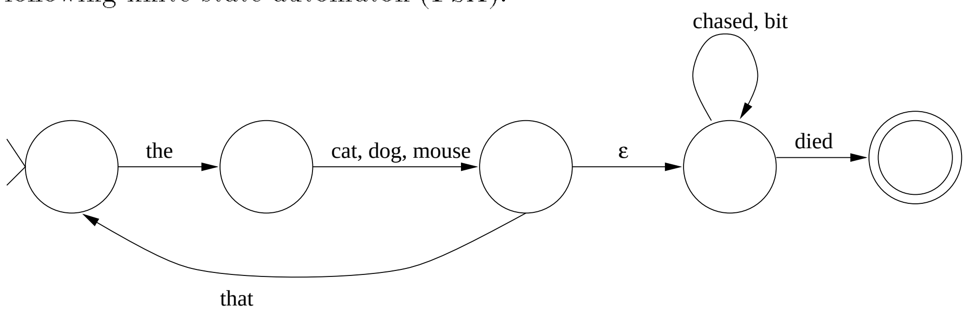
LING645/CMSC723 Midterm 1
October 8, 1997

Name:

Please read through the <u>*entire*</u> *exam once before you start. Read carefully and make sure to answer every part of each question. (You can use the checklist on the last page to make sure.)*

---

1. Consider the following finite-state automaton (FSA):



As you can see, this FSA accepts an infinite language; let's call it L:

the cat died

the mouse that the cat bit died

the mouse that the cat that the dog chased bit died

⋮

(a) Give an example showing that the above FSA overgenerates, i.e. show that L contains strings that are not grammatical sentences of English. [5 pts]

(b) *Briefly* explain the difference between recognition and parsing. [5 pts]

(c) Consider just the strings in L that are grammatical English sentences — for example, the three examples given above. Could a finite-state machine provide a useful syntactic analysis for all of those strings? Explain why or why not. [10 pts]

(d) It has been argued that English is not a regular (finite-state) language. Would this argument hold true if you didn't make a distinction between competence and performance? Explain why or why not. [10 pts]

2. Consider the following infinite language:

the mouse died

the cat died

the dog died

the mouse that the cat bit died

the cat that the mouse bit died

⋮

the mouse that the cat that the dog chased bit died

the cat that the mouse that the dog chased bit died

⋮

Note that the language does not contain grammatical English sentences like

the cat bit the mouse that died

the cat that died bit the mouse

and also notice that it does not contain ungrammatical strings like

*the cat chased

*the cat that the dog died chased

*the cat that the dog bit chased

(a) Give a context-free grammar (CFG) that generates this language. (Note that this must be a *true* context-free grammar — not an RTN, and not a grammar containing regular expressions within the rules.) [25 pts] *Hint: you might want to consider question 2(b) first.*

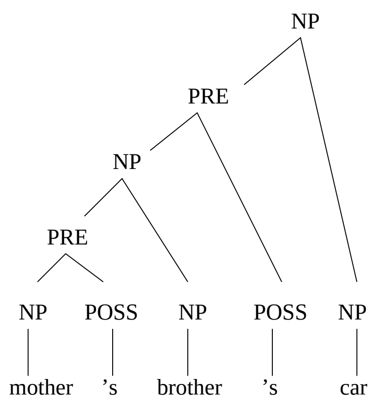
(b) Give a syntactic analysis, in the form of a parse tree, for *the mouse that the cat bit died*, according to the grammar you wrote in question 2a. (If you were unable to complete 2a, give a reasonable parse tree consistent with the constituent structure of English.) [5 pts]

---

3. Consider the following grammar:

NP	$\rightarrow$	PRE NP
PRE	$\rightarrow$	NP POSS
POSS	$\rightarrow$	's
NP	$\rightarrow$	mother | brother | father | car

And consider the following analysis of *mother 's brother 's car*:



(a) Show how *one* of the following algorithms would work when run on the input *mother 's brother 's car*, analyzing it consistent with the structure above. [20 pts]

Whichever algorithm you choose, you should follow the format we used for that algorithm in class and on the homework assignments:

Algorithm	What to show
CKY	Filled-in table
Earley's algorithm	State sets, with operations
Top down (recursive descent)	Series of configurations, with operations
Bottom up (shift reduce)	Series of configurations, with operations
Left corner	Series of configurations, with operations

(b) Which parsing algorithm did you use in question 3a? ____________________
Explain *briefly* how nondeterminism is dealt with when using that algorithm. [10 pts]

Checklist:

- ☐ 1a ____________
- ☐ 1b ____________
- ☐ 1c ____________
- ☐ 1d ____________

- ☐ 2a ____________
- ☐ 2b ____________

- ☐ 3a ____________
- ☐ 3b ____________