Proximity-Based Error Correction in a Speech Recognition Transcript

LTS5 / Sony

Célia BEL - Prof. Andrzej DRYGAJLO (Supervisor) – Dr. Wilhelm HAGG (Advisor)
Celia.bel@epfl.ch

Introduction

→ Problem

ASR System

Noisy transcript = with semantic errors

→ Goal

Automatic Correction of semantic errors

Theory

"The sailing ship..." → "The sailing sheep..."

1. N-best List
2. Scores

1 The sailing ship 2 The sailing sheep 3 The say lie sheep

Gain in Word Error Rate
(WERfinal-WERinit)

8 Audiobooks' transcripts
21 000 words / book
WERinit := 13%

Results

→ Metric / Corpus

Results for 8 Books

Method: baseline baseline_cutoff web bleu bleu_utterance final system

Results Distribution for 8 Books

→ Analysis

Baseline: "Cut-off" helps. Average = 0.81
Web System: Median better, distribution increased, selection helps. Average = 1.01
Final: Median better, distribution more concentrated. Average = 1.14

→ Problem

ASR System

Noisy transcript = with semantic errors

→ Goal

Automatic Correction of semantic errors

Method

Training of a new Language Model to attribute new scores

→ Which Language Model?

N-Grams

P(word | N previous words)
Unigram, bigram, trigram...

RNNLM

Doesn’t limit history

Chosen!

→ Which training data?

Remove less frequent words (« cut-off », optional)

Make a LM out of it

Baseline System

Final System

Keywords Extraction + WEB Search

Web System

Noisy transcript

Web pages selection (optional)

BLEU Score

Results

→ Results

1 The sailing ship
2 The sailing sheep
3 The say lie sheep

Scores

Train a new Language Model

→ Use of a Language Model (LM)


Conclusion

Best results with final System (Gain in WER = 1.14%)

Perspectives

• Better selection process (other metric…)
• More books for better statistical analysis

Acknowledgements

Prof. Andrzej DRYGAJLO (supervisor), Dr. Wilhelm HAGG (advisor)

References