Non-Sentential Utterances in Dialogue: Experiments in classification and interpretation

Paolo Dragone
Sapienza University of Rome (Italy)
dragone.paolo@gmail.com

Pierre Lison
University of Oslo (Norway)
plison@ifi.uio.no

Abstract

We present two ongoing experiments related to the classification and interpretation of non-sentential utterances (NSUs). Extending the work of Fernández et al. (2007), we first show that the classification performance of NSUs can be improved through the combination of new linguistic features and active learning techniques. We also describe a new, hybrid approach to the semantic interpretation of NSUs based on probabilistic rules.

1 Introduction

In dialogue, utterances do not always take the form of complete, well-formed sentences. Many utterances – often called *non-sentential utterances*, or NSUs for short – are indeed fragmentary and lack an overt predicate, as in the following examples from the British National Corpus:

A: How do you actually feel about that?
B: **Not too happy.** [BNC: JK8 168-169]

A: They wouldn't do it, no.
B: **Why?** [BNC: H5H 202-203]

A: [...] then across from there to there.
B: **From side to side.** [BNC: HDH 377-378]

Although these types of NSUs are extremely common, their semantic content is often difficult to extract automatically. NSUs are indeed intrinsically dependent on the dialogue context – for instance, the meaning of ”why” in the example above is impossible to decipher without knowing the statement that precedes it.

We report here on two ongoing experiments. The first experiment focuses on the automatic classification of NSUs according to the taxonomy of Fernández et al. (2007), while the second experiment develops a new approach to the semantic interpretation of NSUs using the probabilistic rules formalism developed by Lison (2015)

2 Classifying NSUs

Non-sentential utterances can serve several types of pragmatic functions, such as providing feedback, asking for clarifications, answering questions or correcting/extending previous utterances.

Fernández et al. (2007) provide a taxonomy of NSUs based on 15 classes as well as a small corpus of annotated NSUs extracted from dialogue transcripts of the British National Corpus. They also present classification experiments using the above-mentioned corpus and taxonomy. We extend their approach through a combination of feature engineering and semi-supervised learning. Semi-supervised learning is used to cope with the scarcity of labelled data for this task. This lack of sufficient training data is especially problematic due to the strong class imbalance between the NSU classes. Furthermore, the most infrequent classes are often the most difficult ones to discriminate. Fortunately, the BNC also contains a large amount of unlabelled NSUs that can be extracted from the raw dialogue transcripts using simple heuristics (syntactic patterns to select utterances that are most likely non-sentential).

One particular technique that we employed in this empirical study is Active Learning. The objective of Active Learning (AL) is to interactively query the user to annotate novel data by selecting the most informative instances (that is, the ones that are most difficult to classify) and avoiding redundant ones.[1] In practice, we applied the active learning algorithm to extract and annotate 100 new instances of NSUs, which were subsequently added to the existing training data.

[1] We used the Java library JCLAL for this purpose, cf. http://sourceforge.net/projects/jclal/.

In order to determine the baseline for our study, we replicated the classification experiment described in Fernández et al. (2007) using the same feature set. This initial set comprised a total of 9 linguistic features extracted from the NSU and its antecedent. We then developed an extended feature set, adding 23 new syntactic and similarity features on top of the ones used in the baseline. Weka's SMO package (based on SVMs) was used to train the classifiers for all experiments.[2]

The empirical results were extracted through 10-fold cross-validation (for the active learning case, the newly annotated instances were added to the training set of each fold). The results demonstrate that the above approach is able to provide modest but significant improvements over the baseline, as illustrated in Table 1. Using a paired t-test with a 95% confidence interval between the baseline and the final result, the improvement in classification accuracy is statistically significant with a p-value of 6.9×10^{-3}.

Experimental setting	**Accuracy**
Train-set (initial features)	0.881
Train-set (extended features)	0.899
Train-set + AL (initial features)	0.883
Train-set + AL (extended features)	**0.907**

Table 1: Summary of the classification accuracy for the baseline and new approach.

The evaluation results illustrate that the active learning approach is only beneficial when combined with the extended (more informative) feature set, while it does not provide any significant improvement on the set of baseline features.

Our experiments demonstrate the potential of the combination of linguistically-informed features and larger amounts of training data for the classification of non-sentential utterances. Of special interest would be the annotation and analysis of NSUs in other dialogue domains than the ones covered in the current corpus.

3 Interpreting NSUs

Non-sentential utterances cannot be interpreted in isolation from their surrounding context. As argued by e.g. (Fernández, 2006; Ginzburg, 2012), NSUs are best described in terms of *update rules* on the current dialogue state. Their framework is however purely logic-based, making it difficult to account for the fact that many state variables are only partially observed (due to e.g. imperfect understanding of the dialogue and its context).

To remedy this shortcoming, we are currently rewriting the update rules for NSUs detailed in (Ginzburg, 2012) using the probabilistic rules formalism of (Lison, 2015). Probabilistic rules indeed share many commonalities with Ginzburg's framework, as both approaches rely on update rules expressed in terms of conditions and effects operating on a rich dialogue state. However, probabilistic rules can also operate on uncertain (probabilistic) knowledge, making them more robust than traditional logical rules.

We are using the OpenDial toolkit[3] to implement the above approach. Crucially, the approach integrates in its pipeline the classifier presented in the previous section in order to derive the most likely class for each NSU. We plan to use a portion of the COMMUNICATOR corpus (Walker et al., 2001) to evaluate the performance of the interpretation rules on real-world dialogues.

4 Conclusion

This abstract presented two ongoing experiments related to the automatic processing of non-sentential utterances in dialogue. The first experiment shows how the use of more expressive linguistic features and active learning can improve the classification accuracy of NSUs. The second experiment focuses on the robust interpretation of NSUs in context based on probabilistic rules.

References

R. Fernández, J. Ginzburg, and S. Lappin. 2007. Classifying non-sentential utterances in dialogue: A machine learning approach. *Computational Linguistics*, 33(3):397–427.

R. Fernández. 2006. *Non-Sentential Utterances in Dialogue: Classification, Resolution and Use*. Ph.D. thesis, King's College London.

J. Ginzburg. 2012. *The Interactive Stance*. Oxford University Press.

P. Lison. 2015. A hybrid approach to dialogue management based on probabilistic rules. *Computer Speech & Language*, 34(1):232 – 255.

M. Walker, B. Pellom, J. Polifroni, A. Potamianos, P. Prabhu, A. Rudnicky, S. Seneff, and D. Stallard. 2001. DARPA Communicator dialog travel planning systems: The June 2000 data collection. In *Proceedings of EUROSPEECH*, pages 1371–1374.

[2] http://www.cs.waikato.ac.nz/ml/weka/

[3] http://opendial-toolkit.net