# Improving Machine Recognition of Collaborative Dialog Acts via Sentence Embeddings

**Students: Justice Parham, Vondarrian Patrick**
**Advisor: Dr. Jung Hee Kim**
**Department of Computer Science**
**North Carolina Agricultural & Technical State University**

## Abstract

### Introduction

- **This project attempts to train a computer to recognize linguistic dialogue acts within transcripts of students working together. In COMPS (Computer-Mediated Problems Solving) exercises students work together via typed-chat, solving problems in small groups in a computer science class.**

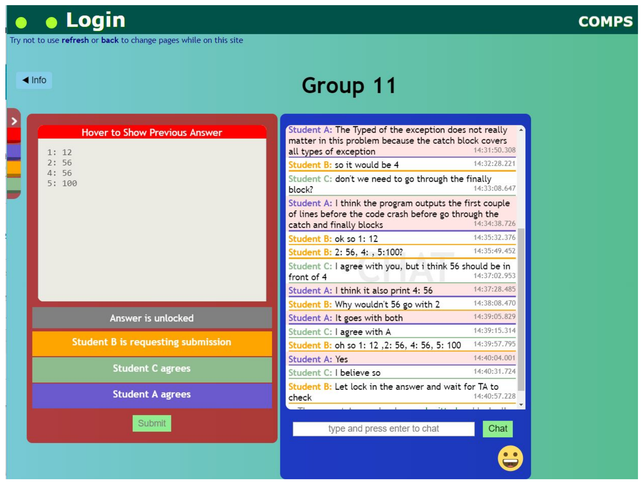


**Figure 1: Chat Engineer Interface**

- **Student dialogue turns are classified according to four categories of collaborative utterance: sharing ideas, negotiating ideas, regulating problem-solving, and maintaining communication.**

### Methods

- **Pre-process dialogue text into set of numbers for each dialogue turn, then train a linear classifier.**
- **Atempt to improve classifier accuracy by changing the way sentences are converted into numerical features**
- **Use Doc2Vec, which is word order sensitive, in place of Topic modeling**
- **Utilize dialogue context by presenting several preceding dialogue turns to the classifier**

### Goal

- **This research advances toward promoting better student collaborative problem-solving exercises, more fully using student group cognition and collaboration skills, and developing computer monitoring of the student conversation groups.**

## Annotation

**Figure 2: Student Dialogue Act Categories**

| Category | Category Label | Category Description |
|---|---|---|
| Sharing Ideas | A | Student shares task relevant idea that contributes information to the process. |
| Negotiating Ideas | B | The student listens and responds to a previously stated idea. They will either agree or disagree. |
| Regulation of Problem Solving | C | The student attempts to direct or regulate workflow. |
| Maintaining Communication | D | The participant contributes something that is not task relevant to the group. |

### Transcript Annotation

- **Figure 3 demonstrates how student dialogues are analyzed and categorized into one of the four categories.**
- **Some turns show attributes of more than one category and are labeled accordingly.**

### Categorization

- **Student groups worked together via typed-chat**
- **The dialogue turns were then categorized according to the 4 dialogue act categories.**
- **715 hand-categorized dialogue turns were used in this experiment**

**Figure 3: Transcript Categorization**

| Participant | Text | Dialogue Acts |
|---|---|---|
| St1 | public String toStrong(){ String result = null; result = lendingInstitution +' '+ PAmount +' '+ iRate +' '+ etc. | A |
| St2 | lol yall going in i think thats right tho | D, B |
| St1 | we just have to explain the getters and setters now | C |
| St3 | Student 1 can u explain them | C |
| St1 | besides excapsulation, accessors make it easier to change future things mybad on the spelling | A, D |
| St4 | So everything except the setters and getters are explained right? | C |

## Topic Modeling Result

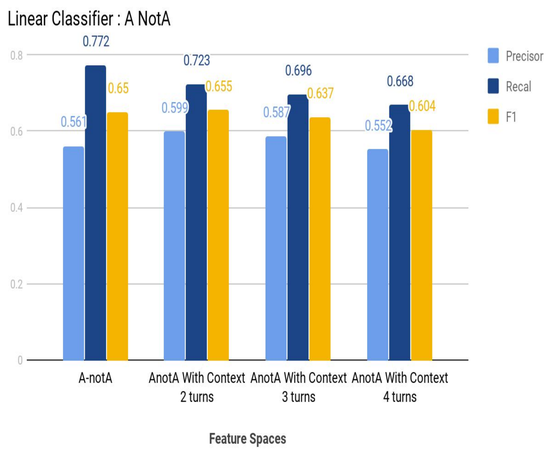


**Figure 4: Topic Model results**

- **The gensim Python library was used to derive the 10 latent topic numbers for each turn**
- **The "AnotA" classifier identified (yes or no) whether dialogue act A was contained within the words of one dialogue turn.**

### Result

- **It achieved F1 accuracy (a combination of precision and recall) ranging from 0.6 for category A) sharing-ideas, to 0.3 for D) maintaining conversation.**
- **Adding 1 turn of context usually slightly improved results, adding more turns made it worse**

## Doc2Vec Result

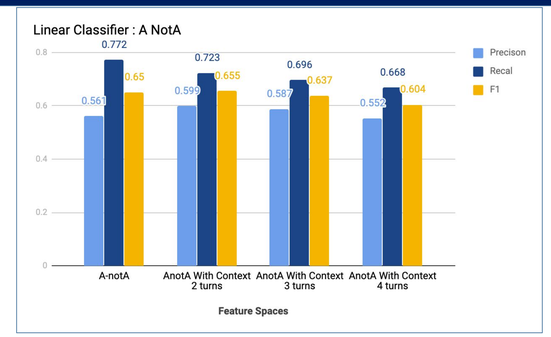


**Figure 5: Doc2Vec Results**

### Sentence Embeddings Approach

- **Two different sentences containing the same words will appear differently to the classifier, making it more likely that the software will be able to recognize the different conversational intentions of the speakers.**

### Result

- **It achieved F1 accuracy ranging from 0.65 for category A) recognizing sharing-ideas, to 0.32 for category D) maintaining conversation.**

## Future Work

- **Expansion on using different classifiers such as decision trees or random forest.**
- **Testing with different combination of features**
- **Tuning the Classifier to make an improvment**

## Acknowledgments

**Partial support for this work was provided by the National Science Foundation's Improving Undergraduate STEM Education (IUSE) program under Award No. 1504918. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.**
**This work was made possible thanks to graduate student Duy Bui and Dr. Michael Glass of Valparaiso University.**