

11 Natural Language Processing (AAC)

- (a) The following text is from a children’s story by Beatrix Potter (with slight modifications):

Mr Jeremy put on a macintosh, and a pair of shiny shoes; he took his fishing rod and basket, and set off with enormous hops to the place where he kept his boat. The boat was round and green, and very like the other lily-leaves. It was tied to a water-plant in the middle of the pond.

Describe six features that are used in pronoun resolution algorithms using classifiers. For each feature, explain the range of values it can take, giving illustrative examples from the text above, and list any resources or systems necessary to derive the feature. [8 marks]

- (b) Describe the data you would need to train and test a classifier for pronoun resolution that made use of these features. Illustrate your answer with examples from the text in (a). [3 marks]
- (c) Given a trained classifier, how could you set up an experiment to evaluate its performance on some new text? [3 marks]
- (d) On conducting such an experiment, you obtain worse performance than the results previously reported for a standard dataset. How would you investigate possible causes of the difference? [6 marks]