A Doxastic Approach to P2P Information Integration

Guido Vetere, Fabrizio Venditti, Alessandro Faraotti

Center for Advanced Studies
IBM Italy

19-05-2008 / Logic In Databases 2008
Outline

1. Motivation
2. Epistemic Approaches to P2P Integration
   - Epistemic Approaches to P2P Integration
3. A Doxastic Approach
   - A Doxastic Approach
   - Knowledge Construction
Motivation

The European project Workpad researches a peer-to-peer data integration system to support crisis management.

Integration takes place at response time, so no much time is alloted to verify / review information sources.

The system should allow peers to take majorities / preferences into account.

Need to modify/extend current approaches.
Motivation

Epistemic Approaches to P2P Integration

A Doxastic Approach

Conclusion

Outline

1 Motivation

2 Epistemic Approaches to P2P Integration
   • Epistemic Approaches to P2P Integration

3 A Doxastic Approach
   • A Doxastic Approach
   • Knowledge Construction
Query Answering in data integration systems based on FOL semantics leads to undecidability [CalvaneseEtAl,2004]

- Query answering is undecidable in presence of cycles in network topology

Need of resorting on epistemic logics

Using Epistemic Logics

- Peers as information agents reasoning about knowledge of facts with respect to the possible accessible worlds.
- Knowledge of a peer as truth in every possible world it can access.
Knowledge Transfer

- Peer to peer integration involves transfer of knowledge from the network to the peer where queries are issued.

- P2P Mappings semantics
  - using S5 axiomatization [CalvaneseEtAl,2004]
    \[
    \forall x (K(\exists y (cq_j(x, y))) \rightarrow K(\exists z (cq_i(x, z))))
    \] (1)
  - using $K45_n$ axiomatization [CalvaneseEtAl,2005]
    \[
    \forall x (K_j(\exists y (cq_j(x, y))) \rightarrow K_i(\exists z (cq_i(x, z))))
    \] (2)
  - implies direct knowledge transfer among peers (other peers’ beliefs are directly transferred to the mapped peer)

- The answer to a query $q(x)$ is the set of Certain epistemic answers provided by the inquired peer
  \[
  ANSW(q, i) = \{ \bar{t} \mid T_K(P) \models_{K45_n} K_i q(\bar{t}) \} \] (3)
Knowledge Routing

- Knowledge of a peer $p_i$ is related to the knowledge of peer $p_j$ without altering the epistemic state. [Majkic, 2006]
- P2P Mappings have the following semantic

$$\forall x \left( \mathcal{K}_j(\exists y \left( cq_j(x, y) \right)) \approx_{in} \mathcal{K}_i(\exists z \left( cq_i(x, z) \right)) \right)$$

(intensional equivalence)

- The answer to a query $q(x)$ over a peer $p_i$ is the union of the answers to the set of all queries intensionally equivalent to $q(x)$
- An epistemological distinction is made between certain answers of the inquired peer $p_i$ and possible answers coming from the other peers.
- The peer doesn’t keep control over returned information (e.g. filtering unreliable sources in presence of inconsistency)
- Effective information integration steps cannot be modeled within this framework
Outline

1. Motivation
2. Epistemic Approaches to P2P Integration
   - Epistemic Approaches to P2P Integration
3. A Doxastic Approach
   - A Doxastic Approach
   - Knowledge Construction
A Doxastic Approach

A new Doxastic approach to:

- Model P2P integration in a descriptive way so to capture essential features of p2p internetworking as it is in real systems
- Overcome some difficulties of merging different knowledge bases, typical of epistemic systems, without limiting the role that each peer can play as integration system

Distinction between knowledge of facts and doxastic knowledge

For each peer:

- Knowledge of facts: local knowledge coming from environments under peer’s control (e.g. databases)
- Doxastic Knowledge: knowledge of information coming from other peers

A variety of merging strategies can be implemented to construct local knowledge.
A Doxastic Approach

- Resort on $KD45_n$ multi-modal epistemic logic, in particular doesn't hold the axiom:

\[ A3': \mathcal{K}_i\phi \rightarrow \phi \]  

(5)

- as a consequence there is **not** knowledge transfer:

\[ \mathcal{K}_i\mathcal{K}_j\phi \rightarrow \mathcal{K}_i\phi \]  

(6)

- Doxastic P2P Mappings follow the semantics:

\[ \forall x (\mathcal{K}_j(\exists y (cq_j(x, y)))) \rightarrow \mathcal{K}_i\mathcal{K}_j(\exists z (cq_i(x, z)))) \]  

(7)

- Doxastic mapping assertions allow to acquire indirect knowledge

- Query answering (same certain epistemic answers of knowledge transfer approach)

\[ ANSW(q, i) = \{ \bar{t} | \mathcal{T}_K(\mathcal{P}) \models_{KD45_n} \mathcal{K}_i q(\bar{t}) \} \]  

(8)

- Semantic isolation: the answer to a query gives same results returned by the inquired peer considered as a standalone system without resorting on p2p mappings

- To avoid isolation Peers have to integrate indirect knowledge into their own knowledge

Knowledge Construction
Outline

1 Motivation

2 Epistemic Approaches to P2P Integration
   - Epistemic Approaches to P2P Integration

3 A Doxastic Approach
   - A Doxastic Approach
   - Knowledge Construction

Guido Vetere, Fabrizio Venditti, Alessandro Faraotti
Knowledge Construction

Knowledge Construction is the process of building direct knowledge from indirect knowledge.

Definition (Indirect knowledge)
Let $\Psi$ be a $KD45_n$-logic formula and let $\phi$ be a first-order formula. We say that $\Psi$ represents indirect knowledge of a peer $p_i$ about $\phi$ (and we denote it with $\Psi_i^\phi$) iff either:

- $\Psi$ is on the form $K_iK_j\phi$ with $1 \leq j \leq |P|$;
- $\Psi$ is on the form $\Psi_1 \land \Psi_2$;
- $\Psi$ is on the form $\Psi_1 \lor \Psi_2$;

where $\Psi_1, \Psi_2$ are such that $\Psi_1^\phi_i$ and $\Psi_2^\phi_i$;

Construction assertions: a set of formulas of the kind

$$\Psi_i^\phi \rightarrow K_i\phi$$

Construction assertions allow to build direct knowledge using a variety of construction strategies.

Guido Vetere, Fabrizio Venditti, Alessandro Faraotti
Example: Knowledge Transfer Strategy
A simple Knowledge Construction assertion for each doxastic mapping assertion of the kind:

$$K_i K_j \phi \rightarrow K_i \phi$$  \hspace{1cm} (9)

Example: Majority Strategy
The peer $p_i$ builds its own knowledge of fact $\phi$ only if the majority of its neighbors directly knows $\phi$

The following assertion models this strategy

$$((K_i K_j \phi \land K_i K_k \phi) \lor (K_i K_j \phi \land K_i K_l \phi) \lor (K_i K_k \phi \land K_i K_l \phi)) \rightarrow K_i \phi$$  \hspace{1cm} (10)

There is not a general rational behavior that every peer is supposed to adopt: every peer can define its own knowledge construction policy

Future work: framework extensions that allow peers to rank their acquaintances based on trust, or to use any other arbitrary knowledge construction and revision strategy

Guido Vetere, Fabrizio Venditti, Alessandro Faraotti
Conclusion

- P2P information integration relevance for semantic integration of distributed systems is crucial.
- Multi-modal logics are at the basis of decidable and tractable integration frameworks.
- Knowledge transfer unable to model arbitrary integration strategies, while Knowledge routing is not suitable for real information integration.
- The doxastic approach
  - Overcomes limitations of current epistemic P2P multi-modal approaches
  - Allows many integration strategies