

Questions for web-based question answering

- What are the major advantages for answering questions using snippets returned by search engines instead of complete documents ?
- Why are the available corpora for QA (which usually consist of elements of form $\langle Q, A, C \rangle$, where Q is a questions, A its answer, and C the context in which A occurs) usefull for Machine Learning based approaches to QA?
- We are using a word-distance statistics for answering factoid questions from search engine snippets. Explain briefly what the word-distance statistic is.
- For the following sentences „John loves Maria. John kisses Maria every night. Maria talks to John in the kitchen. John gives the book to Maria.“
 - Count the word-distance statistics for the word pair (John, Maria).
- For a definition question like „What is Ubuntu?“ and the following two patterns:
 - X (is|are) (a|the) Y
 - X (become|becomes) Y
- Create the possible queries which are send to a search engine to determine usefull snippets.