Exhaustive list of problems in Ontology Learning

The Need for ontology learning :

Ontologies are the fundamental form of knowledge representation. The vast majority of currently used ontologies have been built entirely by hand. This manual development process represents a major knowledge acquisition bottleneck as sometimes hundreds of hours of effort have been involved and there are ongoing teams of people in place to keep the ontologies up to date. So there is a need for Ontology learning.

Mainly the following problems motivates the need for Ontology learning.

- The high manual cost of ontology construction.
- The continuous change in the existing web pages.
- The very large amount of text from the new web pages growing exponentially every year.

First we have to know the process of Ontology learning to figure out the problems in that field.
An ontology learning system allows the input of one or more texts and the output of some form of taxonomy.

The Basics of Ontology Learning :

- Text Collection.
- Term recognition or keyword extraction.
- Relation Extraction and Labelling.
- Hierarchy Construction.

Intermediate work for figuring out all the Challenges :

1. In recent years there has been a large growth of semantic web on the data level, but unfortunately not on the schema level, which contains mostly concept hierarchies. The shortage of schemas makes the semantic web data difficult to be used in many semantic web applications, so schemas learning from semantic web data becomes an increasingly pressing issue.

2. Ontologies are a tool for Knowledge Representation that is now widely used, but the effort employed to build an ontology is high. So, giving a generalised method for automatic ontology learning for domain specific data is also a hurdle.

3. Since the manual construction of ontologies is time-consuming and expensive, an increasing

number of initiatives to ease the construction by automatic or semi-automatic means have been published. However, a challenging issue is to quantitatively evaluate the usefulness or accuracy of the techniques and combinations of techniques when applied to ontology learning.

4. The challenging issue here is to provide a general learning framework in an automated way that make use of different kinds of contextual knowledge, and to make use of that framework to enrich the ontology with new relations, hence facilitate ontology acquisition.

5. The key problem in achieving efficient and user-friendly retrieval is the development of a search mechanism to guarantee delivery of minimal irrelevant information (high precision) while insuring relevant information is not overlooked (high recall). The traditional solution employs keyword-based search. The only documents retrieved are those containing user specified keywords. But many documents convey desired semantic information without containing these keywords.

6. Automatising the approach to identify lexical patterns which represent semantic relationships between concepts.

7. Evaluation metrics of the Ontology framework.

8. To disambiguate complex concepts consisting of two simpler concepts and the relation that holds between them.

9. To handle imprecision and uncertainty involved in user queries.

10. To use the ontology learning to build a automated question answering system based on the user queries., ie to give the related context along with the relevant links. Every answer to the questions must be relevant to the user's query in that context.

Categorising the above mentioned problems :

1. To provide a ***generalised learning framework in an automated way*** that make use of different kinds of contextual knowledge, and to make use of that framework to enrich the ontology with new relations, hence facilitate ontology acquisition.

2. To quantitatively ***evaluate the usefulness or accuracy*** of the techniques and combinations of techniques when applied to ontology learning.

3. ***Nature of the Text*** , handling the imprecision and uncertainty involved in user queries.

4. To use the ontology learning to ***build a automated question answering system*** based on the user queries., ie to give the related context along with the relevant links. Every answer to the questions must be relevant to the user's query in that context.