

**DEPARTMENT OF COMPUTER ENGINEERING**

## CSL804 Computational Lab II

**Eighth Semester, 2021-2022 (Even Semester)**

**Name of Student :** Saurav Kumar

## Roll No. 23

**Division :** BE – CMPN

**Day/ Session :** Monday/Afternoon

**Venue :** SLRTCE Lab 305

## Experiment No. 9

**Title of Experiment :** To study and implement Coreference resolution.

## Date of Conduction :

**Date of Submission :**

|  |  |  |
| --- | --- | --- |
| **Particulars Max. Marks Marks Obtained** | | |
| Preparedness and Efforts(PE) | **3** |  |
| Knowledge of tools(KT) | **3** |  |
| Debugging and results(DR) | **3** |  |
| Documentation(DN) | **3** |  |
| Punctuality & Lab Ethics(PL) | **3** |  |
| **Total** | **15** |  |

**Grades – Meet Expectations (3 Marks), Moderate Expectations (2 Marks), Below Expectations (1 Mark)**

**Checked and Verified by Name of Faculty :** Prof. Neelam Kulkarni

## Signature :

**Date :**

EXPERIMENT NO: 9

COREFERENCE

**AIM:** To study and implement Coreference resolution.

**SOFTWARE:** Python, AllenNLP.

# THEORY:

COREFERENCE RESOLUTION

Coreference resolution is the task of finding all expressions that refer to the same entity in text. It is an important step for a lot of higher level NLP tasks that involve natural language understanding such as documentation summarization, question answering and information extraction.

Coreference resolution is the NLP equivalent of endophoric awareness used in information retrieval systems, conversational agents, and virtual assistants like Amazon’s Alexa. It is the task of clustering mentions in text that refer to the same underlying entities.

For example:

+=================+

| |

“I drove Joe home because he lives close to my apartment”, she said.

| | |

+==========================================+===============+

“I”, “my”, and “she” belong to the same cluster and “Joe” and “he” belong to the same cluster.

# IMPLEMENTATION:

ALLENNLP COREFERENCE RESOLUTION

pip install allennlp

pip install allennlp-models

CODE

from allennlp.predictors.predictor import Predictor

model\_url = "https://storage.googleapis.com/allennlp-public-models/ coref-spanbert-large-2020.02.27.tar.gz"

predictor = Predictor.from\_path(model\_url)

text = "Joseph Robinette Biden Jr. is an American politician who is the 46th and current president of the United States. A member of the Democratic Party, he served as the 47th vice president from 2009 to 2017 under Barack Obama and represented Delaware in the United States Senate from 1973 to 2009."

prediction = predictor.predict(document=text) # get prediction print("CLUSTERS:")

for cluster in prediction['clusters']:

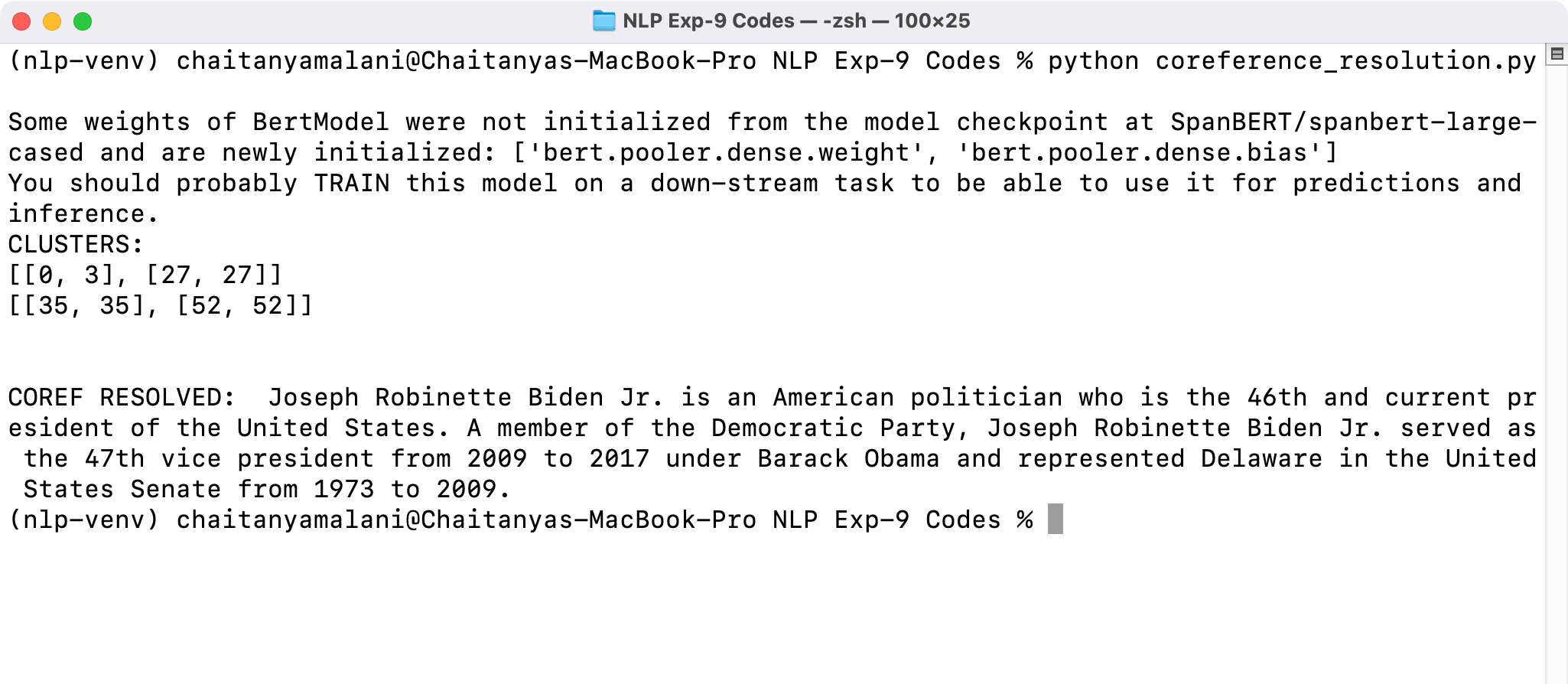
print(cluster) # list of clusters (the indices of spaCy tokens)

# Result: [[[0, 3], [26, 26]], [[34, 34], [50, 50]]]

print('\n') #Newline

print('COREF RESOLVED: ',predictor.coref\_resolved(text)) # resolved text

OUTPUT



# CONCLUSION:

Thus we have studied and implemented Coreference resolution.