START OF QUIZ Student ID: 97713317,Song,Shawn

Topic: Lecture 8 Source: Lecture 8

What information do you think the word tokens on the stack/buffer are providing to the ML SR parser? (1)

Topic: Lecture 5 Source: Lecture 5

Imagine we extended our algorithm to allow for ternary branching (ie, 3 NTs on the RHS). What impact would that have on the complexity of the CYK algorithm? (1)

Topic: Lecture 7 Source: Lecture 7

Briefly describe how the stack changes for a ARC-RIGHT operation. (1)

Topic: Lecture 6 Source: Lecture 6

Describe the purpose of the dot. (1)

Topic: Lecture 7 Source: Lecture 7

Briefly describe how dependency parsing differs from constituency parsing. (1)

Topic: Lecture 6 Source: Lecture 6

What difficulties do you envision when using the Earley parser with a language with large amounts of agreement? (2)

Topic: Lecture 5 Source: Lecture 5

Let's say we wanted to modify PARSEVAL to take ambiguity into account. How might we use a PCFG and two gold references to account for ambiguous parsing? (2)

Topic: Lecture 8 Source: Lecture 8

Imagine that we have a dependency parser that has a very good UAS (90+), but a very bad LAS (50-). Do you think that we could use the output of this parser as input to a neural translation model as is, or do you think that we should first re-train the labeling part of the algorithm to increase LAS? Doing both is probably the best solution, but I'm asking if you think that we could use the output of the existing model, even as we try to improve the quality of the labels. Explain. (2)

Topic: Long

Source: Lecture 6

Often, modern NLP tools work not with words, but with subword units. What modifications would we need to make to the Earley parser in order to work with subword units (for example: "agreement" might get split into "agree" and "-ment"). Where would they need to occur in the parser, and how do you think it might benefit and harm the algorithm? (3)

END OF QUIZ