

10 Natural Language Processing (AAC)

- (a) Use the following text to derive distributions for *rat* and *chased*. Use a five-word window, including open- and closed- class words, ignore case, punctuation and sentence boundaries and weight contexts by frequency.

The cat chased a rat. A big rat chased the big dog.

[4 marks]

- (b) Show unlabelled dependency links for each sentence and give distributions for *rat* and *chased* using contexts derived from single dependency links. [5 marks]
- (c) In general, what type of differences will arise in distributions if dependency links are used instead of word windows as context? [4 marks]
- (d) Outline one or more experiments to determine whether distributional information could be used in conjunction with a syntactic parser to help disambiguate prepositional phrase attachment ambiguities. [7 marks]