

Distributed AMIE+

Preliminary Notes

of the Thesis Project

Omar Trinidad Gutiérrez Méndez

May 25, 2018

1 Introduction

Knowledge bases (KB) have the purpose of representing and store knowledge in a machine-readable format. Some well-known KBs are DBpedia [2], NELL, YAGO [3], or Freebase [1]. A usual task executed in these databases is mining logical rules, that is, find unknown relationships between entities. For example,

$$livesIn(a, x) \text{ and } marriedTo(a, b) \Rightarrow livesIn(b, x)$$

However, these databases are designed under the idea of Open World Assumption (OWA), that means, if the database does not contain a fact, we are not assuming that this fact is false, as happens under the Closed World Assumption (CWA).

Finding these relations in huge datasets, and under the OWA setting is a challenging task. This problem was addressed by Galárraga et al. who proposed Association Rule Mining under Incomplete Evidence (AMIE) and later suggested an improved version of the same method that they simple named AMIE+.

The purpose of the current project is to explore AMIE+ and implement it in a distributed context.

1.1 Theory

Knowledge Bases are collections of facts; every fact is represented by a relation between a subject and object $r(s, o)$.

1.2 AMIE

Inside AMIE, the Partial Completeness Assumption (PCA) was used to guess the so-called counterexamples for rules.

1.3 AMIE plus

In AMIE+ it was aggregated pruning strategies and approximations that allowed to explore the search space more efficiently.

1.4 SANSA

SANSA is a platform whose purpose is...

References

- [1] Kurt Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. Freebase: a collaboratively created graph database for structuring human knowledge. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pages 1247–1250. AcM, 2008.
- [2] Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick van Kleef, Sören Auer, and Christian Bizer. DBpedia - a large-scale, multilingual knowledge base extracted from wikipedia. *Semantic Web Journal*, 6(2):167–195, 2015.
- [3] Fabian M Suchanek, Gjergji Kasneci, and Gerhard Weikum. Yago: a core of semantic knowledge. In *Proceedings of the 16th international conference on World Wide Web*, pages 697–706. ACM, 2007.