Language Technology 2
Summer Semester 2008
Exam preparation questions

1 Multilingual Resources

1. What does the term "hybrid" NLP refer to?
2. What is the Grammar Matrix?
3. Describe the Heart of Gold (HoG) architecture.
4. Describe the ParGram architecture.
5. What are the conceptual differences between grammar engineering with LFG and grammar engineering with HPSG?
6. Describe how multilingual NLP can serve the task of Machine Translation.
7. Give a brief overview of Robust Minimal Recursion Semantics (RMRS).
8. What does the term "parallel grammar development" refer to?
9. What is parallel grammar development good for?

2 Machine Translation

1. Statistical machine translation systems are generally the result of an iterative process in which the system is modified, then tested, then modified again, then tested again, and so on. The testing phase is extensive and time consuming and it is generally agreed that it must therefore be done automatically. Describe in general terms how this automatic testing is typically done.
2. When the producer of a document has it translated, we often speak of "outbound" translation. When the consumer has a document translated, we speak of "inbound" translation. It is often claimed that statistical machine translation is better suited to inbound translation. What reasons might there be for holding this belief?
3. So-called "rule based" machine translation generally involves analysis of the source text on a variety of levels, each more "abstract" the one before. What is meant by saying that one level, say that of syntax, is more "abstract" than another, say morphology?

3 Natural Language Interaction

1. What is a conversation turn? Explain the turn taking rules.
2. Explain (and illustrate) what collaboration in dialogue means. What are collaborative responses?
3. Explain what the concept of push-to-talk is in the implementation of turn taking in a dialogue system.
4. List and explain the Gricean maxims of cooperation.
5. What are adjacency pairs in dialogue, what characteristics do they have, and how is the concept of adjacency pair useful in dialogue modeling?
6. What is the general global structure of a conversation? Give some typical examples of the respective parts.
7. Explain the notion of a speech act.
8. Explain the notion of common ground. When is a piece of information considered common ground between dialogue participants?
9. What is the process of grounding?
10. What levels of interpretation does Clark (1996) distinguish? Explain what they mean and illustrate with examples.
11. What are the various reasons for grounding problems (why can an agent have difficulties grounding an utterance)? Explain and illustrate.
12. What (three) kinds of verification strategies can a grounding model of a dialogue system implement? Discuss advantages and disadvantages.
13. Explain the difference between finite state machine-based, frame-based and information-state update based dialogue models. What are their advantages and disadvantages? Give examples of applications in which you would use them.
14. Explain the concepts behind information state update based dialogue modelling. What kinds of rules have to be defined in such a model?
15. What is the difference between fixed and mixed initiative? What advantages and disadvantaged has fixed system initiative in comparison to mixed initiative model?
16. Explain the concepts of technical evaluation, usability evaluation, and customer evaluation. Give some (at least 3) questions that users may be asked to assess the usability of a system. Discuss the PARADISE and SASSI frameworks?
17. What are basic motivations to carry out WoZ experiments? What are the challenges in conducting a WoZ experiment? What requirements must a WoZ experiment meet in order to serve the purpose of a simulated human-machine interaction? Give examples for methods or tools that are employed to meet the requirements.

18. Reinforcement Learning (RL) is based on the concept of Markov Decision Process (MDP). Give a short and informal description of the components of an MDP and the relevant notions (i.a. the Markov Property and the three kinds of rewards). Explain how MDP is used in RL in general. RL for natural languages interaction uses recorded dialogue data for training. There are two kinds of information derived from the empirical data which is fed to the learning process. What are they?

19. Explain the difference between model-based and simulation-based reinforcement learning.

20. Explain the concepts of exploration vs. exploitation in reinforcement learning.