Preliminary Results on Dialogue Act Classification in Chat-based Online Tutorial Dialogues

Vasile Rus[1], Rajendra Banjade[1], Nabin Maharjan[1], Donald Morrison[1], Steve Ritter[2], Michael Yudelson[3]

University of Memphis[1], Carnegie Learning, Inc[2]., Carnegie Mellon University[3]

Introduction

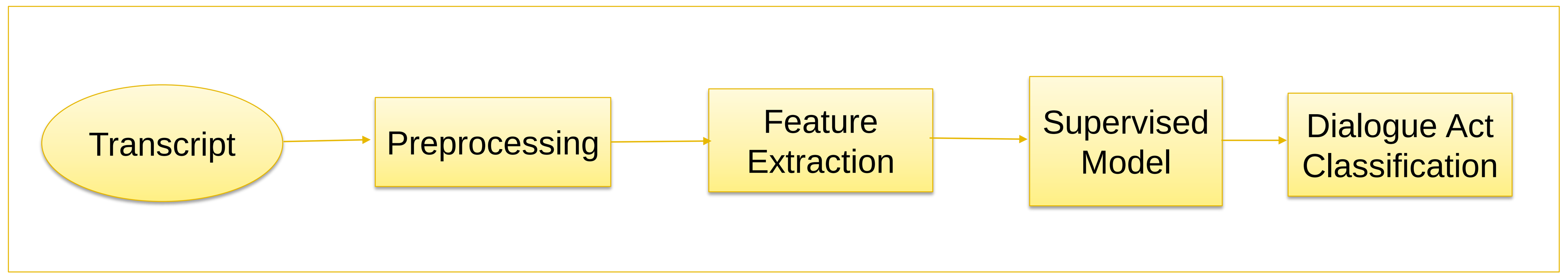
- A key research question in Intelligent Tutoring Systems is understanding the intentions and actions of the speakers in dialogues that relate to various aspects of learning and affect.
- Actions of speakers can be represented by dialogue acts inspired from the **language-as-action** theory [1] : "*when we say something we do something*"
- So, we map all utterances in a tutorial dialogue onto corresponding dialogue acts using a predefined dialogue act **taxonomy**.

Speaker	Utterance	Dialogue Act	Dialogue SubAct
…	…	…	…
Tutor	Have you learnt the "Substitution Method" to solve a system of linear equations?	Request	Confirmation:PriorKnowledge
Student	you take the x from the second equation and substitute it in for the y in the first equation	Assertion	Approach
Tutor	Yes! Do you want to apply this method to solve the current problem?	Request	Confirmation:Approach
Student	Aren't we suppose to or would it be easier to use elimination or what ever that word is	Request	Confirmation:Approach
…	…	…	…

Mapping utterances to dialog acts

Approach

- A supervised machine learning method to automate the process of dialogue act classification
- Instead of using rich feature sets including n-grams, POS and lemma, we use few intuitive and meaningful features such as **first three tokens**, **last token** and **utterance length** based on following assumptions
  - There is one direct speech act per utterance
  - Humans infer speakers' intention after hearing only a few of the leading words of an utterance [2].
  - E.g. *Questions* usually begin with *Wh*-word and *Greeting* use a relatively small set of expressions
  - Punctuation marks can convey intonational clues indirectly in typed dialogues



Future Works

- Annotate more sessions up to 500 and retrain our models
- Use powerful models such as CRF to account for sequential and contextual information. e.g. an utterance in the beginning of a session is more likely a *Greeting*
- Once the accuracy is at acceptable level, we will use the classifiers to automatically tag tens of thousands of sessions with dialogue acts and subacts
- Use these sequences of actions and sub-actions to identify patterns of tutor and student actions that relate to learning and affect.
- Finally, use these identified patterns to develop automated intelligent tutoring system of hybrid systems.

Taxonomy

- Current coding taxonomy builds on an earlier taxonomy built over a large corpus of online tutoring sessions conducted by human tutors in the domains of Algebra and Physics [3]
- More granular than previous schemes such as the one used by Boyer and colleagues [4]
- Top level 16 dialogue act categories: Question, Answer, Assertion, Request, Explanation and so on
- Also includes two more pedagogical categories Prompt and Hint
- Each dialogue act category has 4 to 22 subcategories or subacts

Data, Experiments And Results

Data

- 288 tutorial sessions between professional human tutors and actual college-level, adult students containing 17,537 utterances
- A subset of larger samples of 500 sessions randomly selected from a corpus of 17.711 sessions.
- Students taking two college-level developmental mathematics courses (pre-Algebra and Algebra) were offered these online human tutoring services at no cost.
- The same students had access to computer-based tutoring sessions through Adaptive Math Practice, a variant of Carnegie Learning' Cognitive Tutor.

Expert Annotation Process

- 288 sessions were manually labelled by a team of 6 trained annotators, all of whom were experienced classroom math teachers
- Each session was first manually tagged by two independent annotators
- The tags of the two independent annotators were double-checked by a verifier, who also happens to be the designer of the taxonomy

Overall agreement: 0.77, Kappa: 0.72

Answer	0.46
Assertion	0.81
Clarification	0.33
Confirmation	0.75
Continuation	0.56
Correction	0.59
Directive	0.47
Explanation	0.42
Expressive	0.88
Hint	0.20
LineCheck	0.78
Offer	0.45
Promise	0.74
Prompt	0.74
Question	0.63
Reminder	0.68
Request	0.78
Suggestion	0.68

Inter-annotator agreement

SN	Features	Accuracy (%)	Kappa
1	FT , ST , TT , UL , LT , S , UP	67.27	0.58
2	FT , ST , TT , UL , LT , S	66.77	0.58
3	FT , ST , TT , UL , LT , UP	66.74	0.57
4	FT , ST , TT , UL , LT	66.78	0.57

10-fold cross validation results using Bayes Nets classifier. FT - first token, ST - second token, TT - third token, UL - utterance length, LT - last token, S - speaker, UP - utterance position

- Accuracy dropped below 60% if any of FT, ST, TT, LT and UL features is removed
- We also experimented with other models such as Naïve Bayes and Decision Trees but best results were obtained with Bayes Nets
- First token, Second Token, Third Token, Last Token and the utterance length were the most predictive features

Acknowledgements

- This research was sponsored by a subcontract to The University of Memphis from Carnegie Learning, Inc., under award W911QY-15-C-0070 by Department of Defense
- Institute for Intelligent Systems Student Organization

References

1. Austin, J. L. (1962). *How to do things with words*: Oxford University Press,1962.
2. Moldovan, C., Rus, V., & Graesser, A.C. (2011). *Automated Speech Act Classification for Online Chat*,
3. Morrison, D. M., Nye, B., Samei, B., Datla, V. V., Kelly, C., & Rus, V. (2014). *Building an Intelligent PAL from the Tutor.com Session Database-Phase 1: Data Mining*
4. Boyer, K.E., Phillips, R., Ingram, A., Ha, E.Y., Wallis, M.D.,Vouk, M.A., & Lester, J.C. (2011). Investigating the relationship between dialogue structure and tutoring effectiveness: a hidden Markov modeling approach.

Carnegie Mellon University

Carnegie Learning

INSTITUTE for INTELLIGENT SYSTEMS

THE UNIVERSITY OF MEMPHIS