Using Syntactic and Semantic based Relations for Dialogue Act Recognition

Tina Klüwer, Hans Uszkoreit, Feiyu Xu

Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI)

Projektbüro Berlin

{tina.kluewer,uszkoreit,feiyu}@dfki.de

Abstract

This paper presents a novel approach to dialogue act recognition employing multi-level information features. In addition to features such as context information and words in the utterances, the recognition task utilizes syntactic and semantic relations acquired by information extraction methods. These features are utilized by a Bayesian network classifier for our dialogue act recognition. The evaluation results show a clear improvement from the accuracy of the baseline (only with word features) with 61.9% to an accuracy of 67.4% achieved by the extended feature set.

1 Introduction

Dialogue act recognition is an essential task for dialogue systems. Automatic dialogue act classification has received much attention in the past years either as an independent task or as an embedded component in dialogue systems. Various methods have been tested on different corpora using several dialogue act classes and information coming from the user input.

The work presented in this paper is part of a dialogue system called *KomParse* (Klüwer et al., 2010), which is an application of a NL dialogue system combined with various question answering technologies in a three-dimensional virtual world named *Twinity*, a web-based online product of the Berlin startup company *Metaversum*[1]. The KomParse NPCs provide various services through con-

---

[1] http://www.metaversum.com/

versation with game users such as selling pieces of furniture to users via text based conversation.

The main task of the input interpretation component of the agent is the detection of the dialogue acts contained in the user utterances. This classification is done via a cue-based method with various features from multi-level knowledge sources extracted from the incoming utterance considering a small context of the previous dialogue.

In contrast to existing systems using mainly lexical features, i.e. words, single markers such as punctuation (Verbree et al.,) or combinations of various features (Stolcke et al., 2000) for the dialogue act classification, the results of the interpretation component presented in this paper are based on syntactic and semantic relations. The system first gathers linguistic information coming from different levels of deep linguistic processing similar to (Allen et al., 2007). The retrieved information is used as input for an information extraction component that delivers the relations embedded in the actual utterance (Xu et al., 2007). These relations combined with additional features (a small dialogue context and mood of the sentence) are then utilized as features for the machine-learning based recognition.

The classifier is trained on a corpus originating from a Wizard-of-Oz experiment which was semi-automatically annotated. It contains automatically annotated syntactic relations namely, predicate argument structures, which were checked and corrected manually afterwards. Furthermore these relations are enriched by manual annotation with semantic frame information from VerbNet to gain an additional level of semantic richness. These two representations of relations, the syntax-based re-

lations and the VerbNet semantic relations, were used in separate training steps to detect how much the classifier can benefit from either notations.

A systematic analysis of the data has been conducted. It turns out that a comparatively small set of syntactic relations cover most utterances, which can moreover be expressed by an even smaller set of semantic relations. Because of this observation as well as the overall performance of the classifier the interpretation is extended with an additional rule based approach to ensure the robustness of the system.

The paper is organized as follows: Section 2 provides an overview about existing dialogue act recognition systems and the features they use for classification.
Section 3 introduces the original data used as basis for the annotation and the classification task.
In Section 4 the annotation that provides the necessary information for the dialogue act classification and involves the relation extraction is described in detail. The annotation is split into three main steps: The annotation of dialogue information (section 4.1), the integration of syntactic information (section 4.2) and finally the manual annotation of VerbNet predicate and role information in section 4.3.
Section 5 presents the results of the actual classification task using different feature sets and in Section 6 the results and methods are summarized.
Finally, Section 7 provides a brief description of the rule-based interpretation and presents an outlook on future work.

2 Related Work

Dialogue Acts (DAs) represent the functional level of a speaker's utterance, such as a greeting, a request or a statement. Dialogue acts are verbal or nonverbal actions that incorporate participant's intentions originating from the theory of Speech Acts by Searle and Austin (Searle, 1969). They provide an abstraction from the original input by detecting the intended action of an utterance, which is not necessarily inferable from the surface input (see the two requests in the following example).

Can you show me a red car please?

Please show me a red car!

To detect the action included in an utterance, different approaches have been suggested in recent years which can be clustered into two main classes: The first class uses AI planning methods to detect the intention of the utterance based on belief states of the communicating agents and the world knowledge. These systems are often part of an entire dialogue system e.g. in a conversational agent which provides the necessary information about current beliefs and goals of the conversation participants at runtime. One example is the TRIPS system (Allen et al., 1996). Because of the huge amount of reasoning, systems in this class generally gather as much linguistic information as possible.

The second class uses cues derived from the actual utterance to detect the right dialogue act, mostly using machine learning methods. This class gained much attention due to less computational costs. The probabilistic classifications are carried out via training on labeled examples of dialogue acts described by different feature sets. Frequently used cues for dialogue acts are lexical features such as the words of the utterance or ngrams of words for example in (Verbree et al.,), (Zimmermann et al., 2005) or (Webb and Liu, 2008). Although the performance of the classification task is difficult to compare, because of the variety of different corpora, dialogue act sets and algorithms used, these approaches do provide considerably good results. For example (Verbree et al.,) achieve accuracy values of 89% on the ICSI Meeting Corpus containing 80.000 utterances with a dialogue act set of 5 distinct dialogue act classes and amongst others the features "ngrams of words" and "ngrams of POS information".

Another group of systems utilizes acoustic features derived from Automatic Speech Recognition for automatic dialogue act tagging (Surendran and Levow, 2006), context features like the preceding dialogue act or ngrams of previous dialogue acts (Keizer and Akker, 2006).

However grammatical and semantic information is not that often incorporated into feature sets, with the exception of single features such as the

Dialogue Act	Meaning	Frequency
REQUEST	The utterance contains a wish or demand	449
REQUEST_INFO	The utterance contains a wish or demand regarding information	154
PROPOSE	The utterance serves as suggestion or showing of an object	216
ACCEPT	The utterance contains an affirmation	167
REJECT	The utterance contains a rejection	88
PROVIDE_INFO	The utterance provides an information	156
ACKNOWLEDGE	The utterance is a backchannelling	9

Table 1: The used Dialogue Act Set

type of verbs or arguments or the presence or absence of special operators e.g. wh-phrases (Andernach, 1996). (Keizer et al., 2002) use among others linguistic features like sentence type for classification with Bayesian networks. Although (Jurafsky et al., 1998) already noticed a strong correlation between selected dialogue acts and special grammatical structures, approaches using grammatical structure were not very succesful.

While grammatical and semantic features are not often incorporated into dialogue act recognition, they are a commonly used in related fields like automatic classification of rhetorical relations. For example (Sporleder and Lascarides, 2008) and (Lapata and Lascarides, 2004) extract verbs as well as their temporal features derived from parsing to infer sentence internal temporal and rhetorical relations. Their best model for analysing temporal relations between two clauses achieves 70.7% accuracy. (Subba and Eugenio, 2009) also show a significant improvement of a discourse relation classifier incorporating compositional semantics compared to a model without semantic features. Their VerbNet based frame semantics yield in a better result of 4.5%.

3   The Data

The data serving as the basis for the relation identification as well as the training corpus for the dialogue act classifier is taken from a Wizard-of-Oz experiment (Bertomeu and Benz, 2009) in which 18 users furnish a virtual living room with the help of a furniture sales agent. Users buy pieces of furniture and room decoration from the agent by describing their demands and preferences in a text chat. During the dialogue with the agent, the preferred objects are then selected and directly put to the right location in the apartment. In the exper-

iments, users spent one hour each on furnishing the living room by talking to a human wizard controlling the virtual sales agent. The final corpus consists of 18 dialogues containing 3,171 turns with 4,313 utterances and 23,015 alpha-numerical strings (words). The following example shows a typical part of such a conversation:

> USR.1: And do we have a little side table for the TV?
> NPC.1: I could offer you another small table or a sideboard.
> USR.2: Then I'll take a sideboard that is similar to my shelf.
> NPC.2: Let me check if we have something like that.

Table 2: Example Conversation from the Wizard-of-Oz Experiment

4   Annotation

The annotation of the corpus is carried out in several steps.

4.1   Pragmatic Annotation

The first annotation step consists of annotating discourse and pragmatic information including dialogue acts, projects according to (Clark, 1996), sentence mood, the topic of the conversation and an automatically retrieved information state for every turn of the conversations. From the annotated information the following elements were selected as features in the final recognition system:

- The dialogue acts which carry the intentions of the actual utterance as well as the last preceding dialogue act. The set used for annotation is a domain specific set containing the dialogue acts shown in table 1.

- The sentence mood. Sentence mood was annotated with one of the following values: declarative, imperative, interrogative.

- The topic of the utterance. The topic value is coreferent with the currently discussed object. Topic can consist of an object class (e.g. sofa) or an special object instance (sofa_1836). The topic of the directly preceding utterance was chosen as a feature too.

4.2 Annotation with Predicate Argument Structure

The second annotation step, applied to the utterance level of the input, automatically enriches the annotation with predicate argument structures. Each utterance is parsed with a predicate argument parser and annotated with syntactic relations organized according to PropBank (Palmer et al., 2005) containing the following features: Predicate, Subject, Objects, Negation, Modifiers, Copula Complements.

A single relation mainly consists of a predicate and the belonging arguments. Verb modifiers like attached PPs are classified as "argM" together with negation ("argM_neg") and modal verbs ("argM_modal"). Arguments are labeled with numbers according to the found information for the actual structure. PropBank is organized in two layers, the first one being an underspecified representation of a sentence with numbered arguments, the second one containing fine-grained information about the semantic frames for the predicate comparable to FrameNet (Baker et al., 1998). While the information in the second layer is stable for each verb, the values of the numbered arguments can change from verb to verb. While for one verb the "arg0" may refer to the subject of the verb, another verb may encapsulate a direct object behind the same notation "arg0". This is very complicated to handle in a computational setup, which needs continuous labeling for the successive components. Therefore the arguments were in general named as in PropBank but consistently numbered by syntactic structure. This means for example that the subject is always labeled as "arg1".

Consider the example "Can you put posters or pictures on the wall?". The syntactic relation will yield in the following representation:

```
<predicate: put>
<ArgM_modal: can>
<Arg1: you>
```

```
<Arg2: posters or pictures>
<ArgM: on the wall>
```

Predicate Argument Structure Parser The syntactic predicate argument structure that constitutes the syntactic relations and serves as basis for the VerbNet annotation, is automatically retrieved by a rule-based predicate argument parser. The rules utilized by the parser describe subtrees of dependency structures in XML by means of relevant grammatical functions. For detecting verbs with two arguments in the input, for instance, a rule can be written describing the dependency structure for a verb with a subject and an object. This rule would then detect every occurrence of the structure "Verb-Subj-Obj" in a dependency tree. This sample rule would express the following constraints: The matrix unit should be of the part of speech "Verb" , The structure belonging to this verb must contain a "nsubj" dependency and an "obj" dependency.

The rules deliver raw predicate argument structures, in which the detected arguments and the verb serve as hooks for further information lookup in the input. If a verb fulfills all requirements described by the rule, in a second step all modificational arguments existing in the structure are recursively acquired. The same is done for modal arguments as well as modifiers of the arguments such as determiners, adjectives or embedded prepositions. After the generation of the main predicate argument structure from the grammatical functions, the last step inserts the content values present in the actual input into the structure to get the syntactic relations for the utterance.

Before the input can be parsed with the predicate argument parser, some preprocessing steps of the corpus are needed. These include:

Input Cleaning The input data coming from the users contain many errors. Some string substitutions as well as the external Google spellchecker were applied to the input before any further processing.

Segmentation For clausal separation we apply a simple segmentation via heuristics based on punctuation.

POS Tagging Then the input is processed by

the external part-of-speech tagger TreeTagger (Schmid, 1994).

The embedded dependency parser is the Stanford Dependency Parser (de Marneffe and Manning, 2008), but other dependency parsers could be employed instead. The predicate argument parser is an standalone software and can be used either as a system component or for batch processing of a text corpus.

4.3 VerbNet Frame Annotation

The last step of annotation consists of the manual annotation of semantic predicate classes and semantic roles. Moreover, the automatically determined syntactic relations are checked and corrected if possible. VerbNet (Schuler, 2005) is utilized as a source for semantic information. The VerbNet role set consists of 21 general roles used in all VerbNet classes. Examples of roles in this general role set are "agent", "patient" and "theme".

For the manual addition of the semantic frame information a web-based annotation tool has been developed. The annotation tool shows the utterance which should be annotated in the context of the dialogue including the information from the preceding annotation steps. All VerbNet classes containing the current predicate are listed as possibilities for the predicate classification together with their syntactic frames. The annotators can select the appropriate predicate class and frame according to the arguments found in the utterance. If an argument is missing in the input that is required in the selected frame a null argument is added to the structure. If the right predicate class is existing, but the predicate is not yet a member of the class, it is added to the VerbNet files. In case the right predicate class is found but the fitting frame is missing, the frame is added to the VerbNet files. Thus during annotation 35 new members have been added to the existing VerbNet classes, 4 Frames and 4 new subclasses. Via these modifications, a version of VerbNet has been developed that can be regarded as a domain-specific VerbNet for the sales domain.

During the predicate classification, the annotators also assign the appropriate semantic roles to the arguments belonging to the selected predicate.

The semantic roles are taken from the selected VerbNet frame.

From the annotated semantic structure, semantic relations are inferred such as the one in the following example:

```
<predicate: put-3.1>
<agent: you>
<theme: posters or pictures>
<destination: on the wall>
```

5 Dialogue Act Recognition

Two datasets are derived from the corpus: The dataset containing the utterances of the users (CST) and one dataset containing the utterances of the wizard (NPC), whereas the NPC corpus is cleaned from the "protocol sentences". Protocol sentences are canned sentences the wizard used in every conversation, for example to initialize the dialogue. For the experiments, the two single datasets "NPC" and "CST" as well as a combined dataset called "ALL" are used. Unfortunately from the original 4,313 utterances in total, many utterances could not be used for the final experiments. First, fragments are removed and only the utterances found by the parser to contain a valid predicate argument structure are used. After protocol sentences are taken out too, a dataset of 1702 valid utterances remains. Moreover, 292 utterances are annotated to contain no valid dialogue act and are therefore not suitable for the recognition task. Of the remaining utterances, 171 predicate argument structures were annotated as wrong because of completely ungrammatical input. In this way we arrive at a dataset of 804 instances for the users and 435 for the wizard, summing up to 1239 instances in total.

The features used for dialogue act recognition exploit the information extracted from the different annotation steps:

- Context features: The last preceding dialogue act, equality between the last preceding topic and the actual topic, sentence mood

- Syntactic relation features: Syntactic predicate class, arguments, negation

- VerbNet semantic relation features: VerbNet predicate class, VerbNet frame arguments, negation

- Utterance features: The original utterances without any modifications

Different sets of features for training and evaluation are generated from these:

DATASET_Syn: All utterances of the specified dataset described via syntactic relation and context features.

DATASET_VNSem: All utterances of the specified dataset described via VerbNet semantic relations and context features.

DATASET_Syn_Only: All utterances of the specified dataset only described via the syntactic relations.

DATASET_VNSem_Only: All utterances of the specified dataset only described via the VerbNet semantic relations.

DATASET_Context_Only: All utterances of the specified dataset described via the context features and negation without any information regarding relations.

DATASET_Utterances_Context: The utterances of the specified dataset as strings combined with the whole set of context features without further relation extraction results.

DATASET_Utterances: Only the utterances of the specified dataset as strings. This and the last "Utterances"-set serve as baselines.

Dialogue Act Recognition is carried out via the Bayesian network classifier AOEDsr from the WEKA toolkit. AODEsr augments AODE, an algorithm averaging over all of a small space of alternative naive-Bayes-like models that have weaker independence assumptions than naive Bayes, with Subsumption Resolution (Zheng and Webb, 2006). Evaluation is performed using crossfolded evaluation.

All results of the experiments are given in terms of accuracy.

Results for the dataset "All" comparing the syntactic relations with VerbNet relations as well as the pure utterances and context are shown in table 4.

Dataset	Accuracy
All_Syn	67.4%
All_VNSem	66.8%
All_Utterances_Context	61.9%
All_Utterances	48.1%

Table 4: Dialogue Act Classification Results for the "ALL" Datasets

The best result is achieved with the syntactic information, although the VerbNet information provides an abstraction over the predicate classification. Both the set containing the VerbNet relations as well as the syntactic relations are much better than the set containing only the context and the original utterances. The dataset containing only the utterances could not reach 50%.

Although the experiments show much better results using the relations instead of the original utterance, the overall accuracy is not very satisfying. Several reasons for this phenomenon come into consideration. While it can to a certain extend be the fault of the classifying algorithm (see table 8 for some tests with a ROCCHIO based classifier), the main reason might as well lie in the imprecise boundaries of the dialogue act classes: Several categories are hard to distinguish even for a human annotator as you can see from the wrongly classified examples in table 3. Another possibility can be the comparatively small number of total training instances.

For the NPC dataset the results are slightly better and much better still for the set CST, which is due to a smaller number (6) of dialogue acts: The dialogue act "PROPOSE", which is the act for showing an object or proposing a possibility, was not used by any user, but only by the wizard.

Dataset	Accuracy
CST_Syn	73.1%
NPC_Syn	68.5%

Table 5: Dialogue Act Classification Results for Datasets "CST" and "NPC"

To find out if one sort of features is especially important for the classification we reorga-

Utterance	Right Classification	Classified As
What do you think about this one?	request_info	propose
Let see what you have and where we can put it	request_info	request

Table 3: Wrongly classified instances

nize the training sets to contain only the context features without the relations (All_Context_Only) on the one hand and only the relational information without the context features on the other hand (All_Syn_Only and All_VNSem_Only). Results are shown in table 6.

Dataset	Accuracy
All_Context_Only	56.6%
All_VNSem_Only	53.5%
All_Syn_Only	50.8%

Table 6: Dialogue Act Classification Results for Context and Relation sets

Table 6 shows that the results are considerably worse if only parts of the features are used. The set with context feature performs 3,1% better than the best set with the relations only. Furthermore the VerbNet semantic relation set leads to nearly 3% better accuracy, which may mean that the abstraction of semantic predicates provides a better mapping to dialogue acts after all if used without further features which may be ranked more important by the classifier.

Besides the experiments with the Bayesian networks, additional experiments are performed using a modified ROCCHIO algorithm similar to the one in (Neumann and Schmeier, 2002). Three different datasets were tested (see table 7).

Dataset	Accuracy
All_Utterances	70.1%
All_Utterances_Context	73.2%
All_Syn	74.4%

Table 8: Dialogue Act Classification Results using the ROCCHIO Algorithm

Table 8 shows that the baseline dataset containing only the utterances already provides much bet-

ter results with the ROCCHIO algorithm, delivering 70.1% which is more than 10% more accuracy compared to the 48.1% of the Bayesian classifier. If tested together with the context features the accuracy of the utterance dataset raises to 73.2% and, after including the relational information, even to 74.4%. Thus, the results of this ROCCHIO experiment also prove that the employment of the relation information leads to improved accuracy of the classification.

6   Conclusion

This paper reports on a novel approach to automatic dialogue act recognition using syntactic and semantic relations as new features instead of the traditional features such as ngrams of words.

Different feature sets are constructed via an automatic annotation of syntactic predicate argument structures and a manual annotation of Verb-Net frame information. On the basis of this information, both the syntactic relations as well as the semantic VerbNet-based relations included in the utterances can be extracted and added to the feature sets for the recognition task. Besides the relation information the employed features include information from the dialogue context (e.g. the last preceding dialogue act) and other features like sentence mood.

The feature sets have been evaluated with a Bayesian network classifier as well as a ROCCHIO algorithm. Both classifiers demonstrate the benefits gained from the relations by exploiting the additionally provided information. While the difference between the best baseline feature set and the best relation feature set in the Bayesian network classifier yields a 5,5% boost in accuracy (61.9% to 67.4%), the ROCCHIO setup exceeds the boosted accuracy by another 1,5% , starting from a higher baseline of 73.2%. Based on the observed complexity of the classification task we expect that the benefit of the relational informa-

Predicate	Instances	Example
see-30.1	59	I would like to see a table in front of the sofa
put-9.1	74	Can you put it in the corner?
reflexive_appearance-48.1.2	80	Show me the red one
own-100	137	Do you have wooden chairs?
want-32.1	153	I would like some plants over here

Table 7: The Main Semantic Relations Found in the Data Sorted by Predicate

tion may turn out to be even more significant on larger learning data.

7 Future Work

The results in section 5 show that the pure classification cannot be used as interpretation component in isolation, but additional methods have to be incorporated. In a preceding analysis of the data it was found that certain predicates are very frequently uttered by the users. In the syntactic predicate scenario the total number of different predicates is 80, whereas the semantic predicates build up a total number of 66. The class containing the predicates with one to ten occurrences constitutes 137 of 1239 instances. The remaining 1101 instances are covered by only 21 different predicate classes. These predicates together with their arguments constitute a set of common domain relations for the sales domain. The main domain relations found are shown in table 7.

The figures suggest that the interpretation at least for the domain relations can be established in a robust manner, wherefore the agent's interpretation component was extended to a hybrid module including a robust rule based method. To derive the necessary rules a rule generator was developed and the rules covering the used feature set (including the context features, sentence mood and the syntactic relations) were automatically generated from the given data.

Future work will focus on the evaluation of these automatically derived rules on a recently collected but not yet annotated dataset from a second Wizard-of-Oz experiment, carried out in the same furniture sales setting.

Additional experiments are planned for evaluating the relation-based features in dialogue act recognition on other corpora tagged with different dialogue acts in order to test the overall performance of our classification approach on more transparent dialogue act sets.

Acknowledgements

References

Allen, James F., Bradford W. Miller, Eric K. Ringger, and Teresa Sikorski. 1996. A robust system for natural spoken dialogue. In *Proceedings of ACL 1996*.

Allen, James, Mehdi Manshadi, Myroslava Dzikovska, and Mary Swift. 2007. Deep linguistic processing for spoken dialogue systems. In *DeepLP '07: Proceedings of the Workshop on Deep Linguistic Processing*, Morristown, NJ, USA.

Andernach, Toine. 1996. A machine learning approach to the classification of dialogue utterances. *CoRR*, cmp-lg/9607022.

Baker, Collin F., Charles J. Fillmore, and John B. Lowe. 1998. The berkeley framenet project. In *Proceedings of COLING 1998*.

Bertomeu, Nuria and Anton Benz. 2009. Annotation of joint projects and information states in human-npc dialogues. In *Proceedings of CILC-09*, Murcia, Spain.

Clark, H.H. 1996. *Using Language*. Cambridge University Press.

de Marneffe, Marie C. and Christopher D. Manning. 2008. The Stanford typed dependencies representation. In *Coling 2008: Proceedings of the workshop on Cross-Framework and Cross-Domain Parser Evaluation*, Manchester, UK.

Jurafsky, Daniel, Elizabeth Shriberg, Barbara Fox, and Traci Curl. 1998. Lexical, prosodic, and syntactic cues for dialog acts.

Keizer, Simon and Rieks op den Akker. 2006. Dialogue act recognition under uncertainty using bayesian networks. *Nat. Lang. Eng.*, 13(4).

Keizer, Simon, Rieks op den Akker, and Anton Nijholt. 2002. Dialogue act recognition with bayesian networks for dutch dialogues. In *Proceedings of the 3rd SIGdial workshop on Discourse and dialogue*, Morristown, NJ, USA.

Klüwer, Tina, Peter Adolphs, Feiyu Xu, Hans Uszkoreit, and Xiwen Cheng. 2010. Talking npcs in a virtual game world. In *Proceedings of the System Demonstrations Section at ACL 2010*.

Lapata, Mirella and Alex Lascarides. 2004. Inferring sentence-internal temporal relations. In *Proceedings of the North American Chapter of the Association for Computational Linguistics*, pages 153–160.

Neumann, Günter and Sven Schmeier. 2002. Shallow natural language technology and text mining. *Künstliche Intelligenz. The German Artificial Intelligence Journal*.

Palmer, Martha, Daniel Gildea, and Paul Kingsbury. 2005. The proposition bank: An annotated corpus of semantic roles. *Comput. Linguist.*, 31(1).

Schmid, Helmut. 1994. In *Proceedings of the International Conference on New Methods in Language Processing*.

Schuler, Karin Kipper. 2005. *Verbnet: a broad-coverage, comprehensive verb lexicon*. Ph.D. thesis, Philadelphia, PA, USA.

Searle, John R. 1969. *Speech acts : an essay in the philosophy of language / John R. Searle*. Cambridge University Press, London.

Sporleder, Caroline and Alex Lascarides. 2008. Using automatically labelled examples to classify rhetorical relations: A critical assessment. *Natural Language Engineering*, 14(3).

Stolcke, Andreas, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van, and Ess dykema Marie Meteer. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, 26.

Subba, Rajen and Barbara Di Eugenio. 2009. An effective discourse parser that uses rich linguistic information. In *NAACL '09*, Morristown, NJ, USA.

Surendran, Dinoj and Gina-Anne Levow. 2006. Dialog act tagging with support vector machines and hidden markov models. In *Interspeech*.

Verbree, A.T., R.J. Rienks, and D.K.J. Heylen. Dialogue-act tagging using smart feature selection: results on multiple corpora. In Raorke, B., editor, *First International IEEE Workshop on Spoken Language Technology SLT 2006*.

Webb, Nick and Ting Liu. 2008. Investigating the portability of corpus-derived cue phrases for dialogue act classification. In *Proceedings of COLING 2008*, Manchester, UK.

Xu, Feiyu, Hans Uszkoreit, and Hong Li. 2007. A seed-driven bottom-up machine learning framework for extracting relations of various complexity. In *Proceedings of ACl (07)*, Prague, Czech Republic.

Zheng, Fei and Geoffrey I. Webb. 2006. Efficient lazy elimination for averaged one-dependence estimators. In *ICML*, pages 1113–1120.

Zimmermann, Matthias, Yang Liu, Elizabeth Shriberg, and Andreas Stolcke. 2005. Toward joint segmentation and classification of dialog acts in multiparty meetings. In *Proc. Multimodal Interaction and Related Machine Learning Algorithms Workshop (MLMI05*, page 187.