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

Outline

1 Motivation

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

3 A Doxastic Approach
- A Doxastic Approach
- Knowledge Construction

Epistemic Approaches to P2P Integration

- 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]

$$\forall \overline{x} \ (\mathcal{K}(\exists \overline{y} \ (cq_j(\overline{x}, \overline{y}))) \rightarrow \mathcal{K}(\exists \overline{z}(cq_i(\overline{x}, \overline{z})))) \quad (1)$$

  - using $K45_n$ axiomatization [CalvaneseEtAl,2005]

$$\forall \overline{x} \ (\mathcal{K}_j(\exists \overline{y} \ (cq_j(\overline{x}, \overline{y}))) \rightarrow \mathcal{K}_i(\exists \overline{z}(cq_i(\overline{x}, \overline{z})))) \quad (2)$$

  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

$$ANSW(q, i) = \{\bar{t} \mid \ \mathcal{T}_K(\mathcal{P}) \models_{K45_n} \mathcal{K}_i q(\bar{t})\} \quad (3)$$

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

$$\forall \overline{x} \; (\mathcal{K}_j(\exists \overline{y} \; (cq_j(\overline{x}, \overline{y}))) \approx_{in} \mathcal{K}_i(\exists \overline{z}(cq_i(\overline{x}, \overline{z})))) \tag{4}$$

(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 mergirng 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 enviroments 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 doesnt hold the axiom:

$$\text{A3': } \mathcal{K}_i\phi \rightarrow \phi \tag{5}$$

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

$$\mathcal{K}_i\mathcal{K}_j\phi \rightarrow \mathcal{K}_i\phi \tag{6}$$

- Doxastic P2P Mappings follow the semantics:

$$\forall \overline{x}\ (\mathcal{K}_j(\exists \overline{y}\ (cq_j(\overline{x}, \overline{y}))) \rightarrow \mathcal{K}_i\mathcal{K}_j(\exists \overline{z}(cq_i(\overline{x}, \overline{z})))) \tag{7}$$

- Doxastic mapping assertions allow to acquire indirect knowledge
- Query answering (same certain epistemic answers of knowledge transfer approach)

$$ANSW(q, i) = \{\bar{t} \mid \ \mathcal{T}_K(\mathcal{P}) \models_{KD45_n} \mathcal{K}_i q(\bar{t})\} \tag{8}$$

- 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

Knowledge Construction

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

Definition (Indirect knowledge)

Let Ψ be a $KD45_n$-logic formula and let ϕ be a first-order formula. We say that Ψ *represents indirect knowledge of a peer* p_i *about* ϕ (and we denote it with Ψ_i^{ϕ}) iff either:

- Ψ is on the form $\mathcal{K}_i\mathcal{K}_j\phi$ with $1 \leq j \leq |\mathcal{P}|$;
- Ψ is on the form $\Psi_1 \wedge \Psi_2$;
- Ψ is on the form $\Psi_1 \vee \Psi_2$;

where Ψ_1, Ψ_2 are such that ${\Psi_1}_i^{\phi}$ and ${\Psi_2}_i^{\phi}$;

- Construction assertions: a set of formulas of the kind

$$\Psi_i^{\phi} \rightarrow \mathcal{K}_i\phi$$

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

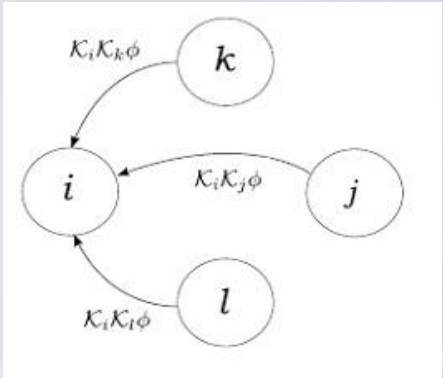
Knowledge Construction

- Example: Knowledge Transfer Strategy
  A simple Knowledge Construction assertion for each doxastic mapping assertion of the kind:

$$\mathcal{K}_i\mathcal{K}_j\phi \rightarrow \mathcal{K}_i\phi \tag{9}$$

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

  The following assertion models this strategy



$$((\mathcal{K}_i\mathcal{K}_j\phi \wedge \mathcal{K}_i\mathcal{K}_k\phi) \vee (\mathcal{K}_i\mathcal{K}_j\phi \wedge \mathcal{K}_i\mathcal{K}_l\phi) \vee (\mathcal{K}_i\mathcal{K}_k\phi \wedge \mathcal{K}_i\mathcal{K}_l\phi)) \rightarrow \mathcal{K}_i\phi \tag{10}$$

- 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

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