

## 2006 Paper 7 Question 12

### Natural Language Processing

The following shows a simple context free grammar (CFG) for a fragment of English.

S → NP VP	Adj → angry	Vbe → is
VP → Vbe Adj	Adj → big	N → dog
NP → Det N	Adj → former	N → cat
N → Adj N	P → at	
Adj → Adj PP	P → on	
PP → P NP	Det → the	

(a) Show the parse tree that this grammar would assign to (1).

(1) the dog is angry at the cat

[3 marks]

(b) One respect in which this grammar overgenerates is that some adjectives, including *former*, occur only before a noun (see (2)) and that PPs do not combine with adjectives occurring before a noun (see (3)).

(2) \* the dog is former

(3) \* the angry at the cat dog is big

Show how the grammar given above could be modified to prevent this type of overgeneration. [4 marks]

(c) The grammar also behaves incorrectly with examples (4), (5) and (6):

(4) \* the dog is big at the cat (*big* does not take a PP)

(5) \* the dog is angry on the cat (*angry* only takes PPs where the P is *at*)

(6) \* the dog is angry at the cat at the cat (adjectives may not combine with multiple PPs)

Show modifications to the grammar which would prevent these types of overgeneration. [5 marks]

(d) Describe how the overgeneration in part (c) could be dealt with in a feature structure (FS) grammar, giving full lexical entries for *angry* and *big* and details of rules and other lexical entries as necessary to explain your account.

[8 marks]