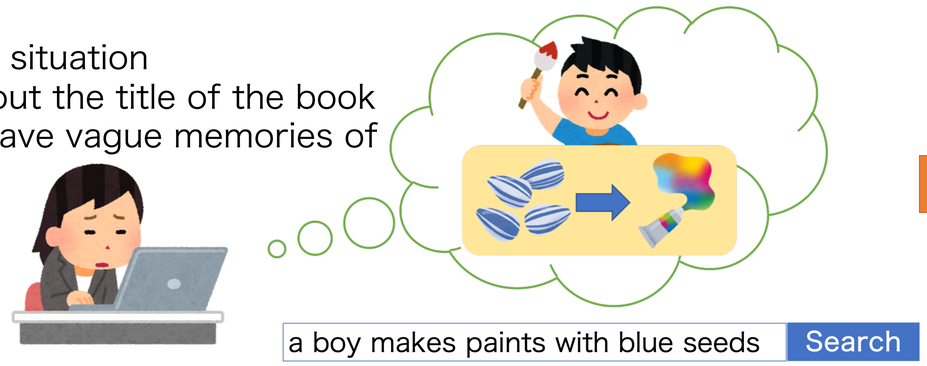
A Ranking Method for Relaxed Queries in Book Search

Momo Kyozuka, Yang Xu, Keishi Tajima (Kyoto University)

Background

- Book search situation
- Not sure about the title of the book
- Users only have vague memories of the stories



Why can't we find the target book?

2 types of query keywords missing
in the description of the book in the book database

- Type 1: the short database description only includes very important keywords
- Type 2: keywords in the query are simply mistaken

Two-step ranking method

A sentence query from a user

In the book,

a boy makes paints with blue seeds.

Generate all subset queries

a boy	make	paint	blue	seed
	make	paint	blue	seed
a boy		paint	blue	seed
⋮		⋮	⋮	⋮
				seed

Step 1: Rank subset queries based on reliability of query words

1	a boy	make	paint	blue	seed
2	a boy		paint		seed
3	a boy		paint	blue	
	⋮		⋮	⋮	⋮
31					seed

[Kyozuka, 2018]

Step 2: Rerank search results of each relaxed query

Original search result of query 1
1. Book A
2. Book B
3. Book C

Search result of query 1 reranked by M(D,W)
1. Book B
2. Book C
3. Book A

$$M(D,W) = \frac{1}{|W|}\sum_{w\in W}\max_{d\in D}\log_2\left(\frac{1}{|H(d)-H(w)|\times H(d)}\right)$$

D: a book description
W: the set of keywords removed by original query
$H(a)$: the number of search result of the word "a"

$M(D,W)$ is based on these two assumptions:

- If a word d in the book description is well-known, it is less likely to be mistaken.
- If a word d in the book description is as well-known as a word w in a original query, people are likely to mistake d for w.

We assume that a well-known word has a large number of its search result.

Experiment

We compared the rank of the target book of the 3 ranking methods:

- only Step 1
- only Step 2
- both Step1 and Step 2 (proposed method)

The title of target books	only Step 1	only Step 2	both
Dwarfs in the house in the woods	13	7	11
Treasure comparison	211	235	207
When the Robbers Came to Cardamom Town	677	256	462
Teacher's report	2	428	2
Terror video game	81	89	77
Adventures of the Polar Cubs	2	269	3
Sebgoegoegoe – An African Story	120	125	76

[Kyozuka, 2018] Momo Kyozuka and Keishi Tajima. Ranking methods for query relaxation in book search. In Proc. of IEEE/WIC/ACM International Conference on Web Intelligence, pp. 466–473, December 2018