CS585 Natural Language Processing - HW2: Parsing CFGs

Group Members: Cristobal Sarome (A20416473), Akhil Suryadevara(A20391322)

Evaluation of the grammar in its current form.

With the current grammar file simple.gr the parser can only parse very simple declarative sentences and yes-no interrogatory sentences (excluding question mark). For example, the current grammar allows to parse the sentences: “he lives”, “he drives” and “she steals” but fails to parse sentences like: “he lives and drives” or “he drives and she steals”.

The terminal symbols available are very limited, so it only can parse a very small number of very simple sentences.

The parser has no tokenization, namely doesn’t normalize the words neither does it handle punctuation marks, so it fails to parse proper sentences that include uppercase words, commas, exclamation marks, etc. It also fail to parse imperative sentences, because they start with a verb.

An additional failure it has is that parse structures that are not constituents because they are grammatically incorrect (for example “he drive”) as actual sentences. That means that this parser does not subject-verb agreement into account.d