

Linguistics 384

Homework 6

Dialogue Systems and iCALL

DUE: Wednesday, November 30, 2005

1. (20 points) There's a transcript from the Callhome corpus I want you to look at.

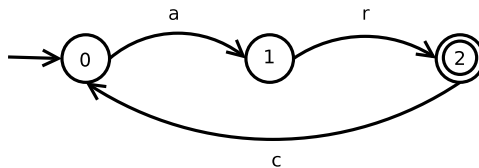
http://ling.osu.edu/~adriane/384/files/en_4248.txt

The Callhome corpus is a collection of transcripts of telephone conversations. You don't have to read the whole thing, but you'll probably want to read everything up to the point mentioned in order to understand the context.

On the left, you'll see the starting time of each speaker (A and B are the speakers) followed by the ending point for that line. e.g. the first line says that B begins speaking at time point 89.22 and ends at 92.90.

- (a) At point 171.25, who/what do you think *we* refers to?
 - (b) Make a list of all the nouns from the previous (contentful) utterance – at point 166.35 – and any other nouns which are somehow still “prominent” from earlier in the dialogue. (Ignore content-less items like *uh so*.)
 - (c) We talked in class about syntactic/semantic constraints on pronoun reference. When trying to figure out what *we* refers to, what constraints rule out the other noun(s) from this list?
2. (20 points)

- (a) Draw a finite-state automaton (FSA) for the regular expression $sha(na)^*$, that is, an FSA which covers *sha*, *shana*, *shanana*, *shananana*, and so on.
- (b) Here's an FSA:



- i. What is the shortest string this FSA will generate?
- ii. What are two other strings it will generate?

- iii. Write a regular expression to capture this FSA.
3. (20 points) We know that in English, the words *guy* and *gentleman* mean more or less the same thing, but they do differ in terms of usage.
- (a) What would you tell a non-native speaker if s/he asks you what the differences are?
 - (b) Go to http://www.lex Tutor.ca/concordancers/concord_e.html and retrieve concordances for each of the two words. Quote a few relevant lines to support your claim.
 - (c) Do you find anything new about how people use these two words that you were not aware of? If yes, describe the new finding and quote those lines that support it. If not, discuss whether concordancers are a good help for learning word usage.
4. (20 points)
- (a) Suppose a learner types in the following in some CALL system designed for people learning English:
I loves the chalkboard.
What is the learner's problem?
 - (b) Now suppose the learner also types in the following:¹
He hates dust.
You runs a lot.
How would keeping a student model help us pinpoint the user's problem more precisely?
 - (c) The correct rule for subject verb agreement in English is that the subject and the verb must agree in *person* and *number*, i.e. a first person singular subject (*I*) must have a first person singular verb (e.g. *am, love, ...*). Describe a mal-rule, or mal-rules, which would allow the system to process the ill-formed input *I loves, I is*,
5. (20 points) In class, we examined the pros and cons of the "Dictation" exercise. We quickly looked at 5 other exercise types (Build a Phrase, Which Word is Different, Word Order Practice, and Fill-in-the-Blank). Pick one other exercise type and give me at least 2 pros and 2 cons of using that type of exercise for language learning. (Think about what skill the exercise tests, what technology it requires, etc.)

¹Just so you're not second-guessing yourself, the *hates* sentence is a correct English sentence.