1.

---------

Sentence: A privately owned Dragon capsule arrived safely at the international space station on Sunday, delivering a ton of supplies, after a shaky, nerve-racking start to its mission.

Matched pattern ng over Dragon capsule

Matched pattern ng over Dragon

Annotating Dragon capsule as ngroup

Matched pattern vg over arrived

Annotating arrived as constitcat=vgroup

Matched pattern ng over the international space

Annotating the international space as ngroup

Matched pattern ng over station

Annotating station as ngroup

Matched pattern ng over Sunday

Matched pattern ng over Sunday

Annotating Sunday as ngroup

Matched pattern ng over a ton

Annotating a ton as ngroup

Matched pattern ng over supplies

Annotating supplies as ngroup

Matched pattern ng over nerve

Annotating nerve as ngroup

Matched pattern ng over start

Annotating start as ngroup

Matched pattern ng over its mission

Annotating its mission as ngroup

“international space” should just be pre-nominal noun, but was tagged as an ngroup. However, this tagging is consistent with the rules defined in chunkPatterns.txt so I am not sure this should be classified as wrong tag.

Similarly, “a ton” should just be a quantifier but was also tagged as an ngroup.

Same applies to “never”, which should go together with “-racking” to form an adj, but was tagged as an ngroup.

---------

Sentence: The Dragon's arrival was one day late but especially sweet—and not because of the fresh fruit on board for the station astronauts who snared the capsule.

Matched pattern ng over Dragon

Matched pattern ng over Dragon

Annotating Dragon as ngroup

Matched pattern ng over arrival

Annotating arrival as ngroup

Matched pattern vg over was

Annotating was as constitcat=vgroup

Matched pattern ng over day

Annotating day as ngroup

Matched pattern ng over sweet—

Annotating sweet— as ngroup

Matched pattern ng over the fresh fruit

Annotating the fresh fruit as ngroup

Matched pattern ng over board

Annotating board as ngroup

Matched pattern ng over the station

Annotating the station as ngroup

Matched pattern ng over astronauts

Annotating astronauts as ngroup

Matched pattern vg over snared

Annotating snared as constitcat=vgroup

Matched pattern ng over the capsule

Annotating the capsule as ngroup

“Dragon” was tagged wrong as an ngroup. It should be that “The Dragon’s arrival” is tagged as an ngroup, since “The Dragon’s” is just a det-pos.

The “sweet—” was tagged wrong. “sweet” should be just an adj.

“station” should just be a pre-nominal noun, but was tagged as an ngroup. However, this tagging is consistent with the rules defined in the original chunkPatterns.txt so I am not sure if this is should be classified as wrong tag.

---------

Sentence: Spending cuts began, clearing the way for budget battles that will consume much of Congress's energy and threaten to eclipse other items on Obama's agenda.

Matched pattern ng over Spending

Annotating Spending as ngroup

Matched pattern ng over cuts

Annotating cuts as ngroup

Matched pattern vg over began

Annotating began as constitcat=vgroup

Matched pattern ng over the way

Annotating the way as ngroup

Matched pattern ng over budget

Annotating budget as ngroup

Matched pattern ng over battles

Annotating battles as ngroup

Matched pattern vg over will consume

Annotating will consume as constitcat=vgroup

Matched pattern ng over Congress

Matched pattern ng over Congress

Annotating Congress as ngroup

Matched pattern ng over energy

Annotating energy as ngroup

Matched pattern to-vg over threaten

Annotating threaten as constitcat=vgroup-inf

Matched pattern to-vg over eclipse

Annotating eclipse as constitcat=vgroup-inf

Matched pattern ng over other items

Annotating other items as ngroup

Matched pattern ng over Obama

Matched pattern ng over Obama

Annotating Obama as ngroup

Matched pattern ng over agenda

Annotating agenda as ngroup

The word “spending” and “budget” should both be just pre-nominal nouns, but were tagged as ngroup. But same as mentioned above, such a tagging is consistent with the rules in chunkPatterns.txt, so I am not sure this should be classified as wrong.

---------

Sentence: China begins a Parliament meeting this week that will unveil the final personnel changes of a once-a-decade leadership transition and provide some clues about the new administration's commitment to address mounting problems.

Matched pattern ng over China

Annotating China as ngroup

Matched pattern vg over begins

Annotating begins as constitcat=vgroup

Matched pattern ng over Parliament

Matched pattern ng over Parliament

Annotating Parliament as ngroup

Matched pattern ng over meeting

Annotating meeting as ngroup

Matched pattern ng over this week

Annotating this week as ngroup

Matched pattern vg over will unveil

Annotating will unveil as constitcat=vgroup

Matched pattern ng over the final personnel

Annotating the final personnel as ngroup

Matched pattern ng over changes

Annotating changes as ngroup

Matched pattern ng over decade

Annotating decade as ngroup

Matched pattern ng over leadership

Annotating leadership as ngroup

Matched pattern ng over transition

Annotating transition as ngroup

Matched pattern to-vg over provide

Annotating provide as constitcat=vgroup-inf

Matched pattern ng over some clues

Annotating some clues as ngroup

Matched pattern ng over the new administration

Annotating the new administration as ngroup

Matched pattern ng over commitment

Annotating commitment as ngroup

Matched pattern to-vg over address

Annotating address as constitcat=vgroup-inf

Matched pattern ng over problems

Annotating problems as ngroup

Similar to the situations in previous sentences, “Parliament” should by means be a pre-nominal noun but was tagged as an ngroup. Even so, it should be that “a Parliament” is tagged as an ngroup, not just “Parliament”.

Same, “the final personnel” should by means be a pre-nominal noun but was tagged as an ngroup.

“decade” was tagged as an ngroup, but in fact it should be included in “once-a-decade” as an adj.

“leadership” should by means be a pre-nominal noun but was tagged as an ngroup.

“provide” should be tagged as vgroup instead of vgroup-inf.

“the new administration” was mistakenly tagged as an ngroup. It should go with ‘s to form the det-pos part of the ng “the new administration’s commitment”.

“mounting problems” by means should be tagged as a whole as an ngroup, but looks like present participles are not recognized as adj by the pattern tagger.

2. The extended patterns are listed below:

// pattern set for noun and verb groups

pattern set chunks;

// patterns for noun groups

ng :=

det-pos? [constit cat=q]? quant-colloq? [constit cat=adj]\* [constit cat=n]+ |

proper-noun |

[constit cat=pro];

// Added possession form for proper nouns. Also added adj before the nouns.

det-pos := [constit cat=det] |

[constit cat=det]? [constit cat=adj]\* [constit cat=n number=singular] "'s" |

[constit cat=det]? [constit cat=adj]\* proper-noun "'s";

proper-noun := ([token case=cap] | [undefinedCap])+;

// Quantifier of colloquial form.

quant-colloq := [constit cat=n] "of";

when ng add [ngroup];

// patterns for active verb groups

// Added “modal do” forms such as “will do” which were missed in the original rules.

vg := [constit cat=tv] |

[constit cat=w] vg-inf |

[constit cat=w] "do" |

tv-vbe vg-ving;

vg-inf := [constit cat=v] |

"be" vg-ving;

vg-ven := [constit cat=ven] |

"been" vg-ving |

"been" [constit cat=ven];

vg-ving := [constit cat=ving] |

"being" [constit cat=ven];

tv-vbe := "is" | "are" | "was" | "were";

tv-vbp := "have" | "has" | "had";

when vg add [constit cat=vgroup];

// patterns for passive verb groups

vg-pass := tv-vbe [constit cat=ven] |

[constit cat=w] "be" [constit cat=ven];

when vg-pass add [constit cat=vgroup-pass];

// pattern for infinitival verb groups

to-vg := vg-inf;

when to-vg add [constit cat=vgroup-inf];

vg-perf := tv-vbp vg-ven |

[constit cat=w] tv-vbp vg-ven;

when vg-perf add [constit cat=vgroup-perf];

Some test cases are listed below:

---------

Sentence: The fish has been being eaten by the cat for ten minutes.

Annotating The fish as ngroup

Annotating has been being eaten as constitcat=vgroup-perf

Annotating the cat as ngroup

Annotating ten minutes as ngroup

---------

Sentence: Last Sunday, I just had my afternoon tea for three times which is rediculous isn't it.

Annotating Sunday as ngroup

Annotating I as ngroup

Annotating had as constitcat=vgroup

Annotating my afternoon tea as ngroup

Annotating three times as ngroup

Annotating is as constitcat=vgroup

Annotating is as constitcat=vgroup

Annotating it as ngroup

---------

Sentence: I have tested the modified pattern rules and they work well.

Annotating I as ngroup

Annotating have tested as constitcat=vgroup-perf

Annotating the modified pattern rules as ngroup

Annotating they as ngroup

Annotating work as constitcat=vgroup

3.

---------

Sentence: A privately owned Dragon capsule arrived safely at the international space station on Sunday, delivering a ton of supplies, after a shaky, nerve-racking start to its mission.

Annotating Dragon capsule as ngroup

Annotating arrived as constitcat=vgroup

Annotating the international space station as ngroup

Annotating Sunday as ngroup

Annotating a ton of supplies as ngroup

Annotating nerve as ngroup

Annotating start as ngroup

Annotating its mission as ngroup

We can see that now “the internatoinal space station” is tagged as an ngroup, as well as the “a ton of supplies”.

---------

Sentence: The Dragon's arrival was one day late but especially sweet—and not because of the fresh fruit on board for the station astronauts who snared the capsule.

Annotating The Dragon's arrival as ngroup

Annotating was as constitcat=vgroup

Annotating one day as ngroup

Annotating sweet— as ngroup

Annotating the fresh fruit as ngroup

Annotating board as ngroup

Annotating the station astronauts as ngroup

Annotating snared as constitcat=vgroup

Annotating the capsule as ngroup

Now “The Dragon’s arrival” and “station astronauts” are both tagged as an ngroup.

---------

Sentence: Spending cuts began, clearing the way for budget battles that will consume much of Congress's energy and threaten to eclipse other items on Obama's agenda.

Annotating Spending cuts as ngroup

Annotating began as constitcat=vgroup

Annotating the way as ngroup

Annotating budget battles as ngroup

Annotating will consume as constitcat=vgroup

Annotating Congress's energy as ngroup

Annotating threaten as constitcat=vgroup-inf

Annotating eclipse as constitcat=vgroup-inf

Annotating other items as ngroup

Annotating Obama's agenda as ngroup

Now “Spending cuts”, “budget battles” and “Congress’s engery” are all tagged as an ngroup.

---------

Sentence: China begins a Parliament meeting this week that will unveil the final personnel changes of a once-a-decade leadership transition and provide some clues about the new administration's commitment to address mounting problems.

Annotating China as ngroup

Annotating begins as constitcat=vgroup

Annotating Parliament as ngroup

Annotating meeting as ngroup

Annotating this week as ngroup

Annotating will unveil as constitcat=vgroup

Annotating the final personnel changes as ngroup

Annotating decade leadership transition as ngroup

Annotating provide as constitcat=vgroup-inf

Annotating some clues as ngroup

Annotating the new administration's commitment as ngroup

Annotating address as constitcat=vgroup-inf

Annotating problems as ngroup

Now “the final personnel changes” is tagged as an ngroup. But “decade leadership transition” is also tagged as an ngroup which is somewhat inappropriate. “the new administration’s commitment” is now also tagged as an ngroup.