# Introduction

Context free grammars can be used to break down sentences and provide a grammar structure or originally unstructured data. The purpose of this assignment is to create enough rules in the context free grammar to give structure to all of the sentences in the sentences.txt file and two sentences in the challenge sentence section. Every time a new sentence has been parsed by the context free grammar, the rules that have been added to the grammar will be explained and the parse tree will displayed.

After all the sentences have been parsed by the grammar, I will create two sentences that show the downside of context free grammars. The first sentence will be a real sentence with words understood by the parser, but the grammar is not understood by the grammar. The second sentence will be a string of words understood by the grammar, and parsed by the grammar, but the sentence does not make sense.

# Part 1 – Creating the CFG

## Trial 1 - No Changes

The program to tag the sentences with the CFG was run before I made any changes to the grammar. Three sentences were recognized by the grammar.

**Sentence 1:** Arthur is the king .

**Sentence 1 Explanation**: The first sentence initially breaks down into a noun phrase, a verb phrase, and an end of sentence tag. The noun phrase is able to be parsed because of the noun phrase rule that a noun phrase only has to consist of a proper noun. In this sentence, the proper noun is "Arthur". The verb phrase is able to be parsed because of the verb phrase rule that a verb phrase could be made of only a transitive verb followed by a noun phrase. The transitive verb in this sentence is "is" and the noun subphrase then deeds to be parsed. The noun phrase rule to parse the noun subphrase is that a noun phrase can be made with only a determiner and another noun phrase. The determiner in the noun subphrase is "the" and the repeated noun subphrase is parsed with the rule that a noun phrase could be made of only a noun.

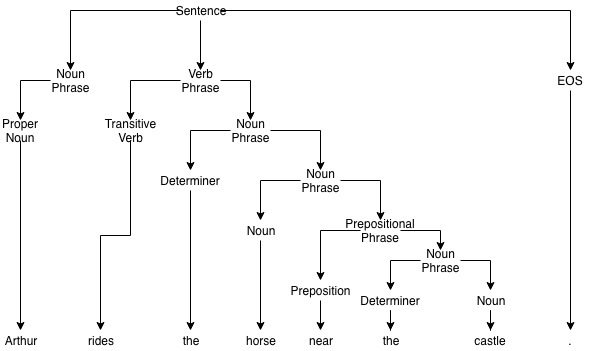
**Parse Tree 1:**

# 

**Sentence 2:** Arthur rides the horse near the castle .

**Sentence 2 Explanation:** The second sentence initially meets the rule that a sentence can consist of only a noun phrase, a verb phrase, and an end of sentence tag. The noun phrase meets the noun phrase rule that only a proper noun is required; in this case: "Arthur". The verb phrase breaks down with the rule that a transitive verb, "rides", and another noun subphrase can create a verb phrase. The noun phrase breaks down as noun subphrases until the noun phrase splits with the rule that a noun phrase can be made of a noun followed by a prepositional phrase, which begins at the word "near". This prepositional phrase then breaks down with the prepositional rules of a prepositional phrase can be a preposition followed by a noun phrase, which in this case is "the castle".

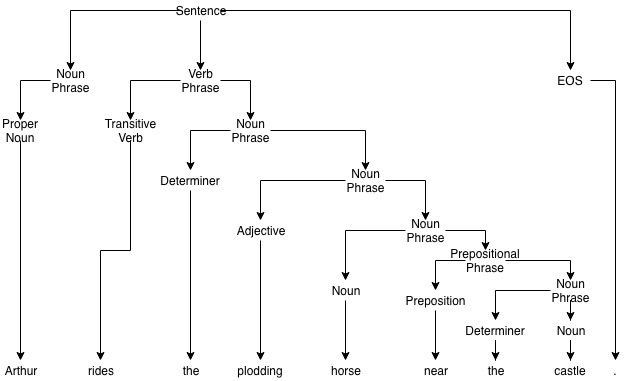
**Parse Tree 2:**

****

**Sentence 3:** Arthur rides the plodding horse near the castle .

**Sentence 3 Explanation:** Initially, the sentence meets the grammar rule that a sentence contains a noun phrase, verb phrase, and an end of sentence tag. The sentence also meets the rule that a sentence contains a noun phrase, verb phrase, prepositional phrase, and an end of sentence tag. This means that there are at least two interpretations of the sentence; either the prepositional phrase breaks down at the sentence level, or within the noun phrase (which is displayed

**Parse Tree 3 (V1 only):**

****

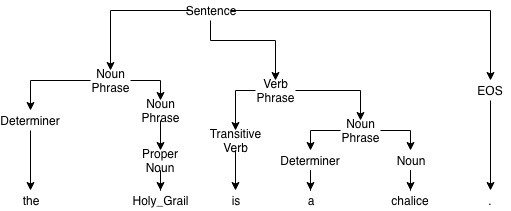
## Rule Changes 1

**Sentence 4:** the Holy\_Grail is a chalice .

**Word Classes Added**: More proper nouns were added to the grammar.

**Sentence 4 Explanation:** Allof the rules that were already in the grammar were enough to parse sentence 4, but the grammar did not know that the "Holy\_Grail" was a proper noun. Once "Holy\_Grail" was added to the list of proper nouns, the grammar was able to understand that "the Holy\_Grail" was a noun phrase that contained a determiner, and another noun phrase, which was the proper noun "Holy\_Grail". The rest of the sentence was parsed using rules that were already programmed into the grammar.

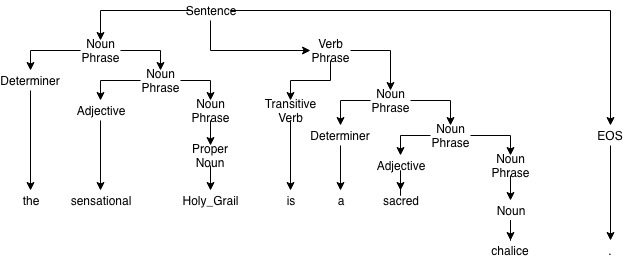
**Parse Tree 4:**

****

**Sentence 5:** the sensational Holy\_Grail is a sacred chalice .

**Sentence 5 Explanation:** The addition of "Holy\_Grail" to the proper noun list allowed the grammar to classify "sensational Holy\_Grail" as a noun phrase that contained an adjective and another noun subphrase. Before the "Holy\_Grail" was added to the proper nouns, the grammar did not know what the word "Holy\_Grail" was, and could not classify any of the part of the noun phrase part of the sentence.

**Parse Tree 5:**

****

## Rule Changes 2

**Sentence 6:** every coconut was carried to the hottest mountains .

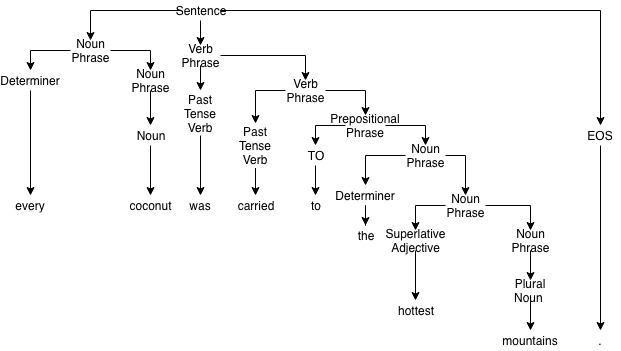
**Word Classes Added:** The first word class added to the grammar was plural nouns. The abbreviation created for plural nouns was NounP. The second word class added to the grammar was superlative adjectives. The class name for superlative adjectives in the grammar was "AdjS". Another word class created allowed for the classification of past tense verbs, such as "was" and "carried". The class name for past tense verbs was VerbD. The final change to the words classes was that the TO class was added to the list of prepositions. While there are many uses for the word "to", use as a preposition is fairly common.

**Rules Added:** The first new rule added to the noun phrases in the grammar allowed a noun phrase to consist of only a plural noun. This rule allowed the word "mountains" to be captured as a noun phrase. Another rule added to the noun phrases in the grammar allowed for a noun phrase to consist of a superlative adjective and another noun phrase. This rule allowed for the parsing of the phrase "hottest mountains" as a superlative adjective and a noun subphrase.

Two rules were added to the verb phrases in the grammar. The first rule change allowed a verb phrase to contain a past tense verb followed by a prepositional phrase. This rule allowed the grammar to capture "carried to the hottest mountains" as a verb phrase. The second rule added allowed a verb phrase to contain a past tense verb followed by another verb phrase. This rule allowed the grammar to catch understand how "was carried" is a verb phrase of "was" followed by another verb phrase.

These changes allowed the grammar to understand the word "to" can also be a preposition, and allowed the grammar catch the prepositional phrase "to the hottest mountains" after the verb "carried".

**Sentence 6 Parse Tree:**



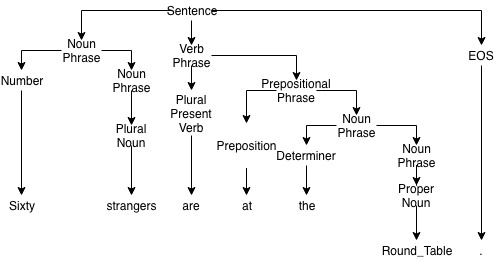
## Rule Changes 3

**Sentence 7:** sixty strangers are at the Round\_Table .

**Word Classes Added:** A word class of numerical tokens was added to the grammar under the class name of Num. Another word class was created for 3rd person, plural, present verbs. The class name for 3rd person, plural, present verbs was chosen as VerbPP.

**Rule Changes:** Two rules were added to the grammar, one to the noun phrases, and one to the verb phrases. The new rule added to the noun phrases allowed a noun phrase to consist of a numerical token followed by a noun phrase. This rule captured the phrase "sixty strangers" as a noun phrase consisting of the numerical token "sixty", followed by the noun subphrase of a plural noun, which is "strangers". The rule added to the verb phrases allowed a verb phrase to contain a plural present verb to be followed by a prepositional phrase. This new rule allowed the grammar to capture the phrase "are at the Round\_Table" as a verb phrase with the plural present verb of "are", followed by the prepositional phrase of "at the Round\_Table".

**Sentence 7 Parse Tree:**

****

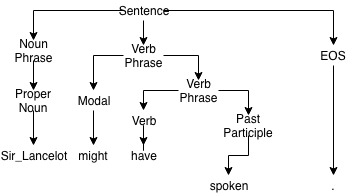
## Rule Changes 4

**Sentence 8:** Sir\_Lancelot might have spoken .

**Word Classes Added:** Three word classes were added to the grammar to capture certain words in the sentence. The first word class added to the grammar was modals. The class name for modals in the grammar is MD. The second class added to the grammar was base verbs, which had the class name of Verb. The final word class added to the grammar was past participles, which had a word class name of VerbPastP.

**Rule Changes:** Two rule changes were added to the verb phrases of the grammar. The first rule added to the verb phrases allowed a verb phrase to contain only a base verb followed by a past participle verb. This rule added to the grammar allowed the grammar to form the verb phrase "have spoken" and to terminate at the phrase "have spoken". The second rule added to the verb phrases allowed a verb phrase to be made up of a modal followed by another verb phrase. Modals are a special kind of verb that indicates modality. This rule allowed the grammar to understand the verb phrase "might have spoken" by first understanding "might" is a modal, and the verb subphrase "have spoken” follows.

**Sentence 8 Parse Tree:**



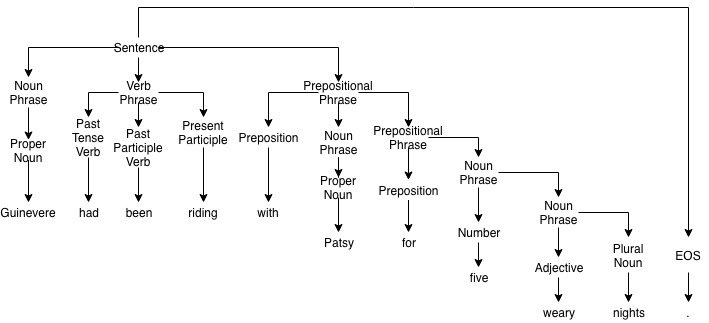
## Rule Changes 5

**Sentence 9:** Guinevere had been riding with Patsy for five weary nights .

**Word Classes Added:** One new word class was added to the grammar to capture present participle verbs such as "having" and "riding". The class name for present participle verbs was "VerbPresP".

**Rule Changes:** Two rules were added to the grammar to allow the grammar to parse the sentence. One rule was added to the verb phrases, and another rule was added to the prepositional phrases. The rule added to the verb phrases allowed for a verb phrase to consist of a past tense verb followed by a past participle and a present participle. This rule was able to capture and terminate on the verb phrase "had been riding". The rule added to the prepositional phrases rules allowed a prepositional phrase to consist of a preposition and a noun phrase, but have another prepositional phrase. This rule allows the grammar to capture prepositional phrases that have more than one prepositional phrase. The new prepositional phrase rule created allowed the grammar to break the prepositional "with Patsy for five weary nights" into two the separate prepositional phrases of "with Patsy" and "for five weary nights". The first branching of noun phrase, verb phrase, and prepositional phrase on the sentence level was also critical in parsing this sentence as the verb phrase is terminated as a verb phrase.

**Sentence 9 Parse Tree:**



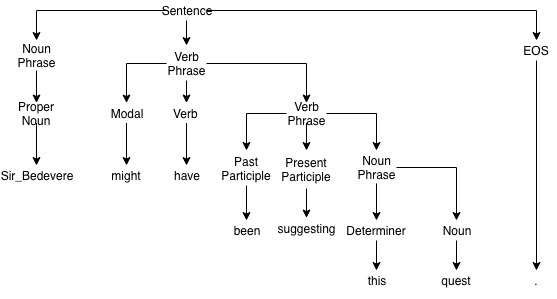
## Rule Change 6

**Sentence 10:** Sir\_Bedevere might have been suggesting this quest .

**Word Classes Added:** No new word classes were added to the grammar for this sentence parsing.

**Rule Changes:** Two rules were added to the grammar to parse sentence 10, both rules were added to the verb phrases. The first rule added to the grammar was that a verb phrase can be a past participle followed by a present participle and a noun phrase. This rule allowed the grammar to break "been suggesting this quest" into the past participle "been", the present participle "suggesting", and the noun phrase "this quest". The second rule added to the grammar allowed a verb phrase to consist of a modal followed by a verb and another verb phrase. This rule allows the grammar to parse the entire sentence because it lets the grammar recognize that "might have" is a verb phrase consisting of a modal and a verb, followed by the verb phrase "been suggesting this quest".

**Sentence 10 Parse Tree:**



## Rule Change 7

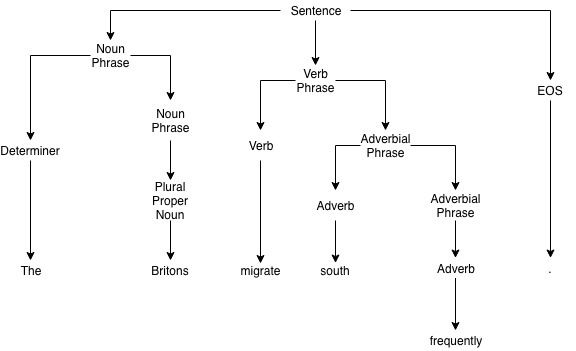
**Sentence 11:** the Britons migrate south frequently .

**Word Classes Added:** Two word classes were added to the grammar to parse this sentence. The first word class added to the grammar was plural proper nouns. The class name for plural proper nouns was "ProperP". The second word class added to the grammar was adverbs, and the word class name for adverbs was "Adv". In addition to the word class of adverbs, a new phrase class was made for adverbial phrases.

**Rule Changes:** The first rules added to the grammar were added to the new adverbial phrases class. The adverbial phrase rules stated that an adverbial phrase can consist of either just an adverb, or an adverb followed by an adverbial phrase. These rules essentially allowed an infinite amount of adverbs in an adverbial phrase. In this sentence, the grammar will first find the two adverbs "south" and "frequently" and realize that these are part of the same adverbial phrase. The first parse will capture the rule "south" + adverbial phrase, and then the adverbial subphrase will be the adverb "frequently". Additionally, plural proper nouns were added to the list of noun phrases, so the grammar would be able to parse "The Britons" as determiner "the" + noun subphrase "Britons".

A new rule was also added to the verb phrase rules. The rule stated that verb phrase can consist of verb and an adverbial subphrase. This rule allowed the grammar to capture the verb "migrate" and the adverbial subphrase "south + adverbial phrase" with a second adverbial subphrase "frequently".

**Sentence 11 Parse Tree:**



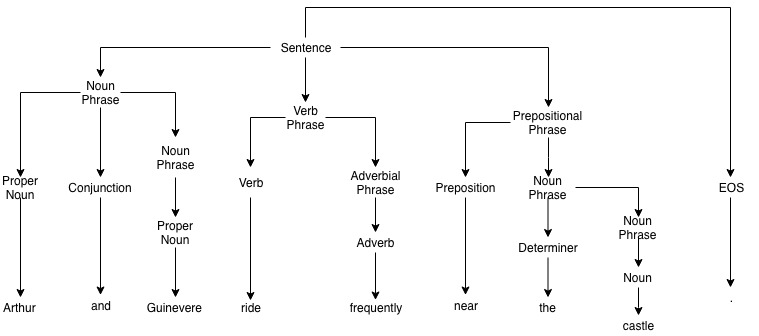
## Rule Change 8

**Sentence 12:** Arthur and Guinevere ride frequently near the castle .

**Word Classes Added:** One new word class was added to the grammar to capture conjunctions. The word class name for conjunctions was "Conj"

**Rule Changes:** One new rule was added to the noun phrases in the grammar to handle conjunctions in the noun phrase. The rule stated that a noun phrase could consist of a proper noun followed by a conjunction and another noun phrase. This rule allowed the grammar to parse the first noun phrase in the conjugation "Arthur and" and then capture the noun subphrase of "Guinevere". The verb phrase is captured by the grammar with the rule that a verb phrase can consist of a verb and an adverbial subphrase. Finally the prepositional phrase is captured at the sentence level with the rule that a sentence can consist of a noun phrase, a verb phrase, a prepositional phrase, and an end of sentence tag.

**Sentence 12 Parse Tree:**



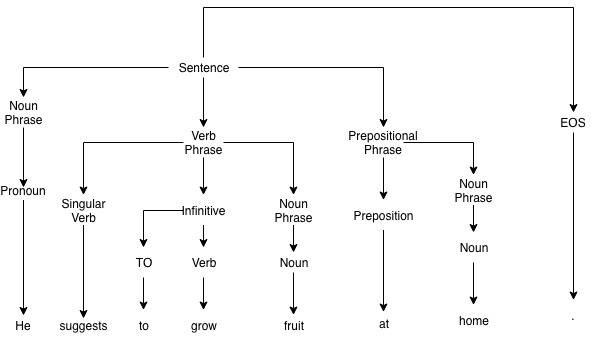
## Rule Change 9

**Sentence 13:** he suggests to grow fruit at home .

**Word Classes Added:** Two new word classes were added to the grammar to capture infinitives and pronouns. Infinitives were defined as a phrase that consisted of the word "to", followed by a verb. Infinitives were given a the word class name of "Inf". Pronouns were added to the grammar to capture pronouns such as "he", "she", and "them". The word class name given to pronouns was "Pro".

**Rule Changes:** Two rules were added to the grammar to parse the sentence. The first rule added stated that a noun phrase can consist of only a pronoun. This rule allowed the grammar to parse the noun phrase to consist of only the pronoun "he". The second rule added to the grammar stated that a verb phrase can consist of a third person singular verb, followed by an infinitive and a noun phrase. This rule allows the grammar to break down the verb phrase initially into "suggests INFINITIVE Noun Phrase". The infinitive subphrase breaks down into "to" and "grow", and the noun phrase is broken down to the noun "fruit". The prepositional phrase is captured by the grammar once again due to the sentence level rule that a sentence can be composed of a noun phrase, a verb phrase, and a prepositional phrase.

**Sentence 13 Parse Tree:**



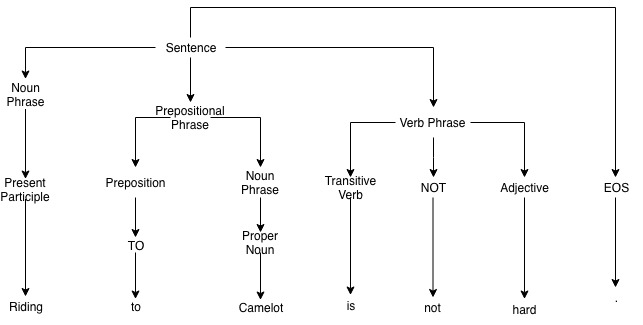
## Rule Change 10

**Word Classes Added:** None

**Sentence 14:** riding to Camelot is not hard .

**Rule Changes:** Three rules were added to the grammar to parse the sentence. The first rule states that a present participle, also known as a gerund, can be used as a noun phrase. These verbs are the only verbs that can be used as the subject in a noun phrase. This rule allows the gerund "riding" to be the subject of the sentence. The second rule added to the grammar allowed for a verb phrase to consist of a transitive verb, the negation "not", and an adjective. This rule is primarily used when a sentence contains a phrase similar to "is not", or "has not" followed by an adjective. This rule will allow the grammar to capture the verb phrase "is not hard". The final rule added to the grammar allowed a sentence to be constructed of a noun phrase, followed by a prepositional phrase, and a verb phrase. This rule allows the sentence to be initially broken down into the noun phrase "Riding”, the prepositional phrase "to Camelot", and the verb phrase "is not hard".

**Sentence 14 Parse Tree:**



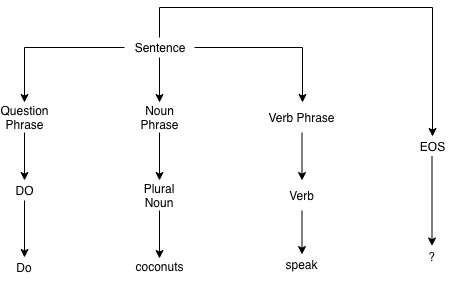
## Rule Change 11

**Sentence 15:** do coconuts speak ?

**Word Classes Added:** One new phrase class was added to the grammar to capture phrasing of questions. The class name for the questioning phrases that start a question was set to be "QP".

**Rule Changes:** To parse the sentence, two new rules were added to the grammar. The first rule states that the a QP can consist of only a word in the "DO" class, such as "do" or "does". The second rule added to the grammar states that a sentence can be made of a question phrase, a noun phrase, a verb phrase, and an end of sentence tag. The QP is able to parse the questioning phrase with only the word "do". The noun phrase follows the QP in the sentence, and is parsed as the plural noun "coconuts", and the verb phrase is parsed as the base verb speak.

**Sentence 15 Parse Tree:**

****

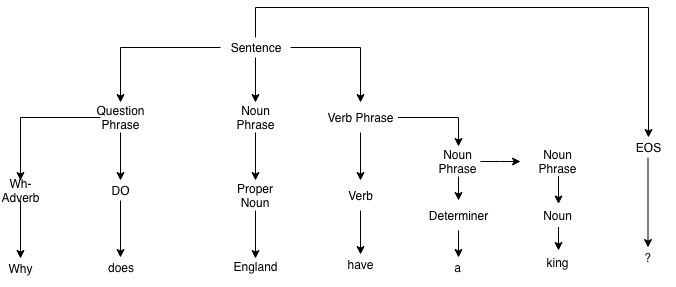
## Rule Change 12

**Word Classes Added:** One word class was added to the grammar to deal with the "wh-" adverbs, such as "why", "where", and "when". The name given to the words class was "AdvW".

**Sentence 16:** why does England have a king ?

**Rule Changes:** One rule was added to QPs in the grammar to parse the sentence. The rule states that the questioning phrase can consist of a "wh-" adverb, and a word in the "DO" class. This rule allows the questioning phrase in the sentence to be parsed as "why does". This sentence is also parsed with the same sentence level rule the previous sentence used; the questioning phrase is followed by a noun phrase and the verb phrase. The noun phrase consists of the proper noun "England", and the verb phase consists of "have a king", with the noun subphrase of "a king".

**Sentence 16 Parse Tree:**



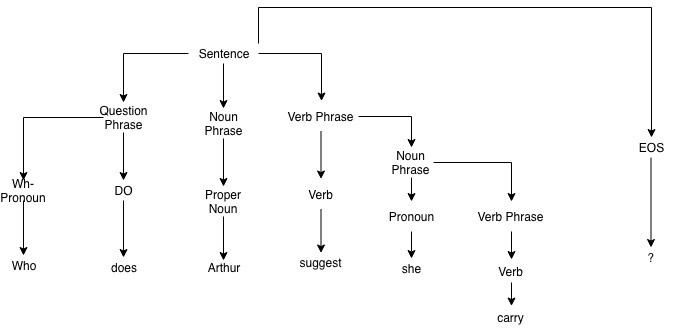
## Rule Change 13

**Challenge Sentence #1:** who does Arthur suggest she carry ?

**Word Classes Added:** One word class was added to the grammar to capture "wh-" pronouns such as "what" and "who". The word class name given to the "wh-" pronouns was called "ProW".

**Rule Changes:** Three rules were given to the grammar to parse the sentence. The first rule stated that the questioning phrase can consist of a "wh-" pronoun followed by a word in the DO class. This rule will capture the phrase "Who does" as a combination of a "wh-" pronoun and a word from the "DO" class. The second rule added to the grammar stated that a noun phrase can consist of a pronoun followed by a verb phrase. This rule allowed the phrase "she carry" be understood as a noun phrase with the pronoun "she", and the verb subphrase "carry". The last rule added to the grammar allowed a verb phrase to consist of a verb followed by a noun phrase. This rule allowed the grammar to understand that "suggest she carry" is a verb phrase that begins with the verb "suggest", and the noun subphrase "she carry".

**Challenge Sentence #1 Parse Tree:**

****

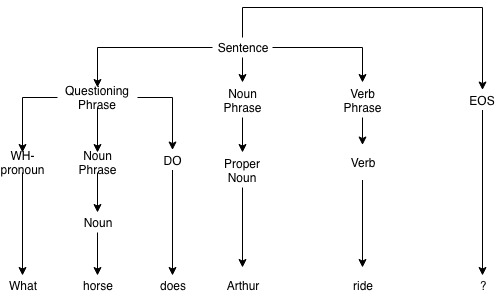
## Rule Change 14

**Challenge Sentence #2:** what horse does Arthur ride ?

**Word Classes Added:** None

**Rule Changes:** One rule was added to question phrases in grammar to parse the sentence. The rule added said that a question phrase can consist of a "wh-" pronoun, followed by a noun subphrase and a DO class word. This rule allows the grammar to break the questioning phrase down into the "wh-" pronoun "What", the noun subphrase with the noun "horse" and the DO class word "does". After the questioning phrase is parsed, the sentence follows a fairly simple structure of Noun Phrase, consisting of the proper noun "Arthur", and the verb phrase consisting of the verb "ride".

**Challenge Sentence #2 Parse Tree:**



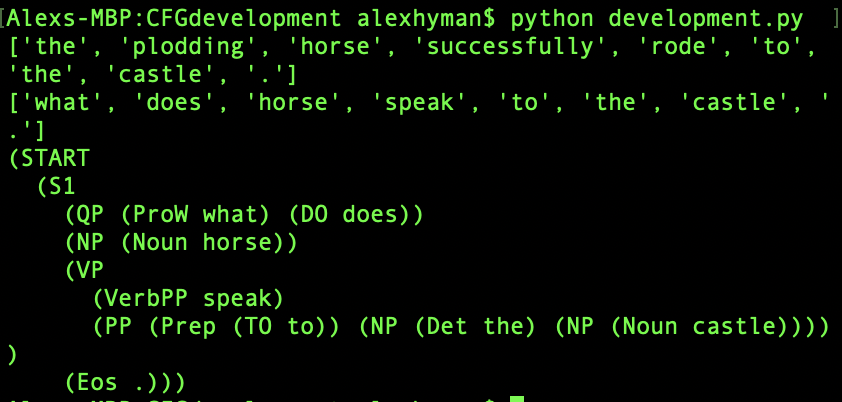
# Part 2

The first sentence I created that used all words known to the grammar, but could not be parsed by the grammar was “**the plodding horse successfully rode to the castle .**” This phrase was unable to be parsed by the grammar, because there is not a rule that allows an adverb to come before a verb; the only adverbial phrases on the grammar occur after the verb. Additionally, adverbs are not known to modify past tense verbs in the grammar, only base form verbs.

The second sentence that I created for the grammar to parse, but was not an actual sentence is “**what does horse speak to the castle.**” The first problem with the sentence is that there is no rule that matches the quantity of the noun in the noun phrase with the tense of the verb in the verb phrase. The other problem with the sentence is that there is not a rule stating when a determiner is required. In the sentence, a determiner is likely needed to specify which horse is being talked about in this instance, it could be “the horse”, “a horse”, or even “your horse”.

These overgeneralizations occur because the sentence only needs a properly constructed noun phrase and a properly constructed verb phrase, but there is an overgeneralization that any constructed verb phrases can go with any constructed noun phrases. This overgeneralization allows does not look at tense or the plurality of the noun phrase or the verb phrase. To fix this, rules would need to be stated either specifying what kind of verb phrase is used in the sentence, or create constructors specifically for certain cases, however it would be quite difficult to develop rules for every scenario.

**Exemplar Sentences Output:**



# Sentence Outputs

**Sentence 1:**

Arthur is the king .

(START

(S1

(NP (Proper Arthur))

(VP (VerbT is) (NP (Det the) (NP (Noun king))))

(Eos .)))

**Sentence 2:**

Arthur rides the plodding horse near the castle .

(START

(S1

(NP (Proper Arthur))

(VP

(VerbT rides)

(NP

(Det the)

(NP

(Noun horse)

(PP (Prep near) (NP (Det the) (NP (Noun castle)))))))

(Eos .)))

(START

(S1

(NP (Proper Arthur))

(VP

(VerbT rides)

(NP (Det the) (NP (Noun horse)))

(PP (Prep near) (NP (Det the) (NP (Noun castle)))))

(Eos .)))

(START

(S1

(NP (Proper Arthur))

(VP (VerbT rides) (NP (Det the) (NP (Noun horse))))

(PP (Prep near) (NP (Det the) (NP (Noun castle))))

(Eos .)))

**Sentence 3:**

Arthur rides the plodding horse near the castle .

(START

(S1

(NP (Proper Arthur))

(VP

(VerbT rides)

(NP

(Det the)

(NP

(Adj plodding)

(NP

(Noun horse)

(PP (Prep near) (NP (Det the) (NP (Noun castle))))))))

(Eos .)))

(START

(S1

(NP (Proper Arthur))

(VP

(VerbT rides)

(NP (Det the) (NP (Adj plodding) (NP (Noun horse))))

(PP (Prep near) (NP (Det the) (NP (Noun castle)))))

(Eos .)))

(START

(S1

(NP (Proper Arthur))

(VP

(VerbT rides)

(NP (Det the) (NP (Adj plodding) (NP (Noun horse)))))

(PP (Prep near) (NP (Det the) (NP (Noun castle))))

(Eos .)))

**Sentence 4:**

the Holy\_Grail is a chalice

(START

(S1

(NP (Det the) (NP (Proper Holy\_Grail)))

(VP (VerbT is) (NP (Det a) (NP (Noun chalice))))

(Eos .)))

**Sentence 5:**

the sensational Holy\_Grail is a sacred chalice .

(START

(S1

(NP (Det the) (NP (Adj sensational) (NP (Proper Holy\_Grail))))

(VP

(VerbT is)

(NP (Det a) (NP (Adj sacred) (NP (Noun chalice)))))

(Eos .)))

**Sentence 6:**

every coconut was carried to the hottest mountains .

(START

(S1

(NP (Det every) (NP (Noun coconut)))

(VP (VerbD was) (VP (VerbD carried)))

(PP

(Prep (TO to))

(NP (Det the) (NP (AdjS hottest) (NP (NounP mountains)))))

(Eos .)))

**Sentence 7:**

sixty strangers are at the Round\_Table .

(START

(S1

(NP (Num sixty) (NP (NounP strangers)))

(VP

(VerbPP are)

(PP (Prep at) (NP (Det the) (NP (Proper Round\_Table)))))

(Eos .)))

**Sentence 8:**

Sir\_Lancelot might have spoken .

(START

(S1

(NP (Proper Sir\_Lancelot))

(VP (MD might) (VP (Verb have) (VerbPastP spoken)))

(Eos .)))

**Sentence 9:**

Guinevere had been riding with Patsy for five weary nights .

(START

(S1

(NP (Proper Guinevere))

(VP (VerbD had) (VerbPastP been) (VerbPresP riding))

(PP

(Prep with)

(NP (Proper Patsy))

(PP

(Prep for)

(NP (Num five) (NP (Adj weary) (NP (NounP nights))))))

(Eos .)))

**Sentence 10:**

Sir\_Bedevere might have been suggesting this quest .

(START

(S1

(NP (Proper Sir\_Bedevere))

(VP

(MD might)

(Verb have)

(VP

(VerbPastP been)

(VerbPresP suggesting)

(NP (Det this) (NP (Noun quest)))))

(Eos .)))

**Sentence 11:**

the Britons migrate south frequently .

(START

(S1

(NP (Det the) (NP (ProperP Britons)))

(VP (Verb migrate) (AP (Adv south) (AP (Adv frequently))))

(Eos .)))

**Sentence 12:**

Arthur and Guinevere ride frequently near the castle .

(START

(S1

(NP (Proper Arthur) (Conj and) (NP (Proper Guinevere)))

(VP (Verb ride) (AP (Adv frequently)))

(PP (Prep near) (NP (Det the) (NP (Noun castle))))

(Eos .)))

**Sentence 13:**

he suggests to grow fruit at home .

(START

(S1

(NP (Pro he))

(VP

(VerbS suggests)

(INF (TO to) (Verb grow))

(NP (Noun fruit) (PP (Prep at) (NP (Noun home)))))

(Eos .)))

(START

(S1

(NP (Pro he))

(VP (VerbS suggests) (INF (TO to) (Verb grow)) (NP (Noun fruit)))

(PP (Prep at) (NP (Noun home)))

(Eos .)))

**Sentence 14:**

riding to Camelot is not hard

(START

(S1

(NP (VerbPresP riding))

(PP (Prep (TO to)) (NP (Proper Camelot)))

(VP (VerbT is) (NOT not) (Adj hard))

(Eos .)))

**Sentence 15:**

do coconuts speak ?

(START

(S1 (QP (DO do)) (NP (NounP coconuts)) (VP (Verb speak)) (Eos ?)))

**Sentence 16:**

why does England have a king ?

(START

(S1

(QP (AdvW why) (DO does))

(NP (Proper England))

(VP (Verb have) (NP (Det a) (NP (Noun king))))

(Eos ?)))

**Challenge Sentence #1:**

who does Arthur suggest she carry ?

(START

(S1

(QP (ProW who) (DO does))

(NP (Proper Arthur))

(VP (Verb suggest) (NP (Pro she) (VP (Verb carry))))

(Eos ?)))

**Challenge Sentence #2:**

what horse does Arthur ride

(START

(S1

(QP (ProW what) (NP (Noun horse)) (DO does))

(NP (Proper Arthur))

(VP (Verb ride))

(Eos ?)))