

Generating Acrostics via Paraphrasing and Heuristic Search – Final Presentation

Bruno Soares Fillmann
Fernando Bombardelli da Silva
Jürgen Bauer
William Bombardelli da Silva

Technische Universität Berlin
Datenbanksysteme und Informationsmanagement
DBPRO – Database Projects (WS 2014/2015)

09.02.2015

Organization

- 1 Goal of the Project
- 2 Subtasks of the original goal
- 3 How we reached the goals?
- 4 Which difficulties occur?
- 5 Unreached Goals
- 6 Example
- 7 References

Goal of the Project

Goal: Implement the methods and techniques to generate acrostics as described in [1] for German language.

Subtasks of the original goal

- Understand the paper and the method
- Design an architecture for the application
- Collect comparable German text corpora, statistics and libraries (TEX Hyphenation, NetSpeakAPI, Thesaurus, etc.)
- Implement the most promising paraphrasing operators (line break, hyphenation, wrong hyphenation, word insertion/deletion, synonym, spelling)
- Implement the A^* -algorithm
- Evaluate the results

How we reached the goals?

- UML-diagram for the application
- Sequence diagram for the A^* -algorithm
- Developed the application in Java 8 with Netbeans
- Use various libraries, web services and databases to implement the operators
- Implemented a test application for evaluating the results, with a timeout after 15 min

Which difficulties occur?

- Synonym operation only inserts synonyms without considering the context (Google N-Gram)
- NetSpeakAPI don't yields many results for German language
- Organization of the input text as line:
- more than one letter at a time could be generated, misfits the theoretical framework of the A^* -algorithm as described in the paper.

Which difficulties occur?

- **Recommendation:** view the text as one single line and apply line break only to generate a letter. When the acrostic is generated, break the lines according to the constraints.

Unreached Goals

- only the most promising operators were implemented
- no real operator to generate the first letter of the acoustic
- no operator **not** mentioned in the paper was implemented

Example

Here it would be great to have an example without wrong hyphenation.

References



Benno Stein, Matthias Hagen, and Christof Bräutigam. *Generating Acrostics via Paraphrasing and Heuristic Search*.

In Junichi Tsujii and Jan Hajic, editors, 25th International Conference on Computational Linguistics (COLING 14), pages 2018-2029, August 2014. Association for Computational Linguistics.



Spiegel Online. *Griechenland und die Euro-Zone: Der fast unmögliche Rausschmiss*.

<http://www.spiegel.de/wirtschaft/soziales/griechenland-euro-austritt-waere-moeglich-aber-kompliziert-a-1011361.html>. Am 07. Januar 2015



Hilko. *Blau ist keine traurige Farbe*.

<https://deutschlich.wordpress.com/2013/03/12/blau-ist-keine-traurige-farbe>. Am 07. Januar 2015



Redis. <http://redis.io/>. Am 07. Januar 2015



Synonym - *OpenThesaurus - Deutscher Thesaurus*.

<https://www.openthesaurus.de/>. Am 07. Januar 2015

Questions?