Annotating Modality and Negation for a Machine Reading Evaluation

Roser Morante and Walter Daelemans

CLiPS - University of Antwerp
Prinsstraat 13, B-2000 Antwerpen, Belgium
{Roser.Morante,Walter.Daelemans}@ua.ac.be

Abstract. In this paper we describe the task *Processing modality and negation for machine reading*, which was organized as a pilot task of the *Question Answering for Machine Reading Evaluation* (QA4MRE) Lab at CLEF 2011. We define the aspects of meaning on which the task focused and we describe the dataset produced.

1 Introduction

Until recently, research on Natural Language Processing (NLP) has focused on propositional aspects of meaning. For example, semantic role labeling, question answering or text mining tasks aim at extracting information of the type "who does what when and where". However, understanding language involves also processing extra-propositional aspects of meaning, such as factuality, uncertainty, or subjectivity, since the same propositional meaning can be presented in a diversity of statements, as exemplified in (1), where the propositional meaning <ADD(earthquake, further threats to the global economy)> is present in multiple statements, none of which has the same meaning.

(1) The earthquake adds further threats to the global economy
The earthquake does not add further threats to the global economy
The earthquake never added further threats to the global economy
Does the earthquake add further threats to the global economy?
The earthquake will never add further threats to the global economy
The earthquake will probably add further threats to the global economy
The earthquake will certainly add further threats to the global economy
The earthquake might have added further threats to the global economy
According to some media sources, the earthquake adds further threats to the global economy
The earthquake will add further threats to the global economy if the right measures are not applied
It is unclear whether the earthquake will add further threats to the global economy
It is expected that the earthquake will add further threats to the global economy
It has been denied that the earthquake adds further threats to the global economy
It is believed that the earthquake adds further threats to the global economy
Why would the earthquake not add further threats to the global economy?

Researchers have started to study phenomena related to extra-propositional meaning such as factuality, belief and certainty, speculative language and hedging, or contradictions and opinions. Modality and negation are two main grammatical devices that allow to express extra-propositional aspects of meaning. Generally speaking, *modality* is a grammatical category that allows to express aspects related to the attitude of the speaker towards her statements in terms of degree of certainty, reliability, subjectivity, sources of information, and perspective. We understand modality in a broad sense, which involves related concepts like *subjectivity* (38), *hedging* (14), *evidentiality* (1), *uncertainty* (31), *committed belief* (6) and *factuality* (33). *Negation* (36) is a grammatical category that allows to change the truth value of a proposition.

Research on modality and negation has been stimulated by a number of data sets annotated with various aspects of modality and negation information, such as the Rubin's certainty corpus (29; 30), the ACE 2008 corpus (17), the BioScope corpus (37), and the FactBank corpus (33).

Two main tasks have been addressed in the NLP community, the detection of various forms of negation and modality and the resolution of the scope of modality and negation cues. For negation detection, a number of rule-based systems have been developed (2; 23; 4), as well as some systems that rely on machine learning (3; 11; 12; 28; 40). Negation has also been incorporated explicitly or implicitly in systems that process contradiction and contrast (13; 27; 16). There are several systems for modality detection (20; 15; 19; 35; 10; 24). The recently introduced scope resolution task is concerned with determining at a sentence level which tokens are affected by negation and modality (22; 21; 25; 24). This task has become very popular after the edition of the CoNLL Shared Task 2010 on *Learning to detect hedges and their scope in natural language texts* (8).

Incorporating information about modality and negation has been shown to be useful for a number of applications, such as biomedical and clinical text processing (9; 18; 23; 4; 35), opinion mining and sentiment analysis (39), recognizing textual entailment (5; 34), or automatic style checking (10). More generally, being able to deal adequately with modality and negation is relevant for any NLP task that requires some form of text understanding and needs to discriminate between factual and non-factual information, including text summarization, question answering, information extraction, and human-computer interaction in the form of dialogue systems.

Machine Reading (MR) is a task that aims at automatic unsupervised understanding of texts(7). Since modality and negation are very relevant phenomena for understanding texts, and, as far as we know, they have not been treated before in machine reading tasks, we proposed a pilot task on *Processing modality and negation for machine reading* as a pilot task of the *Question Answering for Machine Reading Evaluation* (QA4MRE)[1] (REF in this volume) at CLEF 2011.

In Section 2 the task is described, Section 3 introduces the aspects of meaning to be processed, and Section 4 the complexity of the task. Unluckily, no partic-

---

[1] Web site of QA4MRE: `http://celct.isti.cnr.it/ResPubliQA/`.

ipants submitted systems for this task, which means that we cannot describe systems nor provide evaluation results.

2   Task description

The pilot task follows the same set up as the main QA4MRE task. The goal of the QA4MRE evaluation (REF in this volume) is to develop a methodology for evaluating Machine Reading systems through Question Answering and Reading Comprehension Tests. Participating systems should be able to answer multiple choice questions about test documents, which requires a deep knowledge of the documents. Systems are asked to analyse the corresponding test document in conjunction with the background collections provided by the organization. Finding the correct answer might require performing some kind of inference and processing previously acquired background knowledge from reference document collections. Although the additional knowledge obtained through the background collection may be used to assist with answering the questions, the principal answer is to be found among the facts contained in the test documents given. The main characteristics of the reading comprehension tests is that they not only require that systems perform semantic understanding, but they assume also a *cognitive* process that involves using implications and presuppositions, retrieving the stored information, and performing inferences to make implicit information explicit. The organization provides participants with a background collection of about 30,000 unannotated documents related to three topics: music and society, aids, and climate change. Background collections and tests are provided for several languages: English, Spanish, German, Italian, and Romanian.

The task *Processing modality and negation for machine reading*[2] is organised as a pilot task of the QA4MRE Lab. The pilot task aims at evaluating whether machine reading systems understand extra-propositional aspects of meaning beyond propositional content, focusing mostly on phenomena related to modality and negation.

Systems participating in the pilot task are supposed to learn from the background collections provided for the main task, although they will be evaluated on test sets designed specifically for the pilot task. The test documents come from the journal The Economist[3]. The format of the test sets is the same as in the main task, four documents are provided per topic with ten multiple choice questions per document. Each question has five options, from which only one is correct. The options are exclusive. Questions are about how a certain event is presented in the document regarding five aspects of meaning: negation, perspective, certainty, modality and condition. The task can also be seen as a classification task in which systems have to generate the event description. This pilot task will evaluate how systems process the aspects of meaning presented in Section 3.

---

[2] Web site of the pilot task: `http://www.cnts.ua.ac.be/BiographTA/qa4mre.html`.

[3] The Economist kindly made available the texts for non-commercial research purposes.

For example, given a sentence like (2) in the text, possible multiple choice options are listed in (3). The correct option would be (3.d).

(2) Experts consider that it is unclear whether the earthquake will add further threats to the global economy

(3) Event −the earthquake <predicate>add</predicate> further threats to the global economy− is presented in the text as:

    a A negated event

    b A condition for another event

    c An event

    d An uncertain event from the perspective of someone other than the author - CORRECT

    e A purpose event

In order to make the options machine readable, a code will be assigned to them. The aspects of meaning to be coded are presented in Section 3 and the full list of possible code combinations are listed in the Guidelines[4].

A question focuses on an event mentioned in the document. The event and its participants are quoted almost literally in the formulation of the question. The difference with the literal quotation is that only the lemma of the event predicate appears in the question instead of the full form, and that negation and modality marks are also removed. For example, in (3), the lemma of the event predicate is *add*, which substitutes the full form *will add* that occurs in sentence (2). The question does not reproduce the full sentence where the event occurs, but only the event and its participants. In (2) the sentence is *Experts consider that it is unclear whether the earthquake will add further threats to the global economy*, but in the question only the event ADD and its participants are quoted, with the event marked with an xml like tag: *the earthquake <predicate>add</predicate> further threats to the global economy*. For this task, *event* is understood in a broad sense, including actions, processes and states. Events can be expressed by verbs and nouns.

In order to allow participants to tune their systems, two pilot test document were released first[5]. As in the main task, for each document there are ten multiple choice questions, each having five candidate answers, one clearly correct answer and four clearly incorrect answers. The task of a system is to choose one answer for each question, by analysing the corresponding test document in conjunction with the background collection.

We decided to provide test documents from The Economist because they are well written, the style is uniform for all texts, the journal relatively frequently addresses the topics established by the main task organizers, and they not only provide facts, but also opinionated statements, where modality phenomena can

---

[4] The Guidelines of the pilot task can be found at `http://www.clips.ua.ac.be/BiographTA/qa4mre-files/qa4mre-pilot-guidelines.pdf`.

[5] The pilot test documents can be found at `http://www.clips.ua.ac.be/BiographTA/qa4mre-files/qa4mre-pilot-test-examples.zip`.

be found. last, but not least, The Economist agreed to release the text under a Creative Commons license, for which we are very grateful. As a negative aspect of these texts, they do not belong to the same type of texts as provided in the background collection of the main task. However, finding the right answers should be possible by analyzing the document at hand.

Table 1. Test documents from The Economist

Topic	Number	Title	# of words
Aids	1	All colours of the brainbow	915
	2	DARC continent	817
	3	Double, not quits	779
	4	Win some, lose some	1919
Climate change	1	A record-making effort	2841
	2	Are economists erring on climate change?	1412
	3	Climate change and evolution	1256
	4	Climate change in black and white	2850
Music and society	1	The politics of hip-hop	1004
	2	How to sink pirates	773
	3	Singing a different tune	1042
	4	Turn that noise off	677

The test documents can be downloaded from the web site of the of the main task[6]. Since the pilot follows the main task setting of the main task, no annotated training data are provided. Apart from the background collection, systems can use any existing resources and data to solve the task.

As for evaluation, the pilot task is evaluated using the same procedure as the main task. Each question receives one (and only one) of the three following assessments: *correct*, if the system selected the correct answer among the five candidate ones; *incorrect*, if the system selected one of the wrong answers; *NoA*, if the system chose not to answer the question. Two evaluation measures are applied, *c@1* (26), which takes into account the option of not answering certain questions, and *accuracy*. c@1 acknowledges the option of giving NoA answers in the proportion that a system answer questions correctly, which is measured using accuracy.

3 Aspects of meaning to be processed by systems

For this pilot task, five aspects of meaning related to modality and negation were selected. Systems have to choose the answer that best characterises an event along these six aspects described in the following subsections:

---

[6] Test documents available at `http://celct.fbk.eu/ResPubliQA/index.php?page=Pages/pastCampaigns.php`.

- Negation
- Perspective
- Certainty
- Modality
- Condition for another event or conditioned by another event

3.1 Negation

An event can be presented as negated. In (4), the REPLACE event is negated with negation cue *not*. In (5), we consider <PUT the sort of price on carbon use that would drive its emission down> negated by the cue *inability*.

(4) But these new types of climate action do not replace the need to reduce carbon emissions.

(5) In the face of an international inability to put the sort of price on carbon use that would drive its emission down, an increasing number of policy wonks, and the politicians they advise, are taking a more serious look at these other factors as possible ways of controlling climate change.

3.2 Perspective

A statement is presented from the point of view of someone. By default the statement is presented from the perspective of the author of the text, but the author might be mentioning the view from someone else. The task will only evaluate whether systems are able to detect when an event is presented from a different perspective than the auhtor's. This is explicitly indicated in the multiple choice questions as *perspective from someone other than the author*.

For example, in (6) the event <radioactive particles from the Fukushima Dai-ichi nuclear-power plant LEAD this once-prosperous city of 70,000 into a fight for its life> is presented from the perspective of the mayor of Minamisoma.

(6) Yet he [coref: mayor of Minamisoma] believes the radioactive particles from the Fukushima Dai-ichi nuclear-power plant, 25km from his office, have led this once-prosperous city of 70,000 into a fight for its life.

In (7) event <LACK of testing equipment> is presented from the perspective of *traders in this places*, event <tuna that arrived in America SET aside by customs> from the perspective of *an executive at a Japanese trading house*, and event <Japanse food BE off the menu at hotels> from the perspective of *a sake brewer on a sales trip to Las Vegas*.

(7) The European Union has named a dozen prefectures that need radiation tests, yet traders in these places report a lack of testing equipment. In one case, says an executive at a Japanese trading house, tuna that arrived in America was set aside by customs, rotting before it was inspected. A sake brewer on a sales trip to Las Vegas noticed that Japanese food was off the menu at hotels.

3.3   Certainty

Events can be presented with a range of certainty values, including underspecified certainty. Here we include all not certain events under the category of *uncertain* events, without distinguishing degrees. The task focuses only on uncertain events.

In (8) the PROVIDING event is presented as uncertain because of the use of *possible*.

(8) Providing most of that energy from wind, sunshine, plants and rivers, along with a bit of nuclear, is possible.

In (9) event <many of Minamisoma's evacuees COME back> is presented as uncertain and negated because a speculation and a negation cue are used (*may never*).

(9) ... Even though external radiation has since returned to near-harmless levels, Mr Sakurai fears many of Minamisoma's evacuees may never come back.

Event <the investment required to decarbonise power AVERAGE about £30 billion ($42 billion) a year over 40 years> in (10) is uncertain because of the conditional *would*.

(10) The commission says the investment required to decarbonise power would average about £30 billion ($42 billion) a year over 40 years.

In (11) event <you HUNT for every possible deduction for which you're eligible> is uncertain because of the use of *can*, as well as <these alternatives also IMPROVE the content and prospects of other climate action> in (12) because of the use of *could*.

(11) If you are highly motivated to minimise your taxes, you can hunt for every possible deduction for which you're eligible.

(12) As well as having charms that efforts to reduce carbon-dioxide emissions lack, these alternatives could also improve the content and prospects of other climate action.

3.4   Modality

An event can be presented with several modal meanings. For this pilot task we select only the modal meanings listed below, although we are aware that the variety of modal meanings is broader.

Non-modal event This is the default category for events that do not fall under the modal categories below and do not have other modal meanings. In the questions we refer to it as *event*. An event can be in the present, past or future tense.

In (13) the events <A pen-like dosimeter HANG around the neck of Katsunobu Sakurai> and <he EXPOSED during the past two weeks of a four-week nuclear nightmare> are non-modal events.

(13) A pen-like dosimeter hangs around the neck of Katsunobu Sakurai, the tireless mayor of Minamisoma, measuring the accumulated radiation to which he has been exposed during the past two weeks of a four-week nuclear nightmare.

Purpose event An event can be presented as a purpose, aim or goal. In (14) event <MAKE room to store more toxic stuff on land> is presented as the purpose related to the *decision to dump low-level radioactive waste into the sea*. In (15) <DECARBONISE power> is presented as a purpose as well as <PROTECT the ozone layer from similar industrial gases> in (16).

(14) Neighbouring South Korea expressed concern that it was not warned about TEPCOs decision to dump low-level radioactive waste into the sea to make room to store more toxic stuff on land.

(15) The commission says the investment required to decarbonise power would average about £30 billion ($42 billion) a year over 40 years.

(16) For instance, HFC-134a and a whole family of related chemicals could be dealt with by extending the Montreal protocol created to protect the ozone layer from similar industrial gases.

Need event An event might express need or requirement. In (17) event <all that gassy baggage GO> is presented as a need, as well as event <a lot of INVESTMENT in power generation and smarter grids in (18), and <DECARBONISATION> in (19).

(17) By 2050, proposes a "road map" released by the European Commission this week, all that gassy baggage must go.

(18) The plan requires a lot of investment in power generation and smarter grids, best done in the context of –at long last– reformed and competitive energy market.

(19) Broadening climate action can supplement existing efforts on carbon and provide new suppleness to climate politics–both good things. But this does not change the imperative of decarbonisation.

Obligation event In (20) events <global greenhouse-gas emissions FALL by half to limit climate change> and <rich countries CUT the most> are considered to be presented as obligations from the perspective of *Europe*.

(20) Believing that global greenhouse-gas emissions must fall by half to limit climate change, and that rich countries should cut the most, Europe has set a goal of reducing emissions by 80-95% by 2050.

Desire event We consider desires, intentions and plans to be included under this category. In (21) event <DUMP low-level radioactive waste into the sea to make room to store more toxic stuff on land> is presented as a plan (because of *decision*). In (22) events £80 billion GO on buildings and appliances and £150 billion on transport> and <SAVE on fuel costs> are presented as plans.

(21) Neighbouring South Korea expressed concern that it was not warned about TEPCOs decision to dump low-level radioactive waste into the sea to make room to store more toxic stuff on land.

(22) This is one of the cheaper parts of the plan; the total cost is about £270 billion a year, with £80 billion going on buildings and appliances and £150 billion on transport. But the commission's modelling also points to savings on fuel costs, which are low for nuclear and zero for most renewables, of between £175 billion and £320 billion.

3.5 Condition, conditioned by

An event can be presented as a condition for another event or as conditioned by another event. In (23) event <you BE highly motivated to minimise your taxes> is a condition of event <you HUNT for every possible deduction for which you're eligible>, which is conditioned. In (24) event <active measures to remove it from the atmosphere UNDERTAKE at some later date> is considered to be a condition of event <Carbon emitted today CONTINUE to warm the planet for millennia>, which is conditioned.

(23) If you are highly motivated to minimise your taxes, you can hunt for every possible deduction for which you're eligible.

(24) Carbon emitted today will continue to warm the planet for millennia, unless active measures to remove it from the atmosphere are undertaken at some later date.

3.6 Summary of cases to be learned by systems

Systems have to be able to identify for an event the six aspects of meaning described in the previous section. All events are assigned one of the following modality types:

– Event, purpose event, need event, obligation event, desire event

If applicable, events can additionally be described with the following aspects of meaning that systems have to identify:

– Negated
– Perspective of someone other than the author
– Uncertain

- Condition for another event, conditioned by another event

So, an event description consists at least of one modality value and at most of one value per aspect of meaning.

The options provided in the multiple choice characterise and event along this five dimensions. Systems have to choose the answer that best characterises the event mentioned in the question. If no aspect apart from the modality type is mentioned in the possible answer options, we assume that the event is not negated, it is presented from the perspective from the author, it is certain or undefined qua certainty, it is not subject to a condition and it is not the condition for another event. In total there are 120 combinations, although not all of them will be represented in the test set of 12 documents because not all of them are equally frequent. The codes to be assigned to each of the values are:

1. Negated: NEG
2. Perspective of someone other than the author: PERS
3. Uncertain: UNCERT
4. Modality:
   - Event: MOD-NON
   - Purpose event: MOD-PURP
   - Need event: MOD-NEED
   - Obligation event: MOD-MUST
   - Desire event: MOD-WANT
5. Condition:
   - Condition for another event: COND
   - Conditioned by another event: COND-BY

The combinations of codes that conform the answers to the questions can be summarized with the following regular expression:

(25) $[COND|COND - BY]?\ NEG?\ PERS?\ UNCERT?\ MOD\ [-NEED| - NON| - PURP| - MUST| - WANT]$

4 Discussion

Although initially some groups inquired about the setting of the pilot task and declared to be prospective participants, in the end no systems submitted results. One of the reasons can be that the systems that participated in the main task were not designed to answer the type of questions defined for the pilot task and that the timeline of the Lab did not allowed time enough to modify the systems for the pilot task. On the other hand, systems that could be ready to process some aspect of modality and negation, like scope labelers, do not typically participate in machine reading tasks and are not fully prepared to deal with the five aspect of meaning selected for this pilot task.

As we see it, it would be possible to build a baseline system by using some of the existing scope labelers and/or designing a rule-based system. The task can be performed in three steps. First, the event needs to be located in the text and the sentence where the event occurs needs to be extracted. Although not all cases can be solved at sentence level, we could consider that working at this level would be acceptable to produce a baseline. Second, the sentence where the event occurs has to be processed to determine which of the values to assign for each of the five aspects of meaning. Third, an answer combining the tags for all aspects can be generated and it can be checked whether one of the multiple choice options contains the generated answer. Else, the most similar answer can be chosen.

To perform the second step, each of the meaning aspects should be processed apart. To determine whether the event is negated, a negation scope labeler could be use to determine whether the event is within the scope of a negation cue. The same procedure could be use to determine whether an event is uncertain, but using a hedge scope labeler. A factuality profiler like DeFacto (32) could also be used to determine whether an event is uncertain. To determine whether an event is conditioned or conditional, the syntactic structure of the sentence could be exploited to find whether the event is embedded in a conditional structure. To determine whether the event is presented from the perspective of someone other than the author it would be necessary to gather a list of expressions that are used to indicate perspective. Coreference resolution would also be needed to determine whether the pronouns corefer with the pronouns that refer to the author or not. As for the type of event qua modality, a combination of lexical look-up and syntactic analysis could help determining whether the event is presented as a purpose, need, obligation, or desire event.

Obviously, the task is much more complex than that. However, the approach would be sufficient to produce an informed baseline.

Acknowledgments

Bibliography

[1] Aikhenvald, A.: Evidentiality. Oxford University Press, New York, USA (2004)

[2] Aronow, D., Fangfang, F., Croft, W.: Ad hoc classification of radiology reports. JAMIA 6(5), 393–411 (1999)

[3] Averbuch, M., Karson, T., Ben-Ami, B., Maimon, O., Rokach, L.: Context-sensitive medical information retrieval. In: Proceedings of the 11th World Congress on Medical Informatics (MEDINFO-2004). pp. 1–8. IOS Press, San Francisco, CA (2004)

[4] Chapman, W., Bridewell, W., Hanbury, P., Cooper, G., Buchanan, B.: A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform 34, 301–310 (2001)

[5] de Marneffe, M.C., Maccartney, B., Grenager, T., Cer, D., Rafferty, A., Manning, C.: Learning to distinguish valid textual entailments. In: Proceedings of the Second PASCAL Challenges Workshop on Recognising Textual Entailment (2006)

[6] Diab, M., Levin, L., Mitamura, T., Rambow, O., Prabhakaran, V., Guo, W.: Committed belief annotation and tagging. In: ACL-IJNLP 09: Proceedings of the Third Linguistic Annotation Workshop. pp. 68–73 (2009)

[7] Etzioni, O., Banko, M., , Cafarella, M.J.: Machine reading. In: Proceedings of the 21st National Conference on Articial Intelligence (2006)

[8] Farkas, R., Vincze, V., Móra, G., Csirik, J., Szarvas, G.: The CoNLL 2010 shared task: Learning to detect hedges and their scope in natural language text. In: Proceedings of the CoNLL2010 Shared Task. Association for Computational Linguistics, Uppsala, Sweden (2010)

[9] Friedman, C., Alderson, P., Austin, J., Cimino, J., Johnson, S.: A general natural–language text processor for clinical radiology. JAMIA 1(2), 161–174 (1994)

[10] Ganter, V., Strube, M.: Finding hedges by chasing weasels: Hedge detection using wikipedia tags and shallow linguistic features. In: Proceedings of the ACL-IJCNLP 2009 Conference Short Papers. pp. 173–176. Suntec, Singapore (2009)

[11] Goldin, I., Chapman, W.: Learning to detect negation with 'Not' in medical texts. In: Proceedings of ACM-SIGIR 2003 (2003)

[12] Goryachev, S., Sordo, M., Zeng, Q., Ngo, L.: Implementation and evaluation of four different methods of negation detection. Technical report, DSG (2006)

[13] Harabagiu, S., Hickl, A., Lacatusu, F.: Negation, contrast and contradiction in text processing. In: Proceedings of the 21st International Conference on Artificial Intelligence. pp. 755–762 (2006)

[14] Hyland, K.: Hedging in scientific research articles. John Benjamins B.V, Amsterdam (1998)

[15] Kilicoglu, H., Bergler, S.: Recognizing speculative language in biomedical research articles: a linguistically motivated perspective. BMC Bioinformatics 9(Suppl 11), S10 (2008)

[16] Kim, J., Zhang, Z., Park, J., Ng, S.K.: BioContrasts: extracting and exploiting protein-protein contrastive relations from biomedical literature. Bioinformatics 22, 597–605 (2006)

[17] Linguistic Data Consortium: ACE (Automatic Content Extraction) English annotation guidelines for relations. Tech. Rep. Version 6.2 2008.04.28, LDC (2008)

[18] Marco, C., Kroon, F., Mercer, R.: Using hedges to classify citations in scientific articles. In: Croft, W.B., Shanahan, J., Qu, Y., Wiebe, J. (eds.) Computing Attitude and Affect in Text: Theory and Applications, The Information Retrieval Series, vol. 20, pp. 247–263. Springer Netherlands (2006)

[19] Medlock, B.: Exploring hedge identification in biomedical literature. JBI 41, 636–654 (2008)

[20] Medlock, B., Briscoe, T.: Weakly supervised learning for hedge classification in scientific literature. In: Proceedings of ACL 2007. pp. 992–999 (2007)

[21] Morante, R., Daelemans, W.: Learning the scope of hedge cues in biomedical texts. In: Proceedings of BioNLP 2009. pp. 28–36. Boulder, Colorado (2009)

[22] Morante, R., Daelemans, W.: A metalearning approach to processing the scope of negation. In: Proceedings of CoNLL 2009. pp. 28–36. Boulder, Colorado (2009)

[23] Mutalik, P., Deshpande, A., Nadkarni, P.: Use of general-purpose negation detection to augment concept indexing of medical documents. a quantitative study using the UMLS. J Am Med Inform Assoc 8(6), 598–609 (2001)

[24] Øvrelid, L., Velldal, E., Oepen, S.: Syntactic scope resolution in uncertainty analysis. In: Proceedings of the 23rd International Conference on Computational Linguistics. pp. 1379–1387. COLING '10, Association for Computational Linguistics, Stroudsburg, PA, USA (2010)

[25] Özgür, A., Radev, D.: Detecting speculations and their scopes in scientific text. In: Proceedings of EMNLP 2009. pp. 1398–1407. Singapore (2009)

[26] Peñas, A., Rodrigo, A.: A simple measure to assess non-response. In: Proceedings of 49th Annual Meeting of the Association for Computational Linguistics - Human Language Technologies (ACL-HLT 2011). Association for Computational Linguistics (June 2011)

[27] Ritter, A., Soderland, S., Downey, D., Etzioni, O.: It's a contradiction - no, it's not: A case study using functional relations. In: Proceedings of EMNLP 2008. pp. 11–20. Honolulu, Hawai (2008)

[28] Rokach, L., Romano, R., Maimon, O.: Negation recognition in medical narrative reports. Information Retrieval Online 11(6), 499–538 (2008)

[29] Rubin, V.L.: Identifying certainty in texts. Ph.D. thesis, Siracuse University, Syracuse, NY, USA (2006)

[30] Rubin, V.L.: Stating with certainty or stating with doubt: intercoder reliability results for manual annotation of epistemically modalized statements. In: NAACL '07: Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers on XX. pp. 141–144. Association for Computational Linguistics, Morristown, NJ, USA (2007)

[31] Rubin, V., Liddy, E., Kando, N.: Certainty identification in texts: Categorization model and manual tagging results. In: Computing Attitude and Affect in Text: Theory and Applications, Information Retrieval Series, vol. 20, pp. 61–76. Springer-Verlag, New York (2005)

[32] Saurí, R.: A factuality profiler for eventualities in text. Ph.D. thesis, Brandeis University, Waltham, MA, USA (2008)

[33] Saurí, R., Pustejovsky, J.: FactBank: A corpus annotated with event factuality. Language Resources and Evaluation 43(3), 227–268 (2009)

[34] Snow, R., Vanderwende, L., Menezes, A.: Effectively using syntax for recognizing false entailment. In: Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics. pp. 33–40. Association for Computational Linguistics, Morristown, NJ, USA (2006)

14

[35] Szarvas, G.: Hedge classification in biomedical texts with a weakly supervised selection of keywords. In: Proceedings of ACL 2008. pp. 281–289. Association for Computational Linguistics, Columbus, Ohio, USA (2008)

[36] Tottie, G.: Negation in English speech and writing: a study in variation. Academic Press, New York (1991)

[37] Vincze, V., Szarvas, G., Farkas, R., Móra, G., Csirik, J.: The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes. BMC Bioinformatics 9((Suppl 11)), S9 (2008)

[38] Wiebe, J., Wilson, T., Bruce, R., Bell, M., Martin, M.: Learning subjective language. Computational Linguistics 30(3), 277–308 (2004)

[39] Wilson, T., Hoffmann, P., Somasundaran, S., Kessler, J., Wiebe, J., Choi, Y., Cardie, C., Riloff, E., Patwardhan, S.: OpinionFinder: a system for subjectivity analysis. In: Proceedings of HLT/EMNLP on Interactive Demonstrations. pp. 34–35. Association for Computational Linguistics, Morristown, NJ, USA (2005)

[40] Wilson, T., Wiebe, J., Hoffman, P.: Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of HLT-EMNLP. pp. 347 – 354 (2005)