

Scaling up a NLU system from text to dialogue understanding

R. Delmonte, A. Bristot, G. Voltolina
Department of Language Science -
Università Ca' Foscari - 30123 -
VENEZIA
delmont@unive.it

Vincenzo Pallotta
Webster University, Geneva
Switzerland
pallotta@webster.ch

Abstract

In this paper we will present work carried out to scale up the system for text understanding called GETARUNS, and port it to be used in dialogue understanding. We will present the adjustments we made in order to cope with transcribed spoken dialogues like those produced in the ICSI Berkely project. In a final section we present preliminary evaluation of the system on non-referential pronominals individuation.

1 Introduction

Very much like other deep linguistic processing systems (see Allen et al.), our system is a generic text/dialogue understanding system that can be used in connection with an ontology – WordNet – and/or a repository of commonsense knowledge like CONCEPTNET. Word sense disambiguation takes place at the level of semantic interpretation and is represented in the Discourse Model.

Computing semantic representations for spoken dialogues is a particularly hard task which – when compared to written text processing – requires the following additional information to be made available:

- adequate treatment of fragments;
- adequate treatment of short turns, in particular one-word turns;
- adequate treatment of first person singular and plural pronominal expressions;
- adequate treatment of disfluencies, thus including cases of turns made up of just such expressions, or cases when they are found inside the utterance;
- adequate treatment of overlaps;
- adequate treatment of speaker identity for pronominal coreference;

In our system, then, every dialogue turn receives one polarity label, indicating negativity or

positivity, and this is computed by looking into a dictionary of polarity items. This is subsequently used to decide on argumentative automatic classification.

The Berkeley ICSI dialogues are characterized by the need to argument in an exhaustive manner the topics to be debated which are the theme of each multiparty dialogue. The mean length of utterances/turns in each dialogue we parsed was rather long.

2 The System GETARUNS

GETARUNS¹, the system for text understanding developed at the University of Venice, is organized as a pipeline which includes two versions of the system: what we call the Partial and the Deep GETARUNS (Delmonte 2007;2009). The Deep version is equipped with three main modules: a lower module for parsing, where sentence strategies are implemented; a middle module for semantic interpretation and discourse model construction which is cast into Situation Semantics; and a higher module where reasoning and generation takes place.

2.1 The Algorithm for Overlaps

Overlaps are an important component of all spoken dialogue analysis. In all dialogue transcription, overlaps are treated as a separate turn from the one in which they occur, which usually follows it. On the contrary, when computing overlaps we set as our first goal that of recovering the temporal order. This is done because overlaps may introduce linguistic elements which influence the local context. Eventually, they may determine the interpretation of the current utterance.

¹ The system has been tested in STEP competition, and can be downloaded at, <http://project.cgm.unive.it/html/sharedtask/>.

For these reasons, they cannot be moved to a separate turn because they must be semantically interpreted where they temporally belong.

The algorithm we built looks at time stamps, and everytime the following turn begins at a time preceding the ending time of current turn it enters a special recursive procedure. It looks for internal interruption in the current turn and splits the utterance where the interruption occurs. Then it parses it split initial portion of current utterance and continues with the overlapping turn. This may be reiterated in case another overlap follows which again begins before the end of current utterance. Eventually, it returns to the analysis of the current turn with the remaining portion of current utterance.

2.2 The Treatment of Fragments and Short Turns

Fragments and short turns are filtered by a lexical lookup procedure that searches for specific linguistic elements which are part of a list of backchannels, acknowledgements expressions and other similar speech acts. In case this procedure has success, no further computation takes place. However, this only applies to utterances shorter than 5 words, and should be made up only of such special words. No other linguistic element should be present apart from non-words, that is words which are only partially produced and have been transcribed with a dash at the end. Otherwise we proceed as follows:

- graceful failure procedures for ungrammatical sentences, which might be fullfledged utterances but semantically uninterpretable due to the presence of repetitions, false starts and similar disfluency phenomena. Or else they may be just fragments, i.e. partial or incomplete utterances, hence non-interpretable as such; this is done by imposing grammatical constraints of wellformedness in the parser.

We implemented a principled treatment of elliptical utterances and contribute one specific speech act. They may express agreement/ disagreement, acknowledgements, assessments, continuers etc. All these items are computed as being complements of abstract verb SAY which is introduced in the analysis, and has as subject, the name of current speaker.

3 The Experiment

We set up an experiment in order to test the new version of the system, that is detecting referential from nonreferential uses of personal pronouns “you”, “we” and “it”.

In order to take decisions as to whether pronouns are to be interpreted as referential or not a recursive procedure checks the type of governing predicate. Referential pronouns are then passed on to the pronominal binding algorithm that looks for local antecedents if any. Otherwise, the pronouns is labeled as having External coreference in the previous discourse stretch. The Anaphora Resolution module will then take care of the antecedent and a suitable semantic identifier will be associated to it. On the contrary, if the pronouns are judged to be referentially empty or generic, no binding takes place. Here below is a table containing total values for pronouns WE/YOU/IT in all the 10 dialogues analysed.

	Referential	Generic	Total
WE	1186	706	1892
YOU	1045	742	1787
IT	1593	1008	2601
Total	3824	2456	6280

Table 1. Overall count of pronominal expressions

Results for the experiment are as follows

	Recall	Precision	F-Score
WE	98.2%	60.59%	74.94%
YOU	99.3%	70.99%	82.79%
IT	97.6%	64.2%	77.45%

Table 2. Results for pronominal expressions

References

- Allen, J., M. Dzikovska, M. Manshadi, and M. Swift. 2007. Deep linguistic processing for spoken dialogue systems. In *ACL 2007 Workshop on Deep Linguistic Processing*, pp. 49–56.
- Delmonte R. 2007. *Computational Linguistic Text Processing – Logical Form, Semantic Interpretation, Discourse Relations and Question Answering*, Nova Science Publishers, New York.
- Delmonte R. 2009. *Computational Linguistic Text Processing – Lexicon, Grammar, Parsing and Anaphora Resolution*, Nova Science Publishers, New York.