

# L95 Assignment 4

Danyi (Oli) Liu

November 7 2020

(1)

a 1-The 2-old 3-car 4-broke 5-down 6-in 7-the 8-car 9-park

(ncsubj broke:4 car:3 \_)

(ncmod \_ car:3 old:2)

(det car:3 The:1)

(ncmod prt broke:4 down:5)

(iobj broke:4 in:6)

(dojb in:6 park:9)

(ncmod \_ park:9 car:8)

(det park:9 the:7)

c 1-The 2-horses 3-were 4-broken 5-in 6-and 7-ridden 8-in 9-two 10-weeks

(ncsubj and:6 horses:2 obj)

(det horses:2 The:1)

(conj and:6 broken:4)

(conj and:6 ridden:7)

(aux and:6 were:3)

(passive broken:4)

(passive ridden)

(ncmod prt broken:4 in:5)

(iobj and:6 in:8)

(dojb in:8 weeks:10)

(ncmod num weeks:10 two:9)

(2)

a 1-It 2-was 3-my 4-aunt 5-'s 6-car 7-which 8-we 9-sold 10-at 11-auction 12-last 13-year 14-in 15-February

(ncsubj was:2 it:1)

(dojb was:2 car:6)

(ncmod poss car:6 aunt:4)

(det aunt:4 my:3)

(cmod \_ car:6 sold:9)

(obj sold:9 which:7)

(ncsubj sold:9 we:8)

(ncmod \_ sold:9 year:13)

(ncmod \_ year:13 in:14)

(dojb in:14 February:15)

(ncmod num year:13 last:12)

(iobj sold:9 at:10)

(dojb at:10 auction:11)

note: not sure how to deal with expletive it, e.g. how to represent extrapositioned noun phrase or the following modifier clause

b 1-The 2-only 3-rabbit 4-that 5-I 6-ever 7-liked 8-was 9-eaten 10-by 11-my 12-parents 13-one 14-summer

(ncsubj eaten:9 rabbit:3 \_) (ncmod \_ rabbit:3 only:2)

(det rabbit:3 the:1)

(cmod that rabbit:3 liked:7)

(ncsubj liked:7 I:5 \_)  
 (ncmod \_ liked:7 ever:6)  
 (aux eaten:9 was:8)  
 (passive eaten:9)  
 (iobj eaten:9 by:10)  
 (dobj by:10 parents:12)  
 (det parents:12 my:11)  
 (ncmod \_ eaten:9 summer:14)  
 (det summer:14 one:13)

(3)

b 1-It 2-won 3-'t 4-rain 5-but 6-there 7-might 8-be 9-snow 10-on 11-high 12-ground 13-if 14-the 15-temperature 16-stays 17-about 18-the 19-same 20-over 21-the 22-next 23-24 24-hours

(conj but:5 rain:4)  
 (conj but:5 be:8)  
 (ncsubj rain:4 It:1 \_)  
 (aux rain:4 will:2)  
 (ncmod \_ rain:4 not:5)  
 (ncsubj be:8 there:6 \_)  
 (aux be:8 might:7)  
 (dobj be:8 snow:9)  
 (ncmod \_ be:8 on:10)  
 (dobj on:10 ground:12)  
 (ncmod \_ ground:12 high:11)  
 (cmod \_ be:8 if:13)  
 (ccomp if:13 stays:16)  
 (ncsubj stays:16 temperature:15 \_)  
 (det temperature:15 the:14)  
 (iobj stays:16 about:17)  
 (dobj about:17 same:19)  
 (det same:19 the:18)  
 (iobj stays:16 over:20)  
 (dobj over:20 hours:24)  
 (det hours:24 the:21)  
 (ncmod num hours:24 24:23)  
 (ncmod \_ 24:23 next:22)

note: unsure about the relation between stay and about, complement? object?

c 1-My 2-wildest 3-dream 4-is 5-to 6-build 7-a 8-POS 9-tagger 10-which 11-processes 12-10K 13-words 14-per 15-second 16-and 17-uses 18-only 19-1MB 20-of 21-RAM 22-but 23-it 24-may 25-prove 26-too 27-hard

(ncsubj is:4 dream:3 \_)  
 (det dream:3 my:1)  
 (ncmod \_ dream:3 wildest:2)  
 (xcomp to is:4 build:6)  
 (dobj build:6 tagger:9)  
 (det tagger:9 a:7)  
 (ncmod \_ tagger:9 POS:8)  
 (cmod \_ tagger:9 and:16)  
 (conj and:16 processes:11)  
 (conj and:16 uses:17)  
 (dobj processes:11 words:13)  
 (ncmod num words:13 10k:12)  
 (ncmod \_ words:13 second:15)  
 (ncmod part second:15 every:14)  
 (dobj uses:17 1MB:21)  
 (ncmod 1MB:21 of:20)  
 (dobj of:20 RAM:21)

(ta comma is:4 prove:25)

(conj but:22 prove:25)

(ncsubj prove:25 it:23 \_)

(aux prove:25 may:24)

(ncmod \_ prove:25 hard:27)

(ncmod \_ hard:27 too:26)

note: not sure how to represent the relation between "1MB of RAM", i.e. is RAM or 1MB the object?

How to represent the function of but in the second clause? It shouldn't really be the head of 'prove' if 'prove' is the head of the clause?

What's the function of 'too hard'? Seems like complement, but does not belong to any of pcomp, xcomp or ccomp.

(4)

b 1-The 2-Penn 3-Treebank 4-tagset 5-was 6-culled 7-from 8-the 9-original 10-87 11-tag 12-tagset 13-for 14-the 15-Brown 16-Corpus

(ncsubj culled:6 tagset:4 \_)

(det tagset:4 The:1)

(ncmod \_ tagset:4 Treebank:3)

(ncmod \_ Treebank:3 Penn:2)

(aux culled:6 was:5)

(passive culled:6)

(iobj culled:6 from:7)

(dobj from:7 tagset:12)

(det tagset:12 the:8)

(ncmod \_ tagset:12 tag:11)

(ncmod num tag:11 87:10)

(ncmod \_ tagset:12 for:13)

(dobj for:13 Corpus:16)

(ncmod \_ Corpus:16 Brown:15)

(det Corpus:16 the:14)

note: seems to me that 'Penn Treebank' and 'Brown Corpus' should be treated as a whole? but not sure whether there's a mechanism for representing that?

1-For 2-example 3-the 4-original 5-Brown 6-and 7-C5 8-tagsets 9-include 10-a 11-separate 12-tag 13-for 14-each 15-of 16-the 17-different 18-forms 19-of 20-the 21-verbs 22-do 23-e.g. 24-C5 25-tag 26-VDD 27-for 28-did 29-and 30-VDG 31-tag 32-for 33-doing 34-be 35-and 36-have

(ncsubj include:9 tagsets:8 \_)

(ncmod \_ include:9 For:1)

(dobj For:1 example:2)

(ncmod \_ tagsets:8 and:6)

(conj and:6 Brown:5)

(conj and:6 C5-7)

(ncmod \_ tagsets:8 original:4)

(det tagsets:8 the:3)

(dobj include:9 tag:12)

(det tag:12 a:23)

(ncmod \_ tag:12 separate:11)

(idobj include:9 for:13)

(dobj for:13 each:14)

(ncmod \_ each:14 of:15)

(dobj of:15 forms:18)

(ncmod \_ forms:18 different:17)

(det forms:18 the:16)

(ncmod \_ forms:18 of:19)

(dobj of:19 verbs:21)

(det verbs:21 the:20)  
 (ncmod \_ verbs:21 and:35)  
 (conj and:35 do:22)  
 (conj and:35 be:34)  
 (conj and:35 have:36)  
 (ta brack include:9 tag:25)  
 (conj and:29 VDD:26)  
 (conj and:29 tag:31)  
 (ncmod \_ tag:31 VDG:30)  
 (ncsubj tag:25 C5:24 \_)  
 (iobj tag:25 for:27)  
 (iobj tag:25 for:32)  
 (dobj for:27 did:28)  
 (dobj for:32 doing:33)

note: unsure about the relations between 'do/be/have' and 'verbs'; again, not sure how to represent 'e.g.'; unsure to link 'for:27' and 'for:32' to 'and:29' or to 'tag:25')

c 1-The 2-slightly 3-simplified 4-version 5-of 6-the 7-Viterbi 8-algorithm 9-that 10-we 11-present 12-takes 13-as 14-input 15-a 16-single 17-HMM 18-and 19-a 20-sequence 21-of 22-observed 23-words 24-and 25-returns 26-the 27-most 28-probable 29-state 30-or 31-tag 32-sequence 33-together 34-with 35-its 36-probability

(ncsubj and:24 version:4 \_)  
 (conj and:24 takes:12)  
 (conj and:24 returns:25)  
 (det version:4 the:1)  
 (ncmod \_ version:4 simplified:3)  
 (ncmod \_ simplified:3 slightly:2)  
 (ncmod \_ version:4 of:5)  
 (dobj of:5 algorithm:8)  
 (ncmod \_ algorithm:8 Viterbi:7)  
 (det algorithm:8 the:6)  
 (xmod \_ algorithm:8 present:11)  
 (ncsubj present:11 we:10 \_)  
 (iobj takes:12 as:13)  
 (dobj as:13 input:14)  
 (obj2 takes:12 and:18)  
 (conj and:18 HMM:17)  
 (conj and:18 sequence:20)  
 (ncmod num HMM:17 single:16)  
 (det HMM:17 a:15)  
 (det sequence:20 a:19)  
 (ncmod \_ sequence:20 of:21)  
 (dobj of:21 words:23)  
 (ncmod \_ words:23 observed:22)  
 (dobj returns:25 sequence:32)  
 (ncmod \_ sequence:32 or:30)  
 (conj or:30 tag:31)  
 (conj or:30 state:29)  
 (det sequence:32 the:26)  
 (ncmod \_ sequence:32 probable:28)  
 (ncmod \_ probable:28 most:27)  
 (ncmod \_ sequence:32 with:34)  
 (ncmod \_ with:34 together:33)  
 (dobj with:34 probability:36)  
 (det probability:36 its:35)

note: does 'as input' act as a modifier for '... HMM and ... sequence ...' or as an object of 'takes'?