Recommending Source Code Examples via API Call Usages and Documentation

Collin McMillan[1], Denys Poshyvanyk[1], Mark Grechanik[2]
[1]Department of Computer Science, College of William & Mary
[2]Accenture Technology Labs

Recommending Examples

Online software repositories contain source code that already implements certain requirements that developers must fulfill. We propose a new method to recommend source code examples from these repositories to developers. Our approach queries against the Application Programming Interface (API) calls made by those examples.

We use API Calls to Locate Relevant Source Code

Figures 1 through 3 illustrate our approach with an example. In the depicted situation, a programmer wants to view the contents of a compressed archive file. He enters a natural-language query to that effect.

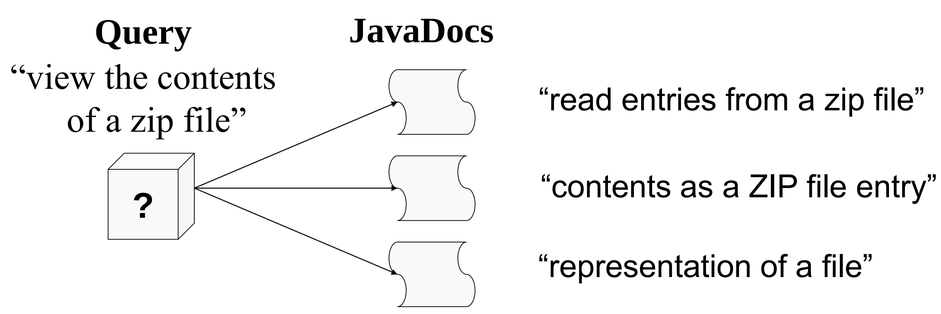


Figure 1. First, we match keywords in the query to words in Java documentation[1].

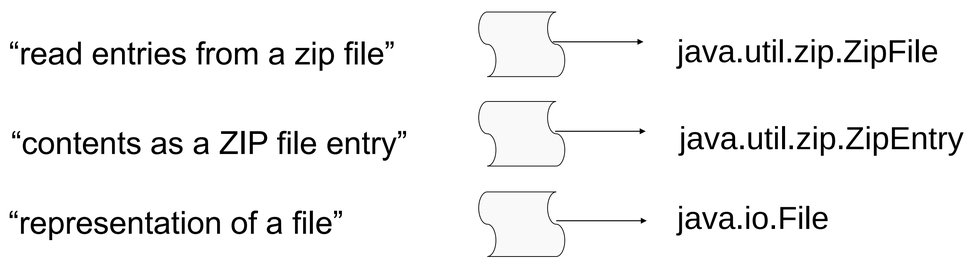


Figure 2. Second, we link the relevant parts of documentation to specific classes in the API.

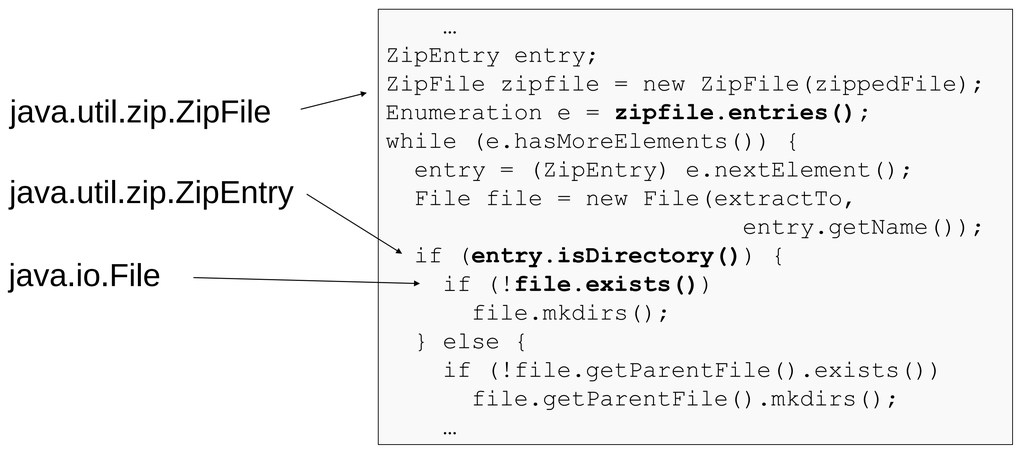


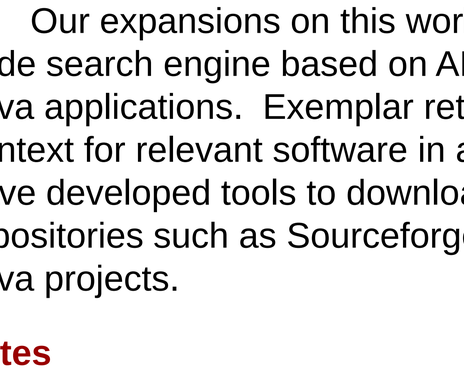
Figure 3. Finally, we recommend examples to the programmer based on the API calls used in those examples.

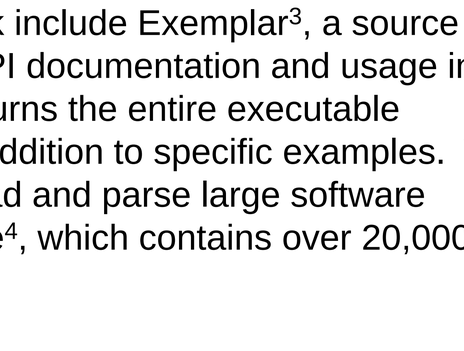
Empirical Evaluation

Our preliminary evaluation focuses on a list of 40 publicly available Java examples[2]. We used the examples' descriptions as queries for our system; this process results in a one-to-one mapping of queries to examples for our study. We define the metrics *accuracy* and *discovery* to evaluate our approach's top 25 recommendations.

$$accuracy = 104 - n * (rank\ of\ correct\ result)$$

$$discovery = \frac{(queries\ with\ positive\ accuracy)}{(total\ number\ of\ queries)}$$





Conclusions

We found that our approach recommends with high accuracy but low discovery – if our system provides the correct result, it will occur within the top three answers. This work is a step towards building effective recommender systems for software reuse that combine source code and usage documentation.

Expanded Work

Our expansions on this work include Exemplar[3], a source code search engine based on API documentation and usage in Java applications. Exemplar returns the entire executable context for relevant software in addition to specific examples. We have developed tools to download and parse large software repositories such as Sourceforge[4], which contains over 20,000 Java projects.

Notes

[1] In our case, the official Sun documentation: http://java.sun.com/j2se/1.5.0/docs/api/

[2] http://www.java2s.com/

[3] http://www.xemplar.org/

[4] http://www.sourceforge.net/

Acknowledgments

We gratefully acknowledge Chen Fu and Qing Xie for their contributions to this and ongoing work.

Publications

McMillan, C., Poshyvanyk, D., and Grechanik, M., "Recommending Source Code Examples via API Call Usages and Documentation", *in Proc. of 2nd ICSE2010 International Workshop on Recommendation Systems for Software Engineering (RSSE'10)*, Cape Town, South Africa, May 4, 2010