VLDB 2010
The 36th International Conference on Very Large Data Bases
13 to 17 September 2010

Singapore

Reviews For Paper

Track	Demonstrations
Paper ID	905
Title	SSQ: Querying Entities Using Shallow Semantics

Masked Reviewer ID: Assigned_Reviewer_1

Review:

Question	
Detailed Comments	This paper describes a prototype system for querying entities that appear in web pages. Several interesting ideas and techniques are implemented in this system. One is that it considers queries that not only can have conditions on entities but also can have predicates on relationships between entities. All conditions/predicates are keyword based. Shallow semantics can be captured by this type of queries. Another novel idea is the entity-centric index, which can speed up the search better than document-based inverted index. I am not aware of similar approaches in published literature. The paper is well written and the demo scenario is carefully described. The described method is simplistic in evaluating query conditions and mistakes can be made easily. For example, using the current method, the condition that “X is a Stanford graduate” would also be satisfied by “a partner of X is a Stanford graduate”, incorrectly. As another example, synonyms are not considered, which will make “X is created by Y” not satisfy “X is founded by Y” even though they may have the same meaning. Discussion on how to address these issues should be provided. There are some typos. Page 3, “Wikiepdia” should “Wikipedia”; not sure what “tupe” is.
Overall Rating	Accept

Masked Reviewer ID: Assigned_Reviewer_2

Review:

Question	
Detailed Comments	This paper demonstrates a shallow semantic query system. Differently from a semantic query system, a shallow semantic query system does not have an RDF knowledgebase. It gets semantic information directly from the indexed documents, it thus has its own advantage. The system should be improved. Many queries work well. However for

	some queries with large number of answers returned, the system responses after a long period which gives a bad user experience.
Overall Rating	Neutral

Masked Reviewer ID: Assigned_Reviewer_3

Review:

Question	
Detailed Comments	The paper describes a system for answering typed queries against entities hidden in text documents. Though I think this can be an interesting application, I don't find the paper convincing as a demo. It mostly reads as a summary of a full technical paper; the demonstrations it proposes are not very impressive. Some more comments: - The claims for the importance of entity search are weak (though they appear as such also in other papers). From my experience, search engines do a very good job in presenting the result snippets. If pages contain more than one result, these are often in tabs, and thus there are no sentences and your search system would fail as well. I would really like to see more convincing and concrete examples. - The difference to methods such as EntityRank are not clear. Can you optimize a ranking such that it is better in the "long range"? What does that mean? If you optimize for precision-at-k, your results should be good at all k. - The IE parts are essentially missing. The way you perform NER in the demo is quite strange. I was expecting more on real NER. - How are values for "credit(o,p,s)" obtained?
Overall Rating	Reject