Paper Review Form

Reviewer Name: Shasha(Amy) Liu


Section I. Overview

A. Reader Interest

1. Which category describes this manuscript?
   ___Practice/Application/Case Study/Experience Report
   _x_Research/Technology
   ___Survey/Tutorial/How-To

B. Content

1. Please explain how this manuscript advances this field of research and/or contributes something new to the literature.
   SOFIE is the very first approach to the ontology-extension problem that integrates logical constraint checking with pattern-based IE (Information Extraction), and is able to provide ontological facts about disambiguated entities in canonical form.

C. Presentation

1. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on?
   _x_Yes
   ___Could be improved
   ___No

2. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic?
   _x_Satisfactory
   ___Could be improved
   ___Poor

3. Please rate and comment on the readability of this manuscript.
   ___Easy to read
   _x_Readable - but requires some effort to understand
   ___Difficult to read and understand
   ___Unreadable

Section II. Evaluation

Please rate the manuscript. Explain your choice.
Section III. Detailed Comments (provide your thoughts/criticism about the ideas in the paper; not only summarize the paper but have a critical look here)
This is a nice paper which presents an approach to the ontology-extension problem. It uses statements and rules to represent the ontology and logical constraint. And by using statements and rules, SOFIE transforms the problem of ontology-extension into the weighted MAX-SAT problem, and uses their new greedy approximation algorithm to solve the original problem. The features of the main algorithm of SOFIE make it source-independent. But there are still many aspects that can be improved, such as the running time, the way to generate the rules, etc.

Additional Comments:
1. Provide one aspect that you liked the most in this paper.
   This paper presents a completely source-independent algorithm which can work well on both semi-structured and unstructured sources.

2. Provide one aspect that you disliked the most in this paper.
   The paper mentions some other related ontology extension systems, but it does not provide any experiment data showing that their work is superior to the others.

Section IV. Discussion Points (provide at least 3 discussion topics/questions related to ideas/techniques described in the paper; these will be used for discussions in the class).
1. It says that the rules that SOFIE uses are designed manually. Would it be possible to generate the rules automatically?
2. We can see from the result table that, in some relation, SOFIE just makes a 80% precision comparing with other 90%+, why this happened.
3. At the end of the paper, authors mentioned that “SOFIE’s main algorithm is complete source-independent, and there is no feature engineering, no learning with cross validation, no parameter estimation, and no tuning of algorithms “, which implies that the system can be improved by adding the features listed above. What aspect of SOFIE can be improved by using features listed in the paper.