John Fields
Dr. Norma Palomino
IST664 - Homework #3
Writing a Context Free Grammar for Camelot Sentences
11/20/19

Figure 1: Camelot
Source: Wikipedia

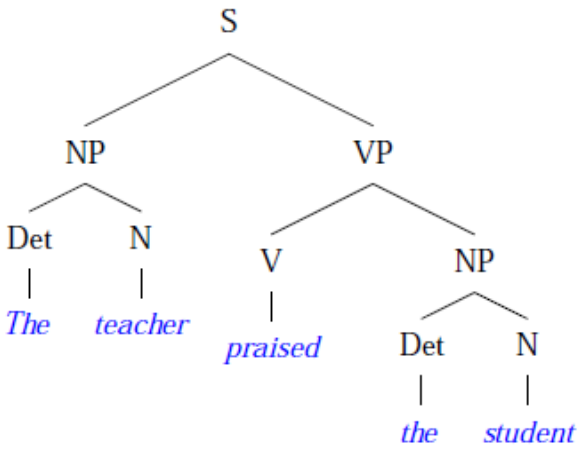

Figure 2: Example Parse Tree
Source: zeromin0.blogspot.com

Introduction

In *Speech and Language Processing* (Jurafsky and Martin, 2018), the authors describes Context Free Grammars as, "The most widely used formal system for modeling constituent structure in English and other natural languages is the Context-Free Grammar, or CFG."[i] For this assignment, a CFG will be developed for English language sentences related to the middle ages story of Camelot. The goal for this assignment is to create the grammar rules and associated vocabulary to parse sentences and identify the structure as shown in the example in Figure 2. The parsing will be completed using the Python NLTK software package for the following text from Camelot:

- 16 standard sentences
- 2 "challenge" sentences
- 2 exemplar sentences

Analysis

About the Data

- For this assignment, the following files were provided:
  - camelot_grammar.cfg
  - sentences.txt
  - challenge.sentences.txt
  - development.py

Methodology

- To find the correct context free grammars in the camelot sentences by developing a grammar.
- The following format will be used throughout this paper to show each new sentence that is evaluated and the results:
  - Sentence #
  - Grammar Rules (additions in bold)
  - Parse Tree
  - Vocabulary Rule Added
  - Description

Results

Sentence 1	Arthur is the king .
Grammar Rules (additions in bold)	# Sentences S1 -> NP VP Eos # Verb phrases VP -> VerbT NP | VerbT NP PP # Noun phrases NP -> Det NP | Proper | Noun PP | Noun # Prepositional phrase PP -> Prep NP
Parse Tree	(S1

	(NP (Proper Arthur)) (VP (VerbT is) (NP (Det the) (NP (Noun king)))) (Eos .)))
Vocabulary Rule Added	None. This was the default rule provided for the assignement.
Description	The parse tree utilizes the S1 sentence structure which already includes the NP VP structure to parse this sentence. The NP Proper matches "Arthur" and VerbT NP matches "is" and Det for "the king" and NP Noun for "king".

Sentence 2	Arthur rides the horse near the castle .
Grammar Rules (additions in bold)	# Sentences S1 -> NP VP Eos # Verb phrases VP -> VerbT NP | VerbT NP PP # Noun phrases NP -> Det NP | Proper | Noun PP | Noun # Prepositional phrase PP -> Prep NP
Parse Tree	(S1 (NP (Proper Arthur)) (VP (VerbT rides) (NP (Det the) (NP (Noun horse) (PP (Prep near) (NP (Det the) (NP (Noun castle))))))) (Eos .)))
Vocabulary Rule Added	None. This was the default rule provided for the assignement.

Description	The NP Proper matches "Arthur" and VP VerbT NP PP matches "ride" for VerbT, NP Det NP for "the horse" and PP "near" and NP Det NP for "the castle".

Sentence 3	Arthur rides the plodding horse near the castle .
Grammar Rules (additions in bold)	# Sentences S1 -> NP VP Eos # Verb phrases VP -> VerbT NP | VerbT NP PP # Noun phrases NP -> Det NP | Proper | Noun PP | Noun | **Det JJ Noun** # Prepositional phrase PP -> Prep NP
Parse Tree	(S1 (NP (Proper Arthur)) (VP (VerbT rides) (NP (Det the) (JJ plodding) (Noun horse)) (PP (Prep near) (NP (Det the) (NP (Noun castle))))) (Eos .)))
Vocabulary Rule Added	# Adjectives JJ -> 'plodding' | 'bloody' | 'weary' | 'unable' | 'trusty' | 'further' | 'sacred' | 'hot' | 'lucky' | 'simple' | 'tiny' | 'hard' | 'sensational' | 'comparable' | 'yellow'
Description	The NP Proper matches "Arthur". The VerbT NP PP uses VerbT for "rides", NP Det JJ Noun for "the plodding horse" and PP Prep NP for "near" with NP Det NP for "the castle".

Sentence 4	the Holy_Grail is a chalice .
Grammar Rules (additions in bold)	# Noun phrases

	NP -> Det NP | Proper | Noun PP | Noun | Det JJ Noun | **Det NNP** # Verb phrases VP -> VerbT NP | VerbT NP PP | **VerbT Det NN**
Parse Tree	(S1 (NP (Det the) (NNP Holy_Grail)) (VP (VerbT is) (NP (Det a) (NP (Noun chalice)))) (Eos .)))
Vocabulary Rule Added	# More proper nouns, not people. NNP -> 'Camelot' | 'England' | 'Holy_Grail' | 'Round_Table'
Description	The NP rule Det NNP finds "the Holy_Grail". VP VerbT Det NN finds "is a chalice".

Sentence 5	the sensational Holy_Grail is a sacred chalice .
Grammar Rules (additions in bold)	# Noun phrases NP -> Det NP | Proper | Noun PP | Noun | Det JJ Noun | Det NNP | **Det JJ NNP** # Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN
Parse Tree	(S1 (NP (Det the) (JJ sensational) (NNP Holy_Grail)) (VP (VerbT is) (NP (Det a) (JJ sacred) (Noun chalice))) (Eos .)))
Vocabulary Rule Added	None
Description	Det JJ NNP was added to the NP to find "the sensational Holy Grail". The VP VerbT NP finds "is" and the NP Det JJ Noun finds "a sacred chalice".

Sentence 6	every coconut was carried to the hottest mountains .
Grammar Rules (additions in bold)	# Noun phrases NP -> ~~Det NP~~ | Proper | Noun PP | Noun | Det JJ Noun | Det NNP | Det JJ NNP | **Det Noun | Det JJS NNS** # Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN | **VBD VP | VBD PP** # Prepositional phrase PP -> Prep NP | **TO NP**
Parse Tree	(S1 (NP (Det every) (NP (Noun coconut))) (VP (VBD was) (VP (VBD carried) (PP (TO to) (NP (Det the) (JJS hottest) (NNS mountains))))) (Eos .)))
Vocabulary Rule Added	# Verbs (past tense). VBD -> 'had' | 'spoke' | 'covered' | 'was' | 'were' | 'rode' | 'drank' | 'grew' | 'carried' | 'suggested' | 'migrated' | 'knew' # Plural nouns. NNS -> 'coconuts' | 'halves' | 'snows' | 'mountains' | 'areas' | 'strangers' | 'inches' | 'ants' | 'nights' # Superlative adjectives. JJS -> 'bloodiest' | 'weariest' | 'trustiest' | 'hottest' | 'simplest' | 'tiniest' | 'hardest'
Description	Det Noun was added as a NP for "every coconut' and Det NP was removed as it was causing duplicates with the new Det Noun phrase. VBD VP was added as a VP for "was" and VBD PP was added for

	"carried to" and TO NP was added for PP. Finally Det JJS NNS was added to NP for "the hottest mountains".

Sentence 7	sixty strangers are at the Round_Table .
Grammar Rules (additions in bold)	# Noun phrases NP -> Det NP | Proper | Noun PP | Noun | Det JJ Noun | Det NNP | Det JJ NNP | Det Noun | Det JJS NNS | **Num NNS** # Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN | VBD VP | VBD PP | **VBP Prep NP**
Parse Tree	(S1 (NP (Num sixty) (NNS strangers)) (VP (VBP are) (Prep at) (NP (Det the) (NNP Round_Table))) (Eos .)))
Vocabulary Rule Added	# Verbs (present, plural, third person). VBP -> 'have' | 'speak' | 'cover' | 'are' | 'ride' | 'drink' | 'grow' | 'carry' | 'suggest' | 'migrate' | 'know'
Description	Num NNS was added as NP for "sixty strangers". VBP Prep NP was added as VP for "are at the Round_Table".

Sentence 8	Sir_Lancelot might have spoken .
Grammar Rules (additions in bold)	# Noun phrases NP -> Proper | Noun PP | Noun | Det NNP | Det JJ NNP | Det Noun | Det JJS NNS | Num NNS | Det JJ Noun | **NNP** # Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN | VBD VP | VBD PP | VBP Prep NP | **MD VP | VB VBN**
Parse Tree	(S1 (NP (Proper Sir_Lancelot))

	(VP (MD might) (VP (VB have) (VBN spoken))) (Eos .)))
Vocabulary Rule Added	# Modals. MD -> 'can' | 'could' | 'may' | 'might' | 'must' | 'ought' | 'shall' | 'should' | 'will' | 'would' # Verbs (base form). VB -> 'have' | 'speak' | 'cover' | 'be' | 'ride' | 'drink' | 'grow' | 'carry' | 'suggest' | 'migrate' | 'know' # Verbs (past participle). VBN -> 'had' | 'spoken' | 'covered' | 'been' | 'ridden' | 'drunk' | 'grown' | 'carried' | 'suggested' | 'migrated' | 'known'
Description	NNP was added to NP for "Sir_Lancelot". MD VP was added to VP for "might" and VB VBN for "have spoken".

Sentence 9	Guinevere had been riding with Patsy for five weary nights .
Grammar Rules (additions in bold)	# Noun phrases NP -> Proper | Noun PP | Noun | Det NNP | Det JJ NNP | Det Noun | Det JJS NNS | Num NNS | Det JJ Noun | NNP | **NNP PP | Proper PP | Num JJ NNS** #Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN | VBD VP | VBD PP | VBP Prep NP | MD VP | VB VBN | **VBN VBG PP** # Prepositional phrase PP -> Prep NP | TO NP
Parse Tree	(S1 (NP (Proper Guinevere)) (VP (VBD had) (VP (VBN been) (VBG riding) (PP

	(Prep with) (NP (Proper Patsy) (PP (Prep for) (NP (Num five) (JJ weary) (NNS nights))))))) (Eos .)))
Vocabulary Rule Added	# Verbs (present participles). VBG -> 'having' | 'speaking' | 'covering' | 'being' | 'riding' | 'drinking' | 'growing' | 'carrying' | 'suggesting' | 'migrating' | 'knowing'
Description	"Guinevere" used the NP proper. "had" used the VBD in VBD VP and then the VP VBN VBG PP was added for "been riding" and the PP used the existing Prep NP for "with" and then Proper PP was added for "Patsy", then the PP Prep NP was used "for" then the NP Num JJ NNS was added for "five weary nights".

Sentence 10	Sir_Bedevere might have been suggesting this quest .
Grammar Rules (additions in bold)	# Noun phrases NP -> Proper | Noun PP | Noun | Det NNP | Det JJ NNP | Det Noun | Det JJS NNS | Num NNS | Det JJ Noun | NNP | NNP PP | Proper PP | Num JJ NNS # Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN | VBD VP | VBD PP | VBP Prep NP | MD VP | VB VBN | VBN VBG PP | **VB VBN VP | VBG Det NP**
Parse Tree	(S1 (NP (Proper Sir_Bedevere)) (VP (MD might) (VP (VB have) (VBN been)

	(VP (VBG suggesting) (Det this) (NP (Noun quest))))) (Eos .)))
Vocabulary Rule Added	None added.
Description	The NP Proper was used for "Sir_Bedevere". Then the VP MD VP was used for the modal "might" and the VP VB VBN VP was added for "have been". Then the NP VBG Det NP was added for "suggesting this quest".

Sentence 11	the Britons migrate south frequently .
Grammar Rules (additions in bold)	# Noun phrases NP -> Proper | Noun PP | Noun | Det NNP | Det JJ NNP | Det Noun | Det JJS NNS | Num NNS | Det JJ Noun | NNP | NNP PP | Proper PP | Num JJ NNS | **Det NNPS** # Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN | VBD VP | VBD PP | VBP Prep NP | MD VP | VB VBN | VBN VBG PP | VB VBN VP | VBG Det NP | **VB RB RB**
Parse Tree	(S1 (NP (Det the) (NNPS Britons)) (VP (VB migrate) (RB south) (RB frequently)) (Eos .)))
Vocabulary Rule Added	# Plural proper nouns. NNPS -> 'Britons' | 'Saxons' # Adverbs. RB -> 'again' | 'already' | 'currently' | 'frequently' | 'precisely' | 'south' | 'successfully' | 'unfortunately'
Description	The NNPS vocabulary was added for the proper noun "Britons" and the RB vocabulary was added for the adverb "south". Next, the NP Det NNPS was added for "the Britons". The VP VB RB RB was added for "migrate south frequently".

Sentence 12	Arthur and Guinevere ride frequently near the castle .
Grammar Rules (additions in bold)	# Noun phrases NP -> Proper | Noun PP | Noun | Det NNP | Det JJ NNP | Det Noun | Det JJS NNS | Num NNS | Det JJ Noun | NNP | NNP PP | Proper PP | Num JJ NNS | Det NNPS | **Proper CC Proper** # Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN | VBD VP | VBD PP | VBP Prep NP | MD VP | VB VBN | VBN VBG PP | VB VBN VP | VBG Det NP | VB RB RB | **VB RB PP**
Parse Tree	(S1 (NP (Proper Arthur) (CC and) (Proper Guinevere)) (VP (VB ride) (RB frequently) (PP (Prep near) (NP (Det the) (Noun castle)))) (Eos .)))
Vocabulary Rule Added	# Coordinating conjunctions. CC -> 'and' | 'but' | 'or' | 'either' | 'nor' | 'neither' | 'so'
Description	The coordinating conjunction CC was added to the vocabulary for "and". Next, the NP Proper CC Proper was added for "Arthur and Guinevere". The VP VB RB PP was added for "ride frequently" and the PP Prep NP was used for "near castle".

Sentence 13	he suggests to grow fruit at home .
Grammar Rules (additions in bold)	# Noun phrases NP -> Proper | Noun PP | Noun | Det NNP | Det JJ NNP | Det Noun | Det JJS NNS | Num NNS | Det JJ Noun | NNP | NNP PP | Proper PP | Num JJ NNS | Det NNPS | Proper CC Proper | **PRP** # Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN | VBD VP | VBD PP | VBP Prep NP | MD VP | VB VBN | VBN VBG PP | VB VBN VP | VBG Det NP | VB RB RB | VB RB PP | **VBZ PP** | **VB NP**
Parse Tree	(S1

	(NP (PRP he)) (VP (VBZ suggests) (PP (TO to) (VP (VB grow) (NP (Noun fruit) (PP (Prep at) (NP (Noun home)))))))) (Eos .)))
Vocabulary Rule Added	# Personal pronouns. PRP -> 'he' | 'her' | 'him' | 'it' | 'one' | 'she' | 'them' | 'they' # More third person singular verbs. VBZ -> 'speaks' | 'grows' | 'goes' | 'migrates' | 'suggests' | 'knows'
Description	Vocabulary Rule PRP was added for "he" and VBZ for "suggests". The NP PRP was added for "he". The VP VBZ PP was added for "suggests" and the PP TO VP was added for "to". The VP VB NP was used for "grow" and the NP Noun PP for "fruit" and "at home" uses the PP Prep NP.

Sentence 14	riding to Camelot is not hard .
Grammar Rules (additions in bold)	# Sentences S1 -> NP VP Eos | **VP VP Eos** # Noun phrases NP -> Proper | Noun PP | Noun | Det NNP | Det JJ NNP | Det Noun | Det JJS NNS | Num NNS | Det JJ Noun | NNP | NNP PP | Proper PP | Num JJ NNS | Det NNPS | Proper CC Proper | PRP | Noun PP # Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN | VBD VP | VBD PP | VBP Prep NP | MD VP | VB VBN | VBN VBG PP | VB VBN VP | VBG Det NP |

	VB RB RB | VB RB PP | VBZ PP | VB NP | **VBG PP | VerbT NOT JJ**
Parse Tree	(S1 (VP (VBG riding) (PP (TO to) (NP (NNP Camelot)))) (VP (VerbT is) (NOT not) (JJ hard)) (Eos .)))
Vocabulary Rule Added	None added.
Description	For the S1 Sentence rule, VP VP Eos was added to allow for two verb phrases in a sentence. Next, the VP VBG PP was added for "riding" and TO NP in PP finds "to" and NNP in NP for "Camelot". Next, the VerbT NOT JJ added to VP finds "is not hard".

Sentence 15	do coconuts speak ?
Grammar Rules (additions in bold)	START -> S1 | **SQ** # Sentences S1 -> NP VP Eos | VP VP Eos | **SQ Eos** **# Clauses and Phrases** **SQ -> DO NNS VB**
Parse Tree	(S1 (SQ (DO do) (NNS coconuts) (VB speak)) (Eos ?)))
Vocabulary Rule Added	None added.
Description	Since this sentence is a question in an inverted yes/no format, the SQ type was added to the S1 Sentences. Next, an SQ section was added with the rule SQ --> DO NNS VB. This finds the sentence "do coconuts speak".

Sentence 16	why does England have a king ?
Grammar Rules (additions in bold)	START -> S1 | SQ | **SBARQ** # Sentences

	S1 -> NP VP Eos | VP VP Eos | SQ Eos |**SBARQ Eos** # Clauses and Phrases SQ -> DO NNS VB | **DO NP VP** **SBARQ -> WRB SQ** # Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN | VBD VP | VBD PP | VBP Prep NP | MD VP | VB VBN | VBN VBG PP | VB VBN VP | VBG Det NP | VB RB RB | VB RB PP | VBZ PP | VB NP | VBG PP | VerbT NOT JJ
Parse Tree	(S1 (SBARQ (WRB why) (SQ (DO does) (NP (NNP England)) (VP (VB have) (NP (Det a) (Noun king))))) (Eos ?)))
Vocabulary Rule Added	# Wh-adverbs WRB -> 'how' | 'when' | 'where' | 'why'
Description	Since the sentence is a direct question introduced by a wh-word, the SBARQ identifier was created with the rule SBARQ -> WRB SQ. The WRB finds the "why" and the SQ rule DO NP VP was added to find "does" and the NP uses the existing NNP for England and Det Noun to find "a king".

Challenge Sentence 1	what horse does Arthur ride ?
Grammar Rules (additions in bold)	# Clauses and Phrases SQ -> DO NNS VB | DO NP VP | **DO NP VB** **WHNP -> WDT NP SQ** SBARQ -> WRB SQ | WHNP
Parse Tree	(S1 (SBARQ (WHNP (WDT what) (NP (Noun horse)) (SQ (DO does) (NP (Proper Arthur)) (VB ride)))) (Eos ?)))
Vocabulary Rule Added	# Wh-determiners. WDT -> 'that' | 'what' | 'which'
Description	The SBARQ clause that was created for Sentence 16 can be utilized for this sentence with the addition of the WHNP for a Wh-noun phrase. This parses by using the WHNP with WDT NP SQ for "what horse", then DO NP VB was added to SQ for "does Arthur ride".

Challenge Sentence 2	who does Arthur suggest she carry ?
Grammar Rules (additions in bold)	# Clauses and Phrases SQ -> DO NNS VB | DO NP VP | DO NP VB WHNP -> WDT NP SQ | **WP SQ** SBARQ -> WRB SQ | WHNP # Verb phrases VP -> VerbT NP | VerbT NP PP | VerbT Det NN | VBD VP | VBD PP | VBP Prep NP | MD VP | VB VBN | VBN VBG PP | VB VBN VP | VBG Det NP | VB RB RB | VB RB PP | VBZ PP | VB NP | VBG PP | VerbT NOT JJ | **VB PRP VB**
Parse Tree	(S1 (SBARQ (WHNP (WP who)

	(SQ (DO does) (NP (Proper Arthur)) (VP (VB suggest) (PRP she) (VB carry))))) (Eos ?)))
Vocabulary Rule Added	# Wh-pronouns. WP -> 'what' | 'who'
Description	The WP vocabulary was first added for "who". Next, grammar for WHNP WP SQ was added for "who" and the DO in SQ finds "does" then the NP finds "Arthur" with Proper. For the VP, VB PRP VB was added for "suggest she carry".

Exemplar Sentence 1	have Sir_Bedevere speak to Guinevere with the coconuts .
Grammar Rules (additions in bold)	None
Parse Tree	From the Stanford Parser: (S (VP (VB have) (S (NP (NNP Sir_Bedevere)) (VP (VB speak) (S (VP (TO to) (VP (VB Guinevere) (PP (IN with) (NP (DT the) (NNS coconuts)))))))))))
Vocabulary Rule Added	None

Description	The camelot_grammar.cfg file was unable to parse Exemplar Sentence 1. This sentence consists of three sentence segments VP NP VP as parsed by the Stanford Parser (see example above). The existing grammar will not work without changing the S1 rules and adding more VP and NP rules.

Exemplar Sentence 2	Patsy drinks the yellow sovereign into the defeater .
Grammar Rules (additions in bold)	None
Parse Tree	(S1 (NP (Proper Patsy)) (VP (VerbT drinks) (NP (Det the) (JJ yellow) (Noun sovereign)) (PP (Prep into) (NP (Det the) (Noun defeater)))) (Eos .)))
Vocabulary Rule Added	None
Description	Exemplar Sentence 2 uses the same structure as Sentence 3 - "Arthur rides the plodding horse near the castle". However, this sentence doesn't make sense because the noun sovereign can't be yellow or consumed as a beverage. The noun "defeater" also doesn't fit at the end of the sentences. To improve the grammar rules, some of the nouns may need to be categorized differently to limit the overgeneralization that occurs in sentences such as Exemplar Sentence 2.

Conclusion

The use of Context Free Grammars is an important tool in Natural Language Processing to develop a grammar that is "tailored" to a specific text or robust enough to work on a variety of texts (eg Penn Tree Bank, Stanford Parser). Modeling CFG's has improved over the last 60 years with the advent of computers and additional research, however, the complex nature of language will ensure that new techniques and improvements will continue to be developed in the future.[ii]

---

[i] Jurafsky, Dan, and James H. Martin. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*., 2018.

[ii] Chen, T. H., Tseng, C. H., & Chen, C. P. (2006, September). Automatic learning of context-free grammar. In *Proceedings of the 18th Conference on Computational Linguistics and Speech Processing* (pp. 53-62).