

Week 8

**Automatic Recognition of Verb Patterns
for Correcting Learner Errors**

a case of e**X**plainable **A**rtificial **I**ntelligence (XAI)

April 17, 2018

Those who want to fight usage with grammar are silly.

—Michel de Montaigne (1987)

It is in vain to set up a language police to stem living development.

—Friedrich Waismann (1952)

An Example of Learner Errors and Correcting Errors

- Right and Wrong:
 - x They discussed about the issues.
 - v They discussed the issue.
 - **Explanation** For *discuss*, use **V n** not **V about n**
- Knowledge required:
 - Parse sentence: They_{np} discussed_{vp} about_{pp} the issue_{np}.
 - Identify (in)correct patterns (i.e., **V about n**):
discussed_V about_{about} the issue_n
 - Correct the pattern using Noisy Channel Model
 - * V n → **V about n**
 - * V for n → **V about n**

– Verify the pattern for the headword

(e.g., legit GPs for **discuss** see)

- * **V n**
- * V wh-to-inf
- * V n with n
- * V wh

Noisy Channel Model of Grammatical Errors

- **Identify verb patterns** in a learner sentence, as well as in the edited sentence in a learner corpus (e.g., in EFCAMDAT)
- **Compute before/after relation of patterns**
 - * Compute relation of (in)correct patterns
 - * **V n → V about n**
DISCUSS, discuss this issue → discuss about this issue
EXPLAIN, explain the process → explain about the process

ANSWER: answer questions → answer about questions
 - * **V for n → V n** (V for n 有 189 次寫錯成為 V n)
WAIT: wait for your answer → wait your answer
HOPE: hope for your answer → hope your answer

APPLY: apply for it → apply it

Shallow Parsing

- Genia tagger (www.nactem.ac.uk/GENIA/tagger/)
- Example

* Input: This liked to discuss about the issues.

* Output:

```
[('They', 'They', 'PRP', 'B-NP'),  
 ('liked', 'like', 'VBD', 'B-VP'), ('to', 'to', 'TO', 'I-  
VP'), ('discuss', 'discuss', 'VB', 'I-VP'),  
 ('about', 'about', 'IN', 'B-PP'),  
 ('the', 'the', 'DT', 'B-NP'), ('issues', 'issue', 'NNS',  
 'I-NP'), ('.', '.', '.', 'O')]
```

* Reformat:

- Remove named entity
- Change from V format to H format
- Separate out "to" in to-infinitive
- Change from BIO to IHO format

```
[('They', 'liked', 'to', 'discuss', 'about', 'the issues',  
'.'),  
(('they', 'like', 'to', 'discuss', 'about', 'the issue', '.'),  
(('PRP', 'VBD', 'TO', 'VB', 'IN', 'DT NNS', '.'),  
(('H-NP', 'H-VB', 'H-TO', 'H-VP', 'H-PP', 'I-NP H-NP',  
'O'))]
```

Verb patterns and Dictionary

- In Collins Cobuild English Dictionary, each sense entry of is annotated with patterns observed in a corpus
- Other publishers followed suit with similar patterns in different format
- Patterns of 5,000 common verbs (in the Bank of English) are available (arts-ccr-002.bham.ac.uk/ccr/patgram/)
- Many learner's writing errors can be attributed to argumentational patterns of verbs, nouns, and adjectives (Macmillan Dictionary for Advanced Learner, IW34)

Source:

(1) Oliver Mason and Susan Hunston (University of Birmingham). The automatic recognition of verb patterns—A feasibility study.

International Journal of Corpus Linguistics 9:2 (2004), 253-270 www.academia.edu/400537/The_Automatic_Recognition_of_Verb_Patterns_A_Feasibility_Study?auto=download

(2) Macmillan Dictionary for Advanced Learners (www.helpforenglish.cz/files/med.pdf)

What is grammar patterns?

- In Collins Cobuild English Dictionary, each sense entry of is annotated with patterns observed in a corpus
- Patterns of 5,000 common verbs (in the Bank of English) are available ()

Source: Oliver Mason and Susan Hunston (University of Birmingham). The automatic recognition of verb patterns—A feasibility study.

International Journal of Corpus Linguistics 9:2 (2004), 253-270 www.academia.edu/400537/The_Automatic_Recognition_of_Verb_Patterns_A_Feasibility_Study?auto=download

Worksheet

– Purpose

1. Compute Noisy Channel Model (of Grammar Pattern) using Learner Corpus
2. Compute Lexical Language Model (given) using Reference Corpus
 - * **discuss**: V n, V wh-to-inf, V n w,
 - * **apply**: V pron-refl, V adv, V to-inf, V, V to n, V for n, V n
 - * **answer**: V with quote, V that, V n as to wh, V, V n, V for n, V to n
 - * **explain**: V with quote, V pron-refl, V that, V wh-to-inf, V, V wh
3. Correct verb-related errors in a given sentence

– Data:

- * ef.small.src.tagged.txt (sentences with errors)
- * ef.small.tgt.tagged.txt (edited sentences)
- * 30 test sentences

– Code:

- * senttag.py (shallow parsing and reformatting)
- * grampat.py (identifying grammar pattern of verbs in given sentences)