Specialist English: Assignment 6

Rebecca J. Stones
rebecca.stones82@nbjl.nankai.edu.cn

Date due: 26 November 2018

This sixth assignment (worth 5% of the final mark) looks at abstracts.

I'll scale the marks on this assignment according to $m \mapsto \min(m, 10)$ for Master's students and $m \mapsto \lceil m/1.3 \rceil$ for Ph.D. students.

My marking will be affected by (a) your English writing, (b) your LaTeX typesetting, (c) your mathematical presentation, and (d) your understanding of the underlying computer science. Basically, I will "peer review" your assignments.

The example abstracts (and snippets) are from published papers as quoted.

Problem 1 These problems relate to the following abstract.

1. Identify the major structural components (if present): introduction, solution, demonstration, and the implications. [2 marks]

2. Ignoring the noise, the introduction component contains the following useful information:

   - social media leads to the increasing use of trajectory data,
   - location-acquisition technology exists, and
   - frequent sub-trajectory data might contain useful information.

   Throwing away the noise, rewrite the introduction component to be more succinct—it should contain these three points of information strung into a single sentence. [3 marks]

ABSTRACT

With the advent of location-based social media and location-acquisition technologies, trajectory data are becoming more and more ubiquitous in the real world. Trajectory pattern mining has received a lot of attention in recent years. Frequent sub-trajectories, in particular, might contain very usable knowledge. In this paper, we define a new trajectory pattern called frequent sub-trajectories with time constraints (FSTTC) that requires not only the same continuous location sequence but also the similar staying time in each location. We present a two-phase approach to find FSTTCs based on suffix tree. Firstly, we select the spatial information from the trajectories and generate location sequences. Then the suffix tree is adopted to mine out the frequent location sequences. Secondly, we cluster all sub-trajectories with the same frequent location sequence with respect to the staying time using modified DBSCAN algorithm to find the densest clusters. Accordingly, the frequent sub-trajectories with time constraints, represented by the clusters, are identified. Experimental results show that our approach is efficient and can find useful and interesting information from the spatio-temporal trajectories.

— Huang, Luo, Wang, UrbComp, 2013.

Problem 2 The following snippets are two examples of demonstration components in abstracts: which is more informative, and why? In your answer, if these are present in the demonstration component, describe: (a) how performance is measured, (b) what is the proposed method compared to, (c) what dataset is used for testing, and (d) how big is the improvement. Otherwise, comment how they are not present. [3 marks]

> ⋮
> mor datasets collected from Weibo and Twitter. The results demonstrate the effectiveness of the proposed end-to-end att-RNN in detecting rumors with multimodal contents.
>
> — Jin et al., MM (2017).

> ⋮
> **of required adders by half. Experimental results show that our proposed scheme improves the energy efficiency by** 45.9%**, the area efficiency by** 93.6% **and achieves 8x of throughput per area compared with the state-of-the-art CMOS-based implementation.**
>
> — Luo et al., ICCAD (2017).

Problem 3 These questions relate to the following abstract.

1. What's wrong with "a large-scale corpora"?
2. How do we fix "and other 12 ones"?
3. The "do" in "We do experiments" is how a child speaks: what is a suitable alternative?
4. What is wrong with using "researches" to mean the plural of "research", and how can we fix this error?
5. The word "different" in "of different years" is difficult to parse (the reader might think "different years to what?"). What is an appropriate replacement for the word "different"?

[5 marks]

> **ABSTRACT**
>
> Online social media yields a large-scale corpora which is fairly informative and sometimes includes many up-to-date entities. The challenging task of expanding entity sets on social media text is to extract more uncommon entities only using several seeds already in hand. In this paper, we present an approach which is able to find novel entities by expanding a small initial seed set on Twitter text. Our method first generates candidate sets on the basis of the semantic similarity feature. Then it jointly utilizes 2 text-based features and other 12 ones which carry social media specific information. With the scores on those features, a ranking model is learned by a supervised algorithm to synthetically score each candidate terms and then the final ranked list is taken as the target expanded set. We do experiments with 24 entity classes on the Twitter corpus and in the expanded sets there come many novel entities which have not been completely detected in previous researches. And the experimental results on the datasets of different years can perfectly consist with the objective law that fresh entities change as time goes on.
>
> — Zhao et al., ICTIR (2018).

Using "researches" as the plural of "research" used to be acceptable (and remains in old dictionaries)—it is almost always considered an error nowadays (as noted in IELTS resources[1,2,3]; see also Figure 1).

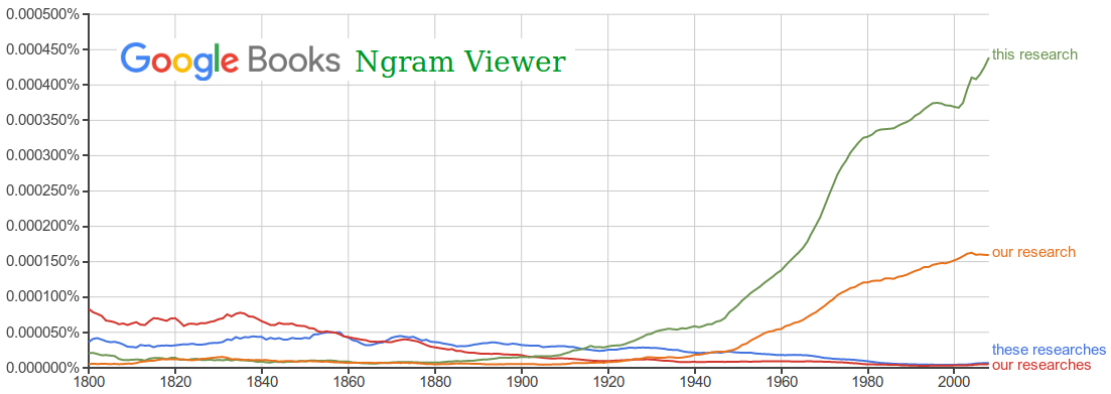

Figure 1: Google Ngram results for "this research", "our research", "these researches", and "our researches".

If you feel like using it, my color scheme is typeset as follows:

- Poor examples: `\begin{tcolorbox} ... \end{tcolorbox}`,
- Good examples: `\begin{tcolorbox}[colback=white,colframe=green!60] ...`,
- Other examples: `\begin{tcolorbox}[colback=blue!15,colframe=blue!50] ...`,
- introduction `\textcolor{red}{introduction}`,
- solution `\textcolor{blue}{solution}`,
- demonstration `\textcolor{green!50!black}{demonstration}`, and
- implications `\textcolor{purple!50!black}{implications}`.

[1] https://ielts.allearsenglish.com/sample-feedback-ielts-writing-task-2/
[2] https://magoosh.com/ielts/5-common-grammar-mistakes-ielts-writing/
[3] http://ielts-simon.com/ielts-help-and-english-pr/2017/08/ielts-grammar-uncountable-nouns.html