

# NLP\_lab-9\_205229118\_Mahalakshmi.S

May 27, 2021

## 0.0.1 Lab9. Building Bigram Tagger

In this lab, you will build a bigram tagger and test it out. We will use the Brown Corpus and its native tagset.

### 0.0.2 EXERCISE-1

```
[1]: import nltk
      from nltk.tokenize import sent_tokenize, word_tokenize
      nltk.download('punkt')
      nltk.download('averaged_perceptron_tagger')
```

```
[nltk_data] Downloading package punkt to
[nltk_data] C:\Users\Mahalakshmi\AppData\Roaming\nltk_data...
[nltk_data] Package punkt is already up-to-date!
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data] C:\Users\Mahalakshmi\AppData\Roaming\nltk_data...
[nltk_data] Package averaged_perceptron_tagger is already up-to-
[nltk_data] date!
```

```
[1]: True
```

```
[2]: text = word_tokenize("And now for something completely different")
      nltk.pos_tag(text)
```

```
[2]: [('And', 'CC'),
      ('now', 'RB'),
      ('for', 'IN'),
      ('something', 'NN'),
      ('completely', 'RB'),
      ('different', 'JJ')]
```

### 0.0.3 EXERCISE-2

Type the following lines and load Brown corpus into the list tagsen.

```
[4]: from nltk.corpus import brown
      tagsen = brown.tagged_sents()
```

```
[6]: len(tagsen)
```

```
[6]: 57340
```

#### 0.0.4 STEP 1: Prepare data sets

There are a total of 57,340 POS-tagged sentences in the Brown Corpus. Among them, assign the first 50,000 to your list of training sentences. Then, assign the remaining sentences to your list of testing sentences. The first of your testing sentences should look like this:

```
[5]: br_train=tagsen[0:50000]
br_test=tagsen[50000:]
br_test[0]
```

```
[5]: [('I', 'PPSS'),
      ('was', 'BEDZ'),
      ('loaded', 'VBN'),
      ('with', 'IN'),
      ('suds', 'NNS'),
      ('when', 'WRB'),
      ('I', 'PPSS'),
      ('ran', 'VBD'),
      ('away', 'RB'),
      (',', ','),
      ('and', 'CC'),
      ('I', 'PPSS'),
      ("haven't", 'HV*'),
      ('had', 'HVN'),
      ('a', 'AT'),
      ('chance', 'NN'),
      ('to', 'TO'),
      ('wash', 'VB'),
      ('it', 'PPO'),
      ('off', 'RP'),
      ('.', '.')]

```

#### 0.0.5 STEP 2: Build a bigram tagger

Following the steps shown in this chapter, build a bigram tagger with two back-off models. The first one on the stack should be a default tagger that assigns ‘NN’ by default.

```
[8]: bt0 = nltk.DefaultTagger('NN')
bt1 = nltk.UnigramTagger(br_train, backoff=bt0)
bt2 = nltk.BigramTagger(br_train, backoff=bt1)
```

#### 0.0.6 STEP 3: Evaluate

Evaluate your bigram tagger on the test sentences. You should be getting the accuracy score of 0.911. If not, something went wrong: go back and re-build your tagger.

```
[9]: bt2.evaluate(br_test)
```

```
[9]: 0.9111006662708622
```

### 0.0.7 STEP 4: Explore

Now, explore your tagger to answer the questions below. ##### 1. How big are your training data and testing data? Answer in terms of the number of total words in them.

```
[12]: total_train=[len(l) for l in br_train]
      sum(total_train)
```

```
[12]: 1039920
```

```
[13]: total_test=[len(l) for l in br_test]
      sum(total_test)
```

```
[13]: 121272
```

2. What is the performance of each of the two back-off taggers? How much improvement did you get: (1) going from the default tagger to the unigram tagger, and (2) going from the unigram tagger to the bigram tagger?

```
[14]: bt1.evaluate(br_test)
```

```
[14]: 0.8897849462365591
```

```
[15]: bt2.evaluate(br_test)
```

```
[15]: 0.9111006662708622
```

3. Recall that ‘cold’ is ambiguous between JJ ‘adjective’ and NN ‘singular noun’. Let’s explore the word in the training data. The problem with the training data, through, is that it is a list of tagged sentences, and it’s difficult to get to the tagged words which are one level below:

```
[16]: br_train[0]
```

```
[16]: [('The', 'AT'),
      ('Fulton', 'NP-TL'),
      ('County', 'NN-TL'),
      ('Grand', 'JJ-TL'),
      ('Jury', 'NN-TL'),
      ('said', 'VBD'),
      ('Friday', 'NR'),
      ('an', 'AT'),
      ('investigation', 'NN'),
      ('of', 'IN'),
      ('Atlanta's', 'NP$'),
      ('recent', 'JJ'),
```

```
(('primary', 'NN'),
 ('election', 'NN'),
 ('produced', 'VBD'),
 ('`', '`'),
 ('no', 'AT'),
 ('evidence', 'NN'),
 ('"', '"'),
 ('that', 'CS'),
 ('any', 'DTI'),
 ('irregularities', 'NNS'),
 ('took', 'VBD'),
 ('place', 'NN'),
 ('.', '.'])]
```

```
[17]: br_train[1277]
```

```
[17]: [(('`', '`'),
 ('I', 'PPSS'),
 ('told', 'VBD'),
 ('him', 'PPO'),
 ('who', 'WPS'),
 ('I', 'PPSS'),
 ('was', 'BEDZ'),
 ('and', 'CC'),
 ('he', 'PPS'),
 ('was', 'BEDZ'),
 ('quite', 'QL'),
 ('cold', 'JJ'),
 ('.', '.'])]
```

```
[18]: br_train[1277][11]
```

```
[18]: ('cold', 'JJ')
```

4. To be able to compile tagged-word-level statistics, we will need a flat list of tagged words, without them being organized into sentences. How to do this? You can use multi-loop list comprehension to construct it:

```
[19]: br_train_f = [(word, tag) for sent in br_train for (word, tag) in sent]
```

```
[20]: br_train_f[:40]
```

```
[20]: [('The', 'AT'),
 ('Fulton', 'NP-TL'),
 ('County', 'NN-TL'),
 ('Grand', 'JJ-TL'),
 ('Jury', 'NN-TL'),
 ('said', 'VBD'),
```

```
(('Friday', 'NR'),
 ('an', 'AT'),
 ('investigation', 'NN'),
 ('of', 'IN'),
 ("Atlanta's", 'NP$'),
 ('recent', 'JJ'),
 ('primary', 'NN'),
 ('election', 'NN'),
 ('produced', 'VBD'),
 ('``', '``'),
 ('no', 'AT'),
 ('evidence', 'NN'),
 ('"', '"'),
 ('that', 'CS'),
 ('any', 'DTI'),
 ('irregularities', 'NNS'),
 ('took', 'VBD'),
 ('place', 'NN'),
 ('.', '.'),
 ('The', 'AT'),
 ('jury', 'NN'),
 ('further', 'RBR'),
 ('said', 'VBD'),
 ('in', 'IN'),
 ('term-end', 'NN'),
 ('presentments', 'NNS'),
 ('that', 'CS'),
 ('the', 'AT'),
 ('City', 'NN-TL'),
 ('Executive', 'JJ-TL'),
 ('Committee', 'NN-TL'),
 ('', '''),
 ('which', 'WDT'),
 ('had', 'HVD')]
```

```
[21]: br_train_f[13]
```

```
[21]: ('election', 'NN')
```

5. Now, exploring this list of (word, POS) pairs from the training data, answer the questions below.

a. Which is the more likely tag for ‘cold’ overall?

```
[24]: fr_dis = nltk.FreqDist(br_train_f)
      con_fd = nltk.ConditionalFreqDist(br_train_f)
```

```
[25]: con_fd['cold'].most_common()
```

```
[25]: [('JJ', 110), ('NN', 8), ('RB', 2)]
```

b. When the POS tag of the preceding word (call it POS<sub>n-1</sub>) is AT, what is the likelihood of ‘cold’ being a noun? How about it being an adjective?

```
[26]: br_train_2gr = list(nltk.ngrams(br_train_f,2))
br_train_cold=[a[1] for (a, b) in br_train_2gr if b[0] == 'cold']
frqdist = nltk.FreqDist(br_train_cold)
[tag for (tag, _) in frqdist.most_common()]
```

```
[26]: ['AT',
      'IN',
      'CC',
      'QL',
      'BEDZ',
      'JJ',
      ',',
      'DT',
      'PP$',
      'RP',
      '`',
      'NN',
      'VBN',
      'VBD',
      'CS',
      'BEZ',
      'DOZ',
      'RB',
      'PPSS',
      'BE',
      'VB',
      'VBZ',
      'NP$',
      'BEDZ*',
      '--',
      'DTI',
      'WRB',
      'BED']
```

c. When POS<sub>n-1</sub> is JJ, what is the likelihood of ‘cold’ being a noun? How about it being an adjective?

```
[27]: br_pos = [(w2+"/"+t2, t1) for ((w1,t1),(w2,t2)) in br_train_2gr]
br_pos_confreqdis = nltk.ConditionalFreqDist(br_pos)
br_pos
```

```
[27]: [('Fulton/NP-TL', 'AT'),
      ('County/NN-TL', 'NP-TL'),
      ('Grand/JJ-TL', 'NN-TL'),
```

('Jury/NN-TL', 'JJ-TL'),  
 ('said/VBD', 'NN-TL'),  
 ('Friday/NR', 'VBD'),  
 ('an/AT', 'NR'),  
 ('investigation/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('Atlanta's/NP\$', 'IN'),  
 ('recent/JJ', 'NP\$'),  
 ('primary/NN', 'JJ'),  
 ('election/NN', 'NN'),  
 ('produced/VBD', 'NN'),  
 ('``/``', 'VBD'),  
 ('no/AT', '``'),  
 ('evidence/NN', 'AT'),  
 ('''/''', 'NN'),  
 ('that/CS', '''),  
 ('any/DTI', 'CS'),  
 ('irregularities/NNS', 'DTI'),  
 ('took/VBD', 'NNS'),  
 ('place/NN', 'VBD'),  
 ('./.', 'NN'),  
 ('The/AT', '.'),  
 ('jury/NN', 'AT'),  
 ('further/RBR', 'NN'),  
 ('said/VBD', 'RBR'),  
 ('in/IN', 'VBD'),  
 ('term-end/NN', 'IN'),  
 ('presentments/NNS', 'NN'),  
 ('that/CS', 'NNS'),  
 ('the/AT', 'CS'),  
 ('City/NN-TL', 'AT'),  
 ('Executive/JJ-TL', 'NN-TL'),  
 ('Committee/NN-TL', 'JJ-TL'),  
 ('/,', 'NN-TL'),  
 ('which/WDI', ','),  
 ('had/HVD', 'WDI'),  
 ('over-all/JJ', 'HVD'),  
 ('charge/NN', 'JJ'),  
 ('of/IN', 'NN'),  
 ('the/AT', 'IN'),  
 ('election/NN', 'AT'),  
 ('/,', 'NN'),  
 ('``/``', ','),  
 ('deserves/VBZ', '``'),  
 ('the/AT', 'VBZ'),  
 ('praise/NN', 'AT'),  
 ('and/CC', 'NN'),

('thanks/NNS', 'CC'),  
 ('of/IN', 'NNS'),  
 ('the/AT', 'IN'),  
 ('City/NN-TL', 'AT'),  
 ('of/IN-TL', 'NN-TL'),  
 ('Atlanta/NP-TL', 'IN-TL'),  
 ("''/'", 'NP-TL'),  
 ('for/IN', "''"),  
 ('the/AT', 'IN'),  
 ('manner/NN', 'AT'),  
 ('in/IN', 'NN'),  
 ('which/WDT', 'IN'),  
 ('the/AT', 'WDT'),  
 ('election/NN', 'AT'),  
 ('was/BEDZ', 'NN'),  
 ('conducted/VBN', 'BEDZ'),  
 ('./.', 'VBN'),  
 ('The/AT', ' '),  
 ('September-October/NP', 'AT'),  
 ('term/NN', 'NP'),  
 ('jury/NN', 'NN'),  
 ('had/HVD', 'NN'),  
 ('been/BEN', 'HVD'),  
 ('charged/VBN', 'BEN'),  
 ('by/IN', 'VBN'),  
 ('Fulton/NP-TL', 'IN'),  
 ('Superior/JJ-TL', 'NP-TL'),  
 ('Court/NN-TL', 'JJ-TL'),  
 ('Judge/NN-TL', 'NN-TL'),  
 ('Durwood/NP', 'NN-TL'),  
 ('Pye/NP', 'NP'),  
 ('to/TO', 'NP'),  
 ('investigate/VB', 'TO'),  
 ('reports/NNS', 'VB'),  
 ('of/IN', 'NNS'),  
 ('possible/JJ', 'IN'),  
 ('``/``', 'JJ'),  
 ('irregularities/NNS', '``'),  
 ("''/'", 'NNS'),  
 ('in/IN', "''"),  
 ('the/AT', 'IN'),  
 ('hard-fought/JJ', 'AT'),  
 ('primary/NN', 'JJ'),  
 ('which/WDT', 'NN'),  
 ('was/BEDZ', 'WDT'),  
 ('won/VBN', 'BEDZ'),  
 ('by/IN', 'VBN'),



('Mayor-nominate/NN-TL', 'IN'),  
 ('Ivan/NP', 'NN-TL'),  
 ('Allen/NP', 'NP'),  
 ('Jr./NP', 'NP'),  
 ('./.', 'NP'),  
 ('``/``', '.'),  
 ('Only/RB', '``'),  
 ('a/AT', 'RB'),  
 ('relative/JJ', 'AT'),  
 ('handful/NN', 'JJ'),  
 ('of/IN', 'NN'),  
 ('such/JJ', 'IN'),  
 ('reports/NNS', 'JJ'),  
 ('was/BEDZ', 'NNS'),  
 ('received/VBN', 'BEDZ'),  
 ('''/''', 'VBN'),  
 ('/,', '``'),  
 ('the/AT', ','),  
 ('jury/NN', 'AT'),  
 ('said/VBD', 'NN'),  
 ('/,', 'VBD'),  
 ('``/``', ','),  
 ('considering/IN', '``'),  
 ('the/AT', 'IN'),  
 ('widespread/JJ', 'AT'),  
 ('interest/NN', 'JJ'),  
 ('in/IN', 'NN'),  
 ('the/AT', 'IN'),  
 ('election/NN', 'AT'),  
 ('/,', 'NN'),  
 ('the/AT', ','),  
 ('number/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('voters/NNS', 'IN'),  
 ('and/CC', 'NNS'),  
 ('the/AT', 'CC'),  
 ('size/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('this/DT', 'IN'),  
 ('city/NN', 'DT'),  
 ('''/''', 'NN'),  
 ('./.', '``'),  
 ('The/AT', '.'),  
 ('jury/NN', 'AT'),  
 ('said/VBD', 'NN'),  
 ('it/PPS', 'VBD'),  
 ('did/DOD', 'PPS'),

('find/VB', 'DOD'),  
 ('that/CS', 'VB'),  
 ('many/AP', 'CS'),  
 ('of/IN', 'AP'),  
 ("Georgia's/NP\$", 'IN'),  
 ('registration/NN', 'NP\$'),  
 ('and/CC', 'NN'),  
 ('election/NN', 'CC'),  
 ('laws/NNS', 'NN'),  
 ('``/``', 'NNS'),  
 ('are/BER', '``'),  
 ('outmoded/JJ', 'BER'),  
 ('or/CC', 'JJ'),  
 ('inadequate/JJ', 'CC'),  
 ('and/CC', 'JJ'),  
 ('often/RB', 'CC'),  
 ('ambiguous/JJ', 'RB'),  
 ("''/''", 'JJ'),  
 ('./.', '``'),  
 ('It/PPS', '.'),  
 ('recommended/VBD', 'PPS'),  
 ('that/CS', 'VBD'),  
 ('Fulton/NP', 'CS'),  
 ('legislators/NNS', 'NP'),  
 ('act/VB', 'NNS'),  
 ('``/``', 'VB'),  
 ('to/TO', '``'),  
 ('have/HV', 'TO'),  
 ('these/DTS', 'HV'),  
 ('laws/NNS', 'DTS'),  
 ('studied/VBN', 'NNS'),  
 ('and/CC', 'VBN'),  
 ('revised/VBN', 'CC'),  
 ('to/IN', 'VBN'),  
 ('the/AT', 'IN'),  
 ('end/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('modernizing/VBG', 'IN'),  
 ('and/CC', 'VBG'),  
 ('improving/VBG', 'CC'),  
 ('them/PP0', 'VBG'),  
 ("''/''", 'PP0'),  
 ('./.', '``'),  
 ('The/AT', '.'),  
 ('grand/JJ', 'AT'),  
 ('jury/NN', 'JJ'),  
 ('commented/VBD', 'NN'),

('on/IN', 'VBD'),  
 ('a/AT', 'IN'),  
 ('number/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('other/AP', 'IN'),  
 ('topics/NNS', 'AP'),  
 ('/,', 'NNS'),  
 ('among/IN', ','),  
 ('them/PPO', 'IN'),  
 ('the/AT', 'PPO'),  
 ('Atlanta/NP', 'AT'),  
 ('and/CC', 'NP'),  
 ('Fulton/NP-TL', 'CC'),  
 ('County/NN-TL', 'NP-TL'),  
 ('purchasing/VBG', 'NN-TL'),  
 ('departments/NNS', 'VBG'),  
 ('which/WDT', 'NNS'),  
 ('it/PPS', 'WDT'),  
 ('said/VBD', 'PPS'),  
 ('``/``', 'VBD'),  
 ('are/BER', '``'),  
 ('well/QL', 'BER'),  
 ('operated/VBN', 'QL'),  
 ('and/CC', 'VBN'),  
 ('follow/VB', 'CC'),  
 ('generally/RB', 'VB'),  
 ('accepted/VBN', 'RB'),  
 ('practices/NNS', 'VBN'),  
 ('which/WDT', 'NNS'),  
 ('inure/VB', 'WDT'),  
 ('to/IN', 'VB'),  
 ('the/AT', 'IN'),  
 ('best/JJT', 'AT'),  
 ('interest/NN', 'JJT'),  
 ('of/IN', 'NN'),  
 ('both/ABX', 'IN'),  
 ('governments/NNS', 'ABX'),  
 ('''/''', 'NNS'),  
 ('./.', '''),  
 ('Merger/NN-HL', '.'),  
 ('proposed/VBN-HL', 'NN-HL'),  
 ('However/WRB', 'VBN-HL'),  
 ('/,', 'WRB'),  
 ('the/AT', ','),  
 ('jury/NN', 'AT'),  
 ('said/VBD', 'NN'),  
 ('it/PPS', 'VBD'),

('believes/VBZ', 'PPS'),  
 ('``/``', 'VBZ'),  
 ('these/DTS', '``'),  
 ('two/CD', 'DTS'),  
 ('offices/NNS', 'CD'),  
 ('should/MD', 'NNS'),  
 ('be/BE', 'MD'),  
 ('combined/VBN', 'BE'),  
 ('to/TO', 'VBN'),  
 ('achieve/VB', 'TO'),  
 ('greater/JJR', 'VB'),  
 ('efficiency/NN', 'JJR'),  
 ('and/CC', 'NN'),  
 ('reduce/VB', 'CC'),  
 ('the/AT', 'VB'),  
 ('cost/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('administration/NN', 'IN'),  
 ('''/''', 'NN'),  
 ('./.', '``'),  
 ('The/AT', '``'),  
 ('City/NN-TL', 'AT'),  
 ('Purchasing/VBG-TL', 'NN-TL'),  
 ('Department/NN-TL', 'VBG-TL'),  
 ('/,', 'NN-TL'),  
 ('the/AT', '`,`'),  
 ('jury/NN', 'AT'),  
 ('said/VBD', 'NN'),  
 ('/,', 'VBD'),  
 ('``/``', '`,`'),  
 ('is/BEZ', '``'),  
 ('lacking/VBG', 'BEZ'),  
 ('in/IN', 'VBG'),  
 ('experienced/VBN', 'IN'),  
 ('clerical/JJ', 'VBN'),  
 ('personnel/NNS', 'JJ'),  
 ('as/CS', 'NNS'),  
 ('a/AT', 'CS'),  
 ('result/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('city/NN', 'IN'),  
 ('personnel/NNS', 'NN'),  
 ('policies/NNS', 'NNS'),  
 ('''/''', 'NNS'),  
 ('./.', '``'),  
 ('It/PPS', '``'),  
 ('urged/VBD', 'PPS'),

('that/CS', 'VBD'),  
 ('the/AT', 'CS'),  
 ('city/NN', 'AT'),  
 ('``/``', 'NN'),  
 ('take/VB', '``'),  
 ('steps/NNS', 'VB'),  
 ('to/TO', 'NNS'),  
 ('remedy/VB', 'TO'),  
 ('''/''', 'VB'),  
 ('this/DT', '''),  
 ('problem/NN', 'DT'),  
 ('./.', 'NN'),  
 ('Implementation/NN', '.'),  
 ('of/IN', 'NN'),  
 ('Georgia's/NP\$', 'IN'),  
 ('automobile/NN', 'NP\$'),  
 ('title/NN', 'NN'),  
 ('law/NN', 'NN'),  
 ('was/BEDZ', 'NN'),  
 ('also/RB', 'BEDZ'),  
 ('recommended/VBN', 'RB'),  
 ('by/IN', 'VBN'),  
 ('the/AT', 'IN'),  
 ('outgoing/JJ', 'AT'),  
 ('jury/NN', 'JJ'),  
 ('./.', 'NN'),  
 ('It/PPS', '.'),  
 ('urged/VBD', 'PPS'),  
 ('that/CS', 'VBD'),  
 ('the/AT', 'CS'),  
 ('next/AP', 'AT'),  
 ('Legislature/NN-TL', 'AP'),  
 ('``/``', 'NN-TL'),  
 ('provide/VB', '``'),  
 ('enabling/VBG', 'VB'),  
 ('funds/NNS', 'VBG'),  
 ('and/CC', 'NNS'),  
 ('re-set/VB', 'CC'),  
 ('the/AT', 'VB'),  
 ('effective/JJ', 'AT'),  
 ('date/NN', 'JJ'),  
 ('so/CS', 'NN'),  
 ('that/CS', 'CS'),  
 ('an/AT', 'CS'),  
 ('orderly/JJ', 'AT'),  
 ('implementation/NN', 'JJ'),  
 ('of/IN', 'NN'),

('the/AT', 'IN'),  
 ('law/NN', 'AT'),  
 ('may/MD', 'NN'),  
 ('be/BE', 'MD'),  
 ('effected/VBN', 'BE'),  
 ("''/'", 'VBN'),  
 ('./.', ' '),  
 ('The/AT', '.'),  
 ('grand/JJ', 'AT'),  
 ('jury/NN', 'JJ'),  
 ('took/VBD', 'NN'),  
 ('a/AT', 'VBD'),  
 ('swipe/NN', 'AT'),  
 ('at/IN', 'NN'),  
 ('the/AT', 'IN'),  
 ('State/NN-TL', 'AT'),  
 ('Welfare/NN-TL', 'NN-TL'),  
 ("Department's/NN\$-TL", 'NN-TL'),  
 ('handling/NN', 'NN\$-TL'),  
 ('of/IN', 'NN'),  
 ('federal/JJ', 'IN'),  
 ('funds/NNS', 'JJ'),  
 ('granted/VBN', 'NNS'),  
 ('for/IN', 'VBN'),  
 ('child/NN', 'IN'),  
 ('welfare/NN', 'NN'),  
 ('services/NNS', 'NN'),  
 ('in/IN', 'NNS'),  
 ('foster/JJ', 'IN'),  
 ('homes/NNS', 'JJ'),  
 ('./.', 'NNS'),  
 ('``/``', '.'),  
 ('This/DT', '``'),  
 ('is/BEZ', 'DT'),  
 ('one/CD', 'BEZ'),  
 ('of/IN', 'CD'),  
 ('the/AT', 'IN'),  
 ('major/JJ', 'AT'),  
 ('items/NNS', 'JJ'),  
 ('in/IN', 'NNS'),  
 ('the/AT', 'IN'),  
 ('Fulton/NP-TL', 'AT'),  
 ('County/NN-TL', 'NP-TL'),  
 ('general/JJ', 'NN-TL'),  
 ('assistance/NN', 'JJ'),  
 ('program/NN', 'NN'),  
 ("''/'", 'NN'),

('/,', ' '),  
 ('the/AT', ' '),  
 ('jury/NN', 'AT'),  
 ('said/VBD', 'NN'),  
 ('/,', 'VBD'),  
 ('but/CC', ' '),  
 ('the/AT', 'CC'),  
 ('State/NN-TL', 'AT'),  
 ('Welfare/NN-TL', 'NN-TL'),  
 ('Department/NN-TL', 'NN-TL'),  
 ('``/``', 'NN-TL'),  
 ('has/HVZ', '``'),  
 ('seen/VBN', 'HVZ'),  
 ('fit/JJ', 'VBN'),  
 ('to/TO', 'JJ'),  
 ('distribute/VB', 'TO'),  
 ('these/DTS', 'VB'),  
 ('funds/NNS', 'DTS'),  
 ('through/IN', 'NNS'),  
 ('the/AT', 'IN'),  
 ('welfare/NN', 'AT'),  
 ('departments/NNS', 'NN'),  
 ('of/IN', 'NNS'),  
 ('all/ABN', 'IN'),  
 ('the/AT', 'ABN'),  
 ('counties/NNS', 'AT'),  
 ('in/IN', 'NNS'),  
 ('the/AT', 'IN'),  
 ('state/NN', 'AT'),  
 ('with/IN', 'NN'),  
 ('the/AT', 'IN'),  
 ('exception/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('Fulton/NP-TL', 'IN'),  
 ('County/NN-TL', 'NP-TL'),  
 ('/,', 'NN-TL'),  
 ('which/WDT', ' '),  
 ('receives/VBZ', 'WDT'),  
 ('none/PN', 'VBZ'),  
 ('of/IN', 'PN'),  
 ('this/DT', 'IN'),  
 ('money/NN', 'DT'),  
 ('./.', 'NN'),  
 ('The/AT', ' '),  
 ('jurors/NNS', 'AT'),  
 ('said/VBD', 'NNS'),  
 ('they/PPSS', 'VBD'),

('realize/VB', 'PPSS'),  
 ('``/``', 'VB'),  
 ('a/AT', '``'),  
 ('proportionate/JJ', 'AT'),  
 ('distribution/NN', 'JJ'),  
 ('of/IN', 'NN'),  
 ('these/DTS', 'IN'),  
 ('funds/NNS', 'DTS'),  
 ('might/MD', 'NNS'),  
 ('disable/VB', 'MD'),  
 ('this/DT', 'VB'),  
 ('program/NN', 'DT'),  
 ('in/IN', 'NN'),  
 ('our/PP\$', 'IN'),  
 ('less/QL', 'PP\$'),  
 ('populous/JJ', 'QL'),  
 ('counties/NNS', 'JJ'),  
 ('''/''', 'NNS'),  
 ('./.', '``'),  
 ('Nevertheless/RB', '.'),  
 ('/,', 'RB'),  
 ('``/``', ','),  
 ('we/PPSS', '``'),  
 ('feel/VB', 'PPSS'),  
 ('that/CS', 'VB'),  
 ('in/IN', 'CS'),  
 ('the/AT', 'IN'),  
 ('future/NN', 'AT'),  
 ('Fulton/NP-TL', 'NN'),  
 ('County/NN-TL', 'NP-TL'),  
 ('should/MD', 'NN-TL'),  
 ('receive/VB', 'MD'),  
 ('some/DTI', 'VB'),  
 ('portion/NN', 'DTI'),  
 ('of/IN', 'NN'),  
 ('these/DTS', 'IN'),  
 ('available/JJ', 'DTS'),  
 ('funds/NNS', 'JJ'),  
 ('''/''', 'NNS'),  
 ('/,', '``'),  
 ('the/AT', ','),  
 ('jurors/NNS', 'AT'),  
 ('said/VBD', 'NNS'),  
 ('./.', 'VBD'),  
 ('``/``', '.'),  
 ('Failure/NN', '``'),  
 ('to/TO', 'NN'),



('do/DO', 'TO'),  
 ('this/DT', 'DO'),  
 ('will/MD', 'DT'),  
 ('continue/VB', 'MD'),  
 ('to/TO', 'VB'),  
 ('place/VB', 'TO'),  
 ('a/AT', 'VB'),  
 ('disproportionate/JJ', 'AT'),  
 ('burden/NN', 'JJ'),  
 ('''/''', 'NN'),  
 ('on/IN', '''),  
 ('Fulton/NP', 'IN'),  
 ('taxpayers/NNS', 'NP'),  
 ('./.', 'NNS'),  
 ('The/AT', '.'),  
 ('jury/NN', 'AT'),  
 ('also/RB', 'NN'),  
 ('commented/VBD', 'RB'),  
 ('on/IN', 'VBD'),  
 ('the/AT', 'IN'),  
 ('Fulton/NP', 'AT'),  
 ('ordinary's/NN\$', 'NP'),  
 ('court/NN', 'NN\$'),  
 ('which/WD', 'NN'),  
 ('has/HVZ', 'WD'),  
 ('been/BEN', 'HVZ'),  
 ('under/IN', 'BEN'),  
 ('fire/NN', 'IN'),  
 ('for/IN', 'NN'),  
 ('its/PP\$', 'IN'),  
 ('practices/NNS', 'PP\$'),  
 ('in/IN', 'NNS'),  
 ('the/AT', 'IN'),  
 ('appointment/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('appraisers/NNS', 'IN'),  
 ('./.', 'NNS'),  
 ('guardians/NNS', ','),  
 ('and/CC', 'NNS'),  
 ('administrators/NNS', 'CC'),  
 ('and/CC', 'NNS'),  
 ('the/AT', 'CC'),  
 ('awarding/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('fees/NNS', 'IN'),  
 ('and/CC', 'NNS'),  
 ('compensation/NN', 'CC'),

('./.', 'NN'),  
 ('Wards/NNS-HL', '.'),  
 ('protected/VBN-HL', 'NNS-HL'),  
 ('The/AT', 'VBN-HL'),  
 ('jury/NN', 'AT'),  
 ('said/VBD', 'NN'),  
 ('it/PPS', 'VBD'),  
 ('found/VBD', 'PPS'),  
 ('the/AT', 'VBD'),  
 ('court/NN', 'AT'),  
 ('``/``', 'NN'),  
 ('has/HVZ', '``'),  
 ('incorporated/VBN', 'HVZ'),  
 ('into/IN', 'VBN'),  
 ('its/PP\$', 'IN'),  
 ('operating/VBG', 'PP\$'),  
 ('procedures/NNS', 'VBG'),  
 ('the/AT', 'NNS'),  
 ('recommendations/NNS', 'AT'),  
 ('''/''', 'NNS'),  
 ('of/IN', '''),  
 ('two/CD', 'IN'),  
 ('previous/JJ', 'CD'),  
 ('grand/JJ', 'JJ'),  
 ('juries/NNS', 'JJ'),  
 ('/,', 'NNS'),  
 ('the/AT', ','),  
 ('Atlanta/NP-TL', 'AT'),  
 ('Bar/NN-TL', 'NP-TL'),  
 ('Association/NN-TL', 'NN-TL'),  
 ('and/CC', 'NN-TL'),  
 ('an/AT', 'CC'),  
 ('interim/NN', 'AT'),  
 ('citizens/NNS', 'NN'),  
 ('committee/NN', 'NNS'),  
 ('./.', 'NN'),  
 ('``/``', '.'),  
 ('These/DTS', '``'),  
 ('actions/NNS', 'DTS'),  
 ('should/MD', 'NNS'),  
 ('serve/VB', 'MD'),  
 ('to/TO', 'VB'),  
 ('protect/VB', 'TO'),  
 ('in/IN', 'VB'),  
 ('fact/NN', 'IN'),  
 ('and/CC', 'NN'),  
 ('in/IN', 'CC'),

('effect/NN', 'IN'),  
 ('the/AT', 'NN'),  
 ("court's/NN\$", 'AT'),  
 ('wards/NNS', 'NN\$'),  
 ('from/IN', 'NNS'),  
 ('undue/JJ', 'IN'),  
 ('costs/NNS', 'JJ'),  
 ('and/CC', 'NNS'),  
 ('its/PP\$', 'CC'),  
 ('appointed/VBN', 'PP\$'),  
 ('and/CC', 'VBN'),  
 ('elected/VBN', 'CC'),  
 ('servants/NNS', 'VBN'),  
 ('from/IN', 'NNS'),  
 ('unmeritorious/JJ', 'IN'),  
 ('criticisms/NNS', 'JJ'),  
 ("''/'", 'NNS'),  
 ('/,', '""'),  
 ('the/AT', ','),  
 ('jury/NN', 'AT'),  
 ('said/VBD', 'NN'),  
 ('./.', 'VBD'),  
 ('Regarding/IN', '.'),  
 ("Atlanta's/NP\$", 'IN'),  
 ('new/JJ', 'NP\$'),  
 ('multi-million-dollar/JJ', 'JJ'),  
 ('airport/NN', 'JJ'),  
 ('/,', 'NN'),  
 ('the/AT', ','),  
 ('jury/NN', 'AT'),  
 ('recommended/VBD', 'NN'),  
 ('``/``', 'VBD'),  
 ('that/CS', '``'),  
 ('when/WRB', 'CS'),  
 ('the/AT', 'WRB'),  
 ('new/JJ', 'AT'),  
 ('management/NN', 'JJ'),  
 ('takes/VBZ', 'NN'),  
 ('charge/NN', 'VBZ'),  
 ('Jan./NP', 'NN'),  
 ('1/CD', 'NP'),  
 ('the/AT', 'CD'),  
 ('airport/NN', 'AT'),  
 ('be/BE', 'NN'),  
 ('operated/VBN', 'BE'),  
 ('in/IN', 'VBN'),  
 ('a/AT', 'IN'),

('manner/NN', 'AT'),  
 ('that/WPS', 'NN'),  
 ('will/MD', 'WPS'),  
 ('eliminate/VB', 'MD'),  
 ('political/JJ', 'VB'),  
 ('influences/NNS', 'JJ'),  
 ("''/'", 'NNS'),  
 ('./.', ' '),  
 ('The/AT', '.'),  
 ('jury/NN', 'AT'),  
 ('did/DOD', 'NN'),  
 ('not/\*', 'DOD'),  
 ('elaborate/VB', '\*'),  
 ('/,', 'VB'),  
 ('but/CC', ','),  
 ('it/PPS', 'CC'),  
 ('added/VBD', 'PPS'),  
 ('that/CS', 'VBD'),  
 ('``/``', 'CS'),  
 ('there/EX', '``'),  
 ('should/MD', 'EX'),  
 ('be/BE', 'MD'),  
 ('periodic/JJ', 'BE'),  
 ('surveillance/NN', 'JJ'),  
 ('of/IN', 'NN'),  
 ('the/AT', 'IN'),  
 ('pricing/VBG', 'AT'),  
 ('practices/NNS', 'VBG'),  
 ('of/IN', 'NNS'),  
 ('the/AT', 'IN'),  
 ('concessionaires/NNS', 'AT'),  
 ('for/IN', 'NNS'),  
 ('the/AT', 'IN'),  
 ('purpose/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('keeping/VBG', 'IN'),  
 ('the/AT', 'VBG'),  
 ('prices/NNS', 'AT'),  
 ('reasonable/JJ', 'NNS'),  
 ("''/'", 'JJ'),  
 ('./.', ' '),  
 ('Ask/VB-HL', '.'),  
 ('jail/NN-HL', 'VB-HL'),  
 ('deputies/NNS-HL', 'NN-HL'),  
 ('On/IN', 'NNS-HL'),  
 ('other/AP', 'IN'),  
 ('matters/NNS', 'AP'),

('/,', 'NNS'),  
 ('the/AT', ','),  
 ('jury/NN', 'AT'),  
 ('recommended/VBD', 'NN'),  
 ('that/CS', 'VBD'),  
 (':/:', 'CS'),  
 ('/((', ':'),  
 ('1/CD', '('),  
 (')/)', 'CD'),  
 ('Four/CD', ')'),  
 ('additional/JJ', 'CD'),  
 ('deputies/NNS', 'JJ'),  
 ('be/BE', 'NNS'),  
 ('employed/VBN', 'BE'),  
 ('at/IN', 'VBN'),  
 ('the/AT', 'IN'),  
 ('Fulton/NP-TL', 'AT'),  
 ('County/NN-TL', 'NP-TL'),  
 ('Jail/NN-TL', 'NN-TL'),  
 ('and/CC', 'NN-TL'),  
 ('`/``', 'CC'),  
 ('a/AT', '``'),  
 ('doctor/NN', 'AT'),  
 ('/,', 'NN'),  
 ('medical/JJ', ','),  
 ('intern/NN', 'JJ'),  
 ('or/CC', 'NN'),  
 ('extern/NN', 'CC'),  
 ('be/BE', 'NN'),  
 ('employed/VBN', 'BE'),  
 ('for/IN', 'VBN'),  
 ('night/NN', 'IN'),  
 ('and/CC', 'NN'),  
 ('weekend/NN', 'CC'),  
 ('duty/NN', 'NN'),  
 ('at/IN', 'NN'),  
 ('the/AT', 'IN'),  
 ('jail/NN', 'AT'),  
 ('''/''', 'NN'),  
 ('./.', '''),  
 ('/((', '.'),  
 ('2/CD', '('),  
 (')/)', 'CD'),  
 ('Fulton/NP', ')'),  
 ('legislators/NNS', 'NP'),  
 ('`/``', 'NNS'),  
 ('work/VB', '``'),

('with/IN', 'VB'),  
 ('city/NN', 'IN'),  
 ('officials/NNS', 'NN'),  
 ('to/TO', 'NNS'),  
 ('pass/VB', 'TO'),  
 ('enabling/VBG', 'VB'),  
 ('legislation/NN', 'VBG'),  
 ('that/WPS', 'NN'),  
 ('will/MD', 'WPS'),  
 ('permit/VB', 'MD'),  
 ('the/AT', 'VB'),  
 ('establishment/NN', 'AT'),  
 ('of/IN', 'NN'),  
 ('a/AT', 'IN'),  
 ('fair/JJ', 'AT'),  
 ('and/CC', 'JJ'),  
 ('equitable/JJ', 'CC'),  
 ("''/'", 'JJ'),  
 ('pension/NN', "''"),  
 ('plan/NN', 'NN'),  
 ('for/IN', 'NN'),  
 ('city/NN', 'IN'),  
 ('employees/NNS', 'NN'),  
 ('./.', 'NNS'),  
 ('The/AT', '.'),  
 ('jury/NN', 'AT'),  
 ('praised/VBD', 'NN'),  
 ('the/AT', 'VBD'),  
 ('administration/NN', 'AT'),  
 ('and/CC', 'NN'),  
 ('operation/NN', 'CC'),  
 ('of/IN', 'NN'),  
 ('the/AT', 'IN'),  
 ('Atlanta/NP-TL', 'AT'),  
 ('Police/NNS-TL', 'NP-TL'),  
 ('Department/NN-TL', 'NNS-TL'),  
 ('./.', 'NN-TL'),  
 ('the/AT', ','),  
 ('Fulton/NP-TL', 'AT'),  
 ('Tax/NN-TL', 'NP-TL'),  
 ("Commissioner's/NN\$-TL", 'NN-TL'),  
 ('Office/NN-TL', 'NN\$-TL'),  
 ('./.', 'NN-TL'),  
 ('the/AT', ','),  
 ('Bellwood/NP', 'AT'),  
 ('and/CC', 'NP'),  
 ('Alpharetta/NP', 'CC'),

('prison/NN', 'NP'),  
 ('farms/NNS', 'NN'),  
 ('./.', 'NNS'),  
 ('Grady/NP-TL', ','),  
 ('Hospital/NN-TL', 'NP-TL'),  
 ('and/CC', 'NN-TL'),  
 ('the/AT', 'CC'),  
 ('Fulton/NP-TL', 'AT'),  
 ('Health/NN-TL', 'NP-TL'),  
 ('Department/NN-TL', 'NN-TL'),  
 ('./.', 'NN-TL'),  
 ('Mayor/NN-TL', '.'),  
 ('William/NP', 'NN-TL'),  
 ('B./NP', 'NP'),  
 ('Hartsfield/NP', 'NP'),  
 ('filed/VBD', 'NP'),  
 ('suit/NN', 'VBD'),  
 ('for/IN', 'NN'),  
 ('divorce/NN', 'IN'),  
 ('from/IN', 'NN'),  
 ('his/PP\$', 'IN'),  
 ('wife/NN', 'PP\$'),  
 ('./.', 'NN'),  
 ('Pearl/NP', ','),  
 ('Williams/NP', 'NP'),  
 ('Hartsfield/NP', 'NP'),  
 ('./.', 'NP'),  
 ('in/IN', ','),  
 ('Fulton/NP-TL', 'IN'),  
 ('Superior/JJ-TL', 'NP-TL'),  
 ('Court/NN-TL', 'JJ-TL'),  
 ('Friday/NR', 'NN-TL'),  
 ('./.', 'NR'),  
 ('His/PP\$', '.'),  
 ('petition/NN', 'PP\$'),  
 ('charged/VBD', 'NN'),  
 ('mental/JJ', 'VBD'),  
 ('cruelty/NN', 'JJ'),  
 ('./.', 'NN'),  
 ('The/AT', '.'),  
 ('couple/NN', 'AT'),  
 ('was/BEDZ', 'NN'),  
 ('married/VBN', 'BEDZ'),  
 ('Aug./NP', 'VBN'),  
 ('2/CD', 'NP'),  
 ('./.', 'CD'),  
 ('1913/CD', ','),

('./.', 'CD'),  
 ('They/PPSS', '.'),  
 ('have/HV', 'PPSS'),  
 ('a/AT', 'HV'),  
 ('son/NN', 'AT'),  
 ('./.', 'NN'),  
 ('William/NP', ','),  
 ('Berry/NP', 'NP'),  
 ('Jr./NP', 'NP'),  
 ('./.', 'NP'),  
 ('and/CC', ','),  
 ('a/AT', 'CC'),  
 ('daughter/NN', 'AT'),  
 ('./.', 'NN'),  
 ('Mrs./NP', ','),  
 ('J./NP', 'NP'),  
 ('M./NP', 'NP'),  
 ('Cheshire/NP', 'NP'),  
 ('of/IN', 'NP'),  
 ('Griffin/NP', 'IN'),  
 ('./.', 'NP'),  
 ('Attorneys/NNS', '.'),  
 ('for/IN', 'NNS'),  
 ('the/AT', 'IN'),  
 ('mayor/NN', 'AT'),  
 ('said/VBD', 'NN'),  
 ('that/CS', 'VBD'),  
 ('an/AT', 'CS'),  
 ('amicable/JJ', 'AT'),  
 ('property/NN', 'JJ'),  
 ('settlement/NN', 'NN'),  
 ('has/HVZ', 'NN'),  
 ('been/BEN', 'HVZ'),  
 ('agreed/VBN', 'BEN'),  
 ('upon/RB', 'VBN'),  
 ('./.', 'RB'),  
 ('The/AT', '.'),  
 ('petition/NN', 'AT'),  
 ('listed/VBD', 'NN'),  
 ('the/AT', 'VBD'),  
 ('mayor's/NN\$', 'AT'),  
 ('occupation/NN', 'NN\$'),  
 ('as/CS', 'NN'),  
 ('`/``', 'CS'),  
 ('attorney/NN', '``'),  
 ('''/'''', 'NN'),  
 ('and/CC', '''),



('his/PP\$', 'CC'),  
 ('age/NN', 'PP\$'),  
 ('as/CS', 'NN'),  
 ('71/CD', 'CS'),  
 ('./.', 'CD'),  
 ('It/PPS', '.'),  
 ('listed/VBD', 'PPS'),  
 ('his/PP\$', 'VBD'),  
 ("wife's/NN\$", 'PP\$'),  
 ('age/NN', 'NN\$'),  
 ('as/CS', 'NN'),  
 ('74/CD', 'CS'),  
 ('and/CC', 'CD'),  
 ('place/NN', 'CC'),  
 ('of/IN', 'NN'),  
 ('birth/NN', 'IN'),  
 ('as/CS', 'NN'),  
 ('Opelika/NP', 'CS'),  
 ('./.', 'NP'),  
 ('Ala./NP', ','),  
 ('./.', 'NP'),  
 ('The/AT', '.'),  
 ('petition/NN', 'AT'),  
 ('said/VBD', 'NN'),  
 ('that/CS', 'VBD'),  
 ('the/AT', 'CS'),  
 ('couple/NN', 'AT'),  
 ('has/HVZ', 'NN'),  
 ('not/\*', 'HVZ'),  
 ('lived/VBN', '\*'),  
 ('together/RB', 'VBN'),  
 ('as/CS', 'RB'),  
 ('man/NN', 'CS'),  
 ('and/CC', 'NN'),  
 ('wife/NN', 'CC'),  
 ('for/IN', 'NN'),  
 ('more/AP', 'IN'),  
 ('than/IN', 'AP'),  
 ('a/AT', 'IN'),  
 ('year/NN', 'AT'),  
 ('./.', 'NN'),  
 ('The/AT', '.'),  
 ('Hartsfield/NP', 'AT'),  
 ('home/NR', 'NP'),  
 ('is/BEZ', 'NR'),  
 ('at/IN', 'BEZ'),  
 ('637/CD', 'IN'),

('E./NP', 'CD'),  
 ('Pelham/NP', 'NP'),  
 ('Rd./NN-TL', 'NP'),  
 ('Aj/NN', 'NN-TL'),  
 ('./.', 'NN'),  
 ('Henry/NP', '.'),  
 ('L./NP', 'NP'),  
 ('Bowden/NP', 'NP'),  
 ('was/BEDZ', 'NP'),  
 ('listed/VBN', 'BEDZ'),  
 ('on/IN', 'VBN'),  
 ('the/AT', 'IN'),  
 ('petition/NN', 'AT'),  
 ('as/CS', 'NN'),  
 ('the/AT', 'CS'),  
 ("mayor's/NN\$", 'AT'),  
 ('attorney/NN', 'NN\$'),  
 ('./.', 'NN'),  
 ('Hartsfield/NP', '.'),  
 ('has/HVZ', 'NP'),  
 ('been/BEN', 'HVZ'),  
 ('mayor/NN', 'BEN'),  
 ('of/IN', 'NN'),  
 ('Atlanta/NP', 'IN'),  
 ('/,', 'NP'),  
 ('with/IN', ','),  
 ('exception/NN', 'IN'),  
 ('of/IN', 'NN'),  
 ('one/CD', 'IN'),  
 ('brief/JJ', 'CD'),  
 ('interlude/NN', 'JJ'),  
 ('/,', 'NN'),  
 ('since/IN', ','),  
 ('1937/CD', 'IN'),  
 ('./.', 'CD'),  
 ('His/PP\$', '.'),  
 ('political/JJ', 'PP\$'),  
 ('career/NN', 'JJ'),  
 ('goes/VBZ', 'NN'),  
 ('back/RB', 'VBZ'),  
 ('to/IN', 'RB'),  
 ('his/PP\$', 'IN'),  
 ('election/NN', 'PP\$'),  
 ('to/IN', 'NN'),  
 ('city/NN', 'IN'),  
 ('council/NN', 'NN'),  
 ('in/IN', 'NN'),

('1923/CD', 'IN'),  
 ('./.', 'CD'),  
 ('The/AT', '.'),  
 ("mayor's/NN\$", 'AT'),  
 ('present/JJ', 'NN\$'),  
 ('term/NN', 'JJ'),  
 ('of/IN', 'NN'),  
 ('office/NN', 'IN'),  
 ('expires/VBZ', 'NN'),  
 ('Jan./NP', 'VBZ'),  
 ('1/CD', 'NP'),  
 ('./.', 'CD'),  
 ('He/PPS', '.'),  
 ('will/MD', 'PPS'),  
 ('be/BE', 'MD'),  
 ('succeeded/VBN', 'BE'),  
 ('by/IN', 'VBN'),  
 ('Ivan/NP', 'IN'),  
 ('Allen/NP', 'NP'),  
 ('Jr./NP', 'NP'),  
 ('/,', 'NP'),  
 ('who/WPS', ','),  
 ('became/VBD', 'WPS'),  
 ('a/AT', 'VBD'),  
 ('candidate/NN', 'AT'),  
 ('in/IN', 'NN'),  
 ('the/AT', 'IN'),  
 ('Sept./NP', 'AT'),  
 ('13/CD', 'NP'),  
 ('primary/NN', 'CD'),  
 ('after/CS', 'NN'),  
 ('Mayor/NN-TL', 'CS'),  
 ('Hartsfield/NP', 'NN-TL'),  
 ('announced/VBD', 'NP'),  
 ('that/CS', 'VBD'),  
 ('he/PPS', 'CS'),  
 ('would/MD', 'PPS'),  
 ('not/\*', 'MD'),  
 ('run/VB', '\*'),  
 ('for/IN', 'VB'),  
 ('reelection/NN', 'IN'),  
 ('./.', 'NN'),  
 ('Georgia/NP', '.'),  
 ('Republicans/NPS', 'NP'),  
 ('are/BER', 'NPS'),  
 ('getting/VBG', 'BER'),  
 ('strong/JJ', 'VBG'),

```
( 'encouragement/NN', 'JJ'),
( 'to/TO', 'NN'),
( 'enter/VB', 'TO'),
( 'a/AT', 'VB'),
( 'candidate/NN', 'AT'),
( 'in/IN', 'NN'),
( 'the/AT', 'IN'),
( '1962/CD', 'AT'),
( "governor's/NN$", 'CD'),
( 'race/NN', 'NN$'),
...]
```

d. Can you find any POSn-1 that favors NN over JJ for the following word ‘cold’?

```
[28]: br_pos_confrqdis['cold/NN'].most_common()
```

```
[28]: [('AT', 4), ('JJ', 2), ('', 1), ('DT', 1)]
```

```
[29]: br_pos_confrqdis['cold/JJ'].most_common()
```

```
[29]: [('AT', 38),
( 'IN', 14),
( 'CC', 8),
( 'QL', 7),
( 'BEDZ', 7),
( 'JJ', 4),
( 'DT', 3),
( '', 3),
( 'PP$', 3),
( '``', 2),
( 'NN', 2),
( 'VBN', 2),
( 'VBD', 2),
( 'CS', 1),
( 'BEZ', 1),
( 'DOZ', 1),
( 'RB', 1),
( 'PPSS', 1),
( 'BE', 1),
( 'VB', 1),
( 'VBZ', 1),
( 'NP$', 1),
( 'BEDZ*', 1),
( '--', 1),
( 'RP', 1),
( 'DTI', 1),
( 'WRB', 1),
( 'BED', 1)]
```

6. Based on what you found, how is your bigram tagger expected to tag ‘cold’ in the following sentences?

```
[30]: bi_tagger = nltk.BigramTagger(br_train)
```

a. I was very cold.

```
[31]: txt = word_tokenize("I was very cold.")
      bi_tagger.tag(txt)
```

```
[31]: [('I', 'PPSS'), ('was', 'BEDZ'), ('very', 'QL'), ('cold', 'JJ'), ('.', '.')] 
```

b. I had a cold.

```
[32]: txt1 = word_tokenize("I had cold.")
      bi_tagger.tag(txt1)
```

```
[32]: [('I', 'PPSS'), ('had', 'HVD'), ('cold', None), ('.', None)] 
```

c. I had a severe cold.

```
[33]: txt2 = word_tokenize("I had a severe cold.")
      bi_tagger.tag(txt2)
```

```
[33]: [('I', 'PPSS'),
      ('had', 'HVD'),
      ('a', 'AT'),
      ('severe', 'JJ'),
      ('cold', 'JJ'),
      ('.', '.')] 
```

d. January was a cold month.

```
[34]: txt3 = word_tokenize("January was a cold month")
      bi_tagger.tag(txt3)
```

```
[34]: [('January', None),
      ('was', None),
      ('a', None),
      ('cold', None),
      ('month', None)] 
```

8. Have the tagger tag the following sentences, all of which contain the word ‘so’:

a. I failed to do so.

```
[36]: txt4 = word_tokenize("I failed to do so")
      bi_tagger.tag(txt4)
```

```
[36]: [('I', 'PPSS'), ('failed', 'VBD'), ('to', 'TO'), ('do', 'DO'), ('so', 'RB')] 
```

b. I was happy, but so was my enemy.

```
[37]: txt5 = word_tokenize("I was happy,but so was my enemy")
      bi_tagger.tag(txt5)
```

```
[37]: [('I', 'PPSS'),
      ('was', 'BEDZ'),
      ('happy', 'JJ'),
      (',', ', ', '),
      ('but', 'CC'),
      ('so', 'RB'),
      ('was', 'BEDZ'),
      ('my', 'PP$'),
      ('enemy', 'NN')]
```

c. So, how was the exam?

```
[38]: txt6 = word_tokenize("So, how was the exam?")
      bi_tagger.tag(txt6)
```

```
[38]: [('So', 'RB'),
      (',', ', ', '),
      ('how', 'WRB'),
      ('was', 'BEDZ'),
      ('the', 'AT'),
      ('exam', None),
      ('?', None)]
```

d. The students came in early so they can get good seats.

```
[39]: txt7 = word_tokenize("The students came in early so they can get good seats")
      bi_tagger.tag(txt7)
```

```
[39]: [('The', 'AT'),
      ('students', 'NNS'),
      ('came', 'VBD'),
      ('in', 'IN'),
      ('early', 'JJ'),
      ('so', 'CS'),
      ('they', 'PPSS'),
      ('can', 'MD'),
      ('get', 'VB'),
      ('good', 'JJ'),
      ('seats', 'NNS')]
```

e. She failed the exam, so she must take it again.

```
[40]: txt8 = word_tokenize("She failed the exam, so she must take it again")
      bi_tagger.tag(txt8)
```

```
[40]: [('She', 'PPS'),
      ('failed', 'VBD'),
```

```
('the', 'AT'),  
('exam', None),  
(',', None),  
('so', None),  
('she', None),  
('must', None),  
('take', None),  
('it', None),  
('again', None)]
```

**f. That was so incredible.**

```
[41]: txt9 = word_tokenize("That was so incredible")  
      bi_tagger.tag(txt9)
```

```
[41]: [('That', 'DT'), ('was', 'BEDZ'), ('so', 'QL'), ('incredible', 'JJ')]
```

**g. Wow, so incredible.**

```
[42]: txt10 = word_tokenize("Wow, so incredible")  
      bi_tagger.tag(txt10)
```

```
[42]: [('Wow', None), (',', None), ('so', None), ('incredible', None)]
```