KAMANEH AKHAVAN

COMPUTATIONAL SYNTAX LT2214

ASSIGNMENT 1 – CHAPTER 4:

Who are they ?
1	Who:<PRON>	root
2	are:<AUX>	cop
3	they:<PRON>	nsubj
4	?:<PUNCT>	punct

utt[Utt[QS[IP who cop are comp[NP[PRON they]]] Punct ?]

Utt ::= Utt Punct
Utt ::= QS
QS ::= IP cop neg? Comp
Comp ::= NP
NP ::= Pron

آنها کی هستند ؟
1 آنها<pron> nsubj
2 کی<pron> root
3 هستند<AUX> cop

utt[Utt[QS[comp[NP[PRON آنها]]] IP کی cop هستند] Punct ?]

Utt ::= Utt Punct
Utt ::= QS
QS ::= comp? IP comp? neg? cop #Here we have to change the grammar cause in farsi(persian) the verb comes in the end of the sentence.
Comp ::= NP
NP ::= Pron

A small town with two minarets glides by .

1   A:<DET>
2   small:<ADJ>
3   town:<NOUN>
4   with:<ADP>
5   two:<NUM>
6   minarets:<NOUN>
7   glides:<VERB>
8   by:<ADV>
9   .:<PUNCT>

10 utt[utt [s[NP[NP[DET a CN[AP[A small] CN [N town]]] NP_obl[Prep with NP [Num two NP [CN[N minarets]]]]]][VP[VP [V glids] ADV by]]] Punct .]

Utt ::= Utt Punct   # head punct
Utt ::= S
S ::= NP VP # nsubj head
VP ::= VP comma? Adv # head punct advmod
VP ::= V
NP ::= NP comma? NP_obl # head punct nmod
NP ::= Det? CN # det head
CN ::= AP CN # amod head
CN ::= N
NP_obl ::= Prep NP # case head
NP ::= Num NP # nummod head
NP ::= Det? CN # det head
AP ::= A

یک شهر کوچک با دو مناره می گذرد#
1 یک :   <NUM>
2 شهر :   <NOUN>
3 کوچک: <ADJ>
4 با: <ADP>
5 دو: <NUM>
6 مناره :   <NOUN>
7 می گذرد: <VERB>
8 .:   <PUNCT>

[Utt[s[NP[NP[Num یک NP[CN [CN[N شهر] AP[A کوچک]]]] NP_obl[Prep با NP [Num دو NP [CN[N مناره]]]]]] VP [V می گذرد]]] Punct .]

Utt ::= Utt Punct   # head punct
Utt ::= S
S ::= NP VP # nsubj head
VP ::= V
NP ::= NP comma? NP_obl
NP ::= Num NP
NP ::= Det? CN
CN ::= CN AP# is changed
CN ::= N
NP_obl ::= Prep NP # case head
NP ::= Num NP # nummod head
NP ::= Det? CN # det head
AP ::= A

I was just a boy with muddy shoes .
```
1          I<PRON>  nsubj
2          was<AUX>           cop
3          just<ADV>          advmod
4          a<DET>   det
5          boy<NOUN>          root
6          with<ADP>          case
7          muddy<ADJ>         amod
8          shoes<NOUN>   nmod
9          .  <PUNCT>         punct
```

[Utt[s[NP[Pron I] cop was AdV just Comp[Comp[NP[Det a CN[N boy]]] NP_obl[Prep with NP[CN[AP [A muddy] CN[N shoes]]]]]] Punct .]

Utt ::= Utt punct
Utt ::= S
S ::= NP cop AdV? neg? Comp # nsubj cop advmod advmod head # nsubj cop advmod advmod head
NP ::= Pron
Comp ::= Comp comma? NP_obl # head punct nmod
Comp ::= NP
NP ::= Det? CN # det head
CN ::= N
NP_obl ::= Prep NP # case head
NP ::= Det? CN # det head
CN ::= AP CN # amod head
AP ::= A

من فقط یک پسربا کفشهای گل آلود بودم#
```
1 من <PRON>
2 فقط <ADV>
3  یک  <NUM>
4 پسر <NOUN>
5 با <ADP>
6 کفشهای <NOUN>
7 گلی <ADJ>
8 بودم <AUX>
9 . <PUNCT>
```

[Utt[S[NP[Pron من] AdV فقط Comp [Comp[Num یک NP[CN[N پسر]]] NP_obl [Prep با NP [CN [CN[N کفشهای AP [A آلود گل]]]]] comp_VP[cop بودم]]] punct .]

Utt ::= Utt punct
Utt ::= S
S ::= NP AdV? Comp neg? comp_VP?# this one is changed
comp_VP ::= N? ADJ? cop AUX? # compound VERB (is changed)
NP ::= Pron
Comp ::= Comp comma? NP_obl # head punct nmod
Comp ::= NP
NP ::= Num NP # nummod head
NP ::= Det? CN # det head
CN ::= N
NP_obl ::= Prep NP # case head
CN ::= CN AP # This one must be changed because in Persian the adjectives come after the nouns
AP ::= A

Shenzhen 's traffic police have opted for unconventional penalties before .

1	Shenzhen	<PROPN>	nmod:poss
2	's	<PART>	case
3	traffic	<NOUN>	compound
4	police	<NOUN>	nsubj
5	have	<AUX>	aux
6	opted	<VERB>	root
7	for	<ADP>	case
8	unconventional	<ADJ>	amod
9	penalties	<NOUN>	obl
10	before	<ADV>	advmod
11	..	<PUNCT>	punct

[Utt[s[NP[NP_poss[NP Shenzhen 's] CN[N[N_one traffic N police]]] have VP[VP[V2 opted NP_obl[Prep for NP [CN[AP [A unconventional] CN [N penalties]]]]] ADV before]]] Punct .]

Utt ::= Utt Punct   # head punct
Utt ::= S
S ::= NP have neg? VP # nsubj aux advmod head
NP ::= NP_poss CN # nmod:poss head
NP_poss ::= NP 's # head case
CN ::= N
N ::= N_one N # compound head
VP ::= VP Adv # head punct advmod
VP ::= V2 NP_obl # head obl
NP_obl ::= Prep NP # case head
NP ::= NP comma? NP_obl # head punct nmod
NP ::= Det? CN # det head
CN ::= AP CN # amod head
CN ::= N

‫# ‫.پلیس راهنمایی و رانندگی شنژن برای مجازاتهای غیرمتعارف پیشین انتخاب شده است

1 پلیس <NOUN>
2 راهنمایی <NOUN>
3 و
4 رانندگی <NOUN>
5 شنژن <PROPN>
6 برای <ADP>
7 مجازاتهای <NOUN>
8 غیرمتعارف <ADJ>
9 پیشین <ADJ>
10 انتخاب   ROOT   <Noun>
11 شده   COP   <AUX>
12 است   AUX   <AUX>
13 . <PUNCT>

[Utt[S [NP_poss [NP[ListNP [NP[CN[N[N پلیس N_oneراهنمایی]]]] NP_conj [Conj و NP[CN[NP_obl[NP [PN شنژن] Prep برای] AP[NP_obl [NP [CN [N مجازاتهای]] AP[A غیرمتعارف] AP[A پیشین]]]]]] VP [comp_VP[N انتخاب COP شده AUX   است]] Punct .]

Utt ::= Utt Punct
Utt ::= S
S ::= NP VP
NP ::= NP_poss NP # changed
NP_poss ::= NP ی   # changed
CN ::= NP_obl AP# changed
NP_obl ::= NP Prep? # changed
AP ::= NP_obl? AP AP? # changed
AP ::= A
NP ::= Det? CN
vp::= comp_VP
comp_VP ::= N? ADJ? cop AUX? # compound VERB (is changed)   NP::= PN
NP ::= Conj_pre? ListNP NP_conj   NP_conj ::= comma? Conj NP
ListNP::= NP   CN ::= N
N ::= N N_one# changed

It was conducted just off the Mexican coast from April to June .
```
1        It<PRON>       nsubj
2        was<AUX>              aux
3        conducted<VERB>    root
4        just<ADV>              advmod
5        off<ADP> case
6        the<DET> det
7        Mexican<ADJ>       amod
8        coast<NOUN>        obl
9        from<ADP>             case
10       April<PROPN>         obl
11       to<ADP>               case
12       June<PROPN>          nmod
13       .  <PUNCT>            punct
```

[Utt[NP[Pron It] cop was VP_pass[VP_pass[VP_pass[VP_pass[v2 conducted] Adv just] NP_obl[Prep off NP[Det the CN[AP [A Mexican]
CN[N[coast]]]]]]NP_obl[prep from NP[PN April NP_obl[Prep to NP[PN June]]]]]] Punct .]

Utt ::= Utt Punct   # head punct
Utt ::= S
S ::= NP cop AdV? neg? VP_pass # nsubj:pass aux:pass advmod advmod head
NP ::= Pron
VP_pass ::= VP_pass NP_obl?
VP_pass ::= VP_pass Adv?
VP_pass ::= V2
NP_obl ::= Prep NP # case head
NP ::= Det? CN # det head
CN ::= AP CN
CN ::= N
NP ::= NP comma? NP_obl
VP ::= VP NP_obl # head obl # 0.001
VP ::= VP Adv # head punct advmod
AP ::= A

#این دقیقاً در سواحل مکزیک از آوریل تا ژوئن انجام شد
```
1 این <DET>
2 دقیقاً <ADV>
3  در <ADP>
4 سواحل <NOUN>
5 مکزیک <ADJ>
6 از <ADP>
7 آوریل <PROPN>
8 تا <ADP>
9 ژوئن <PROPN>
10 انجام <NOUN>
11 شد <AUX>
12 .  <PUNCT>
```

[Utt[S [NP[DET این] AdV دقیقاً Comp[Comp[NP_obl[Prep در   NP[CN [CN [N سواحل] AP[A مکزیک]]]]] NP_obl[Prep[از] NP[NP [PN آوریل]
NP_obl [prep تا NP[PN ژوئن]]]]   ] comp_VP[NOUN انجام cop شد]]] Punct .]

Utt ::= Utt Punct   # head punct
Utt ::= S
S ::= NP AdV? Comp neg? comp_VP?# this one is changed
NP ::= Det? CN? # changed
comp_VP ::= N? ADJ? cop AUX? # compound VERB (is changed)
Comp ::= Comp comma? NP_obl # head punct nmod
Comp ::= NP_obl
NP_obl ::= Prep NP # case head
NP ::= Det? CN # det head
CN ::=CN AP # this one is changed
CN ::= N
NP ::= NP comma? NP_obl
NP ::= PN
AP ::= A

Part2:

Testing the grammar Using the gfud evaluation commands:

- Macro:

    LAS: UDScore {udScore = 0.4996882746882746, udMatching = 11, udTotalLength = 138, udSamesLength = 85, udPerfectMatch = 0} o

    UAS: UDScore {udScore = 0.7035715210715211, udMatching = 11, udTotalLength = 138, udSamesLength = 98, udPerfectMatch = 1}

- Micro:

    LAS:UDScore {udScore = 0.6071812080536913, udMatching = 11, udTotalLength = 138, udSamesLength = 85, udPerfectMatch = 0}

    UAS:UDScore {udScore = 0.6948322147651006, udMatching = 11, udTotalLength = 138, udSamesLength = 98, udPerfectMatch = 1}

It seems the evaluation gives better results when performed using a UAS standard rather than LAS.

Part3:

Modifying the file which include the English and Persian rules is in eng_Per.dbnf. Basic word order in persian is SOV .Also the adjectives come after the nouns.So I add some new rules to the file.
The evaluation of the English sentences had the same result and the result of the Persian was :

- Macro:

    LAS: U UDScore {udScore = 0.3434572316171532, udMatching = 18, udTotalLength = 253, udSamesLength = 92, udPerfectMatch = 0}

    UAS: UDScore {udScore = 0.4740667548744654, udMatching = 18, udTotalLength = 253, udSamesLength = 129, udPerfectMatch = 0}

- Micro:

    LAS:

    UDScore {udScore = 0.39036496350364632, udMatching = 18, udTotalLength = 253, udSamesLength = 92, udPerfectMatch = 0}

    UAS:

    UDScore {udScore = 0.45714538467153283, udMatching = 18 udTotalLength = 253, udSamesLength = 129 udPerfectMatch = 0}