Correcting Verb Related Errors

Noah Weber

***Abstract*—Verb related errors are common in the writings of those learning English as a second language (ESL), due to the various uses and forms of verbs. Though much of the research in automatic ESL error correction has focused solely on fixing either determiner or preposition errors, there has been recent interest in developing methods for automatic verb error correction. Previous approaches to this problem have typically relied on data from error annotated learner's corpora to learn common types of verb errors and the situations in which they are most likely to occur in. In this paper, we compare the results of two models for estimating the probability that a certain verb tense and aspect is correct given the sentence context and the original verb tense and aspect. The first model is a generative model which estimates this value by modeling the prior probability distribution as well the probability distribution for a instance of a verb tense and aspect being an error given the intended tense aspect and the sentence context. From these two distributions, we can generate a model for the posterior probability distribution. The second model is a discriminative model which models the posterior distribution directly.**

I. Introduction

Due to the increase of English as a second language (ESL) students, the interest in automated grammar checking systems has increased in the past decade. Much of the previous work in this field has looked at the correction of article and preposition errors [2][7][8], both respectively being the two most common errors in the writings of ESL students [6]. Though verb related errors are the third most common type of error in ESL writings, relatively little research has been done on the topic.

The problem of correcting verb errors also presents several additional difficulties that separate it from article or preposition correction. Due to the various ways in which verbs can be used, identifying verbs in a text is typically a more involved process when compared to article and preposition detection. Automatic verb correction may also involve detecting missing verbs, in addition to correcting the verbs present in the sentence. For example, in the example sentence:

> *The problem is that the same step being repeat for every interval.

the error to be fixed involves both inserting the verb *is* and changing the verb *repeat* from its base form to its past participle form. The correct choice of verb also typically depends on the context in which it is used which further complicates the task at hand. In addition, the form and tense of verbs usually depends not only on the surrounding sentence context, but also on the form and tense of surrounding verbs. This means that verb correction has an additional complication in that correcting a single verb might have an effect on what other verbs in the sentence should be corrected to, if they need to be corrected at all. In this paper we plan to build upon previous research on this topic and develop a new method for correcting verb errors that utilizes both machine learning and linguistic theory.

II. Related Work

Previous research in the field of automated grammar checking has mostly focused on preposition and article related errors. Because prepositions and articles are both closed word classes, solutions to the problem have typically involved the use of multiclass classifiers, with each individual preposition or article making up a single class [2]. Some approaches treat grammar checking as a machine translation problem, and simply use standard statistical machine translation techniques to translate an ungrammatical sentence to a grammatical one [4]. Some recent work has looked into the problem of correcting several types of errors simultaneously. Dahlmeir and Ng (2012) proposed a model that takes corrects article, preposition, and noun count errors using a beam search decoder. Their model consists of proposers for each type of error, as well as expert models for each type of error. The proposers generate new sentences by making small edits to the current sentence. The expert models are used to score the grammaticality of the sentence. Similar to the decoding step done in statistical machine translation, a beam search is done in order to find the sentence that results from the highest rated series of edits. Work done by Wu and Ng (2013) similarly tries to fix article, preposition, and noun count errors simultaneously, but reformulates the problem from a searching problem to an integer programming problem.

Due to many difficulties involved with verbs, the verb correction problem must be approached in a different manner. The existing research on verb correction use several different approaches. The work of Lee and Seneff (2008) deals with verb form errors by looking for parse trees that are likely to be produced from a verb error, along with probabilities derived from n-gram counts to identify and correct errors. While this method worked well for verb agreement and form mistakes, it did not account for verb tense errors. The work of Tajiri et al. (2012) aims to correct verb tense errors by treating the problem as a sequence classifying problem where each verb is labeled with a tense. The label chosen for an individual verb depends on the labels chosen for surrounding verbs. They use conditional random fields for sequence labeling, allowing them to use both syntactic and semantic features to aid in labeling. Recent work by Rozovskaya et al. (2014) looks into correcting tense, agreement, and form errors. Their work takes advantage of several linguistic properties of verbs, most notably verb finiteness, in order to guide their statistical learning method. Their learning method classifies verbs in a text as either having an aspect, tense, form, or no errors. Error correction is also handled using a multiclass classifier, with each class of verb error having a unique error correction model.

III. Methodology

Our proposed method will build upon previous work and utilize linguistic features in tandem with machine learning methods. The method we propose closely resembles the Bayesian noisy channel model used in tasks such as speech recognition and statistical machine translation. In this model, we treat the possibly incorrect sequence of verbs as a corrupted version of some correct sequence of verbs. The goal of this method is to model the probability distribution:

$$P(C|O, S)$$

Where C is a proposed correct tense aspect for a verb instance. O is the tense aspect of the verb instance originally put down by the writer and S is the sentence context. We try and compare two different models to estimate $P(C|O, S)$. The first way is using a generative model. For the generative model we rewrite the distribution $P(C|O, S)$ using Bayes Theorem as:

$$P(C|O, S) = \frac{P(O, S|C)P(C)}{P(O, S)}$$

For our purposes we rewrite this as:

$$\begin{aligned} P(C|O, S) &= \frac{P(O, S, C)}{P(O, S)} \\ &= \frac{P(S)P(C|S)P(O|S, C)}{P(O|S)P(S)} \\ &= \frac{P(C|S)P(O|S, C)}{P(O|S)} \end{aligned}$$

Since the bottom $P(O|S)$ term remains constant we can simply ignore it. Our goal for the generative model is thus to model the distribution of $P(C|S)$ and $P(O|S, C)$. The seeming advantage of using a generative model in this instance is that is allows one to utilize data from both well-formed corpora in the estimation of $P(C|S)$, as well as data from error annotated learner corpora for the estimation of $P(O|S, C)$. However, as the results show this actually may not be the case. To estimate both $P(C|S)$ and $P(O|S, C)$ we treat the problem as a classification problem, as traditionally done in prior research in automated grammar correction systems. For our classifier we use a Maximum Entropy (MaxEnt) classifier. The reason being its use as a classifier in previous studies done in automated grammatical error checking [2]. However, we plan to use other classifiers as well in the future.

Our second model is a discriminative model which estimates the distribution $P(C|O, S)$ directly using a MaxEnt classifier. The data used in the training and testing of this model as well as the model for the $P(O|S, C)$ distribution comes from the error annotated FCE corpus [9]. The data used for the model the $P(C|S)$ distribution comes from several different sources. The data for this model comprises of sections from the Brown corpus, the MASC section of the Open American National Corpus, as well as a fully corrected version of the FCE corpus. Many of the features we utilize come from features used by both Tajiri et al. (2012) as well as Rozovskaya et al. (2014).

Feature	Description
verb lemma	The lemma of the current verb
left/right lemma	The lemma of the words to the right/left of the current verb
subject	The word, pos, person, and number of the sentence subject
determiner	The determiner for the sentence subject
left/right noun	The word, pos, and person of nouns to the left and right of the current verb
first	Whether the verb is the first in a chain of verbs
last	Whether the verb is the last in a chain of verbs
governor	The governor of the verb and the relation type between them
governee	The governee of the verb and the relation type between them
left/right time adverbs	Time adverbs to the left/right of current verb

Table 1: Description of features used in model

IV. Results

To test the models we split off a section of the FCE corpus solely for use in testing. The section contained around 7052 instances of verb sequences, with around 200 of these sequences having some type of error. To test, we used our trained classifier to classify each unlabeled verb instance into a tense and a aspect. We then compared the tense aspects generated by the model with the actual tense aspects put down by the annotator. As our evaluation criteria we use both precision and recall. As our results show, the discriminative model outperforms the generative model by a large margin.

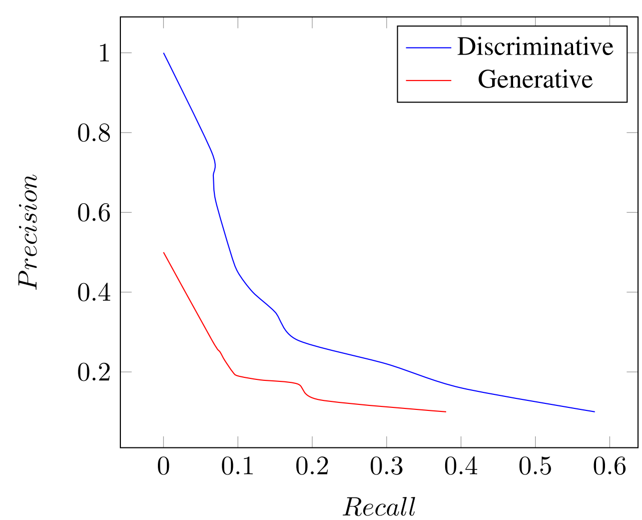

Figure 1: Precision and Recall Curve for both models

References

[1] J. Lee and S. Seneff. 2008. "Correcting misuse of verb forms". *Proceedings of the ACL-08*, pg 174−182, Association for Computational Linguistics, Columbus, Ohio, June.

[2] J. Tetreault and M. Chodorow. 2008. "The ups and downs of preposition error detection in ESL writing". *In Proceedings of the 22nd International Conference on Computational Linguistics"*, pg. 865−872, Manchester, UK, August.

[3] T. Tajiri, M. Komachi, and Y. Matsumoto 2012. "Tense and aspect error correction for esl learners using global context". *In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers), pg 198−202, Association for Computational Linguistics. Jeju Island, Korea, July.

[4] C. Brockett, W. Dolan, and M. Gamon. 2006. "Correcting ESL Errors Using Phrasal SMT Techniques". *In Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, pg 249256, Association for Computational Linguistics. Sydney, Australia, July.

[5] A. Rozovskaya, D. Roth, and V Srikumar. 2014. "Correcting grammatical verb errors". *In Proceedings of EACL*, pg 358−357.

[6] C. Leacock et al. 2010. "Automated grammatical error detection for language learners." *Synthesis Lectures on Human Language Technologies*, pg. 15−27

[7] D. Dahlmeier and H. Ng. 2012. "A beamsearch decoder for grammatical error correction". *In Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pg 568−578.

[8] Y. Wu and H. Ng. 2013. "Grammatical error correction using integer linear programming". *In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, pg 1456−1465.

[9] H. Yannakoudakis, T. Briscoe, and B. Medlock. 2011. *A new dataset and method for automatically grading esol texts. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pg 180−189, Association for Computational Linguistics. Portland, Oregon, USA, June.