UDel: Extending Reference Generation to Multiple Entities

Charles Greenbacker and Kathleen McCoy
Dept. of Computer and Information Sciences
University of Delaware
Newark, Delaware, USA
[charlieg|mccoy]@cis.udel.edu

Abstract

We report on an attempt to extend a reference generation system, originally designed only for main subjects, to generate references for multiple entities in a single document. This endeavor yielded three separate systems: one utilizing the original classifier, another with a retrained classifier, and a third taking advantage of new data to improve the identification of interfering antecedents. Each subsequent system improved upon the results of the previous iteration.

1 Introduction

This paper provides a system report on our submission for the GREC-NEG (Named Entity Generation) Task, one of the two shared task competitions for Generation Challenges 2009. The objective is to select the most appropriate reference to named entities from a given list of alternatives. The corpus consists of introductory sections from approximately 1,000 Wikipedia articles in which single and plural references to all people mentioned in the text have been annotated (Belz and Varges, 2007). The training set contains articles from the categories of Chefs, Composers, and Inventors. GREC-NEG differs from the other challenge task, GREC-MSR (Main Subject References), in that systems must now account for multiple entities rather than a single main subject, and the corpus includes only articles about persons rather than a variety of topics.

2 System Description

Our GREC-NEG systems build upon our work for the GREC-MSR task. Our original approach was to consult findings in psycholinguistic research for guidance regarding appropriate feature selection for the production of referring expressions. We relied upon several common factors recognized by multiple authors (Arnold, 1998; Gordon and Hendrick, 1998), including Subjecthood, Parallelism, Recency, and Ambiguity. We followed (McCoy and Strube, 1999) who stressed the importance of Recency in reference generation. Finally, we made a preliminary attempt at identifying potential interfering antecedents that could affect the Ambiguity of pronouns (Siddharthan and Copestake, 2004).

As an initial attempt (UDel-NEG-1), we simply extended our GREC-MSR submission. By adapting our system to account for multiple entities and the slightly different data format, we were able to use the existing classifier to generate references for GREC-NEG. We suspected that accuracy could be improved by retraining the classifier, so our next system (UDel-NEG-2) added entity and mention numbers as features to train on. Presumably, this could help distinguish between the main subject and secondary entities, as well as plural references. As all named entities are tagged in the GREC-NEG corpus, we leveraged this information to improve our recognition of other antecedents interfering with pronoun usage in a third new system (UDel-NEG-3). As in our GREC-MSR submission, all three of our GREC-NEG systems trained C5.0 decision trees (RuleQuest Research Pty Ltd, 2008) on our set of features informed by psycholinguistic research.

3 Results

System performance, as tested on the development set and scored by the GREC evaluation software,

Proceedings of the 2009 Workshop on Language Generation and Summarisation, ACL-IJCNLP 2009, pages 105–106,
Suntec, Singapore, 6 August 2009.

is offered in Tables 1, 2, and 3. Type accuracy for UDel-NEG-1 remained close to our GREC-MSR submission, and error rate was reduced by over 20% for UDel-NEG-2 and UDel-NEG-3. However, string accuracy was very low across all three systems, as compared to GREC-MSR results.

Table 1: GREC scores for UDel-NEG-1 (unmodified).

Component Score	Value
total pairs	907
reg08 type matches	628
reg08 type accuracy	0.69239250275634
reg08 type precision	0.688699360341151
reg08 type recall	0.688699360341151
string matches	286
string accuracy	0.315325248070562
mean edit distance	1.55126791620728
mean normalised edit dist.	0.657521668367265
BLEU 1 score	0.4609
BLEU 2 score	0.5779
BLEU 3 score	0.6331
BLEU 4 score	0.6678

Table 2: GREC scores for UDel-NEG-2 (retrained).

Component Score	Value
total pairs	907
reg08 type matches	692
reg08 type accuracy	0.762954796030871
reg08 type precision	0.749466950959488
reg08 type recall	0.749466950959488
string matches	293
string accuracy	0.323042998897464
mean edit distance	1.4773980154355
mean normalised edit dist.	0.64564100951858
BLEU 1 score	0.4747
BLEU 2 score	0.6085
BLEU 3 score	0.6631
BLEU 4 score	0.6917

4 Conclusions

The original classifier performed well when extended to multiple entities, and showed marked improvement when retrained to take advantage of new data. All three systems yielded poor scores for string accuracy as compared to GREC-MSR results, suggesting an area for improvement.

Table 3: GREC scores for UDel-NEG-3 (interference).

Component Score	Value
total pairs	907
reg08 type matches	694
reg08 type accuracy	0.7651598676957
reg08 type precision	0.752665245202559
reg08 type recall	0.752665245202559
string matches	302
string accuracy	0.332965821389195
mean edit distance	1.46306504961411
mean normalised edit dist.	0.636499985162561
BLEU 1 score	0.4821
BLEU 2 score	0.6113
BLEU 3 score	0.6614
BLEU 4 score	0.6874

References

Jennifer E. Arnold. 1998. *Reference Form and Discourse Patterns*. Doctoral dissertation, Department of Linguistics, Stanford University, June.

Anja Belz and Sabastian Varges. 2007. Generation of repeated references to discourse entities. In *Proceedings of the 11th European Workshop on NLG*, pages 9–16, Schloss Dagstuhl, Germany.

Peter C. Gordon and Randall Hendrick. 1998. The representation and processing of coreference in discourse. *Cognitive Science*, 22(4):389–424.

Kathleen F. McCoy and Michael Strube. 1999. Generating anaphoric expressions: Pronoun or definite description. In *Proceedings of Workshop on The Relation of Discourse/Dialogue Structure and Reference, Held in Conjunction with the 38th Annual Meeting*, pages 63 – 71, College Park, Maryland. Association for Computational Linguistics.

RuleQuest Research Pty Ltd. 2008. Data mining tools See5 and C5.0. http://www.rulequest.com/see5-info.html.

Advaith Siddharthan and Ann Copestake. 2004. Generating referring expressions in open domains. In *Proceedings of the 42th Meeting of the Association for Computational Linguistics Annual Conference*, pages 408–415, Barcelona, Spain.