# Distributed Instance Retrieval in $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$ Representation Framework

George M. Santipantakis[1] and George A. Vouros[2]

[1] ICSD, University of the Aegean, Greece
[2] Digital Systems, University of Piraeus, Greece

**Abstract.** While there has been a great deal of work concerning distributed reasoning with ontologies, in most cases either only TBox reasoning is concerned, or ABox reasoning is supported for languages of limited expressivity. To a greater extent than other representation frameworks, the proposed $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$ framework allows peers to establish concept-to-concept semantic correspondences to acquaintances' ontologies via bridge rules, and relate individuals by equivalence correspondences or link assertions for the $\mathcal{SHIQ}$ fragment of Description Logics. The paper presents $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$ and proposes an algorithm for the retrieval of individuals in a distributed setting.

## 1 Introduction

In this paper we study settings where heterogeneous, connected peers (i.e. peers with distinct and independently developed ontologies) aim to combine their conceptual and assertional knowledge towards the distributed retrieval of instances. To do so, we need to establish semantic correspondences between concepts and interlink instances from the different ontologies. Several applications can benefit from the combination of the distinct ontologies located in peers. This may also be the case for peers that may decide to modularize their knowledge, so as to enhance their reasoning performance.

Intrigued to provide a solution for such settings, we propose the $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$ representation framework. This framework, to a greater extent than other frameworks/languages for distributed reasoning with distinct ontologies, allows subjective semantic correspondences to be applied between concepts, allows equivalence correspondences between pairs of individuals and it provides a special type of roles, namely links, which relate individuals in distinct ontologies. Links can be transitive or applied to cardinality and qualitative restrictions, be applied to existential and universal restrictions, as well as hierarchically related to other relations. $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$ has been inspired by Distributed Description Logics, originally introduced by [1] and $\mathcal{E}$-connections [4]. The framework therefore naturally inherits the semantic constructors available in these frameworks for the conceptual and assertional part of the ontology and places further restrictions so as to preserve decidability, and distinguish clearly between representation cases.

I. Maglogiannis, V. Plagianakos, and I. Vlahavas (Eds.): SETN 2012, LNAI 7297, pp. 141–148, 2012.

This paper presents the $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$ framework and discusses idiosyncrasies of that framework for distributed information retrieval. Finally, it presents an algorithm for retrieving instances in distributed settings of interconnected heterogeneous peers.

## 2 A Motivating Scenario

To clearly consider the issues that the proposed framework aims to address, we assume the ontology for the centralized semantic information system (SIS) presented in [11]: SIS aims to support the location of markets that trade computational resources in a democratized grid or cloud environment. The concepts in the ontology describe the different types of markets that agents may participate in, the different types of agents that participate in these markets, as well as the types of resources. Aiming to a distributed implementation of SIS, participating peers may possess this ontology, or parts of this ontology according to subjective priorities, expertise or interests. This may result to any arbitrary decomposition of the formalized information. Alternative conceptualizations of the domain elements are also possible between different peers.

In this setting we require peers to combine their knowledge about markets and tradable resources (i.e. resources advertised and requested in markets) and retrieve instances with specific characteristics. Thus, the problem is that, given a query Q formalized as a concept in any of the peers, retrieve all instances of Q that are implied by the distributed knowledge base (i.e. implied by all of its models). The problem that this work addresses, concerns combining heterogenous knowledge bases so as peers to retrieve instances effectively.

## 3 Related Work

There are several prominent works presented for efficient reasoning over large (trillions of triples) knowledge bases. The majority of the methods apply the same idea: construct an abstraction from the specifications of these triples, and use it for reasoning. The abstraction can be either a product of Map-Reduce tasks e.g. [8], or be a summary ABox [2]. Such methods can be complementary to the proposed framework, so as to make reasoning efficient, locally to each peer.

Among the frameworks that can apply distributed reasoning at the assertional level, we distinguish [12] which separates the assertional part of the ontology into modules. The method however cannot support qualified cardinality restrictions and the expressivity has upper bound to $\mathcal{SHI}$. DDL is another framework that clearly supports instance retrieval for $\mathcal{SHIQ}$ [10]. Reasoning is performed by tableaux algorithms in each peer, and subsumptions, as well as instances' specifications are propagated via semantic correspondences. This means that the framework is applicable to ontologies with overlapping domains, where there are no further relations between pairs of instances beyond (unrestricted) correspondences. As shown in Section 6, extending the framework with relations between instances in different ontologies, and restricting correspondences to equalities between instances, as it is done in $E_{HQ^+}^{DDL}$, the retrieval task is getting more complicated. Furthermore, as far as we know, there is not any work describing distributed instance retrieval for $\mathcal{E}$-connections: This totally resides on local means, which is also justified by the assumption that ontologies cover distinct domains.

In addition to the above, we emphasize on the "locality" of knowledge: Peers must not be forced to share their axioms and assertions with others, and distributed reasoning must result from combinations of local reasoning chunks performed in distinct ontology units.

## 4 Introduction to $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$

**Preliminaries :** In this section we present preliminaries on $\mathcal{SHIQ}$, Distributed Description Logics and $\mathcal{E}$-connections.

Let $\mathcal{N_C}$ be a set of concept names, $\mathcal{N_R}$ be a set of role names and $\mathcal{N_O}$ the set of individual names. Let $Inv(R)$ denote the inverse role of $R$ and $(\mathcal{N_R} \cup \{Inv(R) | R \in \mathcal{N_R}\})$ be the set of $\mathcal{SHIQ}$-roles. The set of $\mathcal{SHIQ}$-concepts is the smallest set constructed by the constructors in Table 1. In order to preserve decidability, number restrictions are restricted to *simple* roles only, i.e. roles that are neither transitive nor they have any transitive sub-roles. An interpretation $\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle$ consists of a non empty domain $\Delta^{\mathcal{I}}$ and the interpretation function $\cdot^{\mathcal{I}}$ which maps every concept to a subset of $\Delta^{\mathcal{I}}$ and every role to a subset of $\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$.

**Table 1.** $\mathcal{SHIQ}$ fragment of Description Logics

| | | |
|---|---|---|
| Atomic Concept | $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ | |
| Universal Concept | $\top^{\mathcal{I}} = \Delta^{\mathcal{I}}$ | |
| Bottom Concept | $\bot^{\mathcal{I}} = \emptyset$ | |
| Atomic Role | $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ | |
| Conjunction | $(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$ | $\mathcal{S}$ |
| Disjunction | $(C \sqcup D)^{\mathcal{I}} = C^{\mathcal{I}} \cup D^{\mathcal{I}}$ | |
| Negation | $(\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$ | |
| Existential Restriction | $(\exists R.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \exists y \in \Delta^{\mathcal{I}}, (x,y) \in R^{\mathcal{I}}, y \in C^{\mathcal{I}}\}$ | |
| Value Restriction | $(\forall R.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \forall y \in \Delta^{\mathcal{I}}, (x,y) \in R^{\mathcal{I}} \rightarrow y \in C^{\mathcal{I}}\}$ | |
| Transitive Role | $\mathcal{I} \models Trans(R) \leftrightarrow R^{\mathcal{I}} = (R^{\mathcal{I}})^+$ | |
| Role Hierarchy | $\mathcal{I} \models (P \sqsubseteq R)^{\mathcal{I}} \leftrightarrow P^{\mathcal{I}} \subseteq R^{\mathcal{I}}$ | $\mathcal{H}$ |
| Inverse Role | $(Inv(R))^{\mathcal{I}} = \{(x,y) \mid (y,x) \in R^{\mathcal{I}}\}$ | $\mathcal{I}$ |
| Qualified | $(\geq nS.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}}, \|\|y, (x,y) \in S^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\|\| \geq n\}$ | $\mathcal{Q}$ |
| Number Restrictions | $(\leq nS.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}}, \|\|y, (x,y) \in S^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\|\| \leq n\}$ | |

Let $C$ and $D$ possibly complex concepts, $C \sqsubseteq D$ is called a general concept inclusion (GCI) axiom. A finite set of GCIs is called a TBox (denoted by $\mathcal{T}$). An interpretation $\mathcal{I}$ satisfies a GCI $C \sqsubseteq D$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$. $\mathcal{I}$ satisfies a TBox if it satisfies each GCI in it. In this case $\mathcal{I}$ is a model of this TBox. A concept $C$ is satisfiable w.r.t. a role hierarchy $\mathcal{R}$ and a TBox if there is a *model* $\mathcal{I}$ of TBox and $\mathcal{R}$ with $C^{\mathcal{I}} \neq \emptyset$. A concept $C$ subsumes a concept $D$ w.r.t TBox $\mathcal{T}$ and $\mathcal{R}$ if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$ holds in every model of $\mathcal{T}$ and $\mathcal{R}$.

For a concept expression $C$, a role name $R$, and the individual names $a, b$, assertions are either *instance assertions* of the form $a : C$ or *role assertions* $(a, b) : R$. Also, $a \doteq b$, $a \dot{\neq} b$ are *individual equality*, *inequality assertions*, respectively. The finite set of assertions w.r.t. a TBox $\mathcal{T}$ and a role hierarchy $\mathcal{R}$ is called ABox.

Distributed Description Logics (DDL) is a framework which allows distributed reasoning over $\mathcal{SHIQ}$ [1], [10]. The intuition of the framework is that concepts and individuals in an ontology can be mapped to corresponding elements of remote ontologies. Given a finite index set $I$, the correspondences between concepts are expressed as *onto*-bridge rules $i : C \overset{\sqsupseteq}{\rightarrow} j : G$, or *into*-bridge rules $i : C \overset{\sqsubseteq}{\rightarrow} j : G$, where $i, j$ in $I$ denote distinct ontologies, $i : C$, $j : G$ concepts in these ontologies and the direction of the arrow denotes the subjectiveness of the correspondence (i.e. correspondences are under the subjective view of $j$). For the assertional part, subjective individual correspondences can be either *partial* or *complete*. Given an instance name $a_i$ in a local ABox $\mathcal{A}_i$ and $b_j^1, b_j^2, ..., b_j^n$ instances names in the ABox $\mathcal{A}_j$, a partial individual correspondence (PIC) is an expression of the form $a_i \mapsto b_j^k, k = 1, ..., n$, while a complete individual correspondence (CIC) is an expression of the form $a_i \overset{=}{\mapsto} \{b_j^1, ..., b_j^n\}$. The distributed knowledge base is constructed by the local knowledge bases, and the set of bridge rules and correspondences between mapped elements.

On the other hand, $\mathcal{E}$-connections is a framework which combines representations in different logics ([5]). Concerning DL, as originally proposed in [4], $\mathcal{E}$-connections are intended for modelling scenarios where the respective domains of the ontology units are mutually disjoint, however this assumption has been relaxed in [3]. $\mathcal{E}$-connections $\mathcal{C}^{\mathcal{E}}_{\mathcal{HQ}+}(\mathcal{SHIQ}, \mathcal{SHOQ}, \mathcal{SHIO})$ may combine ontology units in any of the $\mathcal{SHIQ}$, $\mathcal{SHOQ}$, $\mathcal{SHIO}$ fragments of DL using link-properties. Link-properties can be hierarchically related, be transitive and, in case they are simple, be restricted by qualitative restrictions.

For a finite index set $I$, a set of link-properties connecting concepts in the $i$ and $j$ units, $i \neq j \in I$, is defined to be the set $\mathcal{E}_{ij} = \epsilon_{ij}$. In case $i = j$ is the set $\mathcal{E}_{ij} = \epsilon_{ij} \cup \{Inv(E) | E \in \epsilon_{ji}\}$, where $\epsilon_{ij}$ are the sets of link-property or role names that are not pairwise disjoint, but are disjoint with respect to the sets of concept names. An $ij$-property axiom is an assertion of the form $E_{ij}^n \sqsubseteq E_{ij}^m$, where the superscript distinguishes link-properties in $\mathcal{E}_{ij}$. An $ij$-property box $R_{ij}$ includes a finite set of $ij$-property inclusion axioms ($R_{ii}$ is the local RBox $\mathcal{R}_i$). The sets of i-concepts (i.e. concepts specified in the $i$-th unit) are inductively defined as the smallest sets constructed using the constructors provided by the local DL fragment, as well with the link-property specifications' constructors [4].

The $\mathcal{E}-$connections definition of a combined TBox is a family of TBoxes $\mathbf{T} = \{\mathcal{T}_i\}_{i \in I}$, where $\mathcal{T}_i$ is a finite set of $i$-concept inclusion axioms. A combined knowledge base $\Sigma = \langle \mathbf{T}, \mathcal{R} \rangle$ is composed by the combined TBox $\mathbf{T}$, and the combined RBox $\mathcal{R}$. In addition to instance assertions, assertions between individuals can be of the form $a \cdot E_{ij} \cdot b$, where $E_{ij}$ is a property in $\mathcal{E}_{ij}$.

Both DDL and $\mathcal{E}$-connections apply distributed tableau algorithms for deciding concepts' satisfiability. The algorithms can be found in [9] and [6].

$E_{HQ^+}^{DDL}\mathcal{SHIQ}$ combines the features of DDL and $\mathcal{E}-$connections as follows:

**Definition 1.** *( $E_{HQ^+}^{DDL}\mathcal{SHIQ}$ Syntax) Let $I$ be a non empty set of indexes, for $i, j \in I$, $N_{C_i}$ are sets of concept names and the set of $ij$-properties' names is denoted by $\epsilon_{ij}$, not necessarily pairwise disjoint, but disjoint with respect to the sets $N_{C_i}$. For $i, j \in I$, the set of ij-properties connecting concepts in the $i$ and $j$ units, is defined as $\mathcal{E}_{ij} = \epsilon_{ij}$, and in case $i = j$ is the set $\mathcal{E}_{ij} = \epsilon_{ij} \cup \{Inv(E)|E \in \epsilon_{ji}\}$. An ij-property axiom is an assertion of the form $E_{ij}^{n} \sqsubseteq E_{ij}^{m}$, where $E_{ij}^{*}$ are distinct properties in $\mathcal{E}_{ij}$. Transitive axioms are of the form $Trans(E; (i, j))$, where $E$ is a property name defined for the pair of ontology units $i, j \in I$. Such an axiom is a shorthand for the axiom $Trans(E; (i, i), (i, j))$ as defined in [6], meaning that $E$ is a transitive role in $i$ and a transitive link-property connecting $i$ and $j$.*

*An $ij$-property box $R_{ij}$ includes a finite set of $ij$-property inclusion axioms, plus all transitivity axioms concerning $ij$-properties. The combined property box RBox $\mathcal{R}$ contains each of the $ij$-property boxes. The set of $i$-concepts, $i \in I$, are inductively defined as in $\mathcal{E}-$connections and cardinality restrictions may hold only for simple properties.*

*Finally, semantic correspondences are denoted as bridge rules of concept onto concept, or concept into concept rules, as already defined.*

**Definition 2.** *($E_{HQ^+}^{DDL}\mathcal{SHIQ}$ Distributed Knowledge Base) A combined TBox is a family of TBoxes $\boldsymbol{T}= \{\mathcal{T}\}_{i\in I}$, where each $\mathcal{T}_i$ is a finite set of i-concept inclusion axioms. A distributed knowledge base $\Sigma = \langle \boldsymbol{T}, \mathcal{R}, \mathcal{A}, \mathcal{B}\rangle$ is composed by the combined TBox $\boldsymbol{T}$, the combined RBox $\mathcal{R}$, and a collection of bridge rules $\mathcal{B} = \{\mathfrak{B}_{ij}\}_{i\neq j\in I}$ between ontology units. A distributed ABox $\mathcal{A} = \langle\{A\}_{i\in I}, \mathcal{C}, \mathfrak{L}\rangle$ consists of the family of ABoxes $\{A\}_{i\in I}$, a collection of individual correspondences $\mathcal{C} = \{\mathfrak{C}_{ij}\}_{i\neq j\in I}$ of the form $i : a \overset{=}{\mapsto} j : b$, and property assertions $\mathfrak{L} = \{\mathfrak{L}_{ij}\}_{i,j\in I}$ of the form $a \cdot E_{ij} \cdot b$, where $E_{ij}$ is a property in $\mathcal{E}_{ij}$.*

Each TBox $\mathcal{T}_i$, $i \in I$ is locally interpreted by a local, possibly hole interpretation $\mathcal{I}_i$ that consists of a domain $\Delta^{\mathcal{I}_i}$, a valuation function $\cdot^{\mathcal{I}_i}$ which maps every concept to a subset of $\Delta^{\mathcal{I}_i}$. The ij-property boxes $R_{ij}$ with $i, j \in I$, are interpreted by valuation functions $\cdot^{\mathcal{I}_{ij}}$ that map every $ij$-property to a subset of $\Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_j}$. Let $\mathcal{I}_{ij} = \langle\Delta^{\mathcal{I}_i}, \Delta^{\mathcal{I}_j}, \cdot^{\mathcal{I}_{ij}}\rangle$, $i, j \in I$. A hole interpretation maps any concept (including the top and bottom ones) to the domain or to the empty set.

**Definition 3.** *($E_{HQ^+}^{DDL}\mathcal{SHIQ}$ Domain relation) A domain relation $r_{ij}, i \neq j$ from $\Delta^{\mathcal{I}_i}$ to $\Delta^{\mathcal{I}_j}$ is a subset of $\Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_j}$, s.t. for each $d$ in $\Delta^{\mathcal{I}_i}$, $r_{ij}(d) = \{d'|d' \in \Delta^{\mathcal{I}_j}$, with $i : d \overset{=}{\mapsto} j : d'\}$ and it holds that in case $r_{ij}(d_1) = d'$ and $r_{ij}(d_2) = d'$, then $d_1 = d_2$. Also, given a subset $D$ of $\Delta^{\mathcal{I}_i}$, $r_{ij}(D)$ denotes $\cup_{d\in D} r_{ij}(d)$.*

**Definition 4.** *($E_{HQ^+}^{DDL}\mathcal{SHIQ}$ Distributed Interpretation) Given the index $I$ and $i, j \in I$, a distributed interpretation $\mathfrak{I}$ of a distributed knowledge base $\Sigma$ is the tuple formed by the set of local interpretations $\mathcal{I}_i = \langle\Delta^{\mathcal{I}_i}, \cdot^{\mathcal{I}_i}\rangle$ for each $\mathcal{T}_i$, the family $\{\mathcal{I}_{ij}\}$, and a set of domain relations $r_{ij}$. Formally, $\mathfrak{I} = \langle\{\mathcal{I}_i\}_{i\in I}, \{\mathcal{I}_{ij}\}_{i,j\in I}, \{r_{ij}\}_{i\neq j\in I}\rangle$.*

A local interpretation $\mathcal{I}_i$ satisfies an $i$-concept $C$ w.r.t. a distributed knowledge base $\Sigma$ , i.e. $\mathcal{I}_i \vDash C$ iff $C^{\mathcal{I}_i} \neq \emptyset$. $\mathcal{I}_i$ satisfies an axiom $C \sqsubseteq D$ between i-concepts ( i.e. $\mathcal{I}_i \vDash C \sqsubseteq D$) if $C^{\mathcal{I}_i} \subseteq D^{\mathcal{I}_i}$. Also, $\mathcal{I}_{ij}$ satisfies an ij-property axiom $R \sqsubseteq S$ ($\mathcal{I}_{ij} \vDash R \sqsubseteq S$) if $R^{\mathcal{I}_{ij}} \subseteq S^{\mathcal{I}_{ij}}$. A transitivity axiom $Trans(E; (i,j))$ is satisfied by $\mathcal{I}_i$ iff $E^{\mathcal{I}_{ii}} \cup E^{\mathcal{I}_{ij}}$ is transitive.

The distributed interpretation $\mathfrak{I}$ *satisfies* ($\vDash_d$) the elements of a distributed knowledge base, if the following conditions hold:

1. $\mathfrak{I} \vDash_d i : C \sqsubseteq D$, if $\mathcal{I}_i \vDash C \sqsubseteq D$
2. $\mathfrak{I} \vDash_d \mathcal{T}_i$ if $\mathfrak{I} \vDash i : C \sqsubseteq D$ for all $C \sqsubseteq D$ in $\mathcal{T}_i$
3. $\mathfrak{I} \vDash_d i : C \overset{\sqsubseteq}{\rightarrow} j : D$, if $r_{ij}(C^{\mathcal{I}_i}) \subseteq D^{\mathcal{I}_j}$
4. $\mathfrak{I} \vDash_d i : C \overset{\sqsupseteq}{\rightarrow} j : D$, if $r_{ij}(C^{\mathcal{I}_i}) \supseteq D^{\mathcal{I}_j}$
5. $\mathfrak{I} \vDash_d \mathfrak{B}_{ij}$, if $\mathfrak{I}$ satisfies all bridge rules in $\mathfrak{B}_{ij}$
6. $\mathfrak{I} \vDash_d R \sqsubseteq S$, if $\mathcal{I}_{ij} \vDash R \sqsubseteq S$, where $R \sqsubseteq S$ in $R_{ij}$
7. $\mathfrak{I} \vDash_d Trans(E; (i,j))$ if $\mathcal{I}_i \vDash Trans(E; (i,j))$, where $Trans(E; (i,j))$ in $R_{ij}$
8. $\mathfrak{I} \vDash_d R_{ij}$ if $\mathfrak{I} \vDash_d R \sqsubseteq S$ and $\mathfrak{I} \vDash_d Trans(E; (i,j))$ for all inclusion and transitivity axioms in $R_{ij}$
9. $\mathfrak{I} \vDash_d \Sigma$ if for every $i, j \in I, \mathfrak{I} \vDash_d \mathcal{T}_i$, $\mathfrak{I} \vDash_d \mathcal{R}_{ij}$ and $\mathfrak{I} \vDash_d \mathfrak{B}_{ij}$.

**Definition 5.** *($E_{HQ^+}^{DDL}\mathcal{SHIQ}$ Distributed entailment and satisfiability) $\Sigma \vDash_d X \sqsubseteq Y$ if for every $\mathfrak{I}$, $\mathfrak{I} \vDash_d \Sigma$ implies $\mathfrak{I} \vDash_d X \sqsubseteq Y$, where $X$ and $Y$ are either i-concepts, or ij-properties, $i, j \in I$. $\Sigma$ is satisfiable if there exists a $\mathfrak{I}$ such that $\mathfrak{I} \vDash_d \Sigma$. A concept $i : C$ is satisfiable with respect to $\Sigma$ if there is a $\mathfrak{I}$ such that $\mathfrak{I} \vDash_d \Sigma$ and $C^{I_i} \neq \emptyset$.*

The worst case complexity is 2NexpTime w.r.t. the size of the combined TBox and RBox. Further details on $E_{HQ^+}^{DDL}\mathcal{SHIQ}$ can be found in [7].

## 5 Distributed Instance Retrieval in $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$

The task of instance retrieval in any fragment of Description Logics, is defined as the computation of the set of individuals that instantiate a given concept. For the proposed framework $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$, into-bridge rules and individual correspondences provide the means through which information is "translated" and transferred between peers. Intuitively, an into-bridge rule $i : A \overset{\sqsubseteq}{\rightarrow} j : B$ means that if concept $i : A$ has an individual $a$, then there should exist an individual $j : b$ such that $(a^{I_i}, b^{I_j}) \in r_{ij}$. Formally, given a distributed knowledge base $\Sigma = \langle \mathbf{T}, \mathcal{R}, \mathcal{A}, \mathcal{B} \rangle$, and $i : A \overset{\sqsubseteq}{\rightarrow} j : B \in \mathfrak{B}_{ij}$, $\langle i : a \overset{=}{\mapsto} j : b \rangle \in \mathfrak{C}_{ij}$ then $\Sigma \vDash i : A(a) \Longrightarrow \Sigma \vDash j : B(b)$ (locally to $i, j$).

According to the semantics of a complete individual correspondence (CIC) such as $\langle i : a \overset{=}{\mapsto} j : b \rangle$, the pair of individuals $\langle a^{I_i}, b^{I_j} \rangle$ belongs to the domain relation $r_{ij}$ and $a^{I_i}, b^{I_j}$ are the same real-world object. Since equality is a transitive relation, given Definition 3 the following holds: if $\langle i : x \overset{=}{\mapsto} j : u \rangle$, $\langle i : y \overset{=}{\mapsto} j : v \rangle$, $x^{I_i} = y^{I_i}$ then $u^{I_j} = r_{ij}(x^{I_i}) = r_{ij}(y^{I_i}) = v^{I_j}$ under the subjective point of view of $j$. Similarly to the instance retrieval algorithm proposed in [10], we need a transformation function $f_{ij}$, which transforms individuals from $i$ to individuals of $j$, such that their interpretations respect the semantics of domain relation $r_{ij}$. Formally, given a bridge rule $\{i : A \overset{\sqsubseteq}{\rightarrow} j : B\}$, the distributed knowledge base $\Sigma = \langle \mathbf{T}, \mathcal{R}, \mathcal{A}, \mathcal{B} \rangle$, and $f_{ij}$, then: $\Sigma \models i : A(a) \Longrightarrow \Sigma \models j : B(f_{ij}(b))$. Specifically, $f_{ij}(x)$, in case an individual correspondence $\langle i : x \overset{=}{\mapsto} j : y \rangle$ exists, the individual $x$ is mapped to $y$, else injects a new individual $f_{ij}(x)$ and asserts the respective correspondence.

The representation framework $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$ combines CIC and bridge rules, with link-properties and link assertions. These constructors interact to derive new knowledge. The instance retrieval algorithm presented in [10] can be applied only on those logics where the original ABox can be partitioned into a set of separate ABoxes, where the properties of any individual are specified locally. Obviously, the presence of link assertions do not allow ABoxes to be independently processed.

Similarly to DDL, given a set of bridge rules $\{i : A \overset{\sqsupseteq}{\rightarrow} j : G,\ i : B_k \overset{\sqsubseteq}{\rightarrow} j : H_k\}$ for $1 \leq k \leq n$, $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$ can propagate subsumption relations from $i$-th to $j$-th module: $\Sigma \models_d i : A \sqsubseteq \bigsqcup_{k=1}^{n} B_k \Rightarrow \Sigma \models_d j : G \sqsubseteq \bigsqcup_{k=1}^{n} H_k$. But, additionally to DDL, reasoning on the conceptual part of the distributed ontology results to a distributed concept taxonomy that takes into account link-property specifications as well (a specific example is shown below). The method $DTax(C, \mathcal{T}_i)$ computes the set of concepts that are either equal or subsumed by concept $C$ in $i$. The proposed retrieval algorithm for a concept $j : Q$ is as follows:

$InstRetrieve_j(Q)$
  compute the set $DTax(Q, \mathcal{T}_j)$
  retrieve the set of local individuals $S_Q$ of $Q$ w.r.t. the distributed taxonomy
  for every concept $C$ in $DTax(Q, \mathcal{T}_j)$
    for each bridge rule of the form $i : D \overset{\sqsubseteq}{\rightarrow} j : C$
      compute the set of individuals $S_D = InstRetrieve_i(D)$
      for each individual $x$ in $S_D$ compute $S_Q \leftarrow S_Q \cup \{f_{ij}(x)\}$
  return $S_Q$

As an example we consider a distributed knowledge base constructed by the following information:

$\mathcal{T}_j : \{CapableToReason \sqsubseteq \forall hasProcessor.QuadCoreCPU\}$,
$\mathcal{A}_j : \{CapableToReason(myPC), hasProcessor(myPC, i7.sn001)\}$,
$\mathcal{T}_i : \{QuadCoreCPU \sqsubseteq CPU\}$,
$\mathfrak{B}_{ij} : \{i : CPU \overset{\sqsubseteq}{\rightarrow} j : Processor\}$,
$\mathfrak{C}_{ij} : \{i : i7.sn001 \overset{=}{\mapsto} j : i7.myPC\}$.

We want to retrieve all the individuals of the concept $Processor$. The process starts by computing the distributed taxonomy. $DTax_j(Processor, \mathcal{T}_j)$ returns the set $\{Processor\}$. Local reasoning will return the empty set of individuals, and the process will propagate the query to peer $i$ through the bridge rule, invoking $InstRetrieval_i(CPU)$. Peer $i$ will reply with the set of individuals $\{i7.sn001\}$: This is the case, since additionally to DDL, the application of $DTax_j(Processor, \mathcal{T}_j)$ for the specification in $\mathcal{T}_j$ and the assertion *hasProcessor(myPC,i7.sn001)* imply that $QuadCoreCPU(i7.sn001)$, which further implies $CPU(i7.sn001)$ in $i$. The instance $i7.sn001$ is translated by the $f_{ij}$ function to the $i7.myPC$, and returned by the algorithm.

## 6 Discussion

The proposed algorithm for the representation framework $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$ allows reasoning with collective knowledge bases with non-empty ABoxes for the retrieval of individuals. The method inherits from DDL the capability to propagate concept subsumptions across heterogeneous data repositories and in the same time allows object assertions between distinct ontology units. This leads to the conclusion that $E_{HQ^+}^{DDL}$ $\mathcal{SHIQ}$ supports more expressive queries than DDL, while extending $\mathcal{E}$-connections to reasoning with ontologies covering overlapping domains.

Further work on this framework requires investigating the combination of ontology units, with more expressive fragments of DL. The framework is implemented using Pellet 2.2.2 and experimental results will be available upon stabilization of the software. Also, the overall system's performance will be measured for a variety of peer architectures and organizations.

**Acknowledgement.** This research project is being supported by the project "IRAKLITOS II" of the O.P.E.L.L. 2007 - 2013 of the NSRF (2007 - 2013), co-funded by the European Union and National Resources of Greece.

## References

1. Borgida, A., Luciano Serafini, L.: Distributed Description Logics: Assimilating Information from Peer Sources. Data Semantics, 153–184 (2003)
2. Dolby, J., Fokoue, A., Kalyanpur, A., Schonberg, E., Srinivas, K.: Scalable highly expressive reasoner (SHER). Web Semantics 7, 357–361 (2009)
3. Grau, C., Kutz, O.: Modular Ontology Languages Revisited. In: Proc. of SWeCKa 2007 (2007)
4. Grau, B.C., Parsia, B., Sirin, E.: Working with Multiple Ontologies on the Semantic Web. In: McIlraith, S.A., Plexousakis, D., van Harmelen, F. (eds.) ISWC 2004. LNCS, vol. 3298, pp. 620–634. Springer, Heidelberg (2004)
5. Kutz, O., Lutz, C., Wolter, F., Zakharyaschev, M.: E-connections of abstract description systems. Artif. Intell. 156(1), 1–73 (2004)
6. Parsia, B., Grau, C.: Generalized Link Properties for Expressive epsilon-Connections of Description Logics. In: Proc. of AAAI 2005 (2005)
7. Santipantakis, G., Vouros, G.: $E_{HQ^+}^{DDL}\mathcal{SHIQ}$ Representation framework. Technical Report (2011), `http://ai-lab-webserver.aegean.gr/gsant/TechnicalReports.html`
8. Schlicht, A., Stuckenschmidt, H.: MapResolve. In: Rudolph, S., Gutierrez, C. (eds.) RR 2011. LNCS, vol. 6902, pp. 294–299. Springer, Heidelberg (2011)
9. Serafini, L., Borgida, A., Tamilin, A.: Aspects of distributed and modular ontology reasoning. In: Proc. of IJCAI 2005 (2005)
10. Serafini, L., Tamilin, A.: Distributed Instance Retrieval in Heterogeneous Ontologies. In: Proc. of the 2nd Italian Semantic Web Workshop (2005)
11. Vouros, G., et al.: A Semantic Information System for Services and Traded Resources in Grid e-Markets. FGCS 26(7), 916–933 (2010)
12. Wandelt, S., Müller, R.: Sound and Complete SHI Instance Retrieval for 1 Billion ABox Assertions. In: 10th Int. Semantic Web Conf. on Scalable Semantic Web Knowledge Base Systems Workshop (2011)