Topic 1: Evaluation methods (Dr. Wiebe)

Themes:
Evaluation methods in Classification
Evaluation methods in Clustering
Evaluation methods in NLP
  Corpus creation
  Evaluating MT
  Evaluating GEC
  Evaluating Parsers
  Evaluating Paraphrasing
  Evaluating Summarization

Evaluation methods in Classification
   a. This paper is related to chapter 4 of the book “Empirical methods for AI”. Since it is newer and highly cited, I picked it.
   a. This paper discusses evaluation methods in ordinal regression (also known as ordinal classification). This task is a type of multi-class classification where there is ordering relationship between the classes but there is not meaningful numeric differences between them.

Evaluation methods in Clustering
This sections mentions some of the clustering evaluation methods. I used some of these methods in my fragmentation project to evaluate the accuracy of fragmentation methods with the gold standard fragments.

Evaluation methods in NLP
In this section, I will specifically consider evaluation methods in various NLP tasks. The goal is to learn the different ideas of evaluations and how to come up with an evaluation method when dealing with a slightly different NLP task.

Corpus creation
Creating corpora by manually annotating data.

Evaluating MT


Evaluating GEC

15. Madnani, Nitin; Tetreault, Joel R.; Chodorow, Martin; Rozovskaya, Alla. They can help: using crowdsourcing to improve the evaluation of grammatical error detection systems. ACL, 2011.

Evaluating parsers


Evaluating paraphrasing


Evaluating summarization

Topic 2: Education in NLP (Dr. Litman)

In this topic, I first investigate general areas of AIED and then consider sub-areas with the focus of NLP methods to help different Educational tasks.

Themes:
- Artificial Intelligence in Education
  - Educational Data Mining
  - Collaborative Learning
  - Educational Games
  - Intelligent Tutoring Systems
- NLP for Educational applications
  - Essay assessment
  - Tutorial Dialogue Systems
    - Spoken Dialogue Tutoring
    - Typed Dialogue Tutoring
- NLP techniques for Education
  - Discourse Analysis

Artificial Intelligence in Education
The goal of AIED is to design computer based learning systems to help learning. Following is the list of sub-areas in this field along with a few papers that discuss some techniques.

Educational Data Mining
Data mining techniques such as classification, clustering, outlier detection and pattern mining can be applied on educational systems to discover knowledge.

1. Baker, Ryan SJD and Yacef, Kalina. The state of educational data mining in 2009: A review and future visions. JEDM-Journal of Educational Data Mining, 2009. This is a survey paper that discusses the trend and shift of educational data mining.

Collaborative Learning
This set of papers briefly shows some techniques for collaborative learning.

Educational Games
This section contains some samples of educational games.


### Intelligent Tutoring Systems


### NLP for Educational applications

This section focuses on use of NLP techniques in educational applications.

#### Essay assessment

It is among the first educational applications that uses NLP methods to automatically evaluate essays. Since the evaluating readability and grammar error detection related papers are covered in my special topic I will not consider those papers in this sub-area of education.


   a. This paper gives a good introduction over the field of NLP in education. It uses two essay assessment systems as examples and focuses on their NLP aspects. I think it would be helpful to read this paper because the author is one of the major researchers in this area and she made good discussion over NLP part of the two sample systems.


#### Tutorial Dialogue Systems

This section shows a number of techniques to utilize dialogue in educational applications. The interaction with the systems could be either via speech or typing.


   This paper presents a systematic study to test the hypothesis of effectiveness of learning by dialogues or reading. I liked the way they designed their study. They considered tutoring systems with typed
dialogues and also spoken human tutoring and typed human tutoring conditions.

Spoken Dialogue Tutoring
   This paper address a brief history of speech technology for education and some of its main issues. I think that reading this paper will give me a good overview of the field.


Typed Dialogue Tutoring
   This paper discuss an intelligent tutoring systems and the dialogues are typed by the student.

   This paper studies human-human tutoring using textual dialogue messages.

NLP techniques for Education
In this section, I will focus on some common NLP methods that are applied on educational applications.

Discourse Analysis
Discourse analysis contains a number of methods to analyze text or speech. In this section, I just put samples of discourse analysis approaches.


   The corpora that they used contains essays written in easyforum. They mentioned that it is an active community that provides feedback for different kinds of essays. For instance students post their essays there. So, I guess the approach that they propose in this paper can also be considered for educational purposes.
Topic 3: Grammar Error Correction & Robust Parsing (Dr. Hwa)

Themes:
- **Grammatical Errors**
  - ESL errors and corpora
  - MT errors
  - Social media language
  - Artificial errors
- **Error detection and correction**
  - GEC in general
  - GEC using ML
  - GEC using MT
  - GEC using rules
  - GEC evaluation
- **Robust Parsing**
  - Judging grammatically
  - Parsing ungrammatical sentences

Grammatical Errors
In order to detect and then correct errors, it would be nice to have a survey on different types of errors in disfluent sentences (like MT and ESL) and also mentioning some of common annotated standards and corpora.

**ESL errors and corpora**
   This is a technical report which demonstrates the ESL categories of FCE corpus. This paper can be considered as a supplementary reading list for the FCE(above) paper.

   It introduces UIUC error categories. I think it can be considered as one of the references in ESL annotating survey.

   This paper both introduces FCE corpus and some learning methods for assessing ESL essays.

**MT errors**
   This is an early work on error analysis of MT outputs. They introduced an error category that is used frequently in the community. I think this paper should be mentioned in a survey of MT error types.
Social media language
   It shows challenges of social media for the traditional NLP methods. Since social media is another domain of disfluent sentences, I guess it would be interesting to compare it with ESL and MT sentences.

Artificial errors
   This paper is an example of generating artificial errors. The authors have published several papers on this topic and I just picked this paper as a representative of their work.


Error detection and correction
After discussing different error types and some annotated corpora, this set of papers are about approaches of detecting and correcting grammatical errors.

GEC in general
   This book is a comprehensive introduction over topics of ESL errors. Reading this book could be helpful on getting more idea about history and related works of this area.

GEC using ML
    This paper shows the helpfulness of parse tree features for preposition detection and correction. The features that they used might be useful to detect other error types.


GEC using MT
a. This paper introduces MT based approach for GEC. The idea of using MT methods to solve other NLP problems has been recently used, so it would be nice to mention this method as one of the correction approaches. This paper is among the first works on using SMT. The main property of using MT is to correct the whole sentence together. Following papers show the recent attempts:

a. This paper borrows noisy channel model from MT to solve GEC. Their approach can be considered as a working method in the GEC survey.

17. Daniel Dahlmeier and Ng, H. T. A beam-search decoder for grammatical error correction. EMNLP 2012.
a. This paper also uses the idea of decoding in MT for GEC. Again this idea is interesting and worthwhile to be mentioned as one of the GEC methods.

a. This paper uses round trip translations with different pivot languages for GEC. The idea is interesting and I guess could be mentioned in the survey.

GEC using rules
a. This paper introduces a novel approach for preposition error correction by defining error case frames which are extracted by comparing native and learner corpora. Their system can also give learners feedback. I liked the idea of making error case frames as a new approach for GEC.

GEC evaluation
a. This paper first gives an introduction to different GEC methods then discusses the effect of annotation and evaluation of ESL error detection systems. I can put this paper in the evaluation topic, but I also like the background section on GEC systems.

Robust Parsing
The goal of GEC methods are to correct all errors of a sentence. But they are not perfect and there are still some errors that they are not able to handle. So, parsing these erroneous ungrammatical sentences might be problematic for parsers. This set of papers are basically about parsing ungrammatical sentences. The first step of parsing or even correcting grammatical errors is to check whether the sentences is ungrammatical. If it is ungrammatical, then it needs to be corrected or parsed in a special way.
Judging grammatically


   a. This paper both describes an automatic error creation method and a classification to distinguish grammatical and ungrammatical sentences. The features are XLE parser output, POS n-grams and n-grams.

   a. This paper introduces a syntactic language model using parse trees. Treelets are overlapping windows of trees with depth at most 3 containing CFG rules and non-terminals. Since these treelets are similar to fragments, I mentioned this paper in the reading list.

   a. This paper uses extracted TSG rules as features for binary classification of grammaticality. Since TSG rules are fragments of parse trees, I mentioned this paper. Also, this paper is one of the papers that explains fluency check task using parse features.

Parsing ungrammatical sentences

   a. This paper investigates the output probability of parsers when parsing ungrammatical sentences.

   a. This paper presents creating an ungrammatical treebank by artificially adding errors to sentences and changing parse trees. This idea is similar to our fragmentation method to generate gold standards fragments. So, it is one of our references for changing parse trees. I have put this paper in the Evaluation topic.

   a. This paper presents a robust parser which is trained on an ungrammatical treebank. They also used a classifier to choose whether a sentence has grammatical errors before parsing it.

   a. Parsing the web
   a. Parsing tweets

