Semantic Features for Dialogue Act Recognition

Pavel Král[1,2] & Ladislav Lenc[1,2] & Christophe Cerisara[3]

[1]Dept. of Computer Science & Engineering
Faculty of Applied Sciences
University of West Bohemia
Plzeň, Czech Republic
[2]NTIS - New Technologies for the Information Society
Faculty of Applied Sciences
University of West Bohemia
Plzeň, Czech Republic
[3]LORIA-UMR7503, Nancy, France
{pkral,llenc}@kiv.zcu.cz,cerisara@loria.fr

November 24, 2015

Table of Content

Definition of the Dialogue Acts (DAs)

- introduced by Austin in 1962 [Aus62]
- developed by Hary Bunt in [Bun94]
  - DA = meaning of an utterance in the context of a dialogue
- this work
  - DA = function of an utterance, or its part, in the dialogue

Example:

- question → requesting of some information
- answer → providing this information

Applications

- dialogue systems
- machine translation
- automatic speech recognition
- topic tracking
- talking head animation
- etc.

Objectives

- to propose semantic features and integrate them into a dialogue act recognition task to improve the recognition score in Czech
- three different feature computation approaches proposed, evaluated and compared:
  - Latent Dirichlet Allocation (LDA)
  - Hyperspace Analogue to Language (HAL)
  - Correlated Occurrence Analogue to Lexical Semantics (COALS)

Related Work

Features

- lexical [J*et al.*97] (and syntactic [KC14])
- prosodic [SB98]
- dialogue history [S*et al.*00]
- **semantic** [KUX10] (few work × our focus)

Models *supervised machine learning*

- Bayesian Networks [KRN02]
- Discriminative Dynamic BN [JB05]
- Maximum Entropy [ALS05]
- Conditional Random Fields [QIR11]
- Neural Networks [LLL$^+$03]
- ...

Latent Dirichlet Allocation (LDA)

- unsupervised topic model → a topic to each word in the sentence
- semantically close words ≈ similar topics (e.g. synonyms)

- standard LDA model → a sentence topic for each word
- used together with word labels for DA recognition

Semantic Spaces

- words ≈ high dimensional semantic vectors
- semantically close words → similar vectors
- opportunity to use a clustering method to create word clusters
- two semantic space models:
  - Hyperspace Analogue to Language (HAL [LB96])
  - Correlated Occurrence Analogue to Lexical Semantic (COALS [RGP04])
- additive composition of *word-level* vectors (by HAL or COALS methods) → *sentence-level* semantic vectors
- = additional semantic information for DA recognition

Note

- never used for dialogue act recognition before

Some Assumptions about the Models

LDA

- full sentences as a context
- $\rightarrow$ long word dependencies
- $\rightarrow$ information about the topic of the conversation

HAL and COALS

- (relatively) short context window
- $\rightarrow$ local dependencies between words
- $\rightarrow$ syntactic structure information
- $\rightarrow$ important for DA recognition

Expectation

- *HAL and COALS will give better results than LDA*

Dialogue Act Recognition

- W .. sequence of n words w_i in the sentence
- F .. sequence of semantic features f_i ($i \in [1; n]$)
- C .. dialogue act class

Two classifiers used:

Naive Bayes [Ris01]

- sometimes also referred as an *unigram*
- modelling of $P(W|C)$ *(first baseline)*

Maximum Entropy (ME) [BPP96]

- modelling of $P(C|W)$ in lexical case *(second baseline)*
- $P(C|W, F)$ in semantic case

DA Corpus

- evaluation of our approach on both types of transcriptions: manual and automatic
- automatic transcription with the jLASER [PE07] recognizer
  - training on about nine hours (6234 sent.)
  - testing on about three hours (2173 sent.)

DA corpus			
DA	No.	Example	English translation
S	566	Chtěl bych jet do Písku.	I would like to go to Písek.
O	125	Najdi další vlak do Plzně!	Look at for the next train to Plzeň!
Q[y/n]	282	Řekl byste nám další spojení?	Do you say next connection?
Q	1200	Jak se dostanu do Šumperka?	How can I go to Šumperk?
Sent.	2173		

Tools & Parameters I.

LDA

- MALLET [McC02] tool-kit
- Dirichlet distributions parameters initially set to (see [GS04])
  - $\alpha = 50/K$ (K = topic number)
  - $\beta = 0.1$

HAL and COALS semantic space models

- S-Space package [JS10] for implementation
  - four-words context window in both directions
  - matrix composed of 1,000 columns
  - dimensionality reduction not used

Tools & Parameters II.

- LDA and both semantic spaces trained on the training part of the Railways corpus (i.e. 6234 sentences)
- Brainy [Kon14] tool-kit for implementation of Maximum Entropy classifier
- 10-fold cross-validation used (10% of the corpus for testing)
- confidence interval of $\pm$ 1%
- ASR Accuracy (ACC)
  - Sentence ACC = 39.78%
  - Word ACC = 83.36%

Results with Manual Word Transcription

	Accuracy in [%]				
Approach/ Classifier	S	O	Q[y/n]	Q	Glob.
1. Lexical information (baselines)					
NB	93.5	77.6	96.5	89.9	**91.0**
ME	90.3	88.0	97.2	96.5	**94.6**
2. Semantic information					
LDA + ME	93.3	87.2	96.5	98.5	**96.4**
HAL + ME	95.1	96.0	97.9	97.9	**97.2**
COALS + ME	96.1	97.6	99.3	99.2	**98.4**

Table: Dialogue acts recognition accuracy for different approaches/classifiers and their combination with manual word transcription

Results with Automatic Speech Recognition

	Accuracy in [%]				
Approach/ Classifier	S	O	Q[y/n]	Q	Glob.
1. Lexical information (baselines)					
NB	93.1	68.8	94.7	86.3	**88.2**
ME	87.5	77.6	89.7	95.2	**91.6**
2. Semantic information					
LDA + ME	88.3	80.8	89.0	96.3	**92.5**
HAL + ME	92.2	86.4	93.6	96.9	**94.8**
COALS + ME	95.9	96.8	97.5	99.0	**98.0**

Table: Dialogue acts recognition accuracy for different approaches/classifiers and their combination with word transcription by ASR

Conclusions & Perspectives

- three approaches to create semantic features proposed, implemented and evaluated on the Czech Railways corpus
- semantic features **important** for dialogue act recognition
- semantic spaces, HAL & COALS, significantly outperform the LDA model
- explanation: **semantic spaces** $\approx$ modelling of local dependencies between words $\times$ **LDA** $\approx$ global word dependency

Perspectives

- adaptation and evaluation of the proposed methods on larger corpora and in other languages with more dialogue acts
- evaluation of the other classifiers

Acknowledgements

- This work has been partly supported by the project LO1506 of the Czech Ministry of Education, Youth and Sports.

NTIS

FACULTY OF APPLIED SCIENCES UNIVERSITY OF WEST BOHEMIA

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND
INVESTING IN YOUR FUTURE

2007-13
OP Research and Development for Innovation

[ALS05] J. Ang, Y. Liu, and E. Shriberg. Automatic dialog act segmentation and classification in multiparty meetings. In *Proc. ICASSP*, Philadelphia, USA, March 2005.

[Aus62] J. L. Austin. How to do Things with Words. *Clarendon Press, Oxford*, 1962.

[BPP96] Adam L Berger, Vincent J Della Pietra, and Stephen A Della Pietra. A maximum entropy approach to natural language processing. *Computational linguistics*, 22(1):39–71, 1996.

[Bun94] Harry Bunt. Context and Dialogue Control. *Think Quarterly*, 3:19–31, May 1994.

[GS04] Thomas L. Griffiths and Mark Steyvers. Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl 1):5228–5235, April 2004.

[JB05] G. Ji and J. Bilmes. Dialog Act Tagging Using Graphical Models. In *Proc. ICASSP*, volume 1, pages 33–36, Philadelphia, USA, March 2005.

[*Jet al.*97] D. Jurafsky *et al.* Automatic Detection of Discourse Structure for Speech Recognition and Understanding. In *IEEE Workshop on Speech Recognition and Understanding*, Santa Barbara, 1997.

[JS10] D. Jurgens and K. Stevens. The s-space package: An open source package for word space models. *In System Papers of the Association of Computational Linguistics*, 2010.

[KC14] P. Král and C. Cerisara. Automatic dialogue act recognition with syntactic features. *Language Resources and Evaluation*, 48(3):419–441, 8 February 2014.

[Kon14] Michal Konkol. Brainy: A machine learning library. In Leszek Rutkowski, Marcin Korytkowski, Rafał Scherer, Ryszard Tadeusiewicz, Lotfi A. Zadeh, and Jacek M. Zurada, editors, *Artificial Intelligence and Soft Computing*, volume 8468 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2014.

[KRN02] S. Keizer, Akker. R., and A. Nijholt. Dialogue Act Recognition with Bayesian Networks for Dutch Dialogues. In

3rd ACL/SIGdial Workshop on Discourse and Dialogue, pages 88–94, Philadelphia, USA, July 2002.

[KUX10] Tina Klüwer, Hans Uszkoreit, and Feiyu Xu. Using syntactic and semantic based relations for dialogue act recognition. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, pages 570–578. Association for Computational Linguistics, 2010.

[LB96] K. Lund and C. Burgess. Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods Instruments and Computers*, 28(2):203–208, 1996.

[LLL+03] L. Levin, C. Langley, A. Lavie, D. Gates, D. Wallace, and K. Peterson. Domain Specific Speech Acts for Spoken Language Translation. In *4th SIGdial Workshop on Discourse and Dialogue*, Sapporo, Japan, 2003.

[McC02] Andrew Kachites McCallum. Mallet: A machine learning for language toolkit. 2002.

[PE07] Tomáš Pavelka and Kamil Ekštein. JLASER: An automatic speech recognizer written in Java. In *XII International Conference Speech and Computer (SPECOM'2007)*, pages 165–169, Moscow, Russia, 2007.

[QIR11] S. Quarteroni, A. V. Ivanov, and G. Riccardi. Simultaneous dialog act segmentation and classification from human-human spoken conversations. In *Proc. ICASSP*, Prague, Czech Republic, May 2011.

[RGP04] D. L. T. Rohde, L. M. Gonnerman, and D. C. Plaut. An improved method for deriving word meaning from lexical co-occurrence. *Cognitive Psychology*, 7:573–605, 2004.

[Ris01] Irina Rish. An empirical study of the naive bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, volume 3, pages 41–46. IBM New York, 2001.

[SB98] E. Shriberg and R. Bates. *Language and Speech*, volume 41 of *Special Double Issue on Prosody and Conversation*, chapter Can Prosody Aid the Automatic Classification of Dialog Acts in Conversational Speech?, pages 439–487. 1998.

[S*et al.*00] A. Stolcke *et al.* Dialog Act Modeling for Automatic Tagging and Recognition of Conversational Speech. In *Computational Linguistics*, volume 26, pages 339–373, 2000.