Extending English ACE 2005 Corpus Annotation with Ground-truth Links to Wikipedia

Luisa Bentivogli
FBK-Irst
bentivo@fbk.eu

Pamela Forner
CELCT
forner@celct.it

Claudio Giuliano
FBK-Irst
giuliano@fbk.eu

Alessandro Marchetti
CELCT
amarchetti@celct.it

Emanuele Pianta
FBK-Irst
pianta@fbk.eu

Kateryna Tymoshenko
FBK-Irst
tymoshenko@fbk.eu

Abstract

This paper describes an on-going annotation effort which aims at adding a manual annotation layer connecting an existing annotated corpus such as the English ACE-2005 Corpus to Wikipedia. The annotation layer is intended for the evaluation of accuracy of linking to Wikipedia in the framework of a coreference resolution system.

1 Introduction

Collaboratively Constructed Resources (CCR) such as Wikipedia are starting to be used for a number of semantic processing tasks that up to few years ago could only rely on few manually constructed resources such as WordNet and SemCor (Fellbaum, 1998). The impact of the new resources can be multiplied by connecting them to other existing datasets, e.g. reference corpora. In this paper we will illustrate an on-going annotation effort which aims at adding a manual annotation layer connecting an existing annotated corpus such as the English ACE-2005 dataset[1] to a CCR such as Wikipedia. This effort will produce a new integrated resource which can be useful for the coreference resolution task.

Coreference resolution is the task of identifying which mentions, i.e. individual textual descriptions usually realized as noun phrases or pronouns, refer to the same entity. To solve this task, especially in the case of non-pronominal coreference, researchers have recently started to exploit semantic knowledge, e.g. trying to calculate the semantic similarity of mentions (Ponzetto and Strube, 2006) or their semantic classes (Ng, 2007; Soon et al., 2001). Up to now, WordNet has been one of the most frequently used sources of semantic knowledge for the coreference resolution task (Soon et al., 2001; Ng and Cardie, 2002). Researchers have shown, however, that WordNet has some limits. On one hand, although WordNet has a big coverage of the English language in terms of common nouns, it still has a limited coverage of proper nouns (e.g. Barack Obama is not available in the on-line version) and entity descriptions (e.g. president of India). On the other hand WordNet sense inventory is considered too fine-grained (Ponzetto and Strube, 2006; Mihalcea and Moldovan, 2001). In alternative, it has been recently shown that Wikipedia can be a promising source of semantic knowledge for coreference resolution between nominals (Ponzetto and Strube, 2006).

Consider some possible uses of Wikipedia. For example, knowing that the entity mention "Obama" is described on the Wikipedia page `Barack_Obama`[2], one can benefit from the Wikipedia category structure. Categories assigned to the `Barack_Obama` page can be used as semantic classes, e.g. "21st-century presidents of the United States". Another example of a useful Wikipedia feature are the links between Wikipedia pages. For instance, some Wikipedia pages contain links to the `Barack_Obama` page. Anchor texts of these links can provide alterna-

[1] `http://projects.ldc.upenn.edu/ace/`

[2] The links to Wikipedia pages are given displaying only the last part of the link which corresponds to the title of the page. The complete link can be obtained adding this part to `http://en.wikipedia.org/wiki/`.

Beijing, August 2010
Proceedings of the 2nd Workshop on "Collaboratively Constructed Semantic Resources", Coling 2010, pages 19–27,

tive names of this entity, e.g. "Barack Hussein Obama" or "Barack Obama Junior".

Naturally, in order to obtain semantic knowledge about an entity mention from Wikipedia one should link this mention to an appropriate Wikipedia page, i.e. to disambiguate it using Wikipedia as a sense inventory. The accuracy of linking entity mentions to Wikipedia is a very important issue. For example, such linking is a step of the approach to coreference resolution described in (Bryl et al., 2010). In order to evaluate this accuracy in the framework of a coreference resolution system, a corpus of documents, where entity mentions are annotated with ground-truth links to Wikipedia, is required.

The possible solution of this problem is to extend the annotation of entity mentions in a coreference resolution corpus. In the recent years, coreference resolution systems have been evaluated on various versions of the English Automatic Content Extraction (ACE) corpus (Ponzetto and Strube, 2006; Versley et al., 2008; Ng, 2007; Culotta et al., 2007; Bryl et al., 2010). The latest publicly available version is ACE 2005[3].

In this paper we present an extension of ACE 2005 non-pronominal entity mention annotations with ground-truth links to Wikipedia. This extension is intended for evaluation of accuracy of linking entity mentions to Wikipedia pages. The annotation is currently in progress. At the moment of writing this paper we have completed around 55% of the work. The extension can be exploited by coreference resolution systems, which already use ACE 2005 corpus for development and testing purposes, e.g. (Bryl et al., 2010). Moreover, English ACE 2005 corpus is multi-purpose and can be used in other information extraction (IE) tasks as well, e.g. relation extraction. Therefore, we believe that our extension might also be useful for other IE tasks, which exploit semantic knowledge.

In the following we start by providing a brief overview of the existing corpora annotated with links to Wikipedia. In Section 3 we describe some characteristics of the English ACE 2005 corpus, which are relevant to the creation of the extension. Next, we describe the general annotation principles and the procedure adopted to carry out the annotation. In Section 4 we present some analyses of the annotation and statistics about Inter-Annotator Agreement.

2 Related work

Recent approaches to linking terms to Wikipedia pages (Cucerzan, 2007; Csomai and Mihalcea, 2008; Milne and Witten, 2008; Kulkarni et al., 2009) have used two kinds of corpora for evaluation of accuracy: (i) sets of Wikipedia pages and (ii) manually annotated corpora. In Wikipedia pages links are added to terms "only where they are relevant to the context"[4]. Therefore, Wikipedia pages do not contain the full annotation of all entity mentions. This observation applies equally to the corpus used by (Milne and Witten, 2008), which includes 50 documents from the AQUAINT corpus annotated following the same strategy[5]. The corpus created by (Cucerzan, 2007) contains annotation of named entities only[6]. It contains 756 annotations, therefore for our purposes it is limited in terms of size.

Kulkarni et al. (2009) have annotated 109 documents collected from homepages of various sites with as many links as possible[7]. Their annotation is too extensive for our purposes, since they do not limit annotation to the entity mentions. To tackle this issue, one can use an automatic entity mention detector, however it is likely to introduce noise.

3 Creating the extension

The task consists of manually annotating the non-pronominal mentions contained in the English ACE 2005 corpus with links to appropriate Wikipedia articles. The objective of the work is to create an extension of ACE 2005, where all the mentions contained in the ACE 2005 corpus are disambiguated using Wikipedia as a sense repository to point to. The extension is intended for the

[3] http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T06

[4] http://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style

[5] http://www.nzdl.org/wikification/docs.html

[6] http://research.microsoft.com/en-us/um/people/silviu/WebAssistant/TestData/

[7] http://soumen.cse.iitb.ac.in/~soumen/doc/CSAW/

evaluation of accuracy of linking to Wikipedia in the framework of a coreference resolution system.

3.1 The English ACE 2005 Corpus

The English ACE 2005 corpus is composed of 599 articles assembled from a variety of sources selected from broadcast news programs, newspapers, newswire reports, internet sources and from transcribed audio. It contains the annotation of a series of entities (person, location, organization) for a total of 15,382 different entities and 43,624 mentions of these entities. A mention is an instance of a textual reference to an object, which can be either named (e.g. Barack Obama), nominal (e.g. the president), or pronominal (e.g. he, his, it). An entity is an aggregate of all the mentions which refer to one conceptual entity. Beyond the annotation of entities and mentions, ACE 05 contains also the annotation of local co-reference for the entities; this means that mentions which refer to the same entity in a document have been marked with the same ID.

3.2 Annotating ACE 05 with Wikipedia Pages

For the purpose of our task, not all the ACE 05 mentions are annotated, but only the named (henceforth NAM) and nominal (henceforth NOM) mentions. The resulting additional annotation layer will contain a total of 29,300 mentions linked to Wikipedia pages. As specifically regards the annotation of NAM mentions, information about local coreference contained in ACE 05 has been exploited in order to speed up the annotation process. In fact, only the first occurrence of the NAM mentions in each document has been annotated and the annotation is then propagated to all the other co-referring NAM mentions in the document.

Finally, it must be noted that in ACE 05, given a complex entity description, both the full extent of the mention (e.g. president of the United States) and its syntactic head (e.g. "president") are marked. In our Wikipedia extension only the head of the mention is annotated, while the full extent of the mention is available from the original ACE 05 corpus.

3.3 General Annotation Principles

Depending on the mention type to be annotated, i.e. NAM or NOM, a different annotation strategy has been followed. Each mention of type NAM is annotated with a link to a Wikipedia page describing the referred entity. For instance, "George Bush" is annotated with a link to the Wikipedia page `George_W._Bush`.

NOM mentions are annotated with a link to the Wikipedia page which provides a description of its appropriate sense. For instance, in the example "*I was driving Northwest of Baghdad and I bumped into these guys going around the capital*" the mention "capital" is linked to the page which provides a description of its meaning, i.e. `Capital_(political)`. Note that the object of linking is the textual description of an entity, and not the entity itself. In the example, even though from the context it is clear that the mention "capital" refers to Baghdad, we provide a link to the concept of capital and not to the entity Bagdad.

As a term can have both a more generic sense and a more specific one, depending on the context in which it occurs, mentions of type NOM can often be linked to more than one Wikipedia page. Whenever possible, the NOM mentions are annotated with a list of links to appropriate Wikipedia pages in the given context. In such cases, links are sorted in order of relevance, where the first link corresponds to the most specific sense for that term in its context, and therefore is regarded as the best choice. For instance, for the NOM mention head "President" which in the context identifies the United States President George Bush the annotation's purpose is to provide a description of the item "President", so the following links are selected as appropriate: `President_of_the_United_States` and `President`.

The correct interpretation of the term is strictly related to the context in which the term occurs. While performing the annotation, the context of the entire document has always been exploited in order to correctly identify the specific sense of the mention.

3.4 Annotation Procedure

The annotation procedure requires that the mention string is searched in Wikipedia in order to find the appropriate page(s) to be used for annotating the mention. In the annotation exercise, the annotators have always taken into consideration the context where a mention occurs, searching for both the generic and the most specific sense of the mention disambiguated in the context. In fact, in the example provided above, not only “President”, but also “President of the United States” has been queried in Wikipedia as required by the context.

Not only the context, but also some features of Wikipedia must be mentioned as they affect the annotation procedure:

a. One element which contributes to the choice of the appropriate Wikipedia page(s) for one mention is the list of links proposed in Wikipedia’s Disambiguation pages. Disambiguation pages are non-article pages which are intended to allow the user to choose from a list of Wikipedia articles defining different meanings of a term, when the term is ambiguous. Disambiguation pages cannot be used as links for the annotation as they are not suitable for the purposes of this task. In fact, the annotator’s task is to disambiguate the meaning of the mention, so one link, pointing to a specific sense, is to be chosen. Disambiguation pages should always be checked as they provide useful suggestions in order to reach the appropriate link(s).

b. In the same way as Disambiguation pages, Wikitionary cannot be used as linking page, as it provides a list of possible senses for a term and not only one specific sense which is necessary to disambiguate the mention.

c. In Wikipedia, terms may be redirected to other terms which are related in terms of morphological derivation; i.e. searching for the term “Senator” you are automatically redirected to “Senate”; or querying “citizen” you are automatically redirected to “citizenship”. Redirections have always been considered appropriate links for the term.

Some particular rules have been followed in order to deal with specific cases in the annotation, which are described below:

1. As explained before in Section 3.2, as a general rule the head of the ACE 05 mention is annotated with Wikipedia links. In those cases where the syntactic head of the mention is a multiword lexical unit, the ACE 05 practice is to mark as head only the rightmost item of the multiword. For instance, in the case of the multiword “flight attendant” only “attendant” is marked as head of the mention, although “flight attendant” is clearly a multiword lexical unit that should be annotated as one semantic whole. In our annotation we take into account the meaning of the whole lexical unit; so, in the above example, the generic sense of “attendant” has not been given, whereas `Flight_attendant` is considered as the appropriate link.

2. In some cases, in ACE 2005 pronouns like “somebody”, “anybody”, “anyone”, “one”, “others”, were incorrectly marked as NOM (instead of PRO). Such cases, which amount to 117, have been marked with the tag “No Annotation”.

3. When a page exists in Wikipedia for a given mention but not for the specific sense in that context the “Missing sense” annotation has been used. One example of “Missing sense” is for instance the term “heart” which has 29 links proposed in the “Disambiguation page” touching different categories (sport, science, anthropology, gaming, etc.), but there is no link pointing to the sense of “center or core of something”; so, when referring to the heart of a city, the term has been marked as “Missing sense”.

4. When no article exists in Wikipedia for a given mention, the tag “No page” has been adopted.

5. Nicknames, i.e. descriptive names used in place of or in addition to the official name(s) of a person, have been treated as NAM. Thus, even if nicknames look like descriptions of individuals (and their reference should not be solved, following the general rule), they are actually used and annotated as

Number of annotated mentions	16310
Number of single link mentions	13774
Number of multi-link mentions	1458
Number of “No Page” annotations	481
Number of “Missing Sense” annotations	480
Number of “No Annotation” annotations	117
Total number of links	16851
Total number of links in multi-link mentions	3077

Table 1: Annotation data

proper names aliases. For example, given the mention “Butcher of Baghdad”, whose head “Butcher” is to be annotated, the appropriate Wikipedia link is Saddam_Hussein, automatically redirected from the searched string “Butcher of Baghdad”. The link Butcher is not appropriate as it provides a description of the mention. It is interesting the fact that Wikipedia itself redirects to the page of Saddam Hussein.

4 The ACE05-WIKI Extension

Up to now, the 55% of the markable mentions have been annotated by one annotator, amounting to 16,310 mentions. This annotation has been carried out by CELCT in a period of two months from February 22 to April 30, 2010, using the on-line version of Wikipedia, while the remaining 45% of the ACE mentions will be annotated during August 2010. The complete annotation will be freely available at: http://www.celct.it/resources.php?id_page=acewiki2010, while the ACE 2005 corpus is distributed by LDC[8].

4.1 Annotation Data Analysis

Table 1 gives some statistics about the overall annotation. In the following sections, mentions annotated with one link are called “single link”, whereas, mentions annotated with more than one link are named “multi-link”.

These data refer to the annotation of each single mention. It is not possible to give statistics at the entity level, as mentions have different ID depending on the documents they belong to, and the information about the cross-document co-reference is not available. Moreover, mentions of type NOM are annotated with different links depending on their disambiguated sense, making thus impossible to group them together.

Annotation	Mention Type	
	NAM	NOM
Single link mentions	6589	7185
Multi-link mentions	79	1379
Missing sense	96	384
No Page	440	41

Table 2: Distinction of NAM and NOM in the annotation

Most mentions have been annotated with only one link; if we consider multi-link mentions, we can say that each mention has been assigned an average of 2,11 links (3,077/1,458).

Data about “Missing sense” and “No page” are important as they provide useful information about the coverage of Wikipedia as sense inventory. Considering both “Missing sense” and “No page” annotations, the total number of mentions which have not been linked to a Wikipedia page amounts to 6%, equally distributed between “Missing sense” and “No page” annotations. This fact proves that, regarded as a sense inventory, Wikipedia has a broad coverage. As Table 2 shows, the mentions for which more than one link was deemed appropriate are mostly of type NOM, while NAM mentions have been almost exclusively annotated with one link only. The very few cases in which a NAM mention is linked to more than one Wikipedia page are primarily due to (i) mistakes in the ACE 05 annotation (for example, the mention “President” was erroneously marked as a NAM); (ii) or to cases where nouns marked as NAM could also be considered as NOMs (see for instance the mention “Marine”, to mean the Marine Corps).

Table 2 provides also statistics about the “Missing sense” and “No page” cases provided on mentions divided among the NAM and NOM type. The “missing sense” annotation concerns mostly the NOM category, whereas the NAM category is hardly affected. This attests the fact that persons, locations and organizations are well repre-

[8] http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T06

sented in Wikipedia. This is mainly due to the encyclopedic nature of Wikipedia where an article may be about a person, a concept, a place, an event, a thing etc.; instead, information about nouns (NOM) is more likely to be found in a dictionary, where information about the meanings and usage of a term is provided.

4.2 Inter-Annotator Agreement

About 3,100 mentions, representing more than 10% of the mentions to be annotated, have been annotated by two annotators in order to calculate Inter-Annotator Agreement.

Once the annotations were completed, the two annotators carried out a reconciliation phase where they compared the two sets of links produced. Discrepancies in the annotation were checked with the aim of removing only the more rough errors and oversights. No changes have been made in the cases of substantial disagreement, which has been maintained.

In order to measure Inter-Annotator Agreement, two metrics were used: (i) the Dice coefficient to measure the agreements on the set of links used in the annotation[9] and (ii) two measures of agreement calculated at the mention level, i.e. on the group of links associated to each mention.

The Dice coefficient is computed as follows:

$$Dice = 2C/(A + B)$$

where C is the number of common links chosen by the two annotators, while A and B are respectively the total number of links selected by the first and the second annotator. Table 3 shows the results obtained both before and after the reconciliation process. Agreement before reconciliation is satisfactory and shows the feasibility of the annotation task and the reliability of the annotation scheme.

	BEFORE reconciliation	AFTER reconciliation
DICE	0.85	0.94

Table 3: Statistics about Dice coefficient

	BEFORE reconciliation	AFTER reconciliation
Complete	77.98%	91.82%
On first link	84.41%	95.58%

Table 4: Agreement at the mention level

Two measures of agreement at the mention level are also calculated. To this purpose, we count the number of mentions where annotators agree, as opposed to considering the agreement on each link separately. Mention-level agreement is calculated as follows:

$$\frac{\text{Number of mentions with annotation in agreement}}{\text{Total number of annotated mentions}}$$

We calculate both ”complete” agreement and agreement on the first link. As regards the first measure, a mention is considered in complete agreement if (i) it has been annotated with the same link(s) and (ii) in the case of multi-link mentions, links are given in the same order. As for the second measure, there is agreement on a mention if both the annotators chose the same first link (i.e. the one judged as the most appropriate), regardless of other possible links assigned to that mention. Table 4 provides data about both complete agreement and first link agreement, calculated before and after the annotators reconciliation.

4.3 Disagreement Analysis

Considering the 3,144 double-annotated mentions, the cases of disagreements amount to 692 (22,02%) before the reconciliation while they are reduced to 257 (8,18%) after that process. It is interesting to point out that the disagreements affect the mentions of type NOM in most of the cases, whereas mentions of type NAM are involved only in 3,8% of the cases.

Examining the two annotations after the reconciliation, it is possible to distinguish three kinds of disagreement which are shown in Table 5 to-

[9] The Dice coefficient is a typical measure used to compare sets in IR and is also used to calculate inter-annotator agreement in a number of tasks where an assessor is allowed to select a set of labels to apply to each observation. In fact, in these cases measures such as the widely used K are not good to calculate agreement. This is because K only offers a dichotomous distinction between agreement and disagreement, whereas what is needed is a coefficient that also allows for partial disagreement between judgments. In fact, in our case we often have a partial agreement on the set of links given for each mention. Also considering only the mentions for which a single link has been chosen, it is not possible to calculate K statistics in a straightforward way as the categories (i.e. the possible Wikipedia pages) in some cases cannot be determined a priori and are different for each mention. Due to these factors chance agreement cannot be calculated in an appropriate way.

Disagreement type	Number of Disagreements
1) No matching in the link(s) proposed	105 (40,85%)
2) No matching on the first link, but at least one of the other links is the same	14 (5,45%)
3) Matching on the first link and mismatch on the number of additional links	138 (53,70%)
Total Disagreements	257

Table 5: Types of disagreements

gether with the data about their distribution. An example of disagreement of type (1) is the annotation of the mention "crossing", in the following context: *"Marines from the 1st division have secured a key Tigris River Crossing"*. Searching for the word "river crossing" in the Wikipedia searchbox, the Disambiguation Page is opened and a list of possible links referring to more specific senses of the term are offered, while the generic "river crossing" sense is missing. The annotators are required to choose just one of the possible senses provided and they chose two different links pointing to pages of more specific senses: {`Ford_%28river%29`} and {`Bridge`}.

Another example is represented by the annotation of the mention "area" in the context : *"Both aircraft fly at 125 miles per hour gingerly over enemy area"*. In Wikipedia no page exists for the specific sense of "area" appropriate in the context. Searching for "area" in Wikipedia, the page obtained is not suitable, and the Disambiguation page offers a list of various possible links to either more specific or more general senses of the term. One annotator judged the more general Wikipedia page `Area_(subnational_entity)` as appropriate to annotate the mention, while the second annotator deemed the page not suitable and thus used the "Missing sense" annotation.

Disagreement of type (2) refers to cases where at least one of the links proposed by the annotators is the same, but the first (i.e. the one judged as the most suitable) is different. Given the following context: *"Tom, You know what Liberals want"*, the two annotation sets provided for the mention "Liberal" are: {`Liberalism`} and {`Liberal_Party`, `Modern_liberalism_in_the_United_States`, `Liberalism`}.

The first annotator provided only one link for the mention "liberal", which is different from the first link provided by second annotator. However, the second annotator provided also other links, among which there is the link provided by the first annotator.

Another example is represented by the annotation of the mention "killer". Given the context: "*He'd be the 11th killer put to death in Texas*", the two annotators provided the following link sets: {`Assassination`, `Murder`} and {`Murder`}. Starting from the Wikipedia disambiguation page, the two annotators agreed on the choice of one of the links but not on the first one.

Disagreement of type (3) refers to cases where both annotators agree on the first link, corresponding to the most specific sense, but one of them also added link(s) considered appropriate to annotate the mention. Given the context: "*7th Cavalry has just taken three Iraqi prisoners*", the annotations provided for the term "prisoners" are: {`Prisoner_of_war`} and {`Prisoner_of_war`, `Incarceration`}. This happens when more than one Wikipedia pages are appropriate to describe the mention.

As regards the causes of disagreement, we see that the cases of disagreement mentioned above are due to two main reasons:

a. The lack of the appropriate sense in Wikipedia for the given mention

b. The different interpretation of the context in which the mention occurs.

In cases of type (a) the annotators adopted different strategies to perform their task, that is:

i. they selected a more general sense (i.e. "area" which has been annotated with `Area_(subnational_entity)`),

ii. they selected a more specific sense (see for example the annotations of the mentions "river crossing").

iii. they selected the related senses proposed by the Wikipedia Disambiguation page (as in the annotation of "killer" in the example above).

Disagreement type (see above)	**Reas. a**	**Reas. b**	**Tot**
1) No match	95	10	105
2) No match on first link	4	10	14
3) Mismatch on additional links		138	138
Total	99 (38,5%)	158 (61,5%)	257

Table 6: Distribution of disagreements according to their cause

iv. they used the tag “Missing sense”.

As Wikipedia is constantly evolving, adding new pages and consequently new senses, it is reasonable to think that the considered elements might find the appropriate specific/general link as time goes by.

Case (b) happens when the context is ambiguous and the information provided in the text allows different possible readings of the mention to be annotated, making thus difficult to disambiguate its sense. These cases are independent from Wikipedia sense repository but are related to the subjectivity of the annotators and to the inherent ambiguity of text.

Table 6 shows the distribution of disagreements according to their cause. Disagreements of type 1 and 2 can be due to both *a* and *b* reasons, while disagreements of type 3 are only due to *b*.

The overall number of disagreements shows that the cases where the two annotators did not agree are quite limited, amounting only to 8%. The analyses of the disagreements show some characteristics of Wikipedia considered as sense repository. As reported in Table 8, in the 61,5% of the cases of disagreement, the different annotations are caused by the diverse interpretation of the context and not by the lack of senses in Wikipedia. It is clear that Wikipedia has a good coverage and it proves to be a good sense disambiguation tool. In some cases it reveals to be too fine-grained and in other cases it remains at a more general level.

5 Conclusion

This paper has presented an annotation work which connects an existing annotated corpus such as the English ACE 2005 dataset to a Collaboratively Constructed Semantic Resource such as Wikipedia. Thanks to this connection Wikipedia becomes an essential semantic resource for the task of coreference resolution. On one hand, by taking advantage of the already existing annotations, with a relatively limited additional effort, we enriched an existing corpus and made it useful for a new NLP task which was not planned when the corpus was created. On the other hand, our work allowed us to explore and better understand certain characteristics of the Wikipedia resource. For example we were able to demonstrate in quantitative terms that Wikipedia has a very good coverage, at least as far as the kind of entity mentions which are contained in the ACE 2005 dataset (newswire) is concerned.

Acknowledgments

The research leading to these results has received funding from the ITCH project (`http://itch.fbk.eu`), sponsored by the Italian Ministry of University and Research and by the Autonomous Province of Trento and the Copilosk project (`http://copilosk.fbk.eu`), a Joint Research Project under Future Internet - Internet of Content program of the Information Technology Center, Fondazione Bruno Kessler.

We thank Giovanni Moretti from CELCT for technical assistance.

References

Bryl, Volha, Claudio Giuliano, Luciano Serafini, and Kateryna Tymoshenko. 2010. Using background knowledge to support coreference resolution. In *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI 2010)*, August.

Csomai, Andras and Rada Mihalcea. 2008. Linking documents to encyclopedic knowledge. *IEEE Intelligent Systems*, 23(5):34–41.

Cucerzan, Silviu. 2007. Large-scale named entity disambiguation based on Wikipedia data. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pages 708–716, Prague, Czech Republic, June. Association for Computational Linguistics.

Culotta, Aron, Michael L. Wick, and Andrew McCallum. 2007. First-order probabilistic models for coreference resolution. In *Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, pages 81–88.

Fellbaum, Christiane, editor. 1998. *WordNet: an electronic lexical database*. MIT Press.

Kulkarni, Sayali, Amit Singh, Ganesh Ramakrishnan, and Soumen Chakrabarti. 2009. Collective annotation of wikipedia entities in web text. In *KDD '09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 457–466, New York, NY, USA. ACM.

Mihalcea, Rada and Dan I. Moldovan. 2001. Ez.wordnet: Principles for automatic generation of a coarse grained wordnet. In Russell, Ingrid and John F. Kolen, editors, *FLAIRS Conference*, pages 454–458. AAAI Press.

Milne, David and Ian H. Witten. 2008. Learning to link with wikipedia. In *CIKM '08: Proceeding of the 17th ACM conference on Information and knowledge management*, pages 509–518, New York, NY, USA. ACM.

Ng, Vincent and Claire Cardie. 2002. Improving machine learning approaches to coreference resolution. In *ACL '02: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pages 104–111.

Ng, Vincent. 2007. Semantic class induction and coreference resolution. In *ACL 2007, Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, June 23-30, 2007, Prague, Czech Republic*, pages 536–543.

Ponzetto, S. P. and M. Strube. 2006. Exploiting semantic role labeling, WordNet and Wikipedia for coreference resolution. *Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, pages 192–199.

Soon, Wee Meng, Hwee Tou Ng, and Daniel Chung Yong Lim. 2001. A machine learning approach to coreference resolution of noun phrases. *Computational Linguistic*, 27(4):521–544.

Versley, Yannick, Simone Paolo Ponzetto, Massimo Poesio, Vladimir Eidelman, Alan Jern, Jason Smith, Xiaofeng Yang, and Alessandro Moschitti. 2008. Bart: a modular toolkit for coreference resolution. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies*, pages 9–12.