

CS5560 Knowledge Discovery and Management

In-Class-Exercise (ICE-1B)

June 5, 2017

Name: SAIJYOTHI GUDIBANDI

ClassID: 18

1. Given the text example 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
2. Summarize and draw a knowledge graph.

CHICAGO (AP) — 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. American Airlines, a unit AMR, immediately matched the move, spokesman Tim Wagner said. United, a unit of UAL, said the increase took effect Thursday night and applies to most routes where it competes against discount carriers, such as Chicago to Dallas and Atlanta and Denver to San Francisco, Los Angeles and New York.

below are the results for sentence detection, Lemmatization, POS tagging/chunking, parsing, Name Entity Recog. & co-reference resolution operations.

CHICAGO-1, Lemma : CHICAGO, POS : NNP ,NE : LOCATION

-LRB--2, Lemma : -lrb-, POS : -LRB- ,NE : O

AP-3, Lemma : AP, POS : NNP ,NE : ORGANIZATION

-RRB--4, Lemma : -rrb-, POS : -RRB- ,NE : O

---5, Lemma : --, POS : : ,NE : O

Citing-6, Lemma : cite, POS : VBG ,NE : O

high-7, Lemma : high, POS : JJ ,NE : O

fuel-8, Lemma : fuel, POS : NN ,NE : O

prices-9, Lemma : price, POS : NNS ,NE : O

,-10, Lemma : ,, POS : , ,NE : O

United-11, Lemma : United, POS : NNP ,NE : ORGANIZATION

Airlines-12, Lemma : Airlines, POS : NNPS ,NE : ORGANIZATION

said-13, Lemma : say, POS : VBD ,NE : O

Friday-14, Lemma : Friday, POS : NNP ,NE : DATE

it-15, Lemma : it, POS : PRP ,NE : O

has-16, Lemma : have, POS : VBZ ,NE : O

increased-17, Lemma : increase, POS : VBN ,NE : O

fares-18, Lemma : fare, POS : NNS ,NE : O

by-19, Lemma : by, POS : IN ,NE : O

\$-20, Lemma : \$, POS : \$,NE : MONEY

6-21, Lemma : 6, POS : CD ,NE : MONEY

per-22, Lemma : per, POS : IN ,NE : O

round-23, Lemma : round, POS : NN ,NE : O

trip-24, Lemma : trip, POS : NN ,NE : O

on-25, Lemma : on, POS : IN ,NE : O

flights-26, Lemma : flight, POS : NNS ,NE : O

to-27, Lemma : to, POS : TO ,NE : O

some-28, Lemma : some, POS : DT ,NE : O

cities-29, Lemma : city, POS : NNS ,NE : O

also-30, Lemma : also, POS : RB ,NE : O

served-31, Lemma : serve, POS : VBD ,NE : O

by-32, Lemma : by, POS : IN ,NE : O

lower-cost-33, Lemma : lower-cost, POS : JJ ,NE : O

carriers-34, Lemma : carrier, POS : NNS ,NE : O

.-35, Lemma : ., POS : . ,NE : O

The Co-reference Resolution of the following text is as follows:

{1=CHAIN1-["AP" in sentence 1, "it" in sentence 1], 2=CHAIN2-["United Airlines" in sentence 1],
3=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], 4=CHAIN4-["high fuel prices" in sentence 1], 5=CHAIN5-["Friday" in sentence 1],
7=CHAIN7-["fares" in sentence 1], 8=CHAIN8-["round trip" in sentence 1], 9=CHAIN9-["flights" in
sentence 1], 10=CHAIN10-["some cities also served by lower-cost carriers" in sentence 1],
11=CHAIN11-["lower-cost carriers" in sentence 1], 12=CHAIN12-["American Airlines" in sentence 2,
"a unit AMR" in sentence 2], 13=CHAIN13-["AMR" in sentence 2], 14=CHAIN14-["American Airlines ,
a unit AMR" in sentence 2], 16=CHAIN16-["the move" in sentence 2], 17=CHAIN17-["spokesman Tim
Wagner" in sentence 2], 18=CHAIN18-["United" in sentence 3, "a unit of UAL" in sentence 3],
19=CHAIN19-["UAL" in sentence 3], 20=CHAIN20-["Dallas" in sentence 3], 21=CHAIN21-["Atlanta" in
sentence 3], 22=CHAIN22-["San Francisco" in sentence 3], 23=CHAIN23-["Los Angeles" in sentence
3], 24=CHAIN24-["New York" in sentence 3], 25=CHAIN25-["United , a unit of UAL" in sentence 3],
27=CHAIN27-["the increase" in sentence 3, "it" in sentence 3], 28=CHAIN28-["Thursday night" in
sentence 3], 29=CHAIN29-["most routes" in sentence 3], 31=CHAIN31-["discount carriers , such as
Chicago to Dallas and Atlanta and Denver to San Francisco , Los Angeles and New York" in sentence
3], 32=CHAIN32-["discount carriers" in sentence 3], 33=CHAIN33-["Chicago to Dallas and Atlanta and
Denver to San Francisco , Los Angeles and New York" in sentence 3], 34=CHAIN34-["Chicago to Dallas
and Atlanta" in sentence 3], 35=CHAIN35-["Dallas and Atlanta" in sentence 3], 36=CHAIN36-["Denver
to San Francisco , Los Angeles and New York" in sentence 3], 37=CHAIN37-["San Francisco , Los
Angeles and New York" in sentence 3]}

2. To draw the knowledge Graph for the given data, firstly we did some operations on the input text data. These operations are sentence tokenize, lemmatization, name of entity, pos tagging and co-reference. First we extract the data by chunking, entity reference and relational extraction. After this, segment them based on the topics. Finally for graph construction, we incorporate ontological constraints and relational patterns and discover relationships within knowledge graph.

