Name: Mudunuri Sri Sai Sarat Chandra Varma

Class ID: 14

Problem Set (PS-1B)

1. **Given the text below, show the following NLP Operations.**
   a. **Sentence detection/Tokenization**
   b. **Lemmatization**
   c. **POS tagging/Chunking**
   d. **Parsing**
   e. **Name Entity Recognition**
   f. **Co-reference Resolution**

Answer:

I have run the text in intellij and below is the output for the given NLP Operations. Which includes Sentence detection/Tokenization, Lemmatization, POS tagging/Chunking, Parsing, Name Entity Recognition, Co-reference Resolution**.**

a) Sentence detection/Tokenization and c) POS tagging/Chunking and d) Parsing:

(ROOT (S (NP (NP (NNP United)) (, ,) (NP (NP (DT a) (NN unit)) (PP (IN of) (NP (NNP UAL)))) (, ,)) (VP (VBD said) (SBAR (S (NP (DT the) (NN increase)) (VP (VP (VBD took) (NP (NN effect)) (NP-TMP (NNP Thursday) (NN night))) (CC and) (VP (VBZ applies) (PP (TO to) (NP (JJS most) (NNS routes))) (SBAR (WHADVP (WRB where)) (S (NP (PRP it)) (VP (VBZ competes) (PP (IN against) (NP (NP (NN discount) (NNS carriers)) (, ,) (PP (JJ such) (IN as) (NP (NP (NP (NNP Chicago)) (PP (TO to) (NP (NNP Dallas) (CC and) (NNP Atlanta)))) (CC and) (NP (NP (NNP Denver)) (PP (TO to) (NP (NP (NNP San) (NNP Francisco)) (, ,) (NP (NNP Los) (NNP Angeles)) (CC and) (NP (NNP New) (NNP York)))))))))))))))) (. .)))

-> said/VBD (root)

-> United/NNP (nsubj)

-> ,/, (punct)

-> unit/NN (appos)

-> a/DT (det)

-> UAL/NNP (nmod:of)

-> of/IN (case)

-> ,/, (punct)

-> took/VBD (ccomp)

-> increase/NN (nsubj)

  -> the/DT (det)

-> effect/NN (dobj)

-> night/NN (nmod:tmod)

  -> Thursday/NNP (compound)

-> and/CC (cc)

-> applies/VBZ (conj:and)

  -> increase/NN (nsubj)

  -> routes/NNS (nmod:to)

    -> to/TO (case)

    -> most/JJS (amod)

  -> competes/VBZ (advcl)

    -> where/WRB (advmod)

    -> it/PRP (nsubj)

    -> carriers/NNS (nmod:against)

      -> against/IN (case)

      -> discount/NN (compound)

      -> ,/, (punct)

      -> Chicago/NNP (nmod:such_as)

        -> such/JJ (case)

          -> as/IN (mwe)

        -> Dallas/NNP (nmod:to)

          -> to/TO (case)

          -> and/CC (cc)

          -> Atlanta/NNP (conj:and)

        -> Atlanta/NNP (nmod:to)

        -> and/CC (cc)

        -> Denver/NNP (conj:and)

          -> Francisco/NNP (nmod:to)

-> to/TO (case)

-> San/NNP (compound)

-> ,/, (punct)

-> Angeles/NNP (conj:and)

-> Los/NNP (compound)

-> and/CC (cc)

-> York/NNP (conj:and)

-> New/NNP (compound)

-> Angeles/NNP (nmod:to)

-> York/NNP (nmod:to)

-> Denver/NNP (nmod:such_as)

-> applies/VBZ (ccomp)

-> ./. (punct)

f) Co-reference Resolution:

[CHAIN1-["AP" in sentence 1], CHAIN2-["United Airlines" in sentence 1, "it" in sentence 1], CHAIN3-["CHICAGO -LRB- AP -RRB- -- Citing high fuel prices , United Airlines said Friday it has increased fares by $ 6 per round trip on flights to some cities also served by lower-cost carriers ." in sentence 1], CHAIN4-["high fuel prices" in sentence 1], CHAIN5-["Friday" in sentence 1], CHAIN7-["fares" in sentence 1], CHAIN8-["round trip" in sentence 1], CHAIN9-["flights" in sentence 1], CHAIN10-["some cities also served by lower-cost carriers" in sentence 1], CHAIN11-["lower-cost carriers" in sentence 1], CHAIN12-["American Airlines" in sentence 2, "a unit AMR" in sentence 2], CHAIN13-["AMR" in sentence 2], CHAIN14-["American Airlines , a unit AMR" in sentence 2], CHAIN16-["the move" in sentence 2], CHAIN17-["spokesman Tim Wagner" in sentence 2], CHAIN18-["United" in sentence 3, "a unit of UAL" in sentence 3], CHAIN19-["UAL" in sentence 3], CHAIN20-["Dallas" in sentence 3], CHAIN21-["Atlanta" in sentence 3], CHAIN22-["San Francisco" in sentence 3], CHAIN23-["Los Angeles" in sentence 3], CHAIN24-["New York" in sentence 3], CHAIN25-["United , a unit of UAL" in sentence 3], CHAIN27-["the increase" in sentence 3, "it" in sentence 3], CHAIN28-["Thursday night" in sentence 3], CHAIN29-["most routes" in sentence 3], CHAIN31-["discount carriers , such as Chicago to Dallas and Atlanta and Denver to San Francisco , Los Angeles and New York" in sentence 3], CHAIN32-["discount carriers" in sentence 3], CHAIN33-["Chicago to Dallas and Atlanta and Denver to San Francisco , Los Angeles and New York" in sentence 3], CHAIN34-["Chicago to Dallas and Atlanta" in sentence 3], CHAIN35-["Dallas and Atlanta" in sentence 3], CHAIN36-["Denver to San Francisco , Los Angeles and New York" in sentence 3], CHAIN37-["San Francisco , Los Angeles and New York" in sentence 3]]

Process finished with exit code 0

b)Lemmatization and **e)Name Entity Recognition:**

CHICAGO-1
Text Annotation
CHICAGO-1:CHICAGO
Lemma Annotation
CHICAGO-1:CHICAGO
POS
CHICAGO-1:NNP
NER
CHICAGO-1:LOCATION

-LRB--2
Text Annotation
-LRB--2:-LRB-
Lemma Annotation
-LRB--2:-lrb-
POS
-LRB--2:-LRB-
NER
-LRB--2:O

AP-3
Text Annotation
AP-3:AP
Lemma Annotation
AP-3:AP
POS
AP-3:NNP
NER
AP-3:ORGANIZATION

-RRB--4
Text Annotation
-RRB--4:-RRB-
Lemma Annotation
-RRB--4:-rrb-
POS

-RRB--4:-RRB-
NER
-RRB--4:O

---5
Text Annotation
---5:--
Lemma Annotation
---5:--
POS
---5::
NER
---5:O

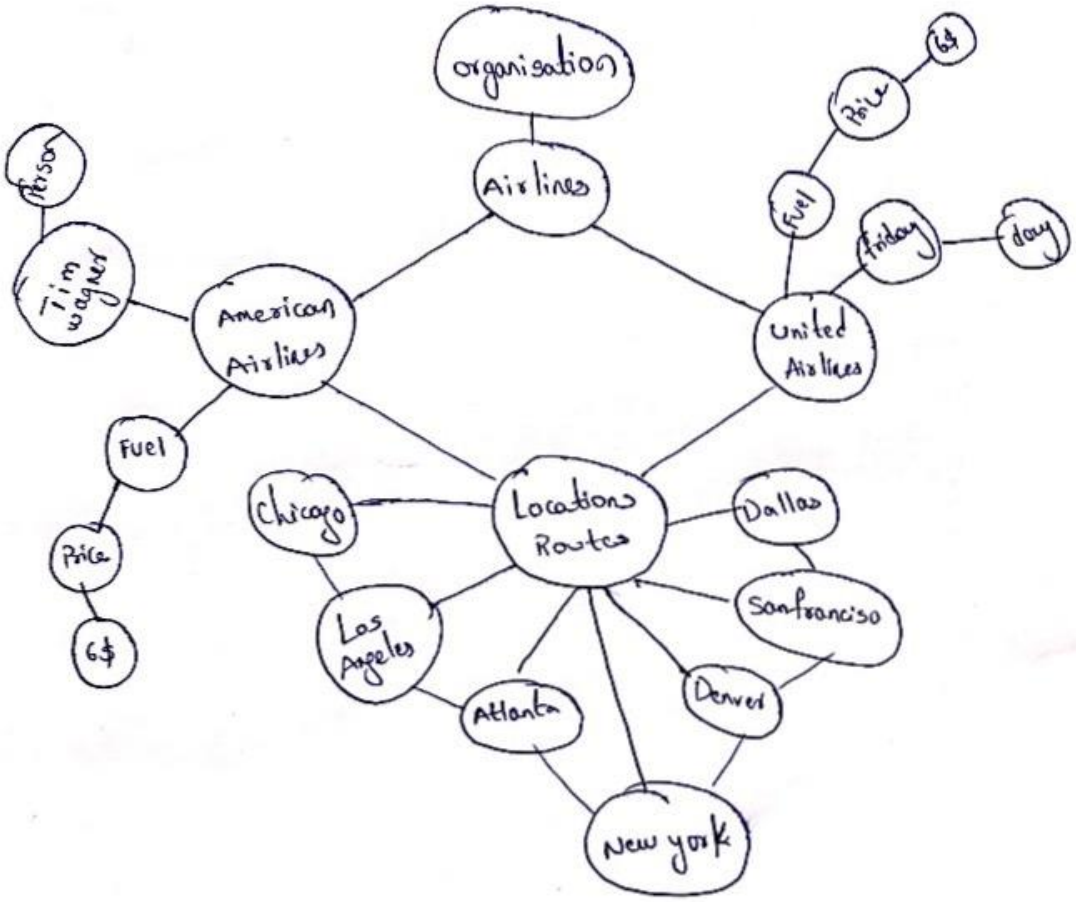
Citing-6
Text Annotation
Citing-6:Citing
Lemma Annotation
Citing-6:cite
POS
Citing-6:VBG
NER
Citing-6:O

2. **Summarize and draw a Knowledge graph.**

Answer:

The entities from the given text are as follows: UNITED AIRLINES, AMERICAN AIRLINES – Organisation; Friday- Date; 6$ - Money; Cities-City; Tim Wagner-Person; Dallas,Denver,Atlanta,San Francisco, Los Angeles, NewYork, Chicago-Locations. Linking these entities with their attributes gives us the knowledge graph for the given input data.

The knowledge graph for the above entities is drawn below:


organisation
Airlines
American Airlines
United Airlines
Locations Routes
Chicago
Dallas
Los Angeles
Sanfrancisco
Atlanta
Denver
New york
Fuel
Price
6$
Person
Tim wagner
Fuel
Price
6$
friday
day