Intensional logic and epistemic independency of intelligent database agents

Zoran Majkić

Dipartimento di Informatica e Sistemistica, University of Roma "La Sapienza"
Via Salaria 113, I-00198 Rome, Italy
majkic@dis.uniroma1.it

Abstract. In the typical Web applications each intelligent database agent can be defined as a Knowledge system (KS) with a global ontology, which integrates a number of source data distributed in Web by traditional extensional mappings, and must be robust enough in order to take in account the incomplete and locally inconsistent information of its sources. The traditional extensional semantics for mappings between the different KSs destroys the epistemic independence of KSs: the beliefs of other KSs are forced into a local knowledge of a given KS, so that its own belief depends directly and automatically from them. Actually we want to find a kind of semantics for external mappings between KSs which is less strong w.r.t. the internal KS's (extensionally based) database mappings. These philosophical considerations motivate the need of a new, alternative semantic characterization, based not on the extension but on the *meaning* of concepts used in the mappings between KSs.

The Cooperative Information Systems has no centralized schema and no central administration. Instead, each intelligent database agent (KS) is an autonomous information system, and information integration is achieved by establishing mappings among various ontologies of these independent KSs. Given the de-centralized nature of the development of the Semantic Web, there will be an explosion in the number of ontologies. Many of these ontologies (that is, KSs) will describe similar domains, but using different terminologies, and others will have overlapping domains. To integrate data from disparate ontologies, we must know the *semantic correspondence* between their elements. Recently are given a number of different architecture solutions [1,2,3,4].
Queries are posed to one KS, and the role of query processing is to exploit both the data that are internal to the KS, and the mappings with other KSs in the system.
In this paper we investigate on the possibility of using the *intensional logic* for both expressing interschema (inter-ontology) knowledge, and reasoning about it. The basic idea of our approach is to propose an intensional logic-based language to express interdependencies between concepts (views defined as conjunctive queries) belonging to different schemas (KS's ontologies). For example, one can assert in our language that the concept represented by the view `GraduateStudent` in the schema S_1 is the same as the concept represented by the view `SeniorStudent` in S_2. Such assertion implies a sort of *intensional equivalence* between the two concepts, but does not imply that the extension (the set of instances) of the former is always the same as the extension of the later.

The existing research papers in the literature share our general goal of representing and using interschema knowledge (for an exhaustive consideration consider [5]), but their approaches does not guarantee the complete epistemic independencies between different KSs.
Let P_i and P_j be the two KSs, denominated by 'Peter' and 'John' respectively, and $q_1(\mathbf{x})$, $q_2(\mathbf{x})$ be the concepts of "the Italian art in the 15'th century" with attributes in $\mathbf{x}$, written in local languages of P_i and P_j respectively. We are able to individuate at least two extreme scenarios, developed from the initial article [6] :
1. The *strongly-coupled* semantics [3] for mappings between different KSs is a direct extension of *extensionally based* database mappings between views of KSs [5] used for a (strong) data integration systems: For any given KS its own knowledge is locally enlarged by extensional knowledge of other KSs: any dynamic change of the knowledge of other KSs is directly reflected into the local knowledge of this KS. As showed in [3], the added knowledge of other KSs is seen as some kind of local 'source' database of data-integration system of a given KS. We can paraphrase this by *imperative* assertion '*John must know all facts about the Italian art in the 15'th century known by Peter*' (also when 'Peter' in his life cycle changes this part of its own knowledge), formally
$K_i q_1(\mathbf{x}) \;\Rightarrow\; K_j q_2(\mathbf{x})$, where $'\Rightarrow'$ is the logic implication.
2. The *weakly-coupled* semantics [7,4]. At a very beginning was my intuition that the real cooperative information systems, where each KS is completely independent entity, with its own epistemic state, which has not to be directly, externally, changed by the mutable knowledge of other independent KSs, needs other *meaning* (approach) to the mapping between their local knowledge. First requirement is that the knowledge of other KSs can not be directly transferred into the local knowledge of a given peer. The second requirement is that, during the life time of a cooperative information system, any local change of knowledge must be independent of the beliefs that can have other KSs: thus, we have not to constrain the *extension* of knowledge which may have different KSs about the same type of real-world concept.
In the example above, 'John' can answer only for a part of knowledge that it really has about Italian art, and not for a knowledge that 'Peter' has. Thus, when somebody (call him 'query-agent') ask 'John' some information about Italian art in the 15'th century, 'John' is able to respond only by facts known by himself (i.e., *certain* answers), and eventually indicate to query-agent that for such question probably 'Peter' is able to give some answer also: so, it is the task of the *query-agent* to reformulate the question (w.r.t. the local language of 'Peter') to 'Peter' in order to obtain some other *possible* answers. We can paraphrase this by the kind of *belief-sentence*-mapping '*John believes that also Peter knows something about Italian art in the 15'th century*', formally
$K_i q_1(\mathbf{x}) \;=_{in}\; K_j q_2(\mathbf{x})$, where $' =_{in}'$ is the believed intensional equivalence.
Such belief-sentence has *referential* (i.e., extensional) *opacity*. In this case we do not specify that the knowledge of 'John' is included in the knowledge of 'Peter' (or viceversa) for the concept 'Italian art in the 15'th century', but only that this concept, $q_2(\mathbf{x})$, for 'John' *implicitly corresponds* to the 'equivalent' concept, $q_1(\mathbf{x})$, for 'Peter'. The 'implicit correspondence between *equivalent* concepts' needs a formal semantic definition for it. It was not easy task, because the mapping defined above deals with the semantics of *natural* language. Motague [8] defined the *intension of a sentence* as a

function from possible worlds to truth values.
In what follows we will use one simplified modal logic framework (we will not consider the time as one independent parameter as in Montague's original work) with a model $\mathcal{M} = (\mathcal{W}, \mathcal{R}, S, V)$, where $\mathcal{W}$ is the set of possible worlds, $\mathcal{R}$ is the accessibility relation between worlds ($\mathcal{R} \subseteq \mathcal{W} \times \mathcal{W}$), S is a non-empty domain of individuals, while V is a function defined for the following two cases:
1. $V : \mathcal{W} \times F \rightarrow \bigcup_{n<\omega} S^{S^n}$, with F a set of functional symbols of the language, such that for any world $w \in \mathcal{W}$ and a functional symbol $f \in F$, we obtain a function $V(w, f) : S^{arity(f)} \rightarrow S$.
2. $V : \mathcal{W} \times P \rightarrow \bigcup_{n<\omega} \mathbf{2}^{S^n}$, with P a set of predicate symbols of the language and $\mathbf{2} = \{t, f\}$ is the set of truth values (true and false, respectively), such that for any world $w \in \mathcal{W}$ and a predicate symbol $p \in P$, we obtain a function $V(w, p) : S^{arity(p)} \rightarrow \mathbf{2}$, which defines the extension $[p] = \{\mathbf{a} | \mathbf{a} \in S^{arity(p)}$ and $V(w, p)(\mathbf{a}) = t\}$ of this predicate p in the world w.
The extension of an expression α, w.r.t. a model $\mathcal{M}$, a world $w \in \mathcal{W}$ and assignment g is denoted by $[\alpha]^{\mathcal{M},w,g}$. Thus, if $c \in F \bigcup P$ then for a given world $w \in \mathcal{W}$ and the assignment function for variables g, $[c]^{\mathcal{M},w,g} = V(w, c)$, while for any formula A, $\mathcal{M} \vDash_{w,g} A \equiv ([A]^{\mathcal{M},w,g} = t)$, means 'A is true in the world w of a model $\mathcal{M}$ for assignment g'. Montague defined the *intension* of an expression α as follows:
$[\alpha]_{in}^{\mathcal{M},g} =_{def} \{w \mapsto [\alpha]^{\mathcal{M},w,g} \mid w \in \mathcal{W}\}$,
i.e., as graph of the function $[\alpha]_{in}^{\mathcal{M},g} : \mathcal{W} \rightarrow \bigcup_{w \in \mathcal{W}_N} [\alpha]^{\mathcal{M},w,g}$.
One thing that should be immediately clear is that intensions are more general that extensions: if the intension of an expression is given, one can determine its extension with respect to a particular world but not viceversa, i.e., $[\alpha]^{\mathcal{M},w,g} = [\alpha]_{in}^{\mathcal{M},g}(w)$.
In particular, if c is a non-logical constant (individual constant or predicate symbol), the definition of the extension of c is, $[c]^{\mathcal{M},w,g} =_{def} V(w, c)$. Hence, the intensions of the non-logical constants are the following functions: $[c]_{in}^{\mathcal{M},g} : \mathcal{W} \rightarrow \bigcup_{w \in \mathcal{W}} V(w, c)$.
The extension of variable is supplied by the value assignment g only, and thus does not differ from one world to the other; if x is a variable we have $[x]_{in}^{\mathcal{M},g} = g(x)$.
Carnap suggested that the intension of an expression is nothing more than all the varying extensions the expression can have. In the next we will take this definition in order to define that two expressions (or concepts) α, β are *intensionally equivalent*, in the following two cases:

Definition 1. *Any two expressions, α, β, are intensionally equivalent (in the* flat-accumulation *or the* world-correspondent *case, respectively) denoted by $\alpha =_{in} \beta$, if and only if :*
1. flat-accumulation case: $lub^{\mathcal{M},g}(\alpha) = lub^{\mathcal{M},g}(\beta)$, where for a given expression δ, its lub (Least Upper Bound) is defined by: $lub^{\mathcal{M},g}(\delta) =_{def} \bigcup_{w \in \mathcal{W}} [\delta]_{in}^{\mathcal{M},g}(w)$.
2. world-correspondent case: $\forall w \exists w'.([\alpha]_{in}^{\mathcal{M},g}(w) = [\beta]_{in}^{\mathcal{M},g}(w'))$,
and viceversa, $\forall w' \exists w.([\alpha]_{in}^{\mathcal{M},g}(w) = [\beta]_{in}^{\mathcal{M},g}(w'))$.

In the context of this work we will consider each temporary instance (in a some time t_k) of the cooperative information system as a particular possible world w: the dynamic changes of any local KS knowledge will result in one other possible world. The intensional mapping between KSs is given by couples of queries $(q_i(x), q_j(x))$ where a conjunctive query $q_i(x)$ over a KS P_i and a conjunctive query $q_j(x)$ over a peer P_j

are both intensionally equivalent to same real-world entity α, w.r.t. the *certain* answers from KSs (we consider that each KS P_i is an epistemic local logic theory with the modal epistemic operator K_i, so that the truth of a modal formula $K_i q_i(x)$ corresponds to the set of certain answers to the conjunctive query $q_i(x)$ only), i.e., $K_i q_i(x) =_{in} \alpha$ and $K_j q_j(x) =_{in} \alpha$, thus, by the symmetry and the transitivity of the relation $=_{in}$, we obtain that holds $K_i q_i(x) =_{in} K_j q_j(x)$.
Notice that for any given world w, both relationships
$[K_i q_i(x)]_{in}^{\mathcal{M},g}(w) \subseteq [K_j q_j(x)]_{in}^{\mathcal{M},g}(w)$, and $[K_j q_j(x)]_{in}^{\mathcal{M},g}(w) \subseteq [K_i q_i(x)]_{in}^{\mathcal{M},g}(w)$ need not to be satisfied. Moreover, if Δ_i and Δ_j are local universes for a KS P_i and P_j respectively (a local universe is the set of all the values that are elements of the domains used in the local schema of a KS), we do not require that for any $c \in \Delta_i \bigcap \Delta_j$, the sentences $K_i q_i(c)$ and $K_j q_j(c)$ have the same truth value as required in [5].

Proposition 1 *Let consider the class of KSs with integrity constraints which does not contain negative clauses of the form* $\neg A_1 \vee ... \vee \neg A_m,\ m \geq 2$. *Then, the intensional equivalence is preserved by conjunction logic operation, that is,*
if $\varphi \equiv (b_1 \wedge ... \wedge b_k)$, $k \geq 1$, *and* $b_i =_{in} c_i,\ 1 \leq i \leq k$, *then* $\varphi =_{in} \psi$
where $\equiv$ *is a logic equivalence and* $\psi \equiv (c_1 \wedge .... \wedge c_k)$.

Thus, for any given conjunctive query (virtual concept) to some intelligent database agent, the query-agent will obtain as answer the set of certain (known) answers from this interrogated database agent, and the set of possible answers from other database agents which are able to express the intensionally equivalent virtual concepts to the original user query.
We believe that the intensional mapping semantics presented in this paper constitutes a sound basis for studying the various issues related to interschema knowledge representation and reasoning, especially for P2P database systems in Web environment, where peers can be considered as complex database agents.

References

1. S.Gribble, A.Halevy, Z.Ives, M.Rodrig, and D.Suciu, "What can databases do for peer-to-peer?," *WebDB Workshop on Databases and the Web*, 2001.
2. L.Serafini, F.Giunchiglia, J.Mylopoulos, and P.A.Bernstein, "The local relational model: Model and proof theory," *Technical Report 0112-23, ITC-IRST*, 2001.
3. D.Calvanese, G. De Giacomo, M.Lenzerini, and R.Rosati, "Logical foundations of peer-to-peer data integration," *PODS 2004, June 14-16, Paris, France*, 2004.
4. Z. Majkić, "Weakly-coupled ontology integration of p2p database systems," *1st Int. Workshop on Peer-to-Peer Knowledge Management (P2PKM), August 22, Boston, USA*, 2004.
5. Tiziana Catarci and Maurizio Lenzerini, "Representing and using interschema knowledge in cooperative information systems," *J. of Intelligent and Cooperative Information Systems*, vol. 2, no. 4, pp. 375–398, 1993.
6. M.Lenzerini and Z. Majkić, "General framework for query reformulation," *Semantic Webs and Agents in Integrated Economies, D3.1, IST-2001-34825, February,*, 2003.
7. Z. Majkić, "Massive parallelism for query answering in weakly integrated p2p systems," *Workshop GLOBE 04, August 30-September 3,Zaragoza, Spain*, 2004.
8. R.Montague, "Formal philosophy. selected papers of richard montague," *in R.Thomason (editor), Yale University Press, New Haven, London*, pp. 108–221, 1974.