

Using Semantic Entity Extraction Method for a New Application

Farhad Abedini, Fariborz Mahmoudi, and Seyedeh Masoumeh Mirhashem

Abstract— Semantic entities extraction method has been used for various aims. Here, it will be used for a new application – computing semantic relatedness of texts. The previous works in the field of computing semantic relatedness of texts convert a text into a set of words by different methods, and then compute semantic relatedness between these words of texts. In this paper a different approach will be proposed in which any text converts to a set of semantic entities to compute semantic relatedness. Comparing the results with the literatures shows that the results of this new approach can be sufficiently reliable.

Index Terms— Semantic entity extraction, YAGO ontology, semantic relatedness, fact extraction, information retrieval.

I. INTRODUCTION

Recognizing of entities in a text is one of important tools in information retrieval methods. Recently, a new entities recognizing method called semantic entities extraction has been introduced [19]. There are many entities extraction systems that each of them has been developed for specific applications. Based on the desired application, the systems extract their required entities from a text including named entities [8] and [15] named entities related biomedical [12] and terms in financial domain [14]. But semantic entities extraction is another matter.

The entities that are extracted by ontologies are semantical. Medelyan et al [2] claim the most appropriate ontology in this field is the YAGO ontology [1]. But YAGO ontology only extracts entities from structured texts such as infoboxes. In this paper, in addition to extracting semantic entities from unstructured texts, a new application of entity extraction is used. This new application is “computing texts semantic relatedness” [20].

The previous works in the semantic relatedness area convert a text to a set of words and then measure the relatedness between these words and the words obtained from another text. But in the present work, it will be suggested that for computing texts semantic relatedness, all of the texts must be converted into a set of semantic entities and then entities of different texts be compared with each

other. To extract semantic information from a text, a resource is needed by which context of information can be extracted from text. This resource is called “background knowledge”. For example, in [10] lexical databases such as WordNet [3] and Roget’s Thesaurus [9] were used as background knowledge.

Creation of lexical resources requires lexicographic expertise, and takes a lot of time and effort. To solve this problem, Corpus-based approaches obtain its background knowledge by performing statistical analysis of large untagged document collections. The most successful and well known of these techniques is Latent Semantic Analysis (LSA) [11], which relies on the tendency for related words to appear in similar contexts. But it can only provide accurate judgments when the corpus is very large, and consequently the pre processing effort required is significant. To solve this problem, Explicit Semantic Analysis (ESA) was introduced [6]. Gabrilovich et al [6] showed that ESA is most successful approach in computing texts semantic relatedness, in which Wikipedia has been used as background knowledge. Such background knowledge must have some properties. One of the most important properties is that in the background knowledge, information about every possible thing should be existed. It is clear that such a resource is not available. But as proved in [2], Wikipedia is the most appropriate existent resource in this field. Using Wikipedia as the background knowledge resource, in addition to its advantages, has two major problems. Firstly, Wikipedia is not completely reliable and then, information of this resource is textual and unstructured. Semantic information can’t easily be extracted from unstructured resources. Suggestion of the present work can solve these problems. For this purpose, it is suggested that, instead of Wikipedia, YAGO ontology be used as background knowledge resource. Since YAGO ontology is obtained from Wikipedia, all its advantages are included. Besides, as YAGO ontology uses WordNet to prove its facts accuracy, so can be relied on [17]. On the other hand, YAGO ontology is a structured knowledgebase, and a set of facts, which can be helpful in easily extracting semantic of entities. Each fact in ontology is as a tuple that includes two entities and a relation between them. These tuples can be used to extract entities from a text, obtain semantic of those entities, and then use the semantics in applications such as computing semantic relatedness.

The contributions of this paper are as follows:

- *Creating a new application for semantic entity extraction.* As previously mentioned, in the present work, it will be shown that how entity extraction can be used for computing text semantic relatedness.

Manuscript received March 18, 2012; revised March 29, 2012. This work was supported in part by the Islamic Azad University, Roudsar and Amlash Branch.

F. Abedini is with the Electrical and Computer Engineering Department, and member of Young Researchers Club, Islamic Azad University, Roudsar and Amlash branch, Roudsar, Iran (e-mail: abedini.ac@gmail.com).

F. Mahmoudi is with the Electrical and Computer Engineering Department, Islamic Azad University, Qazvin Branch, Qazvin, Iran. (e-mail: mahmoudi@qiau.ac.ir).

S. M. Mirhashem is with the Islamic Azad University, Roudsar and Amlash branch, Roudsar, Iran (e-mail: saeedeh.mirhashem@yahoo.com).

- Using YAGO ontology as one of the most appropriate background knowledge resource for semantic entities extraction. In this paper, using of YAGO as background knowledge is proposed and it will be shown that this ontology is one of the most appropriate background knowledge resources for these aims.
- Using disambiguation method to help finding texts semantic relatedness. In the previous works such as [13], texts semantic relatedness methods were used to disambiguate ambiguous words. But here, it will be acted inversely; means that words disambiguation will be used to find texts semantic relatedness. A new algorithm has been introduced for this aim [18].

This paper has been structured as follows. In next section first our solution for semantic entities extraction is described and then by using it, experimental results will be presented. These experimental results are performed on a benchmark dataset, introduced by Lee [7], and is compared with Stanford named entity recognition (NER), one of the best entity extraction systems, and with ESA [6], the most appropriate system in computing semantic relatedness. Finally, conclusions are represented.

II. SOLUTION

The solution for semantic entities extraction from a text using YAGO ontology has been described as follows (Fig.1).

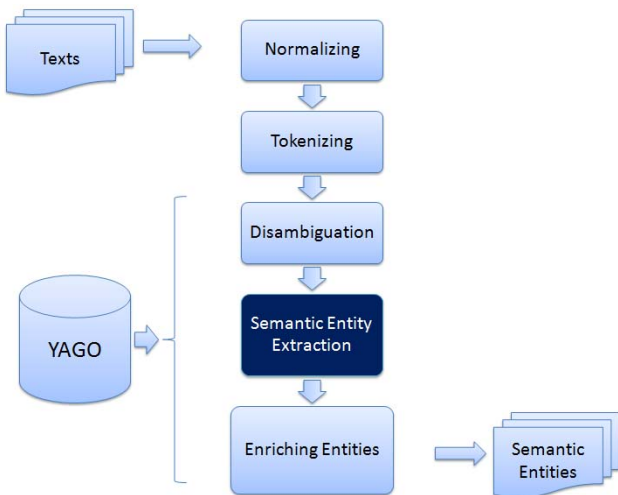


Fig. 1. Steps of suggested method

A. Text Normalization

Since characters, dates and numbers can be an entity, so they can be considered as a semantic entity to be extracted from a text. But each of them can be in different forms to express its purposes. For example, “May 5th, 1983” and “1983-5-5” have a same meaning. So they should have a same structure to present a unique meaning. This work is done by normalization of them.

Different sources come with different encodings. But to have a unique meaning for the same contexts, a unique encoding must be used and other encodings must be changed into it. Here a method is introduced that converts all types of encodings into Unicode. For dates, ISO 6008 format is used

and for numbers all of units are converted into SI units. End step of text normalization is to eliminate additional part of sentences. A same work in this field has been done in LEILA [4], and its idea has been used in this paper.

B. Extracting Tokens from Text

In this step, text is divided into small strings known as tokens. Here the method of SOFIE [5] is used to do this. In this method, a text is given as input and output is a set of tokens with their types.

Assigning each string into one of the token types, types of strings are specified. So unnecessary strings can be ignored and deleted. But since semantic entities will be extracted for special usage of computing semantic relatedness, it must be shown that which of entities can be semantic entities. For this reason, the next part proceeds on finding entities from obtained tokens.

C. Finding entities from tokens by disambiguation, using YAGO ontology

As mentioned earlier, YAGO ontology is a knowledgebase with high coverage and precision that has been obtained from Wikipedia and WordNet [1]. In fact, it can be said that it is the most appropriate available knowledge resource in mining meaning domain [2]. It contains more than 2 million entities and 19 million facts about them and has only 99 unique relations. So it can be appropriate background knowledge for our goal. The entities of YAGO, since all relations of YAGO’s entities with each other are available, are completely semantical. So each of tokens can be matched with one of YAGO entities, one can deduce that a semantic entity has been extracted [18,19]. Here, this matching is introduced as “token disambiguation”.

There are many methods to disambiguate an ambiguous word. For example in [13] and its previous works, texts semantic relatedness methods were used to disambiguate an ambiguous word. But here, it will be acted inversely; means that words disambiguation will be used to find texts semantic relatedness [18]. Here, token is considered as an ambiguous word that can be classified in three statuses. First, if it cannot be matched with YAGO entities, in consequence it is not desired entity and will be ignored. Second, if it can be matched only with one of YAGO entities, in consequence desired entity is found easily. And third, if it can be matched with several YAGO entities, in consequence the token is disambiguated with the method that comes in continue.

This method must select one of the matched entities as the semantic entity. For this aim matched entities is considered as different meaning of token (ambiguous word). These different meaning is shown with e_i . Then all of tokens that obtained from text are matched with YAGO entities. A set of YAGO entities is obtained. This set is shown with $e_set(t)$ that t is text name. Each of YAGO entities that is related with e_i in YAGO ontology, store in $e_set(e_i)$. Then intersection between all values of $e_set(e_i)$ and $e_set(t)$ must be compute. Number of relationships of each e_i with the text entities is shown with $|e_set(t) \cap e_set(e_i)|$. Each of e_i (meanings of ambiguous token) that have more relationship with the text entities is more near to the text and can be resulted that this entity is main meaning of ambiguous token. In fact, the

ambiguate token that was matched with several entities have been disambiguated. And nearest entity is obtained depending on the text. This token disambiguation method is shown in algorithm (1).

Algorithm Token Disambiguation

Input: Token $token$, Text t , YAGO_Ontology o

Output: Semantic Entity for $token$

```

1  $e\_set(t) :=$  set of matched entities in  $o$  with all tokens in  $t$ 
2 IF (Match  $token$  with the entities in  $o$ ) THEN
3    $e_1, \dots, e_n :=$  all of matched entities in  $o$  with  $token$ 
4 ELSE Exit
5 IF ( $n=1$ ) THEN RETURN  $e_1$ 
6 FOR  $i = 1$  TO  $n$ 
7    $e\_set(e_i) :=$  set of entities related to  $e_i$  in  $o$ 
8 FOR  $i = 1$  TO  $n$ 
9    $Number[i] := |e\_set(t) \cap e\_set(e_i)|$ 
10 FOR  $i = 1$  TO  $n$ 
11   IF ( $Number[i] = \text{Max}$ ) THEN RETURN  $e_i$ 

```

(1)

So by this method each of tokens can be matched with one of YAGO entities. Since this ontology is a knowledgebase and its information can be relied (with more than 95% confidence) also each of entity in YAGO has certain relation [1], so it can be claimed that the semantic entities have been obtained.

III. EXPERIMENTAL RESULT

A. Implementation

To implement the project first YAGO ontology has been converted into Mysql database. This work was performed by a computer with 2G RAM and CPU Dual Core with 3M Cache. Its runtime took 22 days. The result was a database of tuple facts with volume 4G. Steps of normalization, token

extraction from text, and finding entities from tokens, have been implemented with java codes on this database.

B. Evaluation

As discussed earlier, the existing entity extraction systems have not been designed to compute texts semantic relatedness. Comparing this system (SESR) with NER (one of the best named entity recognition that is implemented by Stanford Natural Language Processing Group) [8] and ESA (the most benefit computing semantic relatedness system) [6], it will be resulted that SESR is very suitable for computation of texts semantic relatedness.

Since previous works in this field have been used the Lee benchmark dataset [7], in this work also this dataset is used. This dataset contains a collection of 50 documents from the Australian Broadcasting Corporation's news mail service. This datasets have given to some peoples and have requested them to find all entities in these documents. To compare our work with NER, this judgment is used. This means that each of NER or our work is measured with this judgment. And the result of that is shown in Table I.

TABLE I: RESULT OF NER AND SESR COMPARISON

| | Recall | Precision |
|------|--------|-----------|
| SESR | 95% | 98% |
| NER | 90% | 90% |

Precision and recall of NER and SESR compared with the judgments have been compared in this table. The results show that on this dataset SESR can lead to more accurate results. For a case study the text number 12 from the dataset was selected. The results of entity extraction have been shown in Table II. In this table also entities that extract from this text with ESA method is shown that available in [6].

TABLE II: THE RESULT OF THE TEXT NUMBER 12

| NER | SESR | ESA |
|--|--|--|
| LOCATION: U.S - White House , Iraq PERSON: Saddam Hussein , Saddam ORGANIZATION: CIA | Saddam_Hussein/CANONIC White_House/CANONIC Central_Intelligence_Agency/CANONIC Iraq/CANONIC PERSON: BUSH | Iraq disarmament crisis Yellowcake forgery Senate Report of Pre-war Intelligence on Iraq Iraq and weapons of mass destruction Iraq Survey Group Iraq War Scott Ritter Iraq War- Rationale OPERATION DESERT FOX |

It can be seen in table 2, for our purpose in this text the SESR is better than NER. NER does not extract semantic entities and gives only type of entities whereas in SESR entities have matched with synonymous entities in YAGO. but in this method, type of entity obtained in token extraction step. Since the YAGO entities are completely semantical, so we can claim that the entities which obtained with SESR method are "semantic entities".

Comparison with ESA, it can be said that most of the extracted words using ESA is not available in related text. These words are name of articles in Wikipedia such that some

of words in the text are available in context of those articles. These article names are used to compute texts semantic relatedness. But it has been proposed that semantic entities of SESR can be used for computing semantic relatedness. Power of this suggestion has been shown in following.

The method presented here can obtain ESA entities only by one of YAGO relations called FOUNDIN. So, this approach can be more complete than the previous ones. The advantage of SESR is that, the meaning and relations of extracted entities from the text are available in YAGO by which semantic relatedness of texts is computed easily.

Until now, it has been shown that each of texts must be converted into a set of semantic entities. It has been shown in selected step of Fig. 2. After that these sets of semantic entities must be compared together. For example two texts of 1 and 14 is considered. Semantic entities of these texts have been shown in Fig. 2.

| | |
|-----------------------|-----------------------|
| Andrew Bartlett | West Australian |
| Australian Democrats | Brian Greig |
| Wednesday | 7 |
| Thursday | Aden Ridgeway |
| Senator | Greig New York |
| Natasha Stott Despoja | Natasha Stott Despoja |
| Brian Greig | |

Fig. 2. Semantic entities of these texts 14 and 1 of Lee texts

Semantic relatedness of these texts has been judged 100% by humans [7]. But in suggested method this two texts have only 2 same entities. So by this method, semantic relatedness of these texts has been estimated about 26%. It is not a good result.

Totally, in this method numbers of texts that have same entities are few. For example in Table 3 it has been shown that in comparison of 5 texts with other texts how much same entities have.

As it is clear from the table, there are not good results. Many texts don't have same entities with others. So, the texts cannot be compared correctly. It is quite obvious that this method cannot be used for computing semantic relatedness of texts alone. For solving this problem, it is suggested that entities can be enriched with relations of YAGO ontology. The same method has been used in [17].

TABLE III: 5 TEXTS IN COMPARISON WITH OTHERS IN SAME ENTITIES

| Text# | Numbers of texts that same entities have with the text |
|-------|--|
| 1 | 3 |
| 2 | 7 |
| 3 | 11 |
| 4 | 15 |
| 5 | 6 |

For example two texts of 1 and 14 have been developed by *type* relation of YAGO ontology by which more common entities are obtained. Some of these common entities of two texts have been shown in Fig. 2.

| | |
|---|---|
| wikicategory_Australian_bloggers | wikicategory_Australian_Democrats_politicians |
| wikicategory_Australian_Democrats_politicians | wikicategory_Australian_women_in_politics |
| wikicategory_Australian_vegans | wikicategory_British_Classic_Race_winners |
| wikicategory_Australian_vegetarians | wikicategory_British_racehorses |
| wikicategory_Australian_women_in_politics | wikicategory_Federal_Politicians_from_South_Australia |
| wikicategory_Capitals_of_former_nations | wikicategory_Former_students_of_Hale_School |
| wikicategory_Cities_towns_and_villages_in_Yemen | wikicategory_Gay_politicians |
| wikicategory_Coastal_cities | wikicategory_Indigenous_Australian_politicians |
| wikicategory_Days_of_the_week | wikicategory_LGBT_politicians_from_Australia |
| wikicategory_Federal_Politicians_from_South_Australia | wikicategory_LGBT_rights_activists_from_Australia |
| wikicategory_Former_British_colonies | wikicategory_Racehorses_bred_in_Great_Britain |
| wikicategory_Former_students_of_Hale_School | wikicategory_Racehorses_trained_in_Great_Britain |
| wikicategory_Gay_politicians | wikicategory_Thoroughbred_racehorses |
| wikicategory_LGBT_politicians_from_Australia | wikicategory_Towns_in_New_York |
| wikicategory_LGBT_rights_activists_from_Australia | wikicategory_Western_Australian_politicians |
| wikicategory_Liberal_parties | wordnet_administrative_district_108491826 |
| wikicategory_Organisations_based_in_Adelaide | wordnet_adult_109605289 |
| wikicategory_Political_parties_established_in_1977 | wordnet_causal_agent_100007347 |

Fig. 2. Some of enriched entities of two texts 14 and 1

As it is clear in the figure, these two tables have more same entities. Now, total of the same entities between these two texts is 33 (see Fig. 3). By this method semantic relatedness of these two texts is estimated about 41% that is better than before. But with human judgments of Lee are still away.

| | |
|---|---|
| wikicategory_Australian_Democrats_politicians | wordnet_urban_area_108675967 |
| wordnet_causal_agent_100007347 | wikicategory_Australian_women_in_politics |
| wordnet_leader_109623038 | wikicategory_Federal_Politicians_from_South_Australia |
| wordnet_person_100007846 | wordnet_adult_109605289 |
| wordnet_physical_entity_100001930 | wordnet_female_109619168 |
| wordnet_politician_110451263 | wordnet_woman_110787470 |
| wordnet_yagoActorGeo_1 | wikicategory_Gay_politicians |
| wordnet_yagoActor_0 | wikicategory_LGBT_rights_activists_from_Australia |
| wordnet_village_108672738 | wikicategory_Western_Australian_politicians |
| wordnet_administrative_district_108491826 | wikicategory_LGBT_politicians_from_Australia |
| wordnet_district_108552138 | wikicategory_Former_students_of_Hale_School |
| wordnet_geographical_area_108574314 | wordnet_disputant_109615465 |
| wordnet_location_100027167 | wordnet_enrollee_110059162 |
| wordnet_municipality_108626283 | wordnet_militant_110315837 |
| wordnet_region_108630985 | wordnet_reformer_110515194 |
| wordnet_settlement_108672562 | wordnet_student_110665698 |
| wordnet_town_108665504 | |

Fig. 3. Same entities of two texts 1 & 14 after enriching

As was observed, the results can be improved by enriching of entities. In the future work, we are going to obtain benefit amount of entities enrichment for computing semantic relatedness of texts. Then we must find a method for comparing semantic entities of different texts to compute semantic relatedness of them (see Fig. 4).

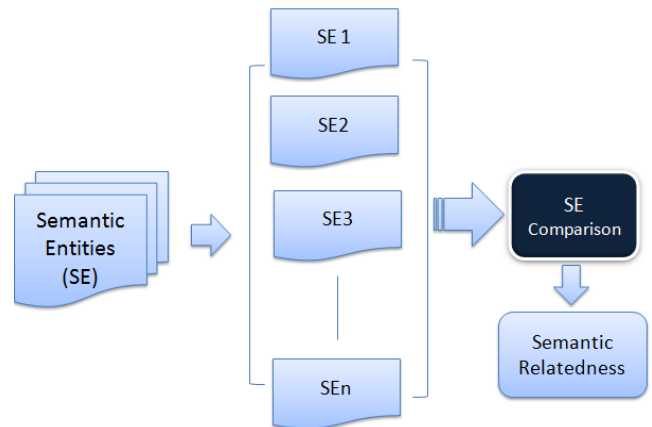


Fig. 4. Comparison of semantic entities

In this figure it is proposed that each text is converted into a set of semantic entities SE_1 to SE_n . Then, to compute semantic relatedness of each two texts semantic entities of a text is compared with semantic entities of another text.

Experimental results show that our method can be compared with related works. Correlation of our method compared with other methods has been shown in Table IV.

TABLE IV: COMPARISON OF OUR METHOD WITH OTHERS

| Method | Correlation with human judgment |
|--------------|---------------------------------|
| Bag of words | 50% |
| LSA | 60% |
| ESA | 72% |
| Our method | 65% |

This table shows that computing semantic relatedness of texts by semantic entities extraction method can offer acceptable results. As mentioned above, in future we are going to improve this method for computing semantic

relatedness of texts.

IV. CONCLUSION

In this paper, a new application of extracting semantic entities from a text was introduced. This application was computing semantic relatedness of texts.

The contributions of this paper was to create a new application for entity extraction, converting an unstructured text into a set of semantic entities, using YAGO ontology as one of the most appropriate background knowledge resources for semantic entities extraction, and using disambiguation to help in finding texts semantic relatedness.

As mentioned in experimental results, the method introduced here can help computing semantic relatedness of texts. For this aim, in our next work we are going to use this method to improve computing semantic relatedness of texts. We also consider using some YAGO relations such as MEANS and TYPE to find upper context for computing semantic relatedness. These relations are available for all entities in YAGO ontology. We are going to make a benefit method for comparing texts entities to improve computing semantic relatedness of texts.

REFERENCES

- [1] F. M. Suchanek, G. Kasneci, and G. Weikum, "YAGO - A Large Ontology from Wikipedia and WordNet," *Elsevier Journal of Web Semantics*, 6(3):203-217, September 2008.
- [2] O. Medelyan, D. Milne, C. Legg, I. H. Witten, "Mining meaning from Wikipedia," *Elsevier Journal of Human-Computer Studies*, Pages 716-754, May 2009.
- [3] C. Fellbaum, "WordNet: An Electronic Lexical Database," *MIT Press*, Cambridge, MA, 1998.
- [4] F. M. Suchanek, G. Ifrim, and G. Weikum, "LEILA: Learning to extract information by linguistic analysis," In *P. Buitelaar, P. Cimiano, and B. Loos, editors, Proceedings of the 2nd Workshop on Ontology Learning and Population (OLP2) at COLING/ACL 2006*, pages 18-25, Sydney, Australia, 2006.
- [5] F. M. Suchanek, M. Sozio, G. Weikum, "SOFIE: a self-organizing framework for information extraction," In: *Proceedings of the WWW 2009 conference*, 2009.
- [6] Gabrilovich, and Markovitch, "Computing Semantic Relatedness using Wikipedia-based Explicit Semantic Analysis," In *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI'07)*, Hyderabad, India, 2007.
- [7] M. D. Lee, B. Pincombe, and M. Welsh, "An empirical evaluation of models of text documents similarity," In *CogSci2005*, pages 1254-1259, 2005.
- [8] The Stanford Natural Language Processing Group, *Stanford Named Entity Recognizer (NER)*, version 1.1.1, 2009-01-16.
- [9] P. Roget, "Roget's Thesaurus of English Words and Phrases," *Longman Group Ltd.*, 1852.
- [10] A. Budanitsky and G. Hirst, "Evaluating wordnet-based measures of lexical semantic relatedness," *Computational Linguistics*, 32(1):13-47, 2006.
- [11] S. Deerwester, S. Dumais, G. Furnas, T. Landauer, and R. Harshman, "Indexing by latent semantic analysis," *JASIS*, 41(6):391-407, 1990.
- [12] Alex, B. Haddow, and C. Grover, "Recognising Nested Named Entities in Biomedical Text," in *BioNLP 2007: Biological, translational, and clinical language processing*, Prague, 2007.
- [13] Sinha, R. and Mihalcea, R., "Unsupervised graph-based word sense disambiguation using measures of word semantic similarity," In *Proc. of ICSC*, 2007.
- [14] F. Xu, D. Kurz, J. Piskorski, S. Schmeier, "Term Extraction and Mining of Term Relations from Unrestricted Texts in the Financial Domain," In *5th International Conference on Business Information Systems*, Poznan, Poland, 2002.
- [15] Finkel, Manning, "Joint parsing and named entity recognition," In *North American Association of Computational Linguistics (NAACL)*, 2009.
- [16] A. Budanitsky and G. Hirst, "Evaluating wordnet-based measures of lexical semantic relatedness," *Computational Linguistics*, 32(1):13-47, 2006.
- [17] Farhad Abedini, Fariborz mahmoudi, Amir Hossein Jadidinejad, "OFE: Ontological Facts Extraction from text for computing Semantic Relatedness," In *proceedings of IEEE 3th International Conference on Machine Learning and Computing (ICMLC 2011)*, Singapore, 26-28 February 2011, pp 84-88.
- [18] Farhad Abedini, Fariborz mahmoudi, Amir Hossein Jadidinejad, "A New Disambiguation Method for Semantic Entity Extraction Using YAGO Ontology," In *proceedings of IEEE 3th International Conference on Machine Learning and Computing (ICMLC 2011)*, Singapore, 26-28 February 2011, pp 79-83.
- [19] Farhad Abedini, Fariborz mahmoudi, Amir Hossein Jadidinejad, "From Text to Knowledge: Semantic Entity Extraction using YAGO Ontology," In *International Journal on Machine Learning and Computing (IJMLC 2011)*, Singapore, June 2011.
- [20] Farhad Abedini, Seyedeh Masoumeh Mirhashem, "SESr: Semantic Entity Extraction for computing Semantic Relatedness," In *proceedings of 4th International Conference on Advanced Computer Theory and Engineering (ICACTE2011)*. Dubai,UAE. 28-30 December 2011. Pp 225-228.



Farhad Abedini was born in Roudsar, Iran in 1983. He received B.E. degree in computer software engineering in 2008 from Tabarestan University Iran, M. Sc. degree in computer engineering in 2011 from Islamic Azad University, Qazvin Branch, Qazvin, Iran (QIAU). He has been a lecturer of computer engineering department at the Islamic Azad University, Roudsar and Amlash Branch, Roudsar, Iran (RAIAU). Now, he is head of Young Researchers Club of Roudsar Branch. He is also member of Young Researchers Club and IACSIT. He was secretary and reviewer of some conferences, seminars and academic competitions. His research interests include Information Retrieval, Text Mining, Question Answering Knowledge Extraction from huge scale Collaboratively Constructed Semantic Resources and using new methods of knowledge representation in semantic processing. CV of Farhad Abedini is available in <http://abedini.org/EnPage/En.htm>.



Fariborz Mahmoudi was born in Tehran, Iran in 1966. He received post diploma in electronic in 1986 from Amirkabir University of Technology Iran, B.E. degree in computer software engineering in 1991 from Shahid Beheshti University Iran, M. Sc. degree in computer engineering in 1994 from Amirkabir University of Technology Iran (Poly-Technique Tehran) and Ph.D. degree in computer engineering from the Science and research campus of Azad University Iran. Since 1996, he has been a lecturer of computer engineering department at the Qazvin Azad University, Qazvin, Iran. As a senior researcher too, he was in the Iran Telecom Research Center (Tehran, Iran) between 2002 and 2006, and in the Mechatronics Research Lab of Qazvin Azad University from 2006 till now. His research interests include machine vision, machine learning and information retrieval in the image and text databases. Dr. Mahmoudi has published more than 100 papers in the international scientific journals and conference proceedings.



Seyedeh Masoumeh Mirhashem was born in Roudsar, Iran in 1983. She received B.A. degree in English language translator in 2011 from Payamnoor University (PNU) of Roudsar Branch, Roudsar, Iran. Now, she is student of M.A. degree in English language teaching at Payamnoor University (PNU) of Rasht Branch, Rasht, Iran. She is member of Young Researchers Club, Islamic Azad University, Roudsar and Amlash Branch, Roudsar, Iran. She also received post diploma in computer software in 2003. His research interests include Linguistics, Information Retrieval, Text Mining, Question Answering, and Motivation in Learning.