AUTOMATIC SHORT ESSAY SCORING USING NATURAL LANGUAGE PROCESSING TO EXTRACT SEMANTIC INFORMATION IN THE FORM OF PROPOSITIONS

I. ABSTRACT

Background: Existing implementations of Common Core assessments across the United States bring with it a shift in essay-based assessments from multiple-choice tests to short essay constructed response tests as measures of student knowledge of a given concept. Possible: Because such tasks are costly and time-consuming to score manually, automated methods of scoring the context of short essay constructed response items need to be developed.

Purpose: This paper introduces a novel technique for using domain-independent, open natural language processing techniques to automatically extract many forms of student responses in the form of propositions and match the extracted propositions to the expected responses.

II. DATA

- Short essay responses by 6th and 8th grade students explaining the hearing process.
- The students replied to the following prompt: “Imagine your friend comes to you with a problem. She has missed the last few months of school and wants your support. How would you explain all about the ear and the hearing process. Think about all the components that make up the ear. What is the shape of your ears? Think about the outer, middle, and inner ears and what they do. Write a short paper to explain how your ears work together. Think of an example to your friend so that she can understand hearing.”
- 10 short essays, containing approximately 3,000 words in 413 sentences.
- Essay content was scored on a 1 to 5 scale by two experts using a rubric with target sentences.

- The target essay consisted of 626 targeted sentences in the rubric.

- The outer ear (tympanic membrane sound wave), sound waves travel through the ear canal until it enters the eardrum. The eardrum, a thin, flexible membrane, is located in the middle ear, which is made up of the tympanic cavity and eardrum. This middle ear cavity contains the ossicles, which are the malleus (hammer), incus (stirrup), and stapes (stirrup). The stapes joins to the oval window, and the oval window is joined to the round window of the cochlea.

- A scoring rubric is used to assess the essays.

- Matching Nodes:
  - Conversion of a form, structure, and function.
  - Sentence matching.
  - Complex matching (with blanks).

- Matching Links:
  - Conversion of verbs to infinitive form.
  - Matching.
  - Syntactic or semantic matching (with blanks).

- Matching Tips:
  - Unique Holes: A interactive microcosm combination of rules and links matching.

III. MATCHING TO TARGET PROPOSITIONS

- Matching Nodes:
  - Conversion of verbs to infinitive form.
  - Syntactic matching.
  - Conversion matching (with blanks).

- Matching Links:
  - Conversion of verbs to infinitive form.
  - Matching.
  - Syntactic or semantic matching (with blanks).

- Matching Tips:
  - Unique Holes: A interactive microcosm combination of rules and links matching.

IV. OVERVIEW. SEMSCAPE PROPOSITION EXTRACTION PROCESS

1. PARSE TREE + NOUNS

2. MAINSCAPE RULES

3. TEXT GRAPH - PROPOSITIONS

4. TEXT GRAPH - NOUNS

5. TEXT GRAPH - LINKS

6. TEXT GRAPH - RULES

V. RESULTS

Precision Extraction Performance: Approximately 15% of the errors observed in our proposition extraction is due to native English sentence paraphrasing. Example: “The sound wave enters the ear. Then it goes down the ear canal.”

Recall Score Alignment Performance: This process achieves a higher correlation and lower MSE than other similar studies, but agreement was still low.

This process accurately scored a majority of student essays, but had difficulty scoring heterogeneous scores in the middle range of the scale.

VI. CONCLUSIONS & FUTURE WORK

- Open Natural Language Processing techniques successfully extract propositions from student essays, achieving an average recall and precision of 65% and an average precision of 75%.
- We successfully included the scores given by human raters, achieving a correlation of 0.374, an MSE of 0.77, and a Cohen’s Kappa of 0.47.
- Future work will focus on increasing the precision and recall of the proposition extractions process by adding additional graph domain rules to identify semantic relationships contained in less common grammatical structures.