

Ontology Extension Towards Analysis of Business News

Inna Novalija and Dunja Mladenic

Department of Knowledge Technologies, Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia

E-mail: {inna.koval, dunja.mladenic}@ijs.si

<http://kt.ijs.si/>

Keywords: ontology extension, text mining, Cyc, semantic web

Received: January 17, 2010

This paper addresses the process of the ontology extension for a selected domain of interest which is defined by keywords and a glossary of relevant terms with descriptions. A new methodology for semi-automatic ontology extension, aggregating the elements of text mining and user-dialog approaches for ontology extension, is proposed and evaluated. We conduct a set of ranking, tagging and illustrative question answering experiments using Cyc ontology and business news collection. We evaluate the importance of using the textual content and structure of the ontology concept in the process of ontology extension. The experiments show that the best results are obtained with giving more weight to ontology concept content and less weight to ontology concept structure.

Povzetek: Prispevek opisuje proces razširitve obstoječe ontologije konceptov.

1 Introduction

This paper explores the process of the ontology extension motivated by usage of the extended ontology for business news analysis. The main contribution of this paper is in proposing a methodology for text-driven semi-automatic ontology extension using ontology content and ontology structure information. Our research also contributes to the analysis of business news by the means of semantic technologies. The new methodology for the semi-automatic ontology extension, aggregating the elements of text mining and user-dialog approaches for ontology extension, is suggested and used for inserting the new financial knowledge into Cyc [1], which maintains one of the most extensive common-sense knowledge bases worldwide.

As the ontology content of a particular concept we consider the available textual representation of the referred concept. The ontology content includes a natural language concept denotation (such as a concept label) and textual comments about the concept. As the ontology structure of a particular concept we consider the neighborhood concepts involved in the hierarchical and non-hierarchical relations with a referred concept. Ontology extension in this paper stands for: adding new concepts to the existing ontology or, augmentation of the existing textual representation of the relevant concepts with new available textual information – extension of the concept comments, changing or adding concept denotation.

The experiments on ranking, business news tagging and simple question answering show that the extended financial ontology allows for a better financial news analysis.

The evaluation of the methodology of the ontology extension shows its ability to fasten the ontology extension process.

The paper is structured as follows: Section 2 presents the information about the existing approaches of ontology extension; the new methodology of ontology extension is discussed in Section 3, Sections 4 describes the experiments and the results, the conclusion is covered in Section 5.

2 Existing approaches to ontology extension

The automatic and semi-automatic ontology extension processes are usually composed of several phases. Most approaches include defining the set of the relevant ontology extension sources, pre-processing the input material, ontology augmentation according to the chosen methodology and ontology evaluating and revision phases. The notable approaches of ontology extension include natural language processing based approach [2, 3], networks/graphs based approach [4, 5] and user interaction approach [6, 7].

The linguistic patterns are used by the authors of Text2Onto [7] framework for ontology learning and SPRAT [8] tool for ontology population.

Several methods of the automatic ontology extension operate with enlarging of Cyc Knowledge Base (Cyc KB). The automated population of Cyc with named entities involves the Web and a framework for validating candidate facts [9]. The semi-automatic approach for Cyc KB extension presented in [6] is based on the user-interactive dialogue system for knowledge acquisition, where, the user is engaged in a natural-language mixed-initiative dialogue. The system contains a natural language generation module, parsing module, post-processing module, dictionary assistant, user interaction agenda and salient descriptor. Medelyan and Legg [10]

describe the methodology for integrating Cyc and Wikipedia, where the concepts from Cyc are mapped onto Wikipedia articles describing correspondent concepts. Sarjant et al. [11] use Medelyan and Legg [10] method to augment Cyc ontology using pattern matching and link analysis.

In the presented research we are using a combination of top-down and bottom-up approaches to the ontology extension and apply it on Cyc ontology. The top-down part involves the user identifying the keywords for extracting relevant data from the ontology, while the bottom-up part involves automatic obtaining of the relevant information available in the ontology. Usage of text mining methods involves data preprocessing, where a chain of linguistic components, such as tokenization, stop-word removal and stemming allows normalizing the textual representation of ontology concepts and a domain relevant glossary of terms with descriptions. Text mining methods are further used for automatically determining candidate concepts in the ontology to relate to the new knowledge from the domain. A list of suggestions is provided to the user for a final decision which allows preventing the inappropriate insertions into the ontology.

3 Methodology

As a part of the research, we propose a new methodology for semi-automatic ontology extension, which combines text mining methods with user-oriented approach and supports the extension of multi-domain ontologies.

The proposed methodology for semi-automatic ontology extension accounts for the following phases:

1. *Domain information identification.* The user identifies the appropriate domain keywords. As well, in this module a domain relevant glossary, containing terms with descriptions is determined. We assume that the glossary terms are the candidate entry concepts for the existing ontology. Consequently, the glossary terms might be in the following relationships with the existing ontology concepts:

- Equivalence relationship: candidate concept represented by a glossary term is equivalent to the existing ontology concept;
- Hierarchical relationship: candidate concept represented by a glossary term is in the superclass-subclass relationship with an existing ontology concept;
- Non-hierarchical relationship: candidate concept represented by a glossary term is in the associative relationship with an existing ontology term. The nature of the relationship is not hierarchical;
- No relationship: candidate concept represented by a glossary term is not related to the existing ontology concept.

2. *Extraction of the relevant domain ontology subset from multi-domain ontology.* In case of large common-sense ontologies, such as Cyc Knowledge Base, the user entering new knowledge very often needs a particular ontology subset of his domain interest. Therefore, the domain keywords are mapped to the natural language representation of the ontology domain information and a set of the relevant domains of interest

is identified. Further, ontology concepts defined in these domains are extracted. By concept extracting we mean obtaining the content and structure of the ontology concept. Correspondently, we find the textual representation (natural language denotation and comments) as content for the particular ontology concept. The ontology structure of the particular concept is represented by the natural language denotations of the hierarchically and non-hierarchically connected ontology concepts. Besides that, the names of the glossary terms are mapped to the natural language denotations of the concepts from other domains and the correspondent concepts are also extracted.

3. *Domain relevant information preprocessing.* The preprocessing phase includes tokenization, stop-word removal and stemming. Textual information is represented using bag-of-words representation with TFIDF weighting and similarity between two text segments is calculated using cosine similarity between their bag-of-words representations, as commonly used in text mining [12]. For each term from the domain relevant glossary we compose bag-of-words aggregating preprocessed textual information from: (1) the glossary term name and (2) the term comment. For each concept from the extracted relevant ontology subset the following information is considered: (1) the ontology concept content consisting of the preprocessed natural language concept denotation and concept comment; (2) the ontology concept structure consisting of the preprocessed natural language concept denotation and natural language denotations of hierarchically and non-hierarchically related concepts.

4. *Composing the list of potential concepts and relationships for ontology extension.* The ranked list of the relevant concepts and possible relationships suitable for ontology extension is composed. Similarity (SIM_{cont}) between glossary term and ontology concept content is calculated and weighted with weight δ ($0 \leq \delta \leq 1$) defined by the user. Similarity (SIM_{str}) between glossary term and ontology concept structure is calculated and weighted with weight $1-\delta$. The combined content and structure similarity (SIM) is used to rank ontology concepts for each glossary term:

$$SIM = \delta * SIM_{cont} + (1 - \delta) * SIM_{str} \quad (1)$$

Ontology concepts with similarity (SIM_c) larger than $SIM_{max} * (1-\beta)$ are suggested to the user, where SIM_{max} represents the highest similarity value between ontology concept and a glossary term for a particular glossary term and β is defined by the user ($0 \leq \beta \leq 1$):

$$SIM_c > SIM_{max} * (1-\beta) \quad (2)$$

5. *User validation.* Furthermore, the user validates the candidate entries results consisting of the glossary terms and relevant existing ontology concepts. In case of the equivalence relationship the user can extend the textual representation of the existing ontology concept by adding comment, adding or changing the natural language denotation. In case of the hierarchical

relationships the user can add subclasses to the existing ontology concepts. If the nature of the relationship is not clear, the user can create an associative relationship or choose any other relationship between a glossary term and existing ontology concept. Moreover, the list with validated entries in the relevant format is created.

6. *Ontology extension* represents adding the new concepts and relationships between concepts into the ontology.

7. *Ontology reuse*. The ontology reuse phase serves as the connection link between separate ontology extension processes.

We have adapted the methodology in order to obtain an exhaustive specific methodology for Cyc knowledge base extension. The main adaptations are based on microtheories (Mt) that Cyc is using to represent thematic subsets of the ontology. Since our motivation is in business news annotation, we have chosen Business and Finances as the domains of primary interest. Given the fact that Cyc Knowledge Base contains common sense knowledge [13], we assume that Cyc KB includes some financial knowledge – financial knowledge base (Cyc FKB).

4 Evaluation

4.1 Experiments

In order to evaluate the proposed methodology we conducted a series of ranking, news tagging and

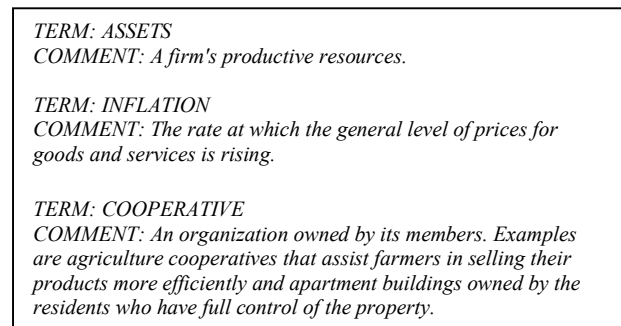


Figure 1: Example Financial Glossary Entries

illustrative question answering experiments on the data sources, described below.

For the data evaluation we have used the RSS feeds data Yahoo! Finance [14] website. The news collection used in the current research accounts for 5812 Yahoo! Finance news.

Following the first phase of the proposed methodology, domain knowledge identification should be made in the initial phase. For these purposes we have selected the Harvey [15] financial glossary which contains around 6000 hyperlinked financial terms.

Figure 1 contains the examples of the typical financial glossary entries.

Tagging experiments give a background for Cyc FKB extension displaying the level of the financial domain representation in Cyc Knowledge Base. We have used a random subset of 100 Yahoo! Finance news to identify the financial terms, occurring most frequently in

the selected news, tagged the terms with Cyc Concept Tagger and checked the precision and recall of news tagging.

For the methodology evaluation, we have conducted ranking experiments on the subset of 500 random Yahoo! Finance news. The most frequent financial terms have been extracted and 100 random financial terms have been chosen. Cyc Financial Knowledge Base is then extended, using the proposed methodology, with concepts corresponding to the chosen financial terms. The efficiency of the automatic concept ranking and the importance of the ontology content and ontology structure in the ontology extension process are measured afterwards.

Illustrative question answering demonstrates the capacity of Cyc to answer simple financial questions before and after the extension of Cyc Financial Knowledge Base. Let us assume that we have a simple question and we want to get an answer using an unextended and extended Cyc Knowledge Base.

4.2 Results

The results of the experiments suggest that the financial ontology extension leads to better business news analysis and confirm the applicability of the suggested methodology for ontology extension to Cyc Knowledge Base augmentation.

We have found 231 financial terms in the random sample of 100 Yahoo! Finance news. The precision and recall of business news tagging with Cyc Concept Tagger accounted for 61% and 46% correspondently. This confirms our hypothesis that the Cyc ontology has still space for extension in the financial domain with terms that are relevant for financial news analysis.

Table 1 shows the quality of automatic concept ranking when using different proportions of ontology concept textual content and ontology concept structure for ranking of the related concepts. We have manually evaluated the automatically suggested ranked related Cyc concepts for every glossary term estimating the proportion of correctly suggested terms among the top 1 suggested terms.

Table 1: Content and Structure Weighting Measures (Financial Glossary).

Weighting Measure	100 Random Terms		
	Top 1 Eqv & Hier Rels	Top 1 Assoc Rels	Top 1 All Rels
Names/Denotation [100%]	18	10	28
Content [0%]	31	30	61
Structure [100%]	32	30	62
Content [10%]	29	31	60
Structure [90%]	30	31	61
Content [20%]	35	33	68
Structure [80%]			
Content [30%]			
Structure [70%]			
Content [40%]			
Structure [60%]			

Content [50%]	35	36	71
Structure [50%]			
Content [60%]	35	37	72
Structure [40%]			
Content [70%]	36	35	71
Structure [30%]			
Content [80%]	36	34	70
Structure [20%]			
Content [90%]	35	33	68
Structure [10%]			
Content [100%]	32	33	65
Structure [0%]			

For this evaluation we explore equivalent, hierarchical and associative relationships between glossary terms and the related Cyc concepts. The best performing proportions are obtained with giving more weight to the similarity calculated between glossary textual representation and Cyc concept content and less weight to the cosine similarity calculated between glossary textual representation and Cyc concept neighborhood. From the row Content [70%] Structure [30%] it is possible to notice that for 36 glossary terms the correct equivalently and hierarchically related Cyc concepts have been found among top 1 suggested concepts. For 71 glossary terms with this weighting measure any related terms have obtained the highest rank among the suggested related concepts.

It means that using the proposed methodology the user is able to compare Cyc and glossary concepts and establish the equivalent, hierarchical and other relations much faster than just using the manual search for the relevant concepts in Cyc.

The following example illustrates the relevance of the proposed Cyc ontology extension for question answering in the financial domain.

For the research purposes we have selected the following simple questions:

*What phase of the business cycle was Egypt in 2008?
Was Indonesia in contraction in 2008?*

<p><i>TERM: BUSINESS CYCLE</i> <i>COMMENT: Repetitive cycles of economic expansion and recession. The official peaks and troughs of the U.S. cycle are determined by the National Bureau of Economic Research in Cambridge, MA.</i></p> <p><i>Phases of Business Cycle:</i></p> <p><i>TERM: CONTRACTION</i> <i>COMMENT: A slowdown in the pace of economic activity.</i></p> <p><i>TERM: TROUGH</i> <i>COMMENT: The lower turning point of a business cycle, where a contraction turns into an expansion.</i></p> <p><i>TERM: EXPANSION</i> <i>COMMENT: A speedup in the pace of economic activity.</i></p> <p><i>TERM: PEAK</i> <i>COMMENT: The upper turning of a business cycle.</i></p>
--

Figure 2: Business Cycle Definition.

Using an unextended Cyc KB we get no appropriate answers because of the insufficient representation of business cycles in Cyc.

Figure 2 presents the textual definition of business cycle and its phases which we use to implement the notion of business cycles in Cyc.

Using the proposed methodology for semi-automatic ontology extension, we obtain a ranked list of related Cyc concepts for the correspondent glossary term (see Table 2).

Table 2: Related Cyc Concepts for Glossary Term “Business Cycle”.

Glossary Term	Ranked Related Cyc Concepts
BUSINESS CYCLE	Cycle-Situation Recession-Economic MacroeconomicEvent Trough (a type of FluidReservoir)

To enter new assertions into Cyc KB we use KE text format which facilitates the knowledge entry process. We select the Cyc concept *Cycle-Situation* as a superclass for glossary term *Business Cycle*:

```

KE text:
Constant: BusinessCycle.
In Mt: UniversalVocabularyMt.
isa: TemporalObjectType.
genls: Cycle-Situation.
comment: "Repetitive cycles of economic expansion and recession. The official peaks and troughs of the U.S. cycle are determined by the National Bureau of Economic Research in Cambridge, MA."
    
```

Furthermore, we create a set of business cycle phases (*Contraction*, *Expansion*, *Peak* and *Trough*) as subclasses for Cyc concept *MacroeconomicEvent*. The following code displays the example of the *Contraction* phase definition:

```

KE text:
Constant: ContractionBusinessCyclePhase.
In Mt: UniversalVocabularyMt.
isa: TemporalObjectType.
genls: MacroeconomicEvent.
comment: "A slowdown in the pace of economic activity".

In Mt: UniversalVocabularyMt.
f:(relationAllExists properSubSituations
BusinessCycle ContractionBusinessCyclePhase).
    
```

In addition, we create a predicate used for answering questions connected to business cycle phases of the specific countries.

```

KE text:
Constant: economyInBusinessCyclePhase.
In Mt: UniversalVocabularyMt.
isa: TernaryPredicate.
arity: 3.
arg1Isa: GeopoliticalEntity.
arg2Isa: TemporalThing.
arg3Isa: MacroeconomicEvent.
    
```

For the illustrative question answering example we estimate the business cycle phases by using the GDP

growth rate - the percentage increase or decrease of Gross Domestic Product (GDP) from the previous measurement cycle. We identify that a term *GDP* is already implemented in Cyc KB as *grossDomesticProduct*.

The following rule defines the conditions of being in the contraction business cycle phase for the particular country in the specified year. We assume that the contraction phase occurs when the real growth rate of GDP in the referred year $GR(GDP)_{Y_n}$ decreases comparatively to the real growth rate of GDP in the previous year $GR(GDP)_{Y_{n-1}}$ but is still higher than the real growth rate of GDP in the following year $GR(GDP)_{Y_{n+1}}$:

$$GR(GDP)_{Y_{n-1}} > GR(GDP)_{Y_n} > GR(GDP)_{Y_{n+1}} \quad (3)$$

KE text:

```
In Mt: UniversalVocabularyMt.
f:
(implies
  (and
    (evaluate ?SUCCESSOR1 (PlusFn ?Y 1))
    (evaluate ?PREDECESSOR1 (DifferenceFn ?Y 1))
    (evaluate ?PREDECESSOR2 (DifferenceFn
      ?PREDECESSOR1 1))
    (grossDomesticProduct ?X (YearFn ?SUCCESSOR1)
      (BillionDollars ?S1GDP))
    (grossDomesticProduct ?X (YearFn ?PREDECESSOR1)
      (BillionDollars ?P1GDP))
    (grossDomesticProduct ?X (YearFn ?PREDECESSOR2)
      (BillionDollars ?P2GDP))
    (grossDomesticProduct ?X (YearFn ?Y)
      (BillionDollars ?YGDP))
    (evaluate ?S1GR (QuotientFn ?S1GDP ?YGDP))
    (evaluate ?YGR (QuotientFn ?YGDP ?P1GDP))
    (evaluate ?P1GR (QuotientFn ?P1GDP ?P2GDP))
    (greaterThan ?P1GR ?YGR)
    (greaterThan ?YGR ?S1GR)
    (isa ?PHASE ContractionBusinessCyclePhase)
    (dateOfEvent ?PHASE (YearFn ?Y)))
    (economyInBusinessCyclePhase ?X (YearFn ?Y)
      ?PHASE)).
```

Country	GDP Growth Rate	Year est.
Egypt	7.1%	2007
Egypt	7.2%	2008
Egypt	4.5%	2009
Indonesia	6.3%	2007
Indonesia	6.1%	2008
Indonesia	4.4%	2009

The expansion, peak and trough phases occur under the following conditions:

Expansion:

$$GR(GDP)_{Y_{n-1}} < GR(GDP)_{Y_n} < GR(GDP)_{Y_{n+1}} \quad (4)$$

Peak:

$$GR(GDP)_{Y_{n-1}} < GR(GDP)_{Y_n} > GR(GDP)_{Y_{n+1}} \quad (5)$$

Trough:

$$GR(GDP)_{Y_{n-1}} > GR(GDP)_{Y_n} < GR(GDP)_{Y_{n+1}} \quad (6)$$

For question answering the information from Cyc KB about the GDP levels of Egypt and Indonesia in 2006-2009 is used:

Cyc KB assertions:

```
(grossDomesticProduct Egypt(YearFn 2009)
  (BillionDollars 470.4))
(grossDomesticProduct Egypt(YearFn 2008)
  (BillionDollars 450.1))
(grossDomesticProduct Egypt(YearFn 2007)
  (BillionDollars 419.9))
(grossDomesticProduct Egypt(YearFn 2006)
  (BillionDollars 392.1))

(grossDomesticProduct Indonesia-TheNation
  (YearFn 2009)(BillionDollars 968.5))
(grossDomesticProduct Indonesia-TheNation
  (YearFn 2008)(BillionDollars 927.7))
(grossDomesticProduct Indonesia-TheNation
  (YearFn 2007)(BillionDollars 874.4))
(grossDomesticProduct Indonesia-TheNation
  (YearFn 2006)(BillionDollars 822.6))
```

After extending Cyc KB with notion of business cycle and business cycle phases, using the information about GDP from Cyc KB, it is possible to get answers for the previously asked questions:

Query:

```
(economyInBusinessCyclePhase Egypt
  (YearFn 2008) ?PHASE)
```

Query result:

```
*[Explain] PeakBCPhase2008
```

Query:

```
(economyInBusinessCyclePhase
  Indonesia-TheNation (YearFn 2008)
  ContractionBCPhase2008)
```

Query result:

```
Query was proven True *[Explain]
```

According to the rules introduced into Cyc KB, Egypt was in the peak business cycle phase and Indonesia was in the contraction phase of the business cycle in 2008. PeakBCPhase2008 and ContractionBCPhase2008 are the correspondent instances of PeakBusinessCyclePhase and ContractionBusinessCyclePhase Cyc collections.

The results obtained in the illustrative question answering experiment are comparable with GDP growth rates in Egypt and Indonesia in 2007-2009 [16].

Table 3: GDP Growth Rates in Egypt and Indonesia.

Extension of Cyc Knowledge Base according to the proposed methodology allows the user to provide Cyc with new concepts and rules and perform a better question answering based on the extended ontology.

5 Conclusion

In this paper the aspects of ontology extension and business news analysis have been explored. The new methodology of ontology extension, combining text mining methods and user-based approach, has been proposed and exposed to the preliminary evaluation.

The evaluation of our methodology has been accomplished in the financial domain. We have tested the importance of using concept textual content and concept structure in the process of ontology extension. The best results are obtained with giving more weight to ontology concept content and less weight to ontology concept structure. In addition, we have illustrated the increase in the effectiveness of simple question answering after Cyc Knowledge Base extension with terms from Harvey [15] financial glossary.

In contrast with many other methodologies for ontology extension, our methodology deals with ontologies and knowledge bases, covering more than one domain. However, it allows restricting the area of ontology extension to a specific domain.

Unlike the developers of Text2Onto [7] and SPRAT [8] tools, we do not use lexico-syntactic patterns for the related concepts identification. The statistically driven techniques used in our methodology make the ontology extension process more language independent.

Furthermore, the user validation helps to avoid adding to the ontology irrelevant concepts and relationships.

The future work should include further extension of Cyc Knowledge Base and using it for more sophisticated news analysis. Furthermore, the proposed methodology for ontology extension should be tested on other domains. In addition, a particular attention should be given to the problem of the disambiguation of the glossary terms and terms extracted from news sources.

Acknowledgement

This work was supported by the Slovenian Research Agency and the IST Programme of the EC under NeOn (IST-4-027595-IP) and ACTIVE (IST-2008-215040).

References

- [1] Cycorp, Inc., <http://www.cyc.com>
- [2] F. Burkhardt, J.A. Gulla, J. Liu, C. Weiss, J. Zhou: Semi Automatic Ontology Engineering in Business Applications, *Workshop Applications of Semantic Technologies*, INFORMATIK. 2008.
- [3] T. Sabrina, A. Rosni, T. Enyakong: Extending Ontology Tree Using NLP Technique, In: *Proceedings of National Conference on Research & Development in Computer Science REDECS 2001*. 2001.
- [4] W. Liu, A. Weichselbraun, A. Scharl, E. Chang: Semi-Automatic Ontology Extension Using Spreading Activation. *Journal of Universal Knowledge Management*, No. 1, pp. 50 – 58. 2005.
- [5] J. McDonald, T. Plate, R. Schvaneveldt: Using pathfinder to extract semantic information from text, In: *Schvaneveldt*, pp. 149–164. 1990.
- [6] M. Witbrock, D. Baxter, J. Curtis, D. Schneider, R. Kahlert, P. Miraglia, P. Wagner, K. Panton, G. Matthews, A. Vizedom: An Interactive Dialogue System for Knowledge Acquisition in Cyc, In: *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence*, 2003.
- [7] P. Cimiano, J. Völker: Text2Onto A Framework for Ontology Learning and Data-driven Change Discovery, In: *Proceedings of NLDB 2005*, 2005, pp.227-238.
- [8] D. Maynard, A. Funk, W. Peters: SPRAT: a tool for automatic semantic pattern-based ontology population, In: *Proceedings of International Conference for Digital Libraries and the Semantic Web*, Trento, Italy, 2009.
- [9] P. Shah, D. Schneider, C. Matuszek, R. C. Kahlert, B. Aldag, D. Baxter, J. Cabral, M. Witbrock, J. Curtis: Automated population of Cyc: Extracting information about named-entities from the web, In: *Proceedings of the Nineteenth International FLAIRS Conference*, 2006, pp. 153-158.
- [10] O. Medelyan, C. Legg: Integrating Cyc and Wikipedia: Folksonomy meets rigorously defined common-sense, In: *Proceedings of Wiki-AI Workshop at the AAAI'08 Conference*, Chicago, US, 2008.
- [11] S. Sarjant, C. Legg, M. Robinson, O. Medelyan: "All You Can Eat" Ontology-Building: Feeding Wikipedia to Cyc, In: *Proceedings of the 2009 IEEE/WIC/ACM International Conference on Web Intelligence*, WI'09, Milan, Italy, 2009.
- [12] M. Grobelnik, D. Mladenic, Knowledge Discovery for Ontology Construction, in: J. Davies, R. Studer, P. Warren (Eds.), *Semantic Web Technologies: Trends and Research in Ontology-Based Systems*, John Wiley & Sons, 2006, pp. 9–27.
- [13] D. Lenat: Cyc: A Large-Scale Investment in Knowledge Infrastructure, *Communic. of the ACM* 38 (11), 1995.
- [14] Yahoo! Finance, <http://finance.yahoo.com>
- [15] C.R. Harvey: Yahoo Financial Glossary, *Fuqua School of Business, Duke University*, 2003.
- [16] Central Intelligence Agency, The World Factbook: <https://www.cia.gov/library/publications/the-world-factbook>