Annals of the University of Craiova, Mathematics and Computer Science Series
Volume 38(3), 2011, Pages 31–36
ISSN: 1223-6934, Online 2246-9958

An Erotetic Logic Approach of Ontology Query

Cristian Kevorchian

Abstract. In this paper we develop a method for integration knowledge base and it's query at the level of representation language. We mention that the answers are returned in same language. The query is obtained through erotetic component of the language.

2010 Mathematics Subject Classification. Primary 68T27; Secondary 68T30.
Key words and phrases. erotetic logic,ontology query, deductive database,Description Logic.

1. Preliminaires

Erotetic[1] logic is an important theoretical framework for process inquiry modelling. One of these type of models is Hintikka's interrogative model of inquiry which extends his logic of questions and answer to an interrogative model of scientific inquiry based on Larry Laudan's theory of science as a problem-solving task. An alternative to Hintikka-Laudan's model is the inferential erotetic logic(IEL) based on set-of-answer methodology(SAM) related to Hamblin's postulate[1958]:

1. An answer to a question is a statement
2. Knowing what counts as an answer is equivalent to knowing the question
3. The possible answer to a question are an exhaustive set of mutually exclusive possibilities

All semantic approaches to question end answer are focused on following core idea. The notation,$?Q$, used here is borrowed from Groenendijk and Stokhof (1997) and represent the question whether the proposition Q is true or false:

I. A yes/no-question is represented through expression $?Q$, where Q is a proposition contained in the question. When asking the question, the questioner wants to be asserted whether the extension of the proposition is true or false. The answer to the question is either Q or $\neg Q$.

II. A *wh*-question can be represented as $?x_1...x_nQ$ where the variables $x_1...x_n$ are *wh*-phrases in the question and occur free in Q. An answer is given by Q where $a_1, ..., a_n$ are instances of $x_1, ..., x_n$

For example, in a text mining approach [1], the answers finding is done by identifying, for each proposition expressed in the question, matching propositional expression expressed in the document collection representation. Wh-phrases are represented by "wild-cards" that match similar with PROLOG variable unification. On the other hand, knowledge representation and reasoning systems use a formal language to represent the knowledge. The system, based on this language, provide an inferencing mechanism that allows deriving new facts. Knowledge is assumed to be made from a set of general rules(axioms), T-Box(Terminology Box) and known facts in the form of assertions, named A-Box(Assertion Box).The knowledge to be queried is represented

Received June 24, 2011. Revision received August 17, 2011.

[1]erotetic: from Greek $\varepsilon\rho\varpi\tau\eta\sigma\iota\varsigma(question)$

as assertions derived from some source of knowledge. The rule knowledge needed for inferencing is provided by an additional external resource, for example computational ontology such as Cyc, possibly improved with additional inference rules.

In general, it is assumed that both the knowledge base and the questions are available in a representation language and that answers are to be returned in the same language. In these languages the type-abstraction hierarchy allows the system to move very easily between general and specific concepts that are taxonomically related, embodying some of the attractive features of description logics in terms of querying.

2. Ontology Query

According to Gruber's[1992][2] approach ontology knowledge is organized based five kinds of components: classes or concepts(a family of entities within a domain), relations(the interaction between concepts), functions(a special case of relations in which the n-th element of the relationship is unique for the n-1 preceding elements), axioms(used to model sentences that are always true- used in an ontology to constrain values of classes) and instances(represent specific individual elements). The classes and relations are organized in taxonomies.

A knowledge base $\mathcal{K} = (T, A)$ model over DL contains two parts: $TBox$ (TerminologicalBox) and $ABox$ (AssertionalBox). Several assertions about concepts and roles belong to the first-while the second consists of specific facts about a particular object obtained from a concept or a particular pair of objects belonging to a particular role. A query to $\mathcal{K}$ is a lambda expression

$$\lambda \overline{x} P_1 \wedge P_2 \wedge ... \wedge P_m \tag{1}$$

where $P_i (i = 1..m)$ are predicates of form C(x) or R(x,y),concept and respectively relation in T and each x and y appear in $\overline{x} = (x_1, x_2, ...x_n)$. We can rewrite the functional form of query over $\mathcal{K}$ with following form[4]:

$$Q(x_1, x_2, ..., x_m) \leftarrow BodyOfQuery \tag{2}$$

where $BodyOfQuery$ has the following expression:

$$\exists x_{m+1}, ..., \exists x_n : (\bigwedge C(x_i) \wedge (\bigwedge R(x_i, x_j)) \tag{3}$$

in which predicate names are associated both concepts and roles in the DL.Let $\{a_1, a_2, ..., a_n\} \subseteq A$ based on define an associate to Q answers space with n^k elements and general form[5]:

$$\theta_i = \{(x_1, a_{i_1}), (x_2, a_{i_2})...(x_k, a_{i_k})\} \tag{4}$$

in which variables of the query are linked with constant symbols a_{i_j}. We can view this link as a substitution $\theta_i = \{(x_1/a_{i_1}), x_2/a_{i_2})...x_k/a_{i_k})\}$ an answer to query Q and we can rewrite as

$$\mathcal{K} \models BodyOfQuery\theta \tag{5}$$

Assume that DL supports nominals-a concept defined by finite enumeration of its elements. Therefor the answer $\theta_i = \{(x_1/a_{i_1}), x_2/a_{i_2})...x_k/a_{i_k})\}$ is a family of paired elements(variables x_j and nominal $\{a_{i_j}\}$) and

$$\mathcal{K} \models \forall x_1, ..., \forall x_k : (x_1\{a_{i_1}\} \wedge ... \wedge x_k\{a_{i_k}\}) \tag{6}$$

[2]An ontology is a specification of a conceptualization.

Nominals allow us to add the dependencies type inclusion as $\{a_i\} \subseteq \mathcal{C}$ or $\{a_i\} \subseteq \exists\mathcal{R}\{a_j\}$ to $TBox$ which correspond to $ABox$ assertion $\mathcal{C}(a_i), \mathcal{R}(a_i, a_j) \in A$ and $ABox$ is reduced to a nominals collection. Considering all concept names from $TBox$ gather with attached relations obtain a taxonomy. More precisely, the term taxonomy means entities classification focused on terms or objects, in a hierarchical structure according to the sub/super class paradigm. There is only one type of relationship relating these entities, namely the $IS - A - relationship$. If we reduce the types of relationships in an ontology to the type $IS - A$ dedicated to represent the concepts, the ontology will be equivalent to a taxonomy.

This approach involves a rich description for the components $x_i(i = 1..k)$ of the k-tuple answer [8], based on $DBox$(**D**escription**Box**) of $\mathcal{K}$ to replace the $ABox$ as a description base. More precisely a $DBox$, D is a set of assertions $A(a)$ and $R(a, b)$ where $A \in \sum_D(C) \subseteq N_C$, $R \in \sum_D(R) \subseteq N_R$ and $a, b \in \sum_D(I) \subseteq N_I$ is a family of individuals. The signature $\sum_D \subseteq \sum$, where $\sum = (N_I, N_C, N_R)$ is a signature over a DL, where N_I is a family of individuals names, N_C is the family of atomic concepts name and N_R is the family of roles.
A terminological interpretation $\mathfrak{I} = (\Delta^{\mathfrak{I}}, \cdot^{\mathfrak{I}})$ is a model of D, $\mathfrak{I} \models D$ iff $a^{\mathfrak{I}} = a$ for every $DBox$ individual $a \in \sum_I$ and for every concept or role name $P \in \sum_D(P)$ and for every $u \in \Delta^{\mathfrak{I}}$ and respectively $(u, v) \in \Delta^{\mathfrak{I}} \times \Delta^{\mathfrak{I}}$ iff $P(u) \in D$ and respectively $(u, v) \in P^{\mathfrak{I}}, P(u, v) \in D$[2][6]. We can say that every model of D extension attached to $DBox$ predicate, identified by:

$$\sum\nolimits_D(P) = \sum\nolimits_D(C) \bigcup \sum\nolimits_D(R) \tag{7}$$

is given by the content of $DBox$ and is the same in every model. We must mention that the domain $\triangle^{\mathfrak{I}}$ of the model D is not fixed.

We must mention that whenever a concept is used as an parameter for a query, is not possible to have all predefined names, because not all of them are available at the time of the system generate.

Semantic query needs, complete information from sources of knowledge where available information is often incomplete. In order to solve such incomplete reasoning problem, we try embedded default logic into the description logic knowledge base, meanwhile it prioritized the default rules, which preferred more specific default rules over more general ones. Then, an original incomplete query could be transformed into a complete query relative to the extended knowledge base, by checking default satisfiability of complex concept according to the query.

3. Erotetic approach of ontology query

Let $\sum = (\mathcal{F}, \mathcal{P})$ a signature of first order language $\mathcal{L}$, where $\mathcal{F}$ is a countable family of functions symbols and $\mathcal{P}$ is a countable family of predicates symbols. Extends $\mathcal{L}$ with questions forming operator ? and brackets $\{,\}$. The resulted language will call $\mathcal{L}^+$.All well formed formulas over $\mathcal{L}$ are declarative well formulas from $\mathcal{L}^+$(on brief d-wff). A question Q is a d-wff of $\mathcal{L}^+$[7]:

$$\underbrace{?(A_1, A_2, ..., A_n)}_{dQ} \tag{8}$$

According to the first and the second Hamblin's postulate a question is a declarative expression which represents a simple yes-no question of the natural language. dQ is the family of direct answers for question Q. If Q is a yes-no question then this question is a abbreviation of $?\{Q, \neg Q\}$. We must mention that from an atomic formula we can obtain a yes-no question. Regarding this issue a conjunctive question $?|A, B|$ is a abbreviation of $?(A \wedge B, \neg A \wedge B, A \wedge \neg B, \neg A \wedge \neg B)$ Suppose that $\mathcal{L}^+$ fulfill the following conditions [9]:

1. $D_{\mathcal{L}^+}$ is the set of all $d - wffs$ from $\mathcal{L}^+$;
2. E_{L^+} is the set of all well-formed erotetic formulas(questions) from $\mathcal{L}^+$, where $D_{L^+} \cap E_{L^+} = \emptyset$;
3. if Q is a question, then there exists a set with least two elements $dQ \subset D_{L^+}$ of direct answer to Q;
4. the declarative part of $\mathcal{L}^+$ will be associated with a semantic capable to define the truth of d-wff.

The inference of questions(erotetic inference)[9] over $\mathcal{L}^+$ which are a syntactic relation of derivability of a question from other question is:

$$Q_1 \xrightarrow{\text{X}} Q_2$$

where for each direct answer A to the question Q_1 $A \bigcup X$ involves the disjunction of all the direct answers to question Q_2. If $X = \emptyset$ then will say that Q_1 implies Q_2

An erotetic query system based on a question Q relative to a family of d-wffs X is a finite tree Ξ is:

I. the nodes of Ξ are elements of $D_{\mathcal{L}^+}$ and $E_{\mathcal{L}^+}$, called d-nodes and respectively, e-nodes[9] ;
II. Q is the root of Ξ
III. each leaf of Ξ is a direct answer of Q
IV. $dQ \bigcap X = \emptyset$
V. each d-node of Ξ ca be:
   1. an element of X, or
   2. a direct answer to an e-node of Ξ different from the root Q, or
   3. involved by d-nodes which precede the d-node in Ξ
VI. for each e-node $\bar{Q} \in \Xi$, $Q \neq \bar{Q}$, where Q is root:
   1. $d\bar{Q} \neq dQ$ and
   2. $\bar{\bar{Q}} \rightarrow \bar{Q}$ where $\bar{\bar{Q}}$ precedes the e-node $\bar{Q}$ or
   3. $\bar{\bar{Q}} \xrightarrow{\text{X}} \bar{Q}$ where $X = Q_1, Q_2, ..., Q_n$ and for some e-nodes $\bar{\bar{Q}}$ and for some d-nodes from $X \subset \Xi$ precedes $\bar{Q}$ in Xi
VII. every d-node has an immediate successor;
VIII. every immediate successor of an e-node different from the root Q is either a direct answer to the e-node, or an e-node. if the immediate successor of an e-node $\bar{Q}$ is not an e-node,then every direct answer is an immediate successor.

A query over a query system is an e-node $\bar{Q} \in \Xi$, where $\bar{Q} \neq Q$, where Q is the Ξ root and immediate successors of $\bar{Q}$ are the direct answers of it.

The orchestration between declarative and erotetic inferential systems provide to IEL [3] an important argumentation potential, comparable with abduction inference.

4. Query Erotetic Framework

A Query Erotetic Framework(QEF) is an extension of first-order logic with object-oriented modeling and support for query. A molecule of QEF is: an "$IS - A$" of the form $C : D$,a "SUBCLASS-OF" assertion of the form $C :: D$ and an eroteic object $(?\epsilon)K(\epsilon)$, which mean that an objects satisfy K condition,a data molecule of the form

[3] Inferential Erotetic Logic

$C[D \twoheadrightarrow E]$, with C,D,E terms. All QEF molecules are ground(it does not contain variables).

An $QFE-structure$ is a tuple $Q = \langle U, \prec_U, \in_U, I_F, I_P, I_{\twoheadrightarrow}, I_? \rangle$, where U is a countable nonempty set, $\prec_U$ is an irreflexive partial order on the domain U and $\in_U$ is a binary relation over U. We have $x \in_U a$ and $a \preceq_U b$ then $x \in_U c$. An n-ary function symbol f is interpreted as a function over domain U: $\mathbf{I}_F(f) : U^n \rightarrow U$. An n-ary predicate symbol $p \in P$ is interpreted as a relation over the domain $U : \mathbf{I}_P(p) \subseteq U^n$ and $I_{\twoheadrightarrow}$ associate to each element of U a partial function $U \rightarrow 2^U$.

A question $Q \in U_?$ is QEF-satisfaction iff:

$$\exists a \in dQ \text{ such us } I \models a$$

where I is the terminological interpretation over QEF.

A query predicate-based ontology language is a first order language in which unary predicates represent classes, n-ary predicates represent properties-relations between objects and over it works an erotetic calculus[5]. Descriptions Logic is a query predicate-based ontology with A a concept identifier, "C", "C'" descriptions, R,R' are role identifiers, $o_1, o_2, ..., o_n$ individuals identifiers and n positive integer. The family of the concepts,role and individual identifiers are disjoint:

$$C, C' \rightarrow A|\top|\bot|C \sqcap C'|C \sqcup C'|\neg C|\{o_1, o_2, ..., o_n\}|\exists R.C|\forall R.C|?\{o_1, o_2, ..., o_n\}|$$

The associate ontology involves the existence of an axioms set with following form:

$$S \rightarrow C \sqsubseteq C' \mid C \equiv C' \mid R \sqsubseteq R' \mid R \equiv R' \mid Trans(R) \mid o_1 \in C \mid < o_1, o_2 > \in R \mid o_1 = o_2 \mid o_1 \neq o_2 \mid ?\{o_1, o_2\}$$

If $o_1, o_2, ..o_n \in C$ then $?o_1, o_2, ..o_n$ is a question more then an arbitrary individual $o \in C$ $(?\forall o)C$,$(?\exists o)C$ are questions.

An Erotetic Query Framework(EQF) is a functional extension of first-order logic which adds explicit support for query. In this section we define a translation from query predicate-based ontology to EQF. Let $\mathcal{L}$ first-order language over the signature $\sum = (\mathcal{F}, \mathcal{P})$ and $\mathcal{L}^+$ be an erotetic language with the same signature. If a first-order theory $\Phi \subseteq \mathcal{L}$ involves $\Phi \subseteq \mathcal{L}^+$ is the corresponding erotetic theory.

Entity	Translation
Class	$\delta(A(t) = t : A$
Property	$\delta(R(t_1, t_2)) = t_1[R \rightarrow t_2]$
Equality	$\delta(t_1 = t_2) = t_1 = t_2)$
n-ary predicate	$\delta(P(\vec{t})) = (\vec{t})$
Universal	$\delta(\forall \vec{x}.f) = \forall \vec{x} \delta(f))$
Existential	$\delta(\exists \vec{x}.f) = \exists(\vec{x} \delta(f))$
Conjunction	$\delta(u \wedge v) = \delta(u) \wedge \delta(v)$
Disjunction	$\delta(u \vee v) = \delta(u) \vee \delta(v)$
Implication	$\delta(u \supset v) = (\delta(u) \supset \delta(v))$
Erotetic implication	$\delta(q_1 \xrightarrow{X} q_2) = \delta(q_1) \xrightarrow{Y} \delta(q_2)$,where $Y = \delta(X)$
Negation	$\delta(\neg u) = \neg \delta(u)$

TABLE 1. Translation of query predicate-based to EQF

An important issue regarding EQF is the direct answer capacity of the framework. Let $Q_1, Q_2 \in EQF_?$ will say that Q_1 is stronger $Q_2(Q_1 \succeq Q_2)$ iff there is a surjection $j : dQ_1 \rightarrow dQ_2$ such that for $a \in dQ_1$, $a \models j(a)$ [7]. If j is bijection $dQ_1 \rightarrow dQ_2$ then $a = j(a)$. A partial answer [7] can be obtained when a declarative d of EQF and a $D \subset dQ$ such that $\| = D$, where $\| = D$ is *multiple conclusion entailment*[7]

5. Conclusions

Starting from applying erotetic logic to build useful analysis of SPARQL and its applicatios by Simon Raboczi. Our propose is to obtain a formal framework which defines an ontology query based on erotetic logic functionalities by including the query concept in the class hierarchy of DL and allowing partial integration of other concepts. Hypothetical answers to not arise in the description logic framework.

References

[1] L. Dehaspe and H. Toivonen Discovery of DATALOG patterns, *Data Mining and Knowledge Discovery* **3** (1999), 7-36.
[2] H. Decker, A Case for Paraconsistent Logic as a Foundation of Future Information Systems *CAISE '05*, Porto, Portugal, 2005, 451–461.
[3] S. Klarman, U. Endriss and S. Schlobach, ABox Abduction in the Description Logic ALC, *Journal of Automated Reasoning*, February 2010.
[4] D. Lembo and M. Ruzzi, Consistent Query Answering over Description Logic Ontologies, *Web Reasoning and Rule Systems First International Conference*, Innsbruck, Austria, June 7-8, (2007), 194–208.
[5] D. Leszczyńska, *The Method of Socratic Proofs for Normal Modal Propositional Logic*, Instytut Filozofii, Zielona Gra, 2006.
[6] J. Mei, L. Ma and Y. Pan, Ontology Query Answering on Databases, *The Semantic Web - ISWC 2006 5th International Semantic Web Conference Proceedings* Athens, GA, USA, November 5-9, 2006, Springer Berlin, Heidelberg, 2006.
[7] M. *Peliš*, *Consequence Relations in Inferential Erotetic Logic*, Research Report of the grant Dynamic Formal Systems, 2008.
[8] I. Seylan, E. Franconi and J. de Bruijn, Effective Query Rewriting with Ontologies over DBoxes, *Proceedings of the 21st International Joint Conference on Artifical Intelligence*, Pasadena, California, USA, 2009, 923–929.
[9] *M.Urbański* and *P.Lupkowski*, Erotetic Search Scenarios: Revealing Interrogator's Hidden Agenda, SemDial 2010, 14th Workshop on the Semantics and Pragmatics of Dialogue, Poznan, Poland, 2010.
[10] *A.Wiśniewski*, *Erotetic Logics. An Introduction to Inferential Erotetic Logic*, UNILOG03, Lisbon, April 2010.

(Cristian Kevorchian) IT Department, Legislative Council, 13 Septembrie Street 1-3, Bucuresti, 71000, Romania
E-mail address: cristi@clr.ro