

CS5560 Knowledge Discovery and Management

In-Class-Exercise (ICE-1B)

June 5, 2017

Name: *Sujitha Pathana*

ClassID: *24*

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.

③ Steps for Knowledge Graph:-

Step 1 - Natural Language Processing

Annotation → Lemmatization → POS Tagging → Named Entity Recognition

CHICAGO → CHICAGO → NNP → LOCATION

LRB → -lrb- → -LRB-

AD → AP → NER → ORGANIZATION

-RRB- → -rrb- → -RRB-

lifting → lift → VBG

high → high → JJ

fuel → fuel → NN

prices → prices → NNS

,

United → United → NNP → Organization

Airlines → Airlines → NNPS → Organization

said → say → VBD

Friday → Friday → NNP → Date

it → it → PRP

has → have → VBZ

increased → increase → VBN

fares → fare → NNS

by → by → IN

\$ → \$ → Money

Annotation → Lemmaization → POS tagging → Named Entity Recognition

Airles → Air lines → NNPS → Organization

6 → 6 → CD → Money

per → per → IN

round → round → NN

trip → trip → NN

On → On → IN

flights → flight → NNS

to → to → TO

some → some → DT

cities → city → NNS

also → also → RB

served → serve → VBD

by → by → IN

lower-cost → lower-cost → JJ

carriers → carrier → NNS

→ .

American → American → NNP → organization

' → ' → ,

a → a → DT

unit → unit → NN

AMR → AMR → NNP

' → ' → ,

immediately → immediately → RB
matched → match → VBD
the → the → DT
move → move → NN

, → , → ,

Spokesman → spokesman → NN

Tim → Tim → NNP → PERSON

Wagner → Wagner → NNP → PERSON

said → say → VBD

United → United → NNP → Organization

a → a → DT

unit → unit → NN

of → of → IN

UAL → UAL → NNP → organization

, → , → ,

said → say → VBD

the → the → DT

increase → increase → NN

took → take → VBD

effect → effect → NN

Thursday → Thursday → NNP → DATE

night → night → NN → TIME

and → and → CC

applies → apply → VBZ

to → to → TO

Annotation → Lemmaization → POS tagging → Named Entity Recognition

most → most → JJ

routes → route → NN

where → where → WRB

it → it → PRP

competes → compete → VBZ

against → against → IN

discont → discont → NN

carriers → carrier → NNS

/ → / → ,

such → such → JJ

as → as → IN

Chicago → Chicago → NNP → Location

to → to → TO

Dallas → Dallas → NNP → Location

and → and → CC

Atlanta → Atlanta → NNP → Location

And → and → CC

Denver → Denver → NNP → Location

to → to → TO

San Francisco → San Francisco → NNP → Location

Los Angeles → Los Angeles → NNP → Location

and → and → CC

New York → New York → NNP → Location

Natural language Processing refers to the use and ability of systems to process sentences in a natural language such as English, rather than in a specialized artificial computer language such as C++.

Steps involved are:-

1. Tokenization
2. Stemming / Lemmatization
3. Parts of Speech tagging
4. Query Expansion
5. Parsing
6. Topic segmentation Recognition / Name Entity recognition.
7. Morphological Segmentation (word/sentence)

Step 2 - Information retrieval and information extraction

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 AAI, immediately matched the move, spokesman Tim Wagner said. United, a unit of UAL, said the increases 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.

Step 3: Topic Discovery

Organization	People	Places
United Airlines American Airlines AAR UAL	Tim Wagner	Chicago Dallas Atlanta Denver San Francisco Los Angeles New York

Step 4: Knowledge Graph Construction:-

