CS5012 Language and Computation
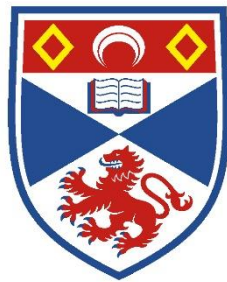

Practical 2: Grammar engineering

Student ID: 170009479

1 Introduction

Well-designed grammar helps translation work and voice processing in natural language processing. In this assignment some context-free grammar and unification grammar are built to identify the offered sentences.

The list of tasks that have been accomplished is as follows:

(1) Make a lexicon for the words showing in the offered sentences. (Refer to the appendix 1.)
(2) Create context-free grammar that allows the offered sentences be derived.
(3) Create unification grammar that not only can derive the original offered sentences but also refuse the offered sentences which have grammar errors.
(4) 4 sentence-extensions have been attempted. (Refer to the appendix 2 to know the specific sentences.)

In this report, firstly, the idea of creating context-free grammar and unification grammar are explained respectively. Secondly, I evaluate derivation result for both types of grammars. Thirdly, critical reflection on the language the grammar accepts is given.

2 The idea of creating context-free grammar and unification grammar

(1) Context-free grammar

Building a good sentence structure is a foundation of creating a powerful grammar. General statements can be divided into noun phrases (NP) and verb phrases (VP). A complex declarative sentence can include multiple sentences connected by conjunctions. Take the sentence 'when Homer drinks milk Bart laughs' for example, the sentence ca be divided into conjunction (CONJ), a sentence(S) and a sentence(S). Using the similar way to analyse sentences. 'When do Lisa and Bart wear shoes' can be divide to adjective (ADV), VP and VP, of which 'do' is regarded as a part of VP. 'Lisa thinks Homer thinks Bart drinks milk' can be divide into NP, VP and S. As my description, a sentence can be derived through several iterations.

NP, VP, and PP (i.e preposition phrase) are derived into part of speech by the similar way as above.

(2) Unification grammar

In unification grammar, to refuse the grammatically incorrect sentences, 'Subcategorisation (SUBCAT)', 'NUM' and 'Form' are used to clear the usage of and decide the form of verbs. For example, a transitive verb should be followed by an object, otherwise it is wrong. Moreover, in a statement sentence, the form of predicate verb is depended on the 'NUM' of the subject, thus, the value of the form of the predicate verb is same as the value of NUM of the subject (expressed in the form of NP). However, notice that if there is auxiliary before the predicate verb, the form of the verb is always base.

a) Basic requirement

Comparing with the context-free grammar, there are some changes in the unification grammar. In addition to add some parameters, such as 'SUBCAT', 'NUM' and 'Form', there are some changes occurring in the structure of sentence, NP and VP and POS of some words.

'When' is tagged as Wh_conj and Wh_adv in order to distinguish with ordinary conjunctions and adverbs because only this kind of conjunctions and adverbs can be accepts in the beginning of a sentence. 'Do' and 'does' are not common verb any more, as the occurrence of them decide the form of the following predicate verb. If a subject is composed by multiple proper nouns, then the verb cannot be the form 'verb-s', thus, a new structure of sentence 'Multi_ProperNoun and VP' are extracted separatedly.

b) Extensions

To realise some extensions, 'whom' and 'what' are added into the category of 'Wh_adv' since the usage of them are similar to 'when' sometimes. Besides, 'drinking' as a new form of present participle is added into verb category.

3 Evaluate derivation result

(1) Context-free grammar

In this grammar, there is no ambiguity because it is created only based on the offered sentences. However, the generalisation of the grammar is predicted to be bad. It means the grammar may not apply well to other sentences. Meanwhile, some grammatically incorrect sentences cannot be detected.

(1) Unification grammar

Ambiguity happened when deriving the sentences through using unification grammar. Firstly, since there are two forms of pronoun, 'NUM=third_singular_present' and 'NUM=third_singular_past', when the pronoun as an object after verb, the proper noun is tagged as the above two types. In addition, if the subject is 'someone and someone', namely the subject is plural, the two proper nouns are tagged as the above two types respectively. Secondly, if there are multiple PP in the end of a sentence, the deviation of PP causes ambiguity. Thirdly, there are two usages of serve, namely 'serve somebody something' and 'serve something' and both somebody and something are derived by NP, thus there occurs ambiguity in the sentence 'Homer serves Lisa'.

4 Critical reflection on the language the grammar accepts

My unification grammar can detect three types of errors. Firstly, subject-predicate inconsistency problem. Secondly, in a question sentence, the form of the predicate verb behind the auxiliary verb is not base. Thirdly, a transitive verb is not followed by NP, and an intransitive verb is followed by NP.

However, there are some flaws in my unification grammar. Firstly, if a subjective is followed by auxiliary, it cannot be derived. Secondly, for the sentence 'Bart likes drinking milk' can be derived, but if the subject is plural, the sentences cannot be derived. Additionally, if 'likes' is change to other third singular present verb, for example, 'wears', the new sentence 'Bart wears drinking milk' can be derived.

Appendix 1

sen1	Bart	laughs							
tag	N	V							
sen2	Homer	laughed							
tag	N	V							
sen3	Homer	serves	Lisa						
tag	N	V	V						
sen4	Bart	wears	blue	shoes					
tag	N	V	ADJ	N					
sen5	Bart	always	drinks	milk					
tag	N	ADV	V	N					
sen6	Bart	and	Lisa	drink	milk				
tag	N	CONJ	N	V	N				
sen7	Lisa	thinks	Homer	thinks	Bart	drinks	milk		
tag	N	V	N	V	N	V	N		
sen8	Lisa	serves	Bart	a	healthy	green	salad		
tag	N	V	N	DET	ADJ	ADJ	N		
sen9	when	Homer	drinks	milk	Bart	laughs			
tag	CONJ	N	V	N	N	V			
sen10	when	do	Lisa	and	Bart	wear	shoes		
tag	ADV	V	N	CONJ	N	V	N		
sen11	when	does	Lisa	drink	the	milk	on	the	table
tag	ADV	V	N	V	DET	N	PREP	DET	N
sen12	Homer	never	drinks	milk	in	the	kitchen	before	midnight
tag	N	ADV	V	N	PREP	DET	N	PREP	N

Appendix 2

(1) Bart likes drinking milk

(2) what does Homer drink

(3) whom does Homer serve salad

(4) whom do Homer and Lisa serve