**Assignment 3**

Dixitha Kasturi

[dkasturi@syr.edu](mailto:dkasturi@syr.edu)

**Topic : Context Free Grammars**

**Note :** To run the .py file I had to create a notebook and use the ‘%run path’ command

**Part 1 :**

1. **Handling sentences from sentences.txt**
2. *Arthur is the king*
3. *Arthur rides the horse near the castle .*

The two sentences get parsed by the default grammar that is given.

Text

Description automatically generated

1. *Arthur rides the plodding horse near the castle .*

We add adject noun preposition phrase to the noun phrase for this sentence to be parsed.

**NP -> JJ Noun NP**

After adding this rule the sentence has the following tree structure.

Graphical user interface, text

Description automatically generated

1. *the Holy\_Grail is a chalice .*

We have a new word, we change misc to NNP. For this NNP to be parsed , we add it as a rule to NP

**NP -> NNP**

*Text

Description automatically generated*

1. *the sensational Holy\_Grail is a sacred chalice .*

For this we didn’t need to add a new rule as the rule used for sentence 3 parsed this sentence as well. The output for this is

Text

Description automatically generated with medium confidence

1. *every coconut was carried to the hottest mountains .*

The word ‘carried’ was in both verb past tense and past participle. Here it is past participle. So to handle this we need to add new rule to the verb phrase. The plural form of adverb i.e, hottest is combined with plural form of noun, this rule is added to the noun phrase. ‘to’ is added to the prepositional phrase because to is considered a special preposition., it made sense to me if I added it to PP instead of NP.

**VP -> VBD VBN PP**

**NP -> JJS NP**

**PP -> TO NP**

**Graphical user interface, text, application

Description automatically generated**

1. *sixty strangers are at the Round\_Table .*

Round\_Table is a proper noun, strangers is a plural noun. To handle these two rules were added to the noun phrase. A verb present, pluralVBP followed by a prepositional phrase is added to the verb phrase to handle the word ‘are’

**NP -> CD NP | NNS**

**VP -> VBP PP**

**A picture containing text

Description automatically generated**

1. *Sir\_Lancelot might have spoken .*

‘might’ is a modal followed by a verb phrase, this rule has to be added to the cfg for verb phrase. 2 rules were added

**VP -> MD VP | VB VBN**

Text

Description automatically generated with medium confidence

1. *Guinevere had been riding with Patsy for five weary nights .*

had here is a past tense verb, been is a past participle and riding is a gerund. All these 3 combined with a preposition phrase were not handled before. So 2 new rules, one for noun phrase with proper noun followed by prepositional phrase and a verb phrase were added

**NP -> Proper PP**

**VP -> VBD VBN VBG PP**

Graphical user interface, text, application, email

Description automatically generated

1. *Sir\_Bedevere might have been suggesting this quest .*

I new rule to the verb phrase was added which takes care of Verbs in third person followed by past tense verbs and gerunds. Here ‘have been suggesting this quest’ is parsed as a verb phrase

**VP -> VBP VBN VBG NP**

Text

Description automatically generated

1. *the Britons migrate south frequently .*

Plural form of proper noun is taken as noun phrase(brightons) and adverbs south and frequently preceded by a verb in third person is taken as a verb phrase

**NP -> NNPS**

**VP -> VBP RB RB**

Text

Description automatically generated

1. *Arthur and Guinevere ride frequently near the castle .*

Two new rules were added. ‘and’ is a coordinating conjunction joining two nouns. To handle ‘ride frequently’ followed by a prepositional phrase, I added a new rule to VP

**NP -> Proper NP| CC NP**

**VP -> VBP RB PP**

**Graphical user interface, text

Description automatically generated with medium confidence**

1. *he suggests to grow fruit at home .*

Two rules were added, to handle pronouns and to get the structure which has ‘to’ in the middle.

**NP -> PRP**

**VP -> VBZ TO VBP NP**

*Text

Description automatically generated with medium confidence*

1. *riding to Camelot is not hard .*

3 new rules were added to the cfg. verb followed by not and adjective and gerund followed by to and noun which is prepositional phrase. Both of these are verb phrases, the overall sentence structure doesn’t handle this, so we add a new rule with twi verb phrases before EOS

**S -> VP VP Eos**

**VP -> VBG PP | VerbT NOT JJ**

A picture containing text

Description automatically generated

1. *do coconuts speak ?*

Do followed by noun was added to noun phrase and verb in third person were added to verb phrase, ‘coconuts’ is a plural noun which is handled as a noun phrase.

**NP -> DO NP**

**VP -> VBZ**

Text

Description automatically generated with medium confidence

1. *why does England have a king ?*

2 new rules were added, which are Verb in 3rd person followed by a Noun phrase and to handle new sentences starting with WRB

**S -> WRB NP VP Eos**

**VP -> VBP NP**

**Text, letter

Description automatically generated**

1. **2 sentences from Challengsentences.txt**
2. *neither Sir\_Lancelot nor Guinevere will speak of it .*

The existing rules already parsed this sentence.

Text

Description automatically generated

1. *Arthur knows Patsy , the trusty servant .*

Two new rules were added to the CFG to parse this,

**NP -> Ps NP**

**VP -> VBZ NP**

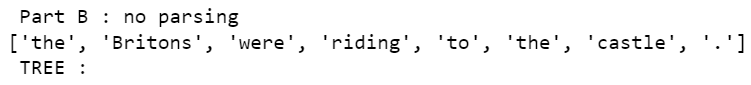
Graphical user interface, text, application, email

Description automatically generated

**Part 2 :**

1. **Made up sentence which doesn’t get parsed**

*‘the Britons were riding to the castle .’*

****

1. **Random sentence that gets parsed**

*‘5,000 and unable castle to coconuts are !’*

**Text

Description automatically generated**

**Takeaways :**

I learnt how to properly write CFG rules. Initially I ended up writing too many rules to parse the sentences, which led to some sentences having additional tree structures as the rules were getting added. This meant that the sentence was being considered in more than one format. While it could be a good thing in few cases, but mostly as I see it, having one sentence be analyzed in 2 different forms creates ambiguity in a way when further analysis must be done looking at the tree structure or to understand the proper context. I am pasting the grammar that I had initially used. Language and parts of speech proper understanding is very important while writing CFGs.

**Initial grammar :**

**# The start symbol is START.**

**START -> S1**

**# This is a very, very, very simple grammar to get you started.**

**# Sentences**

**S1 -> NP VP Eos | NP VP NP Eos | VP VP Eos | WRB NP VP Eos | CC NP VP Eos**

**# Verb phrases**

**VP -> VerbT NP | VerbT NP PP | VBD VBD PP | MD VB VBN VBG | MD VB VBN | MD VB PP |VBP RB PP | VBP PP | VBN VBN VBG PP | VB RB RB | VBZ TO VB NP | VerbT NOT JJ | VBG PP | VB | VB NP | VBZ NP**

**# Noun phrases**

**NP -> Det NP | Proper | Noun PP | Noun | Det JJ NP | Det JJS NP | Proper PP | Proper CC Proper | Num NNS | JJ NP | DO NP | NNP | NNS | NNPS | Num | Proper | PRP | Proper Pau NP**

**# do not use Nominals in RD parser**

**# Nom -> Noun | Noun Nom**

**# Prepositional phrase**

**PP -> Prep NP | TO NP**

**# This is a hand-generated set of lexical rules.**

**# Miscellaneous items that don't fit into any of our simple tags are**

**# given to Misc. You will want to develop your own tags for these!**

**# Singular and mass nouns. Notice that only the**

**# ones that start with consonants get our Noun tag.**

**Noun -> 'castle' | 'king' | 'defeater' | 'sovereign' | 'servant' | 'corner' | 'land' | 'quest' | 'chalice' | 'master' | 'horse' | 'fruit' | 'swallow' | 'sun' | 'winter' | 'coconut' | 'pound' | 'husk' | 'home' | 'weight' | 'story'**

**# Determiners**

**Det -> 'a' | 'another' | 'any' | 'each' | 'every' | 'no' | 'that' | 'the' | 'this'**

**# Prepositions**

**Prep -> 'above' | 'across' | 'at' | 'below' | 'by' | 'for' | 'from' | 'into' | 'near' | 'of' | 'on' | 'over' | 'through' | 'with'**

**# Proper nouns, this tag is just for people**

**Proper -> 'Arthur' | 'Guinevere' | 'Sir\_Lancelot' | 'Sir\_Bedevere' | 'Zoot' | 'Dingo' | 'Patsy'**

**# Verbs (third person singular present).**

**# Note that only transitive verbs get our VerbT tag,**

**# and some of those could also function as intransitives!**

**VerbT -> 'has' | 'covers' | 'is' | 'drinks' | 'carries' | 'rides'**

**# End-of-sentence.**

**Eos -> '!' | '.' | '?'**

**# Pauses.**

**Pau -> ',' | '...' | '--' | ';' | ':'**

**# Coordinating conjunctions.**

**CC -> 'and' | 'but' | 'or' | 'either' | 'nor' | 'neither' | 'so'**

**# Numbers.**

**Num -> 'eight' | 'five' | 'one' | '5.5' | 'sixty' | '5,000'**

**# Expletive.**

**Misc4 -> 'there'**

**# Subordinating conjunctions.**

**IN -> 'that' | 'so' | 'while' | 'because' | 'if'**

**# Modals.**

**MD -> 'can' | 'could' | 'may' | 'might' | 'must' | 'ought' | 'shall' | 'should' | 'will' | 'would'**

**# Adjectives**

**JJ -> 'plodding' | 'bloody' | 'weary' | 'unable' | 'trusty' | 'further' | 'sacred' | 'hot' | 'lucky' | 'simple' | 'tiny' | 'hard' | 'sensational' | 'comparable' | 'yellow'**

**# Comparative adjectives.**

**JJR -> 'bloodier' | 'wearier' | 'trustier' | 'hotter' | 'simpler' | 'tinier' | 'harder'**

**# Superlative adjectives.**

**JJS -> 'bloodiest' | 'weariest' | 'trustiest' | 'hottest' | 'simplest' | 'tiniest' | 'hardest'**

**# Plural nouns.**

**NNS -> 'coconuts' | 'halves' | 'snows' | 'mountains' | 'areas' | 'strangers' | 'inches' | 'ants' | 'nights'**

**# More proper nouns, not people.**

**NNP -> 'Camelot' | 'England' | 'Holy\_Grail' | 'Round\_Table'**

**# Plural proper nouns.**

**NNPS -> 'Britons' | 'Saxons'**

**# Personal pronouns.**

**PRP -> 'he' | 'her' | 'him' | 'it' | 'one' | 'she' | 'them' | 'they'**

**# Possessive personal pronouns**

**PPS -> 'her' | 'his' | 'its' | 'their'**

**# Adverbs.**

**RB -> 'again' | 'already' | 'currently' | 'frequently' | 'precisely' | 'south' | 'successfully' | 'unfortunately'**

**# 'do'/'does'**

**DO -> 'do' | 'does'**

**# 'to'**

**TO -> 'to'**

**# 'not'**

**NOT -> 'not'**

**# Verbs (base form).**

**VB -> 'have' | 'speak' | 'cover' | 'be' | 'ride' | 'drink' | 'grow' | 'carry' | 'suggest' | 'migrate' | 'know'**

**# Verbs (past tense).**

**VBD -> 'had' | 'spoke' | 'covered' | 'was' | 'were' | 'rode' | 'drank' | 'grew' | 'carried' | 'suggested' | 'migrated' | 'knew'**

**# Verbs (present participles).**

**VBG -> 'having' | 'speaking' | 'covering' | 'being' | 'riding' | 'drinking' | 'growing' | 'carrying' | 'suggesting' | 'migrating' | 'knowing'**

**# Verbs (past participle).**

**VBN -> 'had' | 'spoken' | 'covered' | 'been' | 'ridden' | 'drunk' | 'grown' | 'carried' | 'suggested' | 'migrated' | 'known'**

**# More third person singular verbs.**

**VBZ -> 'speaks' | 'grows' | 'goes' | 'migrates' | 'suggests' | 'knows'**

**# Verbs (present, plural, third person).**

**VBP -> 'have' | 'speak' | 'cover' | 'are' | 'ride' | 'drink' | 'grow' | 'carry' | 'suggest' | 'migrate' | 'know'**

**# Wh-determiners.**

**WDT -> 'that' | 'what' | 'which'**

**# Wh-pronouns.**

**WP -> 'what' | 'who'**

**# Wh possessive pronoun.**

**WPS -> 'whose'**

**# Wh-adverbs**

**WRB -> 'how' | 'when' | 'where' | 'why'**

**# Possessive marker.**

**POSS -> "'s"**