NLP Assignment 2 Grammar Writing

Eyal Orbach - Id 015369317
Daniel Juravski - Id 206082323

Part 1:

a)
There are rules that cause recursive generation, such as:
`NP NP PP`
`PP Prep NP`
Those rules have their own LHS symbols on their RHS.
Incrementing the weights of rules with those LHS that have *non recursive* symbols on the RHS will lower the recursion.
If the sum weights of recursive rules is larger than non recursive rules for a LHS than it will likely reach a loop that will reach the stack's limit therefore those weights should be balanced to lean towards non-recursive rules.

b)
The rule that allows that kind of sentences is 'Noun Adj Noun' and this can happen recursively (in other words allow multiple adjectives).
But that LHS (Noun) has 5 terminal state rules with weight 1 each, thus the probability to pick the recursive rule is just 1/6.
For raising the adjectives number we just need to increase that rule's weight.

c)
These are the rules for NP:
`1 NP Det Noun`
`1 NP NP PP`
The first rule results in 2 pre-terminals, which will be replaced with terminals, therefore incrementing its weight will result in shorter sentences.

To increment the usage of adjectives, we will increment the weight for the below rule:
`1 Noun Adj Noun`
Thus incrementing the chances of a Noun being prefixed with an adjective instead of being replaced with a terminal state.

d)
All non-terminal states except 'NP' and 'Noun' and ROOT have exactly one state to be replaced with, hence there is no meaning for their weight, because the probability to choose them is 100%.
Changing the weights in ROOT is not significant since all options are natural
Changing the weights on terminal states can make sentences sound more natural on a semantic level
For example: we decreased weights of some Verbs and Adjectives, since they can refer to less nouns (just a little semantic issue).

Part 2:

We added/modified rules to produce a sentence similar to the ones given, got insights on it and adjusted it to integrate with the other rules.
Below is the modifications required for each sentence

a) We made Noun a non-terminal state (NOUN), and introduced `Nnp` (proper noun, Sally) and Nn (singular noun, president). NP can now transition into `Det Nn` or `Nnp` or `NP PP`

b) We introduced Cc a pre-terminal state (and, or) along with rules that utilize it on `NP` and `VP` and `Verb`

c) We replaced `Verb` with `Vi` (intransitive verb) and `Vt` (transitive verb)

d) We introduced `V5` a pre-terminal conjunctive verb relating to a sub-sentence. `VP` can now be replaced by `V5 SBAR`
We introduced `SBAR` as `Comp S`, where `Comp` is a pre-terminal to a relational word (that)

e)
We introduced `Prp` (personal pronoun) and It can only be placed to the left of a `VP`
We introduced `V6`, verbs that follow a `Prp` and are followed by `NP SBAR` where that `NP` is the subject. This is different than `v5` verbs that are active and therefore the subject NP is to the left of the verb.
"It perplexed Sally that..." compared to "Sally thought that"
We decided these verbs should not be used in an active form like "Sally perplexed the sandwich" though this is debatable

f) We added `Rb` (adverb, 'very') and a rule to replace sometimes an adjective with a adverb and adjective, the adjective can be replaced again many times (very very pickled)

g) We added 2 things to deal with similar sentences
1- a transition `VP –> VP PP`. any verb can happen 'on the floor'
2- `Vprep` -> special verbs that must be followed by a PP. "worked the proposal" is wrong while "worked on the proposal" is correct.
Nouns as before can also be followed by PP

h) We added the lazy adjective, and introduced a verb phrase `Vbzl`.
a pre-terminal present 3rd person linking verb ('is' being the 3rd person present form of 'be'.
be is a linking verb that needs to be followed by an adjective for now).
This introduced an unnatural phenomena that we decided was grammatical,
the 'is adjective' form can sometimes be followed by a PP that sounds wrong
"every sandwich is delicious with the president" but since "a sandwich is delicious with a pickle" is correct we left it as so.

i) We introduced `Vbg`, a pre-terminal present particle (eating) and now `Vbzl` can also be followed by `Vbg` (is eating) or `Vbg NP` (is eating the sandwich).

j) Vbzl is now also followed by a `PP` or a `NP`
"Sally is in the sandwich " and "Sally is a sandwich" respectively

* The problem of handling (b) and (h)/(i):
Issue 1 example: `Sally wanted and is eating a sandwich`
We solved that by not introducing the combination VP NP, but rather adding NP after only specific form verbs therefore we would still generate.
“Sally wanted a sandwich and is eating a pickle” or
“Sally wanted and ate a sandwich”

issue 2 example: `Sally sighed and is eating a sandwich` is un natural because of the time mixtures (and is now eating... will be more natural) though grammatically that is ok.

Part 4:

a) "a" vs. "an" problem:
adding nouns:
Nnv = Nn (Singular noun) with prefix of vowels
Separated Det to: Deta={'a'}, Detan={'an'} and the rest in Det.
adding adjectives:
the problem was Nnv can become Adj Nnv
'it is eating an apple’ => 'it is eating an delicious apple'
the 'a'/'an' should be determined only by the word following following the determiner, here the adjective.
we added Adjv = Adj with prefix of vowels
a) 'it is eating a delicious apple'
b) 'it is eating an excellent apple’

adding adverbs:
Rb splited to Rb and Rbv.
The 'a' , 'an' could be followed by it too.
a) 'it is eating a very delicious apple'
b) 'it is eating an extremely delicious apple'

Then we created NNAPRE (Nn that has 'a' prefix) and NNANPRE (Nn that has 'an' prefix)
Det works with both of them without any limitations.
About the Deta and Detan, we had to insure that the following symbol must have a matching symbol, thus, we created these 2 nonterminals NNAPRE and NNANPRE that their prefix are as follows:
NNAPRE follows Nn or Adj
NNANPRE follows Nnv or Adjv

Then we just did all the possible combinations of the adjectives, adverbs and nouns.
output:
'it is eating an extremely fine sandwich !'
'is it true that Sally wanted a very perplexed eagle ?'

c) Relative clauses:
A noun phrase can be replaced by a noun phrase 'that' something.
Detailed explanation:
NP can transition to NP REL.
REL is a non terminate state that is replaced by Comp RSV (Relative Statement verb) or Comp RSNV (Relative statement noun verb).
RSV is a non terminate state that is replaced by a Verb phase, therefore referring to the initial noun as the subject.
these can be any verb phrase Intransitive/Transitive/Conjunctive Verb
(Sally kissed the president that sighed)
(Sally kissed the president that ate the pickle)
(Sally kissed the president that thought that Sally ate the pickle)

RSNV is replaced by a noun phrase (the subject) and a verb that is either Transitive or Ditransitive, where the initial noun is the direct or indirect object respectively.
(Sally kissed the pickle that the president ate)
(Sally kissed the pickle that Sally thought that the president ate)

part 5

The sentence we used for inspiration was taken from "Rolling Stones" magazine
"*Former first lady Michelle Obama appeared on The Late Show with Stephen Colbert Friday as part of her Becoming book tour, where the conversation ranged from how she adjusted to living in the White House under Secret Service protection to how Barack proposed*"

There are a few phenomenas here but we decided to focus on verbs, dealing with tenses and passive verbs. To handle this, we first decided to deal with tenses of active verbs and then proceed to passive verbs

Tenses:
We decided to implement tenses only for transitive verbs at this point.
Expanding to intransitive or ditransitive verbs should follow the same logic but was left unhandled for now.
We have implemented a rule transitioning Vt (transitive verb) to any of the following
(These non terminal states were used to explicitly state the various tenses)
we also replaced some of our Vbzl rules from the previous part that were now redundant with the present progressive form

TSIMPPAS (Transitive verb past simple)
TSIMPPRE (Transitive verb present simple)
TSIMPFUT (Transitive verb future simple)
TPERFPAS (Transitive verb past perfect)
TPERFPRE (Transitive verb present perfect)
TPERFFUT (Transitive verb future perfect)
TPROGPAS (Transitive verb past progressive)
TPROGPRE (Transitive verb present progressive)
TPROGFUT (Transitive verb future progressive)
TPPPAS (Transitive verb past perfect progressive)
TPPPRE (Transitive verb present perfect progressive)
TPPFUT (Transitive verb future perfect progressive)

for each of the above we created rules that transition them to fixed modals followed by a verb in the correct form where the 5 forms of the verbs are: `Vtprog` (progressive) `Vtper` (perfect) `Vtsf` (simple future) `Vtsp` (simple present) `Vtp` (past)
example rule:

```
1 Tppfut  would have been Vtprog
```

where `vtprog` is any progressive verb
We noticed that these modals can also be dynamically chosen between a few options but we left them fixed at this point

Passive verbs:
We made this sentences as easier targets
"a proposal was offered"
"a proposal was offered to Sally"
"Sally was kissed by the president"

We introduced PVP (passive verb phrase)
We introduced PVT (passive transitive verb)
We implemented the various forms of passive tenses noticing they all utilize the same form of the verb but where in some word it is similar to the active past simple form in others it isn't.

active 'kissed' and passive 'kissed'
active 'ate' and passive 'eaten'

We added `Pprep` (passive prepositions) that are occasionally used with passive verb phrases (by) to explicitly mention the object that have initiated the verb's action (kissed by the president) and created another PVP rule accordingly.