Computational Syntax (2020-2021)

1) CG3 exam questions **(5/20)**

Given the rule in A) and the sentences analysed in B)

1. Explain what the rule does. Then, mark the output in the corresponding readings in B)
2. Is the result satisfactory? How would you modify the rule so that it performs better? Explain.

A)

```
DELIMITERS = "<$.>" "<$!>" "<$?>" "<$\;>" "<$:>" "<$-->" "<$>" "<$start>" "<$START>" ;
MAPPING-PREFIX = @ ;

SETS
LIST DRINK = (<drink>) ; # semantic information in readings

ADD (@WRONG!) TARGET ("take") IF (*1 DRINK);
```

B)

When Jane came, we took a beer to celebrate her birthday.

```
"<when>"
  "when" <*> <interr> <atemp> ADV
  "when" <*> <rel> <atemp> ADV
"<jane>"
  "Jane" <*> <Proper> <hum> <fem> N S NOM
"<came>"
  "come" <move> <va+DIR> <ve> <va+INF>
<v.contact> <v.contact> V IMPF
"<$,>"
  "," PU @PU
"<we>"
  "we" PERS 1P NOM
"<took>"
  "take" <v.contact> <v.contact> V IMPF
"<a>"
  "a" <indef> ART S
"<beer>"
  "beer" <drink> N S NOM
"<to>"
  "to" PRP
  "to" INFM
  "to" ADV
"<celebrate>"
  "celebrate" V PR -3S
  "celebrate" V INF
  "celebrate" V IMP
"<her>"
  "she" PERS FEM 3S ACC
  "she" <fem> <poss> <det> PERS FEM 3S GEN
"<birthday>"
  "birthday" <temp> <ac-cat> N S NOM
"<$.>"
  "." PU <<< @PU
```

It just takes two minutes to prepare a coffee.

```
"<it>"
  "it" <*> PERS NEU 3S NOM
  "it" <*> PERS NEU 3S ACC
"<just>"
  "just" <atemp> ADV
  "just" ADJ POS
"<takes>"
  "take" <act> <mon> N P NOM
  "take" <v.contact> <v.contact> V PR 3S
"<two>"
  "two" <ac-cat> N S NOM
  "two" <fr:99>  <b:186950> <card> NUM P
"<minutes>"
  "minutes" <ac-sign> N P NOM
  "minutes" <sem-r> N P NOM
  "minute" <dur> N P NOM
"<to>"
  "to" PRP
  "to" INFM
  "to" ADV
"<prepare>"
  "prepare" V PR -3S
  "prepare" V INF
  "prepare" V IMP
"<a>"
  "a" <indef> ART S
"<cup>"
  "cup" V PR -3S
  "cup" V INF
  "cup" V IMP
  "cup" <con> <occ> N S NOM
"<of>"
  "of" PRP
"<coffee>"
  "coffee" <col> <drink> N S NOM
"<$.>"
  "." PU <<< @PU
```

2) Language models. **(3/20)**

Given a corpus formed by 5 sentences:

the man walks slow
the girl runs fast
a man runs today
a girl swims fast
some girl walks slow
a man runs slow
a man swims slow
a girl runs slow
a man walks fast

We want to guess the most likely word to complete the following sentence:

a girl * slow

Which one of walks|runs|swims gives the highest probability to the sentence?

p(a girl runs slow) = p(a | *) x p(girl | a) x p(runs | girl) x p(slow | runs) x p(STOP | slow) =

5/8 x 2/6 x 2/4 x 2/4 x 1

p(a girl walks slow) = p(a | *) x p(girl | a) x p(walks | girl) x p(slow | walks) x p(STOP | slow) =

5/8 x 2/6 x 1/4 x 2/3 x 1

p(a girl swims slow) = p(a | *) x p(girl | a) x p(swims | girl) x p(slow | swims) x p(STOP | slow) =

5/8 x 2/6 x 1/4 x 1/2 x 1

3) POS tagging. **(3/20)**

Given the following tables of probabilities obtained from an annotated corpus:

p(word | tag):

	fish
N	0.7
V	0.3

p(tag_i | tag_{i-1}):

t_i / t_{i-1}	**N**	**V**	**STOP**
*****	0.6	0.4	0
N	0.2	0.6	0.1
V	0.5	0.1	0.4

Calculate the most likely sequence of POS tags for the sentence “fish fish”.

Hint: from the chapter on the assignment of POS tags using a bigram model, we obtained the following model:

p(The, man, sleeps, Det, Noun, Verb) =
p(Det | *) x p(Noun | Det) x p(Verb | Noun) x p(STOP | Verb) x
p(The | Det) x p(man | Noun) x p(sleeps | Verb)

p (fish fish N N) = p(N | *) x p(N | N) * p(STOP | N) x p(fish | N) x p(fish | N) =
0.6 x 0.2 x 0.3 x 0.7 x 0.7 = 0.01764

p (fish fish N V) = p(N | *) x p(V | N) * p(STOP | V) x p(fish | N) x p(fish | V) =
0.6 x 0.5 x 0.4 x 0.7 x 0.3 = 0.0252

p (fish fish V N) = p(V| *) x p(N | V) * p(STOP | N) x p(fish | V) x p(fish | N) =
0.4 x 0.1 x 0.1 x 0.7 x 0.3 = 0.00084

p (fish fish V V) = p(V| *) x p(V | V) * p(STOP | V) x p(fish | V) x p(fish | V) =
0.4 x 0.1 x 0.4 x 0.3 x 0.3 = 0.00144

4) CFG. **(3/20)**

Given the following CFG grammar:

S ->	NP VP
NP ->	NP PP
	we
	sushi
	chopsticks
PP ->	IN NP
IN ->	with
VP ->	V NP
	VP PP
V ->	eat

Draw the tree(s) corresponding to the the sentence “we eat sushi with chopsticks”

5) PCFG. **(3/20)**

Given the following PCFG grammar:

S ->	NP VP	1
NP ->	NP PP	1/2
	we	1/4
	sushi	1/8
	chopsticks	1/8
PP ->	IN NP	1
IN ->	with	1
VP ->	V NP	1/3
	VP PP	2/3
V ->	eat	1

Calculate the most likely tree for the sentence “we eat sushi with chopsticks”

(S (NP we) (VP (V eat) (NP (NP sushi) (PP (IN with) (NP chopsticks)))))
1 x 1/4 x 1/3 x 1 x 1/2 x 1/8 x 1 x 1 x 1/8

(S (NP we) (VP (VP (V eat) (NP sushi)) (PP (IN with) (NP chopsticks))))
1 x 1/4 x 2/3 x 2/3 x 1 x 1/8 x 1 x 1 x 1/8 -> higher probability

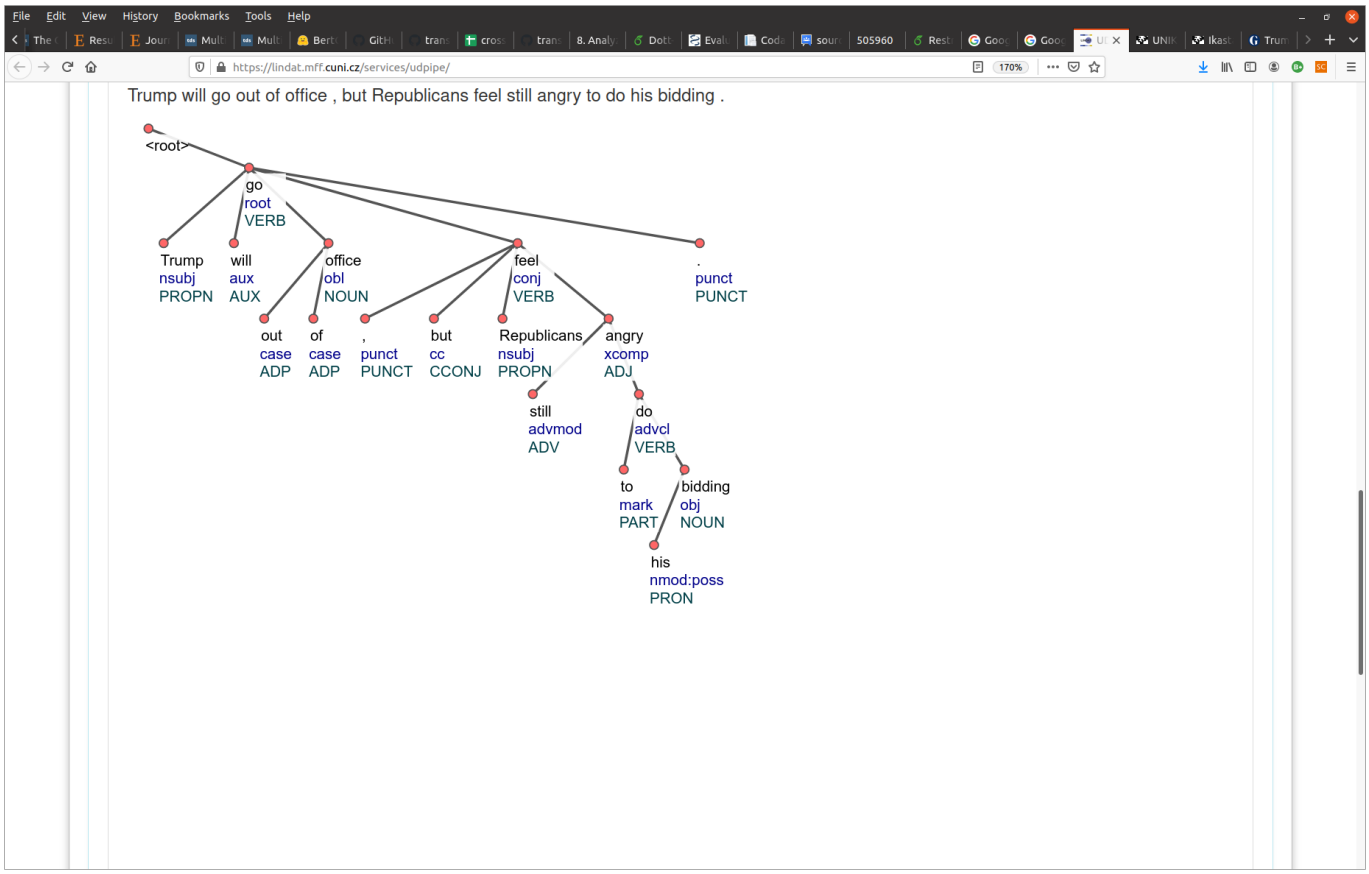
6) Dependency syntax. **(3/20)**

Given the following output for the sentence "Trump will go out of office, but Republicans feel still angry to do his bidding.":

```
# text = Trump will go out of office, but Republicans feel still angry to do his bidding.
1 Trump Trump PROPN NNP Number=Sing 3 nsubj _ TokenRange=0:5
2 will will AUX MD VerbForm=Fin 3 aux _ TokenRange=6:10
3 go go VERB VB VerbForm=Inf 0 root _ TokenRange=11:13
4 out out ADP IN _ 6 case _ TokenRange=14:17
5 of of ADP IN _ 6 case _ TokenRange=18:20
6 office office NOUN NN Number=Sing 3 obl _ SpaceAfter=No|TokenRange=21:27
7 , , PUNCT , _ 10 punct _ TokenRange=27:28
8 but but CCONJ CC _ 10 cc _ TokenRange=29:32
9 Republicans Republicans PROPN NNPS Number=Plur 10 nsubj _ SpacesAfter=\s\s|TokenRange=33:44
10 feel feel VERB VBP Mood=Ind|Tense=Pres|VerbForm=Fin 3 conj _ TokenRange=46:50
11 still still ADV RB _ 12 advmod _ TokenRange=51:56
12 angry angry ADJ JJ Degree=Pos 10 xcomp _ TokenRange=57:62
13 to to PART TO _ 14 mark _ TokenRange=63:65
14 do do VERB VB VerbForm=Inf 12 advcl _ TokenRange=66:68
15 his he PRON PRP$ Gender=Masc|Number=Sing|Person=3|Poss=Yes 16 nmod:poss _ TokenRange=69:72
16 bidding bidding NOUN NN Number=Sing 14 obj _ SpaceAfter=No|TokenRange=73:80
17 . . PUNCT . _ 3 punct _ SpaceAfter=No|TokenRange=80:81
```

a) Draw the corresponding dependecy tree

b) Comment on the analysis given. Is there any error in the analysis?


File Edit View History Bookmarks Tools Help
https://lindat.mff.cuni.cz/services/udpipe/
Trump will go out of office , but Republicans feel still angry to do his bidding .
<root>
go root VERB
Trump nsubj PROPN
will aux AUX
office obl NOUN
. punct PUNCT
out case ADP
of case ADP
, punct PUNCT
but cc CCONJ
feel conj VERB
Republicans nsubj PROPN
angry xcomp ADJ
still advmod ADV
do advcl VERB
to mark PART
bidding obj NOUN
his nmod:poss PRON