

# Questions on “Natural Language and Speech”

4190.408 Artificial Intelligence  
Department of Computer Science and Engineering  
Seoul National University

Prof. Byoung-Tak Zhang

October 11, 2012

1. Why is natural language difficult for computers to understand? Analysis of natural language involves different levels of processing, such as phonetics, morphology, syntax, semantics, and pragmatics. Explain each of the processing levels. Use the sentence “John saw the man with a telescope” as an example. Give parse trees for the sentence. Give phrase-structure grammar rules for generating these and related sentences.
2. What is machine translation? What is Systran? How does Systran measure its performance in translation? Google Translate is based on Systran in part. What is the relation (hint: refer to Wikipedia for Siri and Systran). What is EUROTRA project?
3. What are question answering systems? Explain the BASEBALL system. What is the task? How does the system work? What does the STUDENT system? Why is interesting? What might be the technical difficulties?
4. What is SHRDLU? Use typical dialogue of the system to explain its functions. What does it mean for an NLP system to understand a sentence? How do you know it? How do you know it in SHRDLU?
5. Natural languages queries have been used as front-end systems for accessing databases and computer systems. Explain the following natural language interface systems: LIFER, LADDER, CHAT-80, and TEAM. What kinds of grammars are used in these systems?
6. What is statistical NLP? What are probabilistic context-free grammars (PCFGs)? Give examples of PCFG grammar rules. Give a PCFG parse trees for “John shot elephants in pajamas”. Give some names of grammars developed for natural language processing.
7. What is speech recognition? How is it different from speech understanding? Can we recognize speech without understanding it? What have been the problem domains of speech projects SPEECHLIS and WHIM?
8. Describe the technology underlying the speech recognition systems of Dragon, HARPY, and HEARSAY-II. How did Dragon use hidden Markov models for speech recognition? How did HARPY use a network structure to understand spoken sentences? What are the features of the blackboard architecture developed for HEARSAY-II project?