Natural Language Processing

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

\[
\begin{align*}
S & \rightarrow NP \ VP \\
VP & \rightarrow Vbe \ Adj \\
NP & \rightarrow Det \ N \\
N & \rightarrow Adj \ N \\
Adj & \rightarrow Adj \ PP \\
PP & \rightarrow P \ NP \\
Vbe & \rightarrow is \\
Adj & \rightarrow angry \\
N & \rightarrow dog \\
N & \rightarrow cat \\
Adj & \rightarrow former \\
P & \rightarrow at \\
P & \rightarrow on \\
Det & \rightarrow the
\end{align*}
\]

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

(1) the dog is angry at the cat  

(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.

(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.

(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.