

# Automatically Suggesting Example Sentences of Near-Synonyms for Language Learners

Chieh-Yang Huang<sup>1</sup>, Nicole Peinelt<sup>2</sup>, Lun-Wei Ku<sup>1</sup>

<sup>1</sup> Academia Sinica, Taipei, Taiwan. {appleternity, lwku}@iis.sinica.edu.tw

<sup>2</sup> The Alan Turing Institute, UK. npeinelt@turing.ac.uk



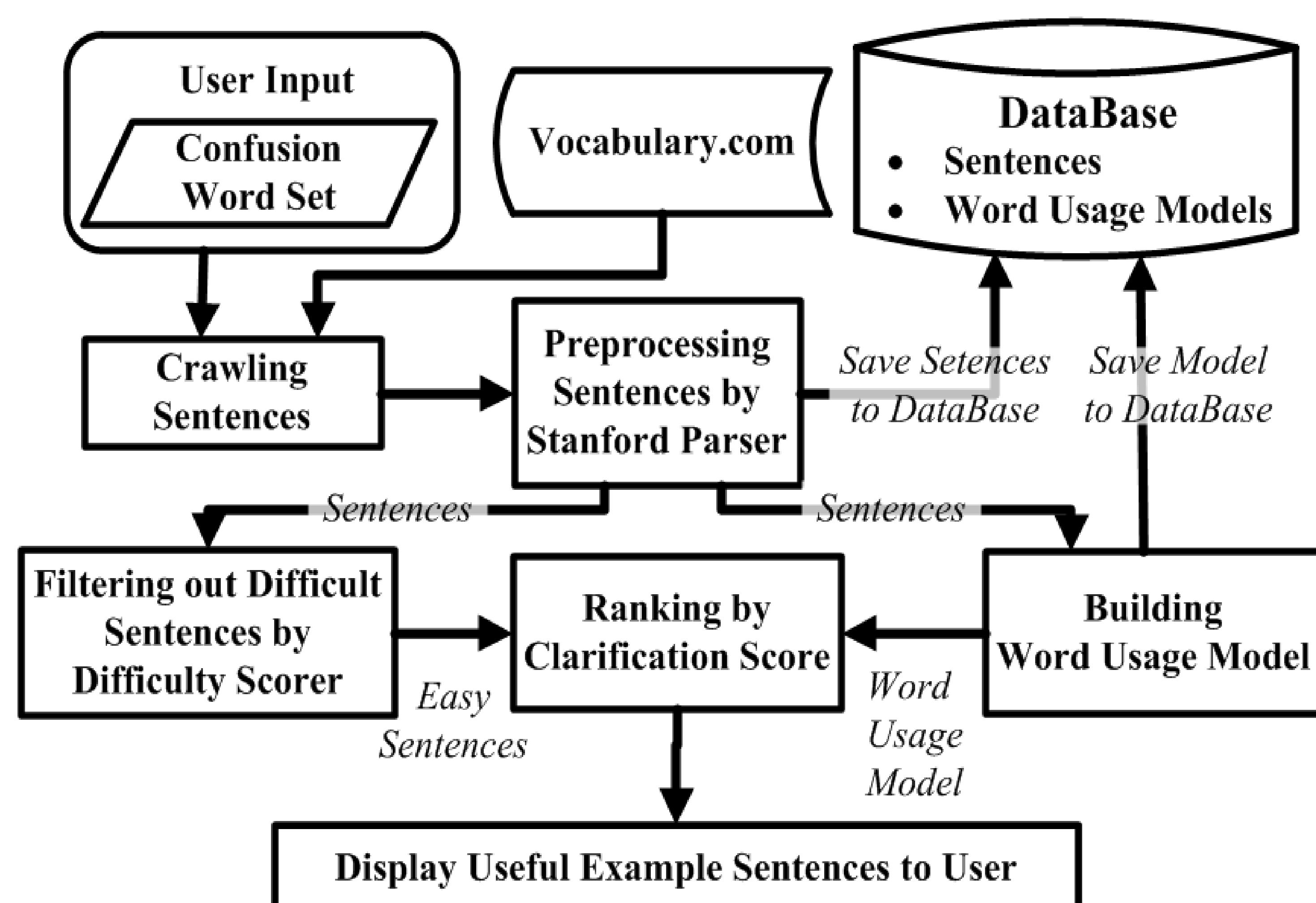
## Introduction

- Previously, we have developed **GiveMeExample**, a system that could automatically suggest example sentences which best highlight the difference between near-synonyms.
- In this paper, we further enhance **GiveMeExample** by
  - Improve word usage model by Bi-LSTM.
  - Add Language support for Chinese.
  - Provide sentences illustration.

## Clarification Score Function

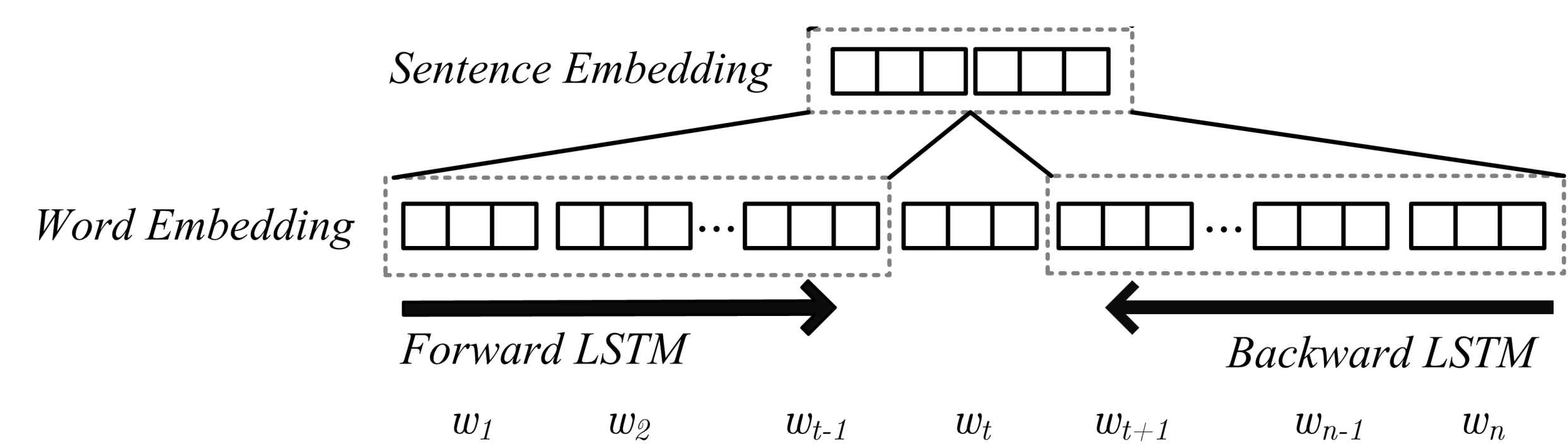
- When searching for the useful example sentences of the target word  $w_i$  in word set  $W$ . There are two related factors.
  - Fitness Score:**  
 $P(s|w_i)$ , measuring appropriateness between  $w_i$  a sentence  $s$ .
  - Relative Closeness:**  
 $\sum_{w_j \in W - w_i} P(s|w_i) - P(s|w_j)$ , the idea is that  $s$  should fit the target word  $w_i$  but be inappropriate for the rest of words in  $W$ .
- We define the clarification scoring function as the multiplication of these two scores:  
$$\text{score}(s|w_i) = P(s|w_i) * \sum_{w_j \in W - w_i} P(s|w_i) - P(s|w_j)$$

## Framework



## Word Usage Model

- Word Usage Model is used to estimate  $P(s|w)$ .
- We propose two model to estimate the probability.
  - Gaussian Mixture Model with context feature
  - Bi-directional Long Short-Term Memory Network



## Evaluation and Discussion

- Evaluate word usage model by **Fill In The Blank** task. GMM and Bi-LSTM achieve 70.26% and 73.05%, respectively. The baseline 5-gram language model is 69.90%.
- Evaluate clarification scoring function by **Example Sentence Suggestion** task.
  - GMM and Bi-LSTM achieve 0.502 and 0.500, respectively.
  - Random-ordered is 0.423. First-seen baseline is 0.429.

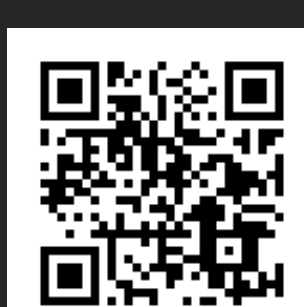
Disagree		Refuse	
1	"You may <b>disagree</b> with members of the community for a while," he told me.	1	The bus boycott began after Parks was arrested for <b>refusing</b> to give up her seat on a Montgomery city bus.
2	Justice Anthony Kennedy <b>disagreed</b> . "Don't make me feel bad because I thought that was pomegranate juice."	2	The document adds: "The school district has <b>refused</b> to pay those amounts."
3	Few Jewish leaders <b>disagree</b> with the notion that anti-Semitism in Europe is on the rise.	3	When Ms. Buchmann demanded a more stable schedule, her employer <b>refused</b> , an experience that is not uncommon.
4	"They threw something together," she said. Volberding <b>disagrees</b> , saying the clemency petition was "thorough."	4	Mr Zhu <b>refused</b> to pay, and said the restaurant owner pulled out a stick and threatened to beat him up.
5	A team of experimental physicists based in the US and China <b>disagrees</b> .	5	But in a letter to investors earlier this month he <b>refused</b> to back down from his conviction.

- Analysis the suggestion result.

Differing Aspect	Near-synonym Pair	Score	Differing Aspect	Near-synonym Pair	Score
abstract vs. concrete	blunder - error	7/10	low vs. high degree	mist - fog	2/10
	维护 - 保护	6/10		经常 - 往往	3/10
formal vs. informal	child - kid	6/10	pejorative vs. favorable	skinny - slim	3/10
	购买 - 买	9/10		产生 - 造成	8/10

## Difficulty Scorer

- Extract syntactic feature and lexical feature.
- Apply linear regression to estimate the difficulty score.
- Data collection
  - English training data is manually labeled by a native speaker, who considers both difficulty of syntactic structure and lexicons.
  - Chinese training data is collected from mock tests for Hanyu Shuiping Kaoshi (HSK). The difficulty degree of a extracted sentence is set to the proficiency degree this sentence comes from.
- GiveMeExample provides three levels for learner, including **beginner**, **medium** and **advanced learner**.



<http://givemeexample.com/GiveMeExample>