CoDeR Model for Continuous Reasoning over OWL 2 RL Ontologies

David Monks

Department of Electronics and Computer Science, University of Southampton, United Kingdom
dm11g08@ecs.soton.ac.uk

1 Expression of Interest

I'm a final year PhD student investigating continuous reasoning over streamed RDF data with the expressivity of the OWL 2 RL profile. This research has expanded upon earlier research that I undertook that involved adapting research into continuous operators of the stream-to-stream algebra for use in continuous reasoning over streamed RDF data. This continuous reasoning was implemented as a multi-threaded production system that would be generated according to Forgy's Rete pattern matching algorithm, and could express any Datalog program by treating unary predicates as class membership statements, binary predicates as RDF triples and *n*-ary predicates as RDF graphs consisting of a blank node, a single incoming RDF triple and *n-1* outgoing RDF triples.

My current research builds upon and broadens the scope of my investigation to the general case of distributed continuous reasoning over streamed RDF data, considering reasoning expressed in both OWL 2 and RIF Core, as well as what continuous windowing means in that context. This has involved a more in depth study of those forms of reasoning currently supported by Semantic Web technologies and considering the effects and requirements of real-time processing on such reasoning tasks, which I believe to be an under-investigated area. My results to date have included a model for distributed continuous stream reasoning, which is what I would like to present in the **research session**. This model draws its abstract elements from both the family of database models originating with the relational model (most recently adapted for stream processing in the stream-to-stream algebra) and the family of Eddy-based models for long-lived querying of federated databases/data streams (most recently adapted for semantic data in CQELS). It unifies these models, making use of similarities in the operator behaviours to describe both within a single operator set, then extends the model with an environment in which results from operator plans may be returned to the system as input, thereby supporting rule-based reasoning to the extent proposed by many semantic stream querying offerings such as C-SPARQL and EP-SPARQL. However, the primary focus of my investigations has been the temporally localised nature of both input streamed data and, crucially, streamed entailments: the processing delay inherent in the production of the latter requires an alternative to the query-centric windowing of stream querying engines such as C-SPARQL and CQELS, else inference cascades and their entailments may remain in the system past the point in time at which the original input data loses relevance according to the reasoning task at hand (this assumes the ability to specify windows of relevance within the TBox of an ontology, or over a set of inference rules).

With my focus on streamed RDF data and the impact of temporal relevance implied thereby on the process of continuous expressive ABox reasoning, I believe that my research is closely related to the interests of the RSP community. In addition, as an investigation into the technologies underpinning RDF stream processing in the general case, my work will be relevant to any in the wider ESWC audience who are interested in reasoning over RDF data in real-time. However, I recognise that the ongoing broadening of the ESWC agenda to include research into the interaction between society and the Semantic Web will leave the specific details of my research outside the scope of many researchers in that particular track. Nevertheless, even those that do not have an interest in the detail may be interested in the decisions made regarding handling of time in the TBoxes/rule-sets, as well as the management of temporally sensitive ABox data.