde{\Phi}_\Sigma(D) \cap \neg \{a \mid a \in N_I\} & \text{otherwise}
\end{cases}
\end{align*}
\]

EQuiKα implements this translation function. Note that a naive implementation of \( \tilde{\Phi}_\Sigma \) might not be feasible in practice. In the next section, we discuss several optimization techniques that we devised in order to ensure the run-time feasibility of EQuiKα.
3 Optimization

A naive implementation of the translation function presented in Definition ?? does not need to be feasible in practice, in particular when dealing with ontologies containing relatively large number of individuals. We came up with several optimization rules. As we will see in the next section, these rules speed up the computation time of EQuIKa in retrieving instances of an epistemic concept. Basically, every rule checks the structure of a given epistemic concept and reduces either the number of \( K \)'s occurring in the concept or the number of calls to the core reasoner during the translation of the concept such that the correctness of the answers is preserved. For the proof of the correctness of the answers via translation based upon the optimization rules, we show that for an epistemic concept \( C \), the extensions of the translation of \( C \) coincide when the translation is done with and without the optimization rules.

In the following we discuss each of these rules and provide the proof of their correctness, where by \( \hat{\Phi}_\Sigma \) we mean the translation function based on the optimization rules.

– **Rule 1 (Nominals):** For individual names \( a_1, \ldots, a_n \) we have

\[
\hat{\Phi}_\Sigma(K\{a_1, \ldots, a_n\}) \mapsto \{a_1, \ldots, a_n\}
\]

**Proof.** For the left to right direction, let \( x \in K\{a_1, \ldots, a_n\}^{I, M(\Sigma)} \). By semantics, therefore,

\[
x \in \varphi_I(\bigcap_{J \in M(\Sigma)} \varphi_J^{-1}(\{a_1, \ldots, a_n\}^J))
\]

In other words, \( x \in \varphi_I(\varphi_J^{-1}(\{a_1, \ldots, a_n\}^J)) \) for each \( J \in M(\Sigma) \). In particular, \( x \in \varphi_I(\varphi_I^{-1}(\{a_1, \ldots, a_n\}^I)) \). Since \( \varphi_I \) is a surjective mapping, we get that \( x \in \{a_1, \ldots, a_n\}^I = \{a_1, \ldots, a_n\}^{I, M(\Sigma)} \).

For the right to left direction, let \( x \in \{a_1, \ldots, a_n\}^{I, M(\Sigma)} \) and suppose that \( x \notin K\{a_1, \ldots, a_n\}^{I, M(\Sigma)} \). Note that by definition of \( \varphi_J \), we have that \( a_i \in \varphi_J^{-1}(a_i^J) \) for each \( J \in M(\Sigma) \) and \( 1 \leq i \leq n \). It means that \( \{a_1, \ldots, a_n\} \subseteq \varphi_J^{-1}(\{a_1, \ldots, a_n\}^J) \) for each \( J \in M(\Sigma) \) i.e.,

\[
\{a_1, \ldots, a_n\} \subseteq \bigcap_{J \in M(\Sigma)} \varphi_J^{-1}(\{a_1, \ldots, a_n\}^J)
\]

Applying \( \varphi_I \) we get

\[
\varphi_I(\{a_1, \ldots, a_n\}) \subseteq \varphi_I\left(\bigcap_{J \in M(\Sigma)} \varphi_J^{-1}(\{a_1, \ldots, a_n\}^J)\right)
\]
In other words,

\[ \varphi_I(\{a_1, \ldots, a_n\}) \subseteq K(a_1, \ldots, a_n)^{I, M(\Sigma)} \]

By assumption, since \( x \notin K(a_1, \ldots, a_n)^{I, M(\Sigma)} \), consequently we get that \( x \notin \varphi_I(\{a_1, \ldots, a_n\}) = \{\varphi_I(a_1), \ldots, \varphi_I(a_n)\} \). By definition of \( \varphi_I \) we have that \( \varphi_I(a) = a^I \) for each \( a \in N_I \). Therefore we get that \( x \notin \{a^I_1, \ldots, a^I_n\} = \{a_1, \ldots, a_n\}^{I, M(\Sigma)} \), which is a contradiction. Hence, \( x \in K(a_1, \ldots, a_n)^{I, M(\Sigma)} \) must hold.

- **Rule 2 (Conjunction):** Let \( C_1, \ldots, C_n \) be concepts where each \( C_i \) is either an epistemic concept of the form \( KD \) for some concept \( D \) or a one-of concept for \( 1 \leq i \leq n \). By \( C'_i \) we denote the concept obtained from \( C_i \) by dropping \( K \) or \( C'_i = C_i \) otherwise. Then,

\[ \hat{\Phi}(C_1 \cap \ldots \cap C_n) \mapsto K(C'_1 \cap \ldots \cap C'_n) \]

**Proof.** For a one-of concept \( \{a_1, \ldots, a_k\} \), it follows from the proof of Rule 1 that \( \{a_1, \ldots, a_k\} \) is equivalent to \( K\{a_1, \ldots, a_k\} \). Hence, we assume that each concept in \( C_1 \cap \ldots \cap C_n \) is of the form \( KD \) for some concept \( D \). Now

\[ x \in [KD_1 \cap \ldots \cap KD_n]^I, M(\Sigma) \]

\[ \iff x \in (KD_1^{I, M(\Sigma)} \cap \ldots \cap KD_n^{I, M(\Sigma)}) \]

\[ \iff x \in \varphi_I(\bigcap_{J \in M(\Sigma)} \varphi_J^{-1}(D^I_J)) \cap \ldots \cap \varphi_I(\bigcap_{J \in M(\Sigma)} \varphi_J^{-1}(D^I_n)) \]

\[ \iff x \in \varphi_I\left(\bigcap_{J \in M(\Sigma)} (\varphi_J^{-1}(D^I_1) \cap \ldots \cap \varphi_J^{-1}(D^I_n))\right) \]

\[ \iff x \in \varphi_I(\bigcap_{J \in M(\Sigma)} \varphi_J^{-1}(\bigcap_{J \in M(\Sigma)} D^I)) \]

\[ \iff x \in [K(D_1 \cap \ldots \cap D_n)]^{I, M(\Sigma)} \]

**Rule 3 (Existential Quantification):**
By Definition ???, for a concept \( \exists K.R.D \) we get that

\[ \exists K.P.D \mapsto \begin{cases} \bigcup_{a \in N_I} \{a \cap \exists P.(\{b \in N_I | \Sigma \models P(a, b) \cap \Phi_S(D)\}) \} \\ \cup \exists P.(\{b \in N_I | \Sigma \models T \subseteq \exists P.\{b\} \cap \Phi_S(D)\}) \\ \cup \{a \in N_I | \Sigma \models T \subseteq \exists P^- \{a\} \cap \exists P.\Phi_S(D)\} \\ \cup \{ \Phi_S(D) \} \quad \text{if } \Sigma \models T \subseteq \exists P.\text{Self} \\ \bot \quad \text{otherwise} \end{cases} \]

Suppose \( \text{Inst}(D) \) returns the set of instances of a concept \( D \) and let \( \text{OUT}_P = \text{Inst}(\exists P.\Phi(D)) \) and \( \text{IN}_P = \text{Inst}(\exists P^-). \) Now Rule 3 is as follows

\[ \hat{\Phi}(\exists K.P.D) \mapsto \begin{cases} \bigcup_{a \in \text{OUT}_P} \{a \cap \exists P.(\{b \in \text{IN}_P | \Sigma \models P(a, b) \cap \Phi_S(D)\}) \} \\ \cup \exists P.(\{b \in \text{IN}_P | \Sigma \models T \subseteq \exists P.\{b\} \cap \Phi_S(D)\}) \\ \cup \{a \in \text{OUT}_P | \Sigma \models T \subseteq \exists P^- \{a\} \cap \exists P.\Phi_S(D)\} \\ \cup \{ \Phi_S(D) \} \quad \text{if } \Sigma \models T \subseteq \exists P.\text{Self} \\ \bot \quad \text{otherwise} \end{cases} \]
Note that without the optimization, translating a concept of the form $\exists K P . D$ requires at least $|N_I| \times |N_I| + 2|N_I| + 1+ (\text{No. of calls needed to translate } D)$ calls to the core reasoner. With the optimization Rule 3, such a translation reduces the number of calls when $|\text{Inst}(\exists P . \tilde{\Phi}(D))| < |N_I|$ or $|\text{Inst}(\exists P^-. \top)| < |N_I|$.

**Proof.** We introduce the following abbreviations,
- $C_1 = \bigcup_{a \in N_I} \{a\} \cap \exists P . (\{b \in N_I \mid \Sigma \models P(a, b) \sqcap \tilde{\Phi}_\Sigma(D)\})$
- $C'_1 = \bigcup_{a \in \text{OUT}_P} \{a\} \cap \exists P . (\{b \in \text{IN}_P \mid \Sigma \models P(a, b) \sqcap \tilde{\Phi}_\Sigma(D)\})$
- $C_2 = \exists P . (\{b \in N_I \mid \Sigma \models T \sqsubseteq \exists P . \{b\} \sqcap \tilde{\Phi}_\Sigma(D)\})$
- $C'_2 = \exists P . (\{b \in \text{IN}_P \mid \Sigma \models T \sqsubseteq \exists P . \{b\} \sqcap \tilde{\Phi}_\Sigma(D)\})$
- $C_3 = \{a \in N_I \mid \Sigma \models \exists P^-. \{a\} \sqcap \exists P . \tilde{\Phi}_\Sigma(D)\}$
- $C'_3 = \{a \in \text{OUT}_P \mid \Sigma \models \exists P^-. \{a\} \sqcap \exists P . \tilde{\Phi}_\Sigma(D)\}$

We first show that for an extended interpretation $\mathcal{I} \in M(\Sigma)$ and $x \in \Delta^\mathcal{I}$, $x \in C_1^{\mathcal{I}, M(\Sigma)}$ if and only if $x \in C_1^{\mathcal{I}, M(\Sigma)}$ if $x \in C_1^{\mathcal{I}, M(\Sigma)}$. For this note that if $b' \in \text{Inst}(\{b \in N_I \mid \Sigma \models P(a', b)\})$ for individuals $a'$ and $b'$ then it immediately follows that $b' \in \text{Inst}(\{b \in \text{IN}_P \mid \Sigma \models P(a', b)\})$ and vice versa. Now for $x \in \Delta^\mathcal{I}$, $x \in C_1^{\mathcal{I}, M(\Sigma)}$ if and only if there is an individual $a' \in N_I$ such that $x = a'^{\mathcal{I}, M(\Sigma)}$ and $x \in \exists P . (\{b \in N_I \mid \Sigma \models P(a, b) \sqcap \tilde{\Phi}_\Sigma(D)\})^{\mathcal{I}, M(\Sigma)}$. This is the case if and only if $x \in \exists P . (\{b \in \text{IN}_P \mid \Sigma \models P(a', b) \sqcap \tilde{\Phi}_\Sigma(D)\})^{\mathcal{I}, M(\Sigma)}$ which is equivalent to $x \in C_1^{\mathcal{I}, M(\Sigma)}$.

In the same way, we can prove that $x \in C_2^{\mathcal{I}, M(\Sigma)}$ if and only if $x \in C_2^{\mathcal{I}, M(\Sigma)}$. Similarly, $x \in C_3^{\mathcal{I}, M(\Sigma)}$ if and only if $x \in C_3^{\mathcal{I}, M(\Sigma)}$. Consequently, this establishes the proof of the correctness of Rule 3.

---

**Rule 4 (Number Restriction):**
Similar to Rule 3, $\tilde{\Phi}_\Sigma$ maps the concept $\geq nK S . D$ to

$$
\left\{ \begin{array}{l}
\bigcup_{a \in \text{OUT}_{P,n}} \{a\} \cap \geq nS . (\{b \in \text{IN}_{P,n} \mid \Sigma \models S(a, b) \sqcap \tilde{\Phi}_\Sigma(D)\}) \\
\cup \{a \in \text{OUT}_{P,n} \mid \Sigma \models T \sqsubseteq \exists S^- . \{a\} \cap \geq nS . \tilde{\Phi}_\Sigma(D)\} \\
\cup \geq nS . (\{b \in \text{IN}_{P,n} \mid \Sigma \models T \sqsubseteq \exists S . \{b\} \sqcap \tilde{\Phi}_\Sigma(D)\}) \\
\cup \tilde{\Phi}_\Sigma(D) \sqcap \neg \{a \mid a \in N_I\} \\
\downarrow \\
\forall \Sigma \models T \sqsubseteq \exists S . \text{Self} \\
\forall \Sigma \models T \sqsubseteq \exists S . \text{Self}
\end{array} \right.
$$

where $\text{OUT}_{P,n} = \geq nP . \tilde{\Phi}(D)$ and $\text{IN}_{P,n} = \geq nP^- . \top$.

**Proof.** Similar to Rule 3.

These rules suffice in the sense that for most of the remaining constructs, we use their dual which correspond to one of the rules discussed above.
After showing that the rules preserve the semantics, we need to show that they indeed improve the run time of EQuIKa. For this we performed several experiments, which are discussed in Section ??.

4 Implementation

The EQuIKa system is implemented on top of the OWL-API. It can be used as an API as well as within Protégé using an epistemic query tab. The following considerations and design decisions underly our implementation:

– Since the standard OWL-API does not support epistemic constructs, we extended several classes of the API. The $K$-operator syntactically behaves similar like the complement construct ($\neg$) for concepts and like the inverse role construct for roles. We therefore followed the same implementation patterns.

– For parsing we created an EpistemicSyntaxParser based on the ManchesterOWLSyntaxOntologyParser. The $K$ operator is expressed by the token KnownConcept for concepts and by the token KnownRole for the roles.

– We implemented the translation function in a recursive fashion. For this, we implemented a visitor pattern by extending the OWLClassExpressionVisitor class in order to handle the epistemic operator.

– In order to support epistemic querying within the Protégé editor, we implemented an additional tab based on the DL Query tab. Figure ?? shows a snapshot of epistemic querying in Protégé.

The class diagram for EQuIKa is given in Figure ???. Note that the new types OWLObjectEpistemicConcept and OWLObjectEpistemicRole are derived from the respective standard types OWLBooleanClassExpression and OWLObjectPropertyExpression to fit the design of the OWL-API. As our translation method depends on intermediate calls to a standard reasoner, the class EQuIKaReasoner implements the OWLReasoner interface. As already mentioned, EQuIKa translates an epistemic concept into a $K$-free one in a recursive fashion using the class Translator that implements the OWLClassExpressionVisitor. Further, since Protégé can utilize any reasoner that implements the OWLReasoner interface, this enables Protégé to use EQuIkaReasoner to answer epistemic queries. Last but not least, EQuIKa has been shared on googlecode for testing purposes.

4 http://owlapi.sourceforge.net/
4 http://code.google.com/p/epistemicdl/
is provided as jar file\textsuperscript{5} that can be installed via the Protégé 4.1 plugin folder.

5 Evaluation

For the purpose of evaluation, we performed several experiments with the following setup:

– We used IBM Thinkpad T60 dual core, 2 GHz each core, with Windows 7 (32-bit) as the operating system. A total of 2 GB memory was available.
– For benchmark tests, we used a populated version of the Wine ontology\textsuperscript{6}, that contains 483 individuals and uses most of the OWL 2 DL constructs. These ontologies can be downloaded along with EQuIKa.
– To evaluate the performance of EQuIKa, we constructed several epistemic concepts and translated them into $K$-free ones. These concepts are given in Table ?? where $r_1, \ldots, r_{108}$ are individuals representing wine regions in the ontologies.

\textsuperscript{5} https://epistemicdl.googlecode.com/svn/EpistemicQueryTab/equika.protege.querytab.jar
\textsuperscript{6} https://code.google.com/p/epistemicdl/source/browse/trunk/EQuIK/wine_1.owl
To the best of our knowledge, EQuIKa is the only reasoner of its nature for epistemic query answering, such that there is no other existing reasoner with these capabilities against which we could compare EQuIKa’s performance. To give an impression about the runtime behavior, we performed two kind of experiments and as a measure, we consider the time required to translate the epistemic concepts (given in Table 3) to K-free equivalent ones and the instance retrieval time of the translated concept. In the first series of experiments, we evaluated the benefit of the optimization rules introduced in Section 2. We implemented two versions of EQuIKa; a naive one called EQuIKa-N implementing the translation function of Definition 2 as is and an optimized one called EQuIKa-O where the optimization rules were used. The corresponding results are

---

**Table 3. Concepts used for instance retrieval experiments.**

| EC1 | K hasWineDescriptor ⊓ K WineDescriptor |
| EC2 | K hasWineDescriptor ⊓ K WineDescriptor ⊓ ∃ K madeFromFruit ⊓ K WineGrape |
| EC3 | K RoseWine |
| EC4 | K RoseWine ⊓ K WhiteWine ⊓ {r₁,...,r₁₀₈} |
| EC5 | K Wine ⊓ ¬∃ K hasSugar. {Dry} ⊓ ¬∃ K hasSugar. {Off Dry} ⊓ ¬∃ K hasSugar. {Sweet} |

---

**Fig. 2.** The EQuIKa-system extending the OWL-API
Table 4. Results of instance retrieval experiments.

<table>
<thead>
<tr>
<th>Concept</th>
<th>EQuIKa-N</th>
<th></th>
<th></th>
<th></th>
<th>EQuIKa-O</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_{\text{trans}}$</td>
<td>$T_{\text{inst}}$</td>
<td>$#_{\text{inst}}$</td>
<td>$T_{\text{trans}}$</td>
<td>$T_{\text{inst}}$</td>
<td>$#_{\text{inst}}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EC_1$</td>
<td>4</td>
<td>192.7</td>
<td>132</td>
<td>21</td>
<td>97.8</td>
<td>132</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EC_2$</td>
<td>9</td>
<td>198.9</td>
<td>3</td>
<td>3</td>
<td>37.5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EC_3$</td>
<td>110</td>
<td>110.1</td>
<td>3</td>
<td>26</td>
<td>26.5</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EC_4$</td>
<td>203</td>
<td>211.7</td>
<td>0</td>
<td>122</td>
<td>122.1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EC_5$</td>
<td>206</td>
<td>400.6</td>
<td>0</td>
<td>121</td>
<td>121.9</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EC_6$</td>
<td>13</td>
<td>–</td>
<td>–</td>
<td>0.5</td>
<td>487.3</td>
<td>119</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

shown in Table ?? where $T_{\text{trans}}$, $T_{\text{inst}}$ and $\#_{\text{inst}}$ represent the translation time, instance retrieval time and the number of instances respectively. A time comparison in seconds between both versions of EQuIKa is presented in Figure ???. For each concept, the first (brown) bar represents the total time taken by EQuIKa-N and the second (blue) bar represents the total time taken by EQuIKa-O. In both bars, the dark-shaded part represented the time taken for the translation and the unshaded part represents the instance retrieval time by the corresponding EQuIKa version. It can be seen in the chart that the optimized EQuIKa is one or several orders of magnitude faster than the non-optimized one. In particular for concept $EC_6$, EQuIKa-N didn’t responded for almost an hour and we stopped it, whereas EQuIKa-O translated $EC_6$ and retrieved its instances in few seconds. Hence, beside being correct, the optimized rules introduced are of high importance toward the feasibility of EQuIKa in practice.

In the second series of experiments, we evaluated the computation time of EQuIKa-O in general to provide an impression of how the cost of epistemic querying relates to standard reasoning tasks. For this purpose, we consider non-epistemic concepts $C_1, \ldots, C_6$ where each $C_i$ is obtained by dropping $K$ in $EC_i$ for $1 \leq i \leq 6$. Note that an epistemic concept $EC_i$ and the corresponding $C_i$ are semantically different concepts. Table ?? shows the results of our experiments. It can be seen that even when comparing to the $K$-free counter part of the epistemic concepts, the computation time of EQuIKa-O is roughly in the same order of magnitude.

6 Conclusion

In this paper, we have motivated the importance of epistemic querying of OWL ontologies for ontology introspection and integrity constraint
EQuIKa is based on a reduction of epistemic queries to standard reasoning. In order to assure its practical feasibility, we have presented, proven and implemented several optimization rules leading to a speed-up of EQuIKa by one to several orders of magnitude. We performed

### Table 5. Evaluation epistemic vs. standard instance retrieval

| Concept | $t_C$ | $|C_i|$ | Concept | $t_{EC}$ | $t_{EC}$ | $|EC_i|$ |
|---------|-------|--------|---------|---------|---------|--------|
| $C_1$   | 2.18  | 159    | $EC_1$  | 20      | 95.7    | 132    |
| $C_2$   | 41.9  | 159    | $EC_2$  | 3       | 36.5    | 3      |
| $C_3$   | 10.7  | 3      | $EC_3$  | 10      | 10.8    | 3      |
| $C_4$   | 2.68  | 0      | $EC_4$  | 2       | 2.9     | 0      |
| $C_5$   | 0.2   | 0      | $EC_5$  | 2       | 2.9     | 0      |
| $C_6$   | 61.1  | 80     | $EC_5$  | 0.5     | 487.3   | 119    |
several experiments checking computation time of EQuIKa and also evaluated EQuIKa against the standard reasoning task of instance retrieval. We found that EQuIKa performs in the same scale as the standard counter part, witnessing that epistemic querying can be efficiently realized in practice.

Avenues for future research are manifold: we will carry out a more extensive evaluation of our system with data stemming from ontology design scenarios from industry project’s use cases. These experiments will provide a clearer view on which – already implemented or still to be defined – optimizations and heuristics in EQuIKa pay off in practice. On the theoretical side, we want to investigate the practical benefits of epistemic constructors as part of the ontology modeling language and try to extent our framework to this case. This is clearly a non-trivial task, since favorable model-theoretic properties get lost.

References
