4.2 Knowledge-enriched Data Management

Diego Calvanese (Free University of Bozen-Bolzano, IT)

(*Moderator:* Diego Calvanese. *Participants:* Andreas Pieris, Carsten Lutz, Claire David, Filip Murlak, Georg Gottlob, Magdalena Ortiz, Marcelo Arenas, Meghyn Bienvenu, Paolo Guagliardo, Reinhard Pichler, Thomas Eiter, Jianwen Su, Ron Fagin, Sudeepa Roy).

The working group identified four important practical challenges:

1. Develop personalized and context-aware data access and management tools. Data in this case is highly heterogenous, multi-model and multi-modal. Here we deal with "small data" at the individual level, tuned to different view points.
2. Providing end users flexible and integrated access to large amounts of complex, distributed, heterogenous data (under different representations and different models). End users should be assumed to be domain experts, not data management experts.
3. Ensure interoperability at the level of systems exchanging data.
4. Bring knowledge to data analytics and data extraction.

It also identified seven relevant theoretical challenges related to the latter:

1. Development of reasoning-tuned DB systems, including new optimizations, new cost models, new/improved database engines optimized for reasoning, approximate answers, distributed evaluation, etc.
2. Choosing/designing the right languages for supporting these tasks. Here we need pragmatic choices motivated by user needs, but also supporting different types of knowledge and data (e.g., mixing CWA+OWA, temporal, spatial, etc.)
3. We need new measures of complexity for understanding easy/difficult cases, that explain better what works in practice. It would be interesting to explore alternative complexity measures, such as parameterized and average complexity, measuring complexity on the Web, smoothed analysis, etc.
4. Building user-friendly interfaces (beyond Protege). In particular, we need tools support geared towards end users (i.e., domain experts, lay people, and not just IT/knowledge engineers), but also support for query formulation, tools for exploration, etc.
5. Developing next-generation reasoning services. Here we need to explore notions related to explanation, abduction, hypothetical reasoning, defeasible reasoning, etc.
6. Reasoning with imperfect data, e.g., reasoning under contradictions and/or uncertainty, reasoning about quality of data, and support for improving data quality.
7. In depth study of temporal and dynamic aspects, such as managing changing data and knowledge, streaming data, reasoning about change, updating data in the presence of knowledge, etc.