Semantic Processing for Text Entailment with *VENSES*

Rodolfo Delmonte, Sara Tonelli, Rocco Tripodi
Department of Language Sciences
Università Ca' Foscari – Ca' Bembo
30123, Venezia, Italy
delmont@unive.it

Abstract

In this paper we present two new mechanisms we created in VENSES, the system for semantic evaluation of the University of Venice. The first mechanism is used to match predicate-argument structures with different governors, a verb and a noun, respectively in the Hypothesis and the Text. It can be defined Augmented Finite State Automata (FSA) which are matching procedures based on tagged words in one case, and dependency relations in another. In both cases, a number of inferences – the augmentation - is fired to match different words. The second mechanism is based on the output of our module for anaphora resolution. Our system produces antecedents for pronominal expressions and equal nominal expressions. On the contrary, no decision is taken for "bridging" expressions. So the "bridging" mechanism is activated by the Semantic Evaluator and has access to the History List and the semantic features associated to each referring expression. If constraint conditions meet, the system looks for a similar association of property/entity in web ontologies like Umbel, Yago and DBPedia. The two mechanisms have been proven to contribute a 5% and 3% accuracy, respectively.

1. Introduction

Venses, the system for semantic processing, represents a linguistically-based approach for semantic inference [2] which is built around a neat division of labour between two main components. The first is a grammatically-driven subsystem which is responsible for the level of predicate-arguments well-formedness and works on the output of a deep parser that produces augmented head-dependency structures. A second subsystem fires allowed logical and lexical inferences on the basis of different types of structural transformations intended to produce a semantically valid meaning correspondence. VENSES has a discourse level anaphora resolution module, coupled to a full-fledged semantic interpreter, which is paramount to allow entailment in pairs where the relevant portion of text contains pronominal expressions.

The system is organized into twelve layers as described below – but see [1,3,5] for a complete description:

• Tokenizer and sentence splitting;
• Tagger from dictionary lookup or from morphological analysis;
• Tag disambiguation with finite-state automata and the aid of lexical information;
• Head-based Chunk building phase;
• Recursive argument/adjunct (A/A) constituent building procedure as a list of syntactic-semantic structures with tentative GFs labels;
• Clause builder that takes as input the A/A vector and tries to split it into separate clauses;
• Recursive clause-level interpretation procedure, that filters displaced or discontinuous constituents;
• Complex sentence organizer which outputs DAG structures;
• Logical Form with syntactic indices and Semantic Roles;
• Transducer from DAGs to AHDSs by recursive calls;
• Pronominal Binding at clause level followed by Anaphora Resolution at intersentential level;
• Semantic Module builds propositional level feature vectors, which also contain discourse relations.

Text entailment in VENSES is interpreted and implemented in four different steps:

- semantic similarity: in general terms by searching entities and events of the T/H pairs and comparing them on the basis of dictionaries, computational lexica and ontologies;
- propositional level constraints: this is what we also call General Consistency checks which are targeted to high level semantic attributes like

presence of modality, negation, and opacity operators, the latter ones are expressed either by the presence of discourse markers of conditionality or by a secondary level relation intervening between the main predicate and a governing higher predicate belonging to the class of non factual verbs. The governing head predicate is responsible for the factitivity of the dependent. Opacity is determined by type of governing predicates, basically those belonging to the class of nonfactive predicates. Modality is revealed by the presence of modal verbs at this level of computation. Modality could also be instantiated at sentence level by adverbials, and be verified by General Consistency Checks. Finally, negation may be expressed locally as adjunct of the verb, but also as negative conjunction and negative adverbial. It may also be present in the determiner of the nominal head and checked separately when comparing referring expressions considered in the inference. Negation may also be lexically incorporated in the verb;

- if general constraints are satisfied, we look for argument and adjuncts correspondences and for their possible paraphrases, taking care of inadmissible cases on the basis of semantic roles;
- inside arguments and adjuncts, we look for quantifiers and generic numerical specification by searching modifiers of nominal heads; we also look for non intersective modifiers;
- at adjunct level we take care of spatio-temporal modifiers if any and check for their semantics.

The problem posed by this year text level RTE which needs to be addressed fully requires a full-fledged system for anaphora resolution both at pronominal and nominal (bridging) level (see [4] for a complete description. What we mean by a full-fledged system will be clarified here below: suffice it to say that in order for bridging to apply we regard it mandatory that anaphora resolution be enacted by a Topic Hierarchy mechanism in the vein of what Grosz and Sidner suggested in their papers [7,8].

2. Two New Mechanisms at Work in Getaruns

In this year version of the Semantic Evaluator we implemented two new mechanisms. The first one – discussed in the next section – uses the output of our system for anaphora resolution which is based on Sidner and Grosz's intuition on the importance of the presence of a Topic Hierarchy in addition to a mechanism for Centering. The system looks for antecedents of pronominal expressions, and anchors for bridging coreference of non identical nominal expressions. The mechanism implemented tries to bridge the gap between linguistic knowledge and commonsense knowledge or knowledge of the world, represented here as UMBEL, the web ontology.

The second one – discussed in this section - uses Augmented FSA for nominal paraphrases, i.e. it searches for nominal paraphrases of the copulative construction contained in the Hypothesis. In some cases the construction may be governed by what we define as a "light verb" construction, that is a verb that may be paraphrased by a preposition in a nominal construction: some such verbs are "locate, situate" which may be represented as "in" in the corresponding NP.

The Augmented FSA are organized either as matching procedures based on tagged words, or on dependency structures. In both cases, additional inferential processes are called for in order to match non identical linguistic descriptions. Here below are some example to illustrate the two techniques.

2.1. Using Augmented FSA for SE with Nominal Paraphrases

The procedure takes as input the tagged list of words making up the Hypothesis and tries to match with the relevant portion of the Text that contains similar words and tags, as follows

match_template(Hypothesis, Text)

If the match succeeds the semantic evaluation outputs a value that is indicative of the type of decision taken. This matching produce is reached by the analysis only after General Consistency Checks have passed at higher propositional level. Consider the first example where we highlight the portions of T/H relevant for the semantic evaluation:

> RTE5 – T/H Pair 364 – Entailment=True
> Trains, trams, cars and buses ground to a halt on Monday after a shoot-out between 18:00 CET and 19:00 CET in the historical ***city of Basel in Switzerland***. The first shots were fired at around 16:00 CET according to the spokesperson of the Government. Special Police Forces were getting prepared to launch an assault against this building.

The authors of the shoot-out are unknown as well as their target. Railway traffic was blocked off at around 18:00 CET, according to the Swiss Federal Railways' (SBB-CFF-FFS) spokesperson, Christian Krìuchi. At around 19:00 CET, railway traffic resumed .
Basel is a European city.

In more detail, Augmented Finite State Automata mean that in addition to equality matching that is at the basis of the whole algorithm, the system looks for inferences and other lexical information to authorize the match. In fact, these procedures as a whole allow the matching to become more general with constraints though. The instructions reported below are expressed in Prolog which treats words constituted by or beginning with upper case letters to be treated as variables. Constants on the contrary are written with lower case letter, as for instance the words "of" and "in" below.

```
match_template   ([A,Is-_,T-_,F-_,G|Hyp],[G,of-_,A,in-_,L-_|Text]):-
   lightsvbs(Is),
   high_rank(T,Lex),
   locwn(L),
   is_in(L,F1),
   (natl(F1,F,_);natl(F1,_,F)), !.
```

where the procedure "lightverbs" looks for copulative verbs, i.e. the verb of the Hypothesis must be a copulative verb; "high_rank" looks for high frequency words like articles; "locwn" verifies that the word present in the variable "L" is a location. Then there are two inferences: the first one is fired by the call "is_in" that recovers the name of the continent in which "L" belongs, thus implicitly requiring "L" to be a name of a nation. Then the second inference looks for the corresponding nationality adjective. Values for the variables are then as follows:
L --> Switzerland
F1 --> Europe
F --> European
A = Basel-np, Is = is, T = a, G = city-n

2.2. Using Dependency Relations for SE with Nominal Paraphrases

In the second mechanism we activated, FSA are used to match dependency relations. Consider the second example:

RTE3 – T/H Pair 173 – Entailment=True
Prince Laurent of Belgium, the youngest son of King Albert II of Belgium, has been questioned last night by the federal police and is attending today's court session in Hasselt in a marine fraud case that has gripped Belgian media since last December.
King Albert II of Belgium is the father of Prince Laurent.

Here below we list the augmented dependency relations for the text and hypothesis where the indices at the end of each term are taken from the corresponding syntactic constituent and identify uniquely in each sentence the semantic heads:

```
T.  'King_Albert' - ncmod / specif - 'Belgium' - sn6,
    'King_Albert' - det - 'II' - sn5,
    'Prince_Laurent' – ncmod/specif - son - sn1,
    son - ncmod / specif - 'King_Albert' - sn5

H.  'King_Albert' - ncmod / specif - 'Belgium' - sn3,
    'King_Albert' - det - 'II' - sn1,
    father - ncmod / specif - 'Prince_Laurent' - sn4
    be - xcomp / prop - father - sn2,
    'King_Albert' - subj / theme_unaff - sn1
```

where we see that relations are reversed and need to be checked carefully. In the text we know that Prince Laurent is the son of King Albert of Belgium; on the contrary in the Hypothesis we are told that King Albert of Belgium is the father of Prince Laurent. Now, what is needed is the possibility to draw inferences about the complementariness in the relation existing between ***father*** and ***son***.
The code related to the semantic evaluator is shown here below:

```
searchmatchtemprels(Text_rels,Hypos_rels):-
   remove(be-xcomp/prop-R-Sn, Hypos_rels,Rests),
   relatives(Rels),
   on(R,Rels),
   remove(R-xcomp/prop-Sn2,Rests,Resto),
   remove(Head-subj/Rol-Sn1,Resto,Rest),
   remove(R1-ncmod/specif-Head1-Sn3,
Text_rels,Rest1),
   on(R1,Rels),
   Head=Head1,
  (compl_rel(R1,R);
    compl_rel(R,R1)),
   …
```

compl_rel(father,son).
compl_rel(father,daughter).
…

The procedure looks for a copulative construction in the Hypothesis and then checks to see whether the property predicated is one of the set of "relatives". If yes, it removes the property and the predicated subject of the property to use it for matching purposes with the corresponding relations in the Text. Then it removes modifier relations of specification in the Text pool of relations where the governing head is identical to the subject head of the Hypothesis and the property head is in a complementary relation with the corresponding head found in the Text. The following portion of the procedure will recursively eliminate all identical relations until the Hypothesis pool is empty.

3. Topic Hierarchy And Bridging

A conspicuous number of T/H pairs of this year datasets is characterized by the need to look for semantic attributes and properties asserted in the Hypothesis in different sentences contained in the Text. So that the possibility to match T/H one sentence at a time and produce a sufficient environment for the semantic evaluation is simply no longer available. One such case is represented by Intersective adjectives and is presented here below:

> RTE5 – TH Pair 81 - Entailment=True
> ***The pope flew by helicopter from Vatican City to a tent camp near the village of Onna***, where he led a prayer in the cold and rainy weather for the hundreds killed in the April 6 quake. ***The camp*** houses hundreds of families left ***homeless*** when the magnitude-6.3 quake destroyed their homes. Residents there welcomed the pope, who kissed and hugged some of the children. "I have come here personally to this splendid and hurt land of yours, which is living days of great pain and precariousness, to express in the most direct way my kind closeness", the pope told residents. "I've followed the news with apprehension, sharing with you your consternation... for the dead, along with your anxious worries about how much you've lost in a brief moment.
> ***The pope reached the homeless camp near Onna by helicopter***.

In order to know precisely where the pope actually landed with his helicopter, the adjective "homeless" is needed. But to recover this piece of information, the system has to be able to bridge the information expressed in the first sentence to the information of the second one. This can only be done with a system of anaphora resolution that can compute the bridging relation [9,10] intervening between the first occurrence of CAMP as an ***indefinite expression***, and the second occurrence in second sentence as ***definite expression***. This will be explained in more details below.
One such system is shown in Fig. 1 below, where we highlight the architecture and main processes undergoing at the anaphora level. First of all, the subdivision of the system into two levels: Clause level – intrasentential pronominal phenomena – where all pronominal expressions contained in modifiers, adjuncts or complement clauses receive their antecedent locally. Possessive pronouns, pronouns contained in relative clauses and complement clauses choose preferentially their antecedents from list of higher level referring expressions. Not so for those pronouns contained in matrix clauses. In particular the ones in subject position are to be coreferred in the discourse. This requires the system to be equipped with a History List of all referring expressions to be used when needed.

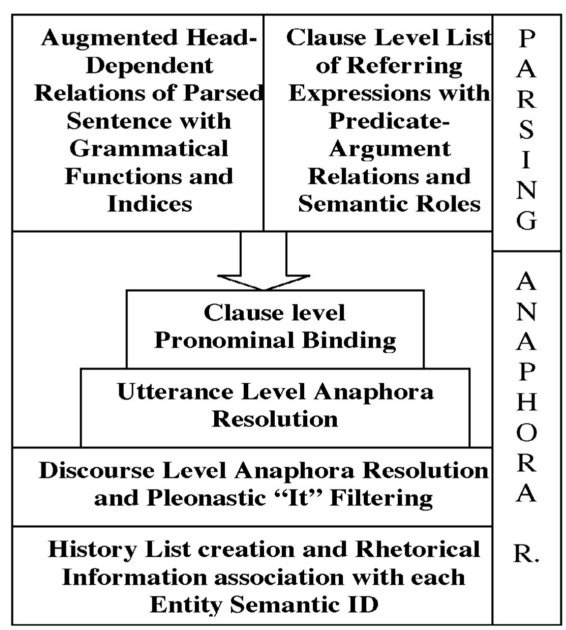

Fig. 1 Anaphoric Processes in VENSES

In the system, three levels are indicated: Clause level, i.e. simple sentences; Utterance level, i.e. complex sentences; Discourse level, i.e. intersententially. Our system computes semantic structures in a sentence by sentence fashion and any information useful to carry out anaphoric processes needs to be made available to the following portion of text, and eventually to the Semantic Evaluation that computes entailment. As a first case, I will present an example of pronominal anaphora resolution which requires the system to identify genre and number in order to reject local possible antecedents. The pronoun we are referring to is the first HER highlighted in the text below:

> RTE5–TestSet - TH Pair 42 - Entailment=True
> DENVER - ***Angie Zapata*** was a tall woman with striking black hair and eyes who would attract the attention of men, even those who knew she was biologically male. But prosecutors say when Allen Andrade found out, he ***beat her to death with a fire extinguisher***. Her sister discovered her battered body under a blanket in her Greeley apartment last July. Andrade, 32, of Thorton, is scheduled to go on trial Tuesday on charges including first-degree murder and a bias-motivated crime, which could add three years to his prison sentence if convicted.
> Angie Zapata has been killed with a fire extinguisher.

Of course the system has to be able to deal with feminine Names as opposed to masculine ones. Then there must be a History List and a Centering mechanisms. Both are needed in order to choose the appropriate antecedent for HE and HER. Obviously syntactic structure needs to be present in order to be able to assign Allen_Andrade as possible antecedent of HE: the pronoun and the proper name are positioned in different clauses and do not undergo have to obey to disjointness constraints. Eventually, the Semantic Evaluator has to be able to compute passive structures and compare it to active structures, and this is done by looking at Semantic Roles and Grammatical Functions. Then the last problem to be solved is the semantic similarity of KILL with BEAT_TO_DEATH that must be defined somewhere if the entailment has to hold.

As commented above, the idea is that every time we have a predication in the Hypothesis that associated a property to a proper noun that we don't find in the Text – that is and the semantic evaluation based on predicate-argument structures has failed - we switch to the Bridging coreference module. The Hypothesis will typically contain two linguistic descriptions associated by means of some "light verb". The module will look for one of the nominal expressions used in the Hypothesis either in definite, indefinite or proper name linguistic description, in a sentence that precedes the one containing the other linguistic descriptions of the Hypothesis. This procedure starts by searching in the History List where referring expressions are listed by sentence number and their rhetorical label in the Topic Hierarchy – i.e. they can be computed as either Main Topic, Secondary Topic or Potential Topic [6,7,8]. We search for Potential or Main Topics because the definite description would constitute a new linguistic item in the History List; on the contrary Secondary Topic would be used only in case the linguistic description has already been asserted as Main Topic in previous stretch of discourse and is accompanied by another Topic.

The mechanism implemented tries to bridge the gap between linguistic knowledge and commonsense knowledge or knowledge of the world, represented here as web ontologies like UMBEL, YAGO and DBPEDIA [11]. In fact, what we do is looking for external knowledge whenever our linguistic procedures require it. The reason for this move is motivated as follows:

- linguistic descriptions to be matched and related by a bridging coreference link are DIFFERENT
- the usual semantic relations made available in WordNet and also by the lexical fields of Moby have failed
- we don't have enough confidence due to the presence of other possible coreferents

In other words, this strategy is used to restrain, reduce, limit access to external knowledge of the world to those cases that really require it. Thus eliminating all T/H pairs which are lacking such enabling conditions and must be regarded as FALSE cases.

We will comment a number of significant examples to clarify the way in which our system operates.

3.1. Topic Hierarchy And Bridging

We chose three examples from the two sets Development and Test, and indicated in a headline the configuration of Bridging Expression and its preceding Coreferent.

Proper Name + Definite Expression
Rte5 – DevSet - TH Pair 10

> (CNN) -- Malawians are rallying behind ***Madonna*** as she awaits a ruling Friday on whether she can adopt a girl from the southern African nation. ***The pop star***, ***who has three children***, adopted a son from Malawi in 2006. She is seeking to adopt Chifundo "Mercy" James, 4. "Ninety-nine percent of the people calling in are saying, let her take the baby," said Marilyn Segula, a presenter at Capital FM, which broadcasts in at least five cities, including the capital, Lilongwe.
> ***Madonna has three children***.

Indefinite Expression + Proper Name
Rte5 – TestSet - TH Pair 83

> ***A Ugandan spy*** who set up a bogus charity and embezzled thousands of dollars of funding meant for Aids patients has been jailed for 10 years. ***Teddy Sseezi Cheeye***, 51, took $56,000 (£38,000) from the Global Fund charity, which aims to prevent HIV, tuberculosis and malaria. He set up an NGO, the Uganda Centre for Accountability, which received cash in 2005 to do HIV/Aids community work. But the High Court in Kampala heard Cheeye siphoned off the funds instead.
> ***Teddy Sseezi Cheeye is an Ugandan spy***.

Definite Expression + Proper Name
Rte5 – TestSet - TH Pair 269

> The eruption happened at around 1:30 PM local time, the United States Geological Survey reported. ***The volcano*** had erupted four times on Friday, billowing ash up to 51,000 feet up into the air. These are the latest in a series of eruptions from ***Mount Redoubt***, which started on March 22. The volcano had not erupted since a four-month period in 1989-90. ***The Alaska Volcano Observatory*** set its alert level at red, the highest possible level, meaning that an eruption is imminent, and that it would send a "significant emission of volcanic ash into the atmosphere."
> ***Mount Redoubt is located in Alaska***.

As can be gathered from the headers and the highlighted portions of texts, the cases to be covered all involve a proper noun which can be either a person's name or a location. There are three different configurations to account for, which require basically a search for the type of definiteness and/or semantic type of the head noun. These information are in our encoded in the vector associated to each referring expression semantic head, as follows:

ref_ex(sn1, Madonna, [+ref, def0, nil, nil, -pro, -ana, -class], 3, fem, sing, [human], subj/theme_unaff)
ref_ex(sn1, existence, [+ref, +def, very, nil, -pro, -ana, +class], 3, neut, sing, [place, state], subj/actor)
ref_ex(sn9, it, [+ref, +def, nil, nil, +pro, -ana, +le], 3, neu, sing, [any], subj/agent)

As can be seen from representations, proper nouns are marked def0, +ref and –class; on the contrary common nouns are marked +def/-def, +ref and +class. Pronouns do not have the attribute CLASS but +/-le which stands for Lexically Expressed. The vector includes Functional Features – Person, Gender, Number – and Semantic Features in the sense of General Nouns or Inherent Features. At the end of the vector or Prolog term, we report grammatical function and semantic role associated to the head noun which can be found in syntactic and dependency representations by means of the index, positioned at the beginning.

4. Evaluation and Ablation test

The evaluation results we present try to give a comprehensive picture of the system performance over the overall datasets made available with RTE. It is worthwhile reminding that the first 2 challenges contained very short Texts if compared to the Text average size of the following challenges. In particular then, Texts contained in RTE4 testset and RTE5 development and test sets are much longer then those contained in RTE3. The difference in treatment of these datasets is quite obvious: modeling a paragraph long text is certainly much harder. In addition, RTE5 texts have a certain number of T/H pairs where the contents of the Hypothesis is scattered amongst a number of sentences in the paragraph. This makes the task much harder than in all those cases in which semantic matching can be concentrated on just one sentence in the Text paragraph.
As will be noticed in the data reported in Table 1. below, there is a remarkable difference in the results obtained in the Development and the Test set: 10 percent point accuracy. A possible reason for this is the fact that RTE5 Testset contains a lot more cases of difficult to spot entailment relations.

It is a fact, that a great number of T/H pairs contain Texts where the relevant relations are scattered in more than one sentence, thus making the semantic matching task harder to perform.

Subtask	Accuracy(%)
ir	61.00
qa	65.00
ie	58.50
Average	**61.50**

Table 1: Official results for Run 1 of our system – No Ranking

Subtask	Precision(%)
ir	58.62
qa	67.14
ie	66.00
Average	**64.45**

Table 2: Official results for Run 2 of our system – Ranking

Results for past RTE datasets as a whole fare on average 63% - but see table 4. below. Results for the Contradiction Dataset, are as follows:
Accuracy measured as ratio of Correct Pairs/All Pairs: 108/131 = 0.8245
Results for the Development and the Test set of RTE 5, are as follows:
DEVELOPMENT set: Accuracy measured as ratio of Correct Pairs/All Pairs: 0.73
It is important to notice that in all cases with no exception whatsoever, the percentage of True T/H pairs found is higher than the percentage of False.

4.1 Ablation Test

We carried out one ablation test where we removed matching procedures related to Grady Ward's MOBY Thesaurus as well as to Roget's Thesaurus. In fact what we eliminated was a procedure which used "lexical fields" as semantic similarity matching in all cases of non identical lemmas. We used this procedure after eliminating cases of antonymy which could degrade the semantic similarity matching. After the filter for antonyms, matching was carried out on lemmas as usual. Access to Thesaura can in some cases contribute important and relevant information, but this is not always guaranteed as shown by the results of the test reported here below. In particular, we may notice that in one case, IR subtask, we improved accuracy by 0.045 points. So, even though in the remaining subtasks there is always a reduction of the overall accuracy, it is interesting to notice that not all task behave in the same way.
This type of "sloppy" semantic similarity matching is fired every time the system needs approximated or fuzzy similarity information. In particular, it is never permitted whenever precise information is required, as for instance in what we call General Consistency Checking procedures. These procedures are carried out to check for the presence of Quantified Expressions, information related to Spatio-Temporal Location, and any kind of numerical information present in the Hypothesis that has to be present also in the Text. On the contrary, whenever we look for attributes, modifiers and other similar adjuncts of the arguments expressed in predicate-argument structure, we allow access to lexical fields contained in thesaura. This may also apply in all copulative constructions, whenever a certain property is being associated to the subject of the predication.
These matching procedure are scattered all over the evaluation algorithm: what we did was simply dummifying the access to the matching procedures, by inserting a dummy couple of values – nil, nil – in place of the two variables that had to be taken into consideration by the matching procedure, and inserted a cut – in Prolog an instruction not to allow recursion and oblige a failure – in place of the procedure itself, which was hidden.

Subtask	Accuracy1	Accuracy2
ir	61.00	65.50
qa	65.00	58.00
ie	58.50	52.50
Average	**61.50**	**58.67**

Table 3: Ablation Test results compared with Run1

5. Conclusions and Future Work

We presented our improvements to VENSES, our system for semantic evaluation, which uses a proprietor complete system of text analysis based on a deep system called GETARUNS. We introduced a number of new modules that take advantage of the output of the anaphora resolution algorithm and exploit its representations to attempt bridging coreference. In case constraints are respected, the system looks for similar relations in web ontologies, to confirm the anaphoric link. We also implemented Augmented FSA both at tagging and at dependency levels. The results are very encouraging and we saw an improvement of 8% overall.

We report here below a table with the overall results the system obtains on all RTE datasets.

Table 4: *VENSES* overall results with RTE datasets

	Corr. False	Corr. True	Prec.False	Prec.True	Tot. Precision	Comput. Time	No. sentences
RTE1 DEV1	96/145	90/142	65.51	64.79	65.16	38m	578
RTE1 DEV2	75/140	85/140	47.86	60.71	54.28	26m	564
RTE1 TEST	234/400	236/400	58.5	59.0	58.75	51m	1603
RTE2 DEV	231/400	256/400	57.75	64.0	60.87	45m	1647
RTE2 TEST	228/400	268/400	56.75	69.5	63.12	58m	1657
RTE3 DEV	280/388	267/412	68.3	69.17	68.75	1h	1900
RTE3 TEST	245/390	267/410	62.82	65.12	64.0	48m	1774
RTE4	354/500	251/500	70.8	50.2	60.5	1h 44m	2688
RTE5 DEV	210/300	221/300	70.0	73.67	71.83	1h 31m	3180
RTE5 TEST	208/300	189/300	69.33	63.0	66.17	1h 29m	3243
CONTRADIC	120/131		92.0		92.0	15m	294
totals false totals true	2281/3363 2130/3404		719.62	641.99	724.74	9h 37m	20828
Mean precs			67.83%	62.57%	63.27%		

References

1. Delmonte R., 2009. Computational Linguistic Text Processing – Lexicon, Grammar, Parsing and Anaphora Resolution, Nova Science Publishers, New York.
2. Delmonte R., A. Bristot, M.A.Piccolino Boniforti, S.Tonelli (2007), Entailment and Anaphora Resolution in RTE3, in Proc. ACL Workshop on Text Entailment and Paraphrasing, Prague, ACL Madison, USA, pp. 48-53.
3. Delmonte R., 2007. Computational Linguistic Text Processing – Logical Form, Semantic Interpretation, Discourse Relations and Question Answering, Nova Science Publishers, New York.
4. Delmonte R., A. Bristot, M. A. Piccolino Boniforti and Sara Tonelli, 2006. Another Evaluation of Anaphora Resolution Algorithms and a Comparison with GETARUNS' Knowledge Rich Approach, ROMAND 2006, 11th EACL, Trento, ACL, 3-10.
5. Delmonte R.(2002), GETARUN PARSER - A parser equipped with Quantifier Raising and Anaphoric Binding based on LFG, Proc. LFG2002 Conference, Athens, at http://cslipublications.stanford.edu/hand/miscpubsonline.html, 130-153.
6. Delmonte, R., Bianchi D.: Binding Pronominals with an LFG Parser. In: Proceeding of the Second International Workshop on Parsing Technologies, Cancun (Mexico), ACL. (1991) 59-72
7. B. J. Grosz and C. L. Sidner. 1986. Attention, intention, and the structure of discourse. Computational Linguistics, 12(3):175–204.
8. B. J. Grosz, A. K. Joshi, and S. Weinstein. 1995. Centering. Computational Linguistics, 21(2):202–225.
9. K. Markert, M. Strube, and U. Hahn. 1996. Inferential realization constraints on functional anaphora in the centering model. In Proc. of 18th Conf. of the Cog. Science Soc., pages 609–614.
10. K. Markert, M. Nissim, and N.. Modjeska. 2003. Using the Web for nominal anaphora resolution. In Proc. of the EACL Workshop on the Computational Treatment of Anaphora, pages 39–46.
11. N. Modjeska, K. Markert, and M. Nissim. 2003. Using the Web in ML for anaphora resolution. In Proc. of EMNLP-03, pages 176–183.