

Natural Language Processing 1 – 2018 – Paper 2

1 Parsing (submit to your TA by 7 December)

The following shows a simple context-free grammar:

```
S -> NP VP      N -> N PP      Det -> the
PP -> P NP      N -> cat       P -> in
NP -> Det N     N -> box       VP -> snored
N -> Adj N     Adj -> big
```

- (a) Describe the concept of a *chart* as used in natural language parsing. [4 marks]
- (b) Assuming that the sample grammar is applied by a bottom-up parsing algorithm without packing, show the chart that would be produced from a complete analysis of the following sentence:

the big cat in the box snored

[5 marks]

- (c) Show what the differences would be between a packed chart and the chart you gave in answer to part (b). [2 marks]
- (d) What is meant by the term *overgeneration*? Why might overgeneration be a problem when using a grammar for parsing grammatical text? [3 marks]
- (e) Outline how techniques from distributional semantics can be used in conjunction with a syntactic parser to help disambiguate prepositional phrase attachment ambiguities. [6 marks]