Unsupervised Extraction of Lexical Simplifications from Wikipedia
Mark Yatskar, Bo Pang, Cristian Danescu-Niculescu-Mizil, Lillian Lee
Cornell University and Yahoo! Research

For the Sake of Simplicity:

Method 1: The simpl method

1. Find all SimpleWiki revision pairs where the comment contains “simpl”.
2. Align sentences between the elements of the pair using TF-IDF
3. Extract possible substitutions from the alignment
4. Rank by PMI(complicated,simple)

Method 2: Edit Mixture Model

P of seeing ‘A’ rewritten into ‘a’
P(a|A) = \sum_\alpha P(a|A)P(\alpha|A, a_0)

An edit operation:

Fix (o_0) – A correction either in language or content
Simplify(o_0) – Lexical content is simplified
No-op(o_0) – Lexical content is unchanged
Spam(o_0) – Inappropriate text

Estimation Details:

Assumptions:
1. EnglishWiki only has fixes
2. The probability of fix being performed in SimpleWiki
   is proportional to the same probability in EnglishWiki

Let f(A) be the fraction of docs in a collection that had revision that rewrote A

P(o_0|A) = af_{simpl}(A).
P(\alpha|A, a_0) = \frac{P(a|A)P(\alpha|A, a_0)}{P(o_0|A) + P(\alpha|A, a_0) - P(o_0|A)P(\alpha|A, a_0)}
P(a|A) = P(o_0|A) \cdot f_{simpl}(A) + \frac{P(\alpha|A, a_0)P(o_0|A)}{P(\alpha|A, a_0) + P(o_0|A)}

RHS can be estimated using statistics from English and Simple Wikipedia.

Rank by either P(a|A, o_0) or P(o_0|A). Latter more robust to infrequent edits.

Examples of lexical simplifications

- “The GDP is computed annually.”
- “Eventually, the “Style Dial”:
  Add hyperlinks from phrases to simplifications.”

Where do we find simple language?

- Adults speaking to children?
- Book titles?
- Simple English Wikipedia

Use SimpleWiki edits to learn to simplify...

But, it’s not that simple

Look at the way people edit SimpleWiki?

- But, maybe they’re correcting spelling or fixing inaccuracies.

Use SimpleWiki as a filtering mechanism for SimpleWiki edits

Get help from EnglishWiki

- Use MT with (EnglishWiki,SimpleWiki) pairs?
- But you need to align revisions in the two wikis
- Maybe some SimpleWiki articles are copied from English Wiki and then simplified?
- But which SimpleWiki revision should be aligned?

Results

Top 100 pairs from each method were manually annotated:
- Manualy assembled dictionary: SpList (by a SimpleWiki author)
- Edit and Simpl produce correct pairs not found in SpList (71% and 62%)

Correct Simplifications

Simpl Method

Incorrect Simplifications

Simpl Method – trusts revisions

Edit Model – misses fixes

Attempt: bootstrapping on the simpl method

1. Find new revisions by getting revisions that contain high ranking substitutions
   - Converges quickly without finding many new revisions
2. Find new comments from revisions that contain high ranking substitutions
   - Example: “reword.” Not many comments to find

Related Work

Simplifying medical text for non-doctors: Delaeger, Zweigenbaum ’09; ElHadad, Sutaria ’07
Wordnet + frequency simplification: Devlin, Tait ’98
Tense and word order simplification: Beigman Klebanov, Knight, Marcu ’04
Identifying simple versus non-simple text: Napoles, Dredze ’10

Future Work

- Estimate probabilities using EM
- Account for word complexity with inherent model of complexity
- Train a model for rewriting into Simple