Wie funktioniert eigentlich Google Translate?

Utz Pöhlmann

Universität Hamburg - Proseminar: Künstliche Intelligenz

4poehlma@informatik.uni-hamburg.de

2. Juli 2015

Gliederung

- Einleitung 2
 - Vorschau
 - Beispiel
- 2 Sprachübersetzungssysteme
 - Übersetzung von Grund auf
 - 🐧 🏅 Schablonentechnik
- 7 Fazit & Awblich
 - Aussicht
 - Fazit
 - Quellen

nicht zwiele Details an dieser Stelle P

an rican

Warum Sprachübersetzungssysteme?

Mit dieser Folie vor der Sliederung starter?

• 7000 Sprachen weltweit

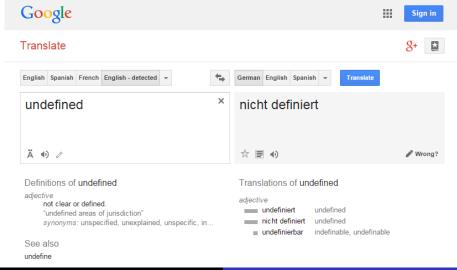
Warum Sprachübersetzungssysteme?

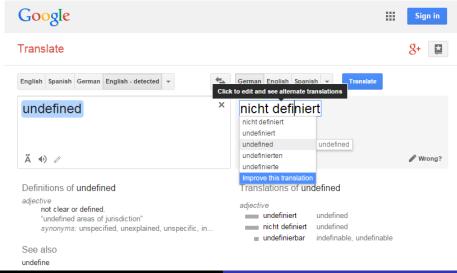
- 7000 Sprachen weltweit
- Reisen sind "IN"

Warum Sprachübersetzungssysteme?

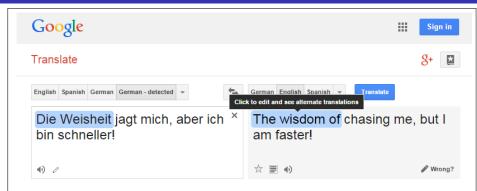
- 7000 Sprachen weltweit
- Reisen sind "IN"
- Übersetzung überall und für jedermann

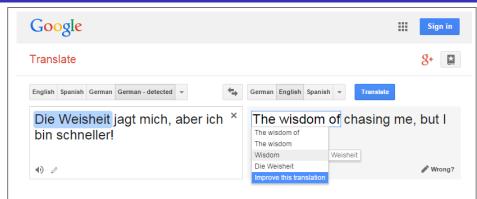


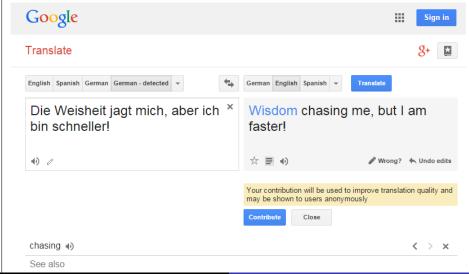




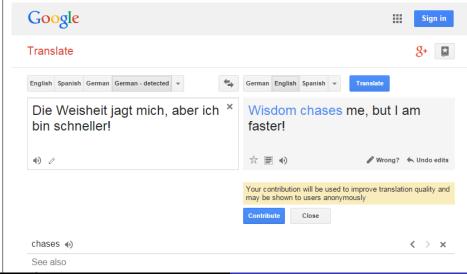


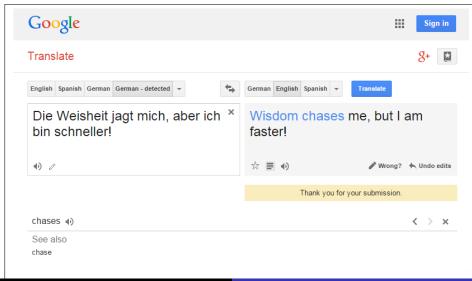












Wort für Wort

Wort f
 ür Wort

Schablonentechnik

- Wort f
 ür Wort
 - Übersetzen Marke Lexikon
- Schablonentechnik

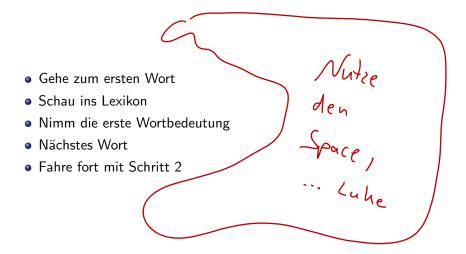
- Wort f
 ür Wort
 - Übersetzen Marke Lexikon
 - Sinnabschnitte und Wahrscheinlichkeiten
- Schablonentechnik

• Gehe zum ersten Wort

- Gehe zum ersten Wort
- Schau ins Lexikon

- Gehe zum ersten Wort
- Schau ins Lexikon
- Nimm die erste Wortbedeutung

- Gehe zum ersten Wort
- Schau ins Lexikon
- Nimm die erste Wortbedeutung
- Nächstes Wort



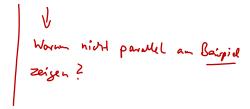
Parallelen

- Parallelen
- Worttypen

- Parallelen
- Worttypen
- Voraussagen

- Parallelen
- Worttypen
- Voraussagen
- Filterung

- Parallelen
- Worttypen
- Voraussagen
- Filterung
- Übersetzen



Textabschnitte in anderen Texten wiederfinden

- Textabschnitte in anderen Texten wiederfinden
- Übersetzung aus dem anderen Text nehmen

Jeberfalls: glick an Beispiel aufbanen V

- Textabschnitte in anderen Texten wiederfinden
- Übersetzung aus dem anderen Text nehmen
- Als Schablone behalten

- Textabschnitte in anderen Texten wiederfinden
- Übersetzung aus dem anderen Text nehmen
- Als Schablone behalten
- Schablonen, die am wenigsten in Konflikt stehen nehmen

- Textabschnitte in anderen Texten wiederfinden
- Übersetzung aus dem anderen Text nehmen
- Als Schablone behalten
- Schablonen, die am wenigsten in Konflikt stehen nehmen
- Diese Schablonen als aktuelle Übersetzung ausgeben

Aussicht

• Diverse Versuche, bessere Übersetzungssysteme herzustellen

Aussicht

Vermei de Meta-talle ⇒ Kanbrete Probleme /Fabiter ⇒ Sechicle Grinde

- Diverse Versuche, bessere Übersetzungssysteme herzustellen للمانية المانية المان
- Systeme aber bisher immer schlechter als Google Translate

Fazit

• Internetübersetzungssysteme recht gut

Fazit

• Internetübersetzungssysteme recht gut

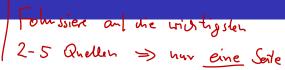
Wird ständig verbessert

Siehe letzle Folie ? >> bonlerete Fahter, vermoide hohle phrasen

Fazit

- Internetübersetzungssysteme recht gut
- Wird ständig verbessert
- Gänzlich neue Lösung jedoch nicht in Sicht

Quellenverzeichnis I





Kamal Nicholas (2015)

Google Translate: Über 80 Sprachen verstehen

http://www.giga.de/webapps/google-translate/ Absatz 3, Mai 2015.



Adak, C. (2014)

A bilingual machine translation system: English amp; Bengali

Automation, Control, Energy and Systems (ACES), 2014 First International Conference on pages 1 – 4, Feb 2014.



Brkic, M. and Mikulic, B.B. and Matetic, M. (2012)

Can we beat google translate?

Information Technology Interfaces (ITI), Proceedings of the ITI 2012 34th International Conference on pages 381 – 386, June 2012.

Quellenverzeichnis II



Klyuev, V. and Haralambous, Y. (2012)

Query translation for CLIR: EWC vs. Google Translate

Information Science and Technology (ICIST), 2012 International Conference on pages 707 – 711, March 2012.



Miltsakaki, Eleni and Robaldo, Livio and Lee, Alan and Joshi, Aravind (2008)

Sense annotation in the penn discourse treebank

Computational Linguistics and Intelligent Text Processing pages 275 – 286, Springer, 2008.



Wan, Xiaojun (2008)

Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis

Proceedings of the Conference on Empirical Methods in Natural Language Processing pages 553 – 561. Association for Computational Linguistics, 2008.



Quellenverzeichnis III



Demberg, Vera (2007)

A language-independent unsupervised model for morphological segmentation *ANNUAL MEETING-ASSOCIATION FOR COMPUTATIONAL LINGUISTICS* volume 45, page 920. Citeseer, 2007.



Toutanova, Kristina and Ilhan, H Tolga and Manning, Christopher D (2002)

Extensions to HMM-based statistical word alignment models

Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10 pages 87 – 94. Association for Computational Linguistics, 2002.



Kaji, Hiroyuki and Kida, Yuuko and Morimoto, Yasutsugu (1992)

Learning translation templates from bilingual text

Proceedings of the 14th conference on Computational linguistics-Volume 2 pages 672 – 678. Assosiacion of Computational Linguistics, 1992.



Ende

Ende

Noch Fragen?



Einleitung Sprachübersetzungssysteme Fazit

Krasser Stuff der mir eventuell noch einfällt.