DHRUV GARG 16 BCE 1190

NLP DIGITAL ASSIGNMENT 2

Question: To take 10 sentences and perform POS Tagging, Shallow Parsing and Named entity recognition using various tools

SENTENCE 1: *European authorities fined Google a record $5.1 billion on Wednesday for abusing its power in the mobile phone market and ordered the company to alter its practices.*

Spacy tagging and NER

```
TOKEN, LEMMA, POS, TAG, DEP
European european ADJ JJ amod
authorities authority NOUN NNS nsubj
fined fin VERB VBD ROOT
Google Google PROPN NNP dobj
a a DET DT det
record record NOUN NN npadvmod
$ $ SYM $ quantmod
5.1 5.1 NUM CD compound
billion billion NUM CD nummod
on on ADP IN prep
Wednesday Wednesday PROPN NNP pobj
for for ADP IN prep
abusing abuse VERB VBG pcomp
its -PRON- DET PRP$ poss
power power NOUN NN dobj
in in ADP IN prep
the the DET DT det
mobile mobile ADJ JJ amod
phone phone NOUN NN compound
market market NOUN NN pobj
and and CCONJ CC cc
ordered order VERB VBD conj
the the DET DT det
company company NOUN NN dobj
to to PART TO aux
alter alter VERB VB xcomp
its -PRON- DET PRP$ poss
practices practice NOUN NNS dobj
. . PUNCT . punct

European NORP
Google ORG
$5.1 billion MONEY
Wednesday DATE
```

NLTK Parsing

```
(S
  European/JJ
  authorities/NNS
  fined/VBD
  Google/NNP
  (NP a/DT record/NN)
  $/$
  5.1/CD
  billion/CD
  on/IN
  Wednesday/NNP
  for/IN
  abusing/VBG
  its/PRP$
  (NP power/NN)
  in/IN
  (NP the/DT mobile/JJ phone/NN)
  (NP market/NN)
  and/CC
  ordered/VBD
  (NP the/DT company/NN)
  to/TO
  alter/VB
  its/PRP$
  practices/NNS)
```

NLTK Named Entity

```
(S
  (GPE European/JJ)
  authorities/NNS
  fined/VBD
  (PERSON Google/NNP)
  a/DT
  record/NN
  $/$
  5.1/CD
  billion/CD
  on/IN
  Wednesday/NNP
  for/IN
  abusing/VBG
  its/PRP$
  power/NN
  in/IN
  the/DT
  mobile/JJ
  phone/NN
  market/NN
  and/CC
  ordered/VBD
  the/DT
  company/NN
  to/TO
  alter/VB
  its/PRP$
  practices/NNS)
```

Stanford core NLP

Part-of-Speech:

	JJ	NNS	VBN	NNP	DT	NN	$	CD	CD	IN	NNP	IN	VBG	PRP$	NN	IN	DT	JJ	NN	NN	CC	VBD	DT	NN	TO	VB	PRP$	NNS	.
1	European	authorities	fined	Google	a	record	$	5.1	billion	on	Wednesday	for	abusing	its	power	in	the	mobile	phone	market	and	ordered	the	company	to	alter	its	practices	.

Named Entity Recognition:

	NATIONALITY			ORGANIZATION			MONEY $5.1E9			DATE 2019-04-03									TITLE										
1	European	authorities	fined	Google	a	record	$	5.1	billion	on	Wednesday	for	abusing	its	power	in	the	mobile	phone	market	and	ordered	the	company	to	alter	its	practices	.

Cognitive Computation Tools

[NP European authorities] [VP fined] [NP Google] [NP a record $5.1 billion] [PP on] [NP Wednesday] [PP for] [VP abusing] [NP its power] [PP in] [NP the mobile phone market] and [VP ordered] [NP the company] [VP to alter] [NP its practices] .

[MISC **European**] authorities fined Google a record $5.1 billion on Wednesday for abusing its power in the mobile phone market and ordered the company to alter its practices.

INFERENCE:

Spacy: It has done a good job at both tagging and named entity recognition. Very similar to Stanford core NLP.

NLTK: It has done a good job at parsing but incorrectly tags Google as a person. Also, it does not account for date and money labels.

Stanford core NLP: It correctly tags the sentence tokens, and determines the named entities correctly, except for having labelled mobile as a title.

Cognitive computation: It correctly parses the entire sentence, but misses out on many named entities like Google, Wednesday.

SENTENCE 2: *A great man once said that many of life's failures are people who did not realize how close they were to success when they gave up.*

Spacy tagging and NER

```
TOKEN, LEMMA, POS, TAG, DEP
A a DET DT det
great great ADJ JJ amod
man man NOUN NN nsubj
once once ADV RB advmod
said say VERB VBD ROOT
that that ADP IN mark
many many ADJ JJ nsubj
of of ADP IN prep
life life NOUN NN poss
's 's PART POS case
failures failure NOUN NNS pobj
are be VERB VBP ccomp
people people NOUN NNS attr
who who PRON WP nsubj
did do VERB VBD aux
not not ADV RB neg
realize realize VERB VB relcl
how how ADV WRB advmod
close close ADJ JJ acomp
they -PRON- PRON PRP nsubj
were be VERB VBD ccomp
to to ADP IN prep
success success NOUN NN pobj
when when ADV WRB advmod
they -PRON- PRON PRP nsubj
gave give VERB VBD advcl
up up PART RP prt
. . PUNCT . punct
```

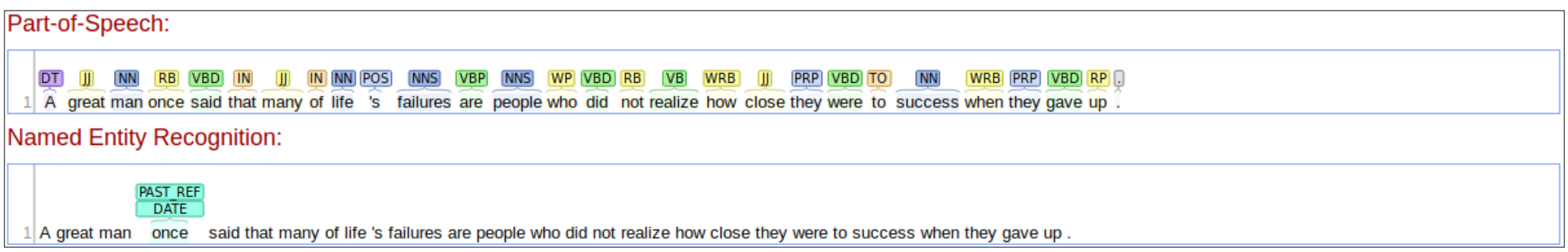
NLTK Parsing

```
(S
  (NP A/DT great/JJ man/NN)
  once/RB
  said/VBD
  that/IN
  many/JJ
  of/IN
  (NP life/NN)
  's/POS
  failures/NNS
  are/VBP
  people/NNS
  who/WP
  did/VBD
  not/RB
  realize/VB
  how/WRB
  close/JJ
  they/PRP
  were/VBD
  to/TO
  (NP success/NN)
  when/WRB
  they/PRP
  gave/VBD
  up/RP
  ./.)
```

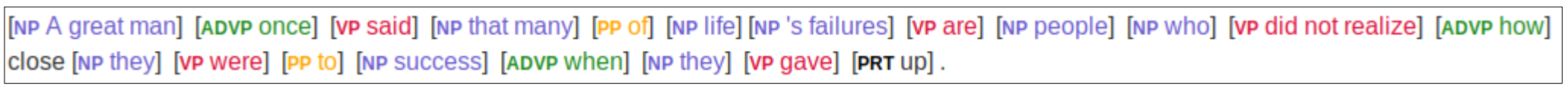
NLTK NER

```
(S
  A/DT
  great/JJ
  man/NN
  once/RB
  said/VBD
  that/IN
  many/JJ
  of/IN
  life/NN
  's/POS
  failures/NNS
  are/VBP
  people/NNS
  who/WP
  did/VBD
  not/RB
  realize/VB
  how/WRB
  close/JJ
  they/PRP
  were/VBD
  to/TO
  success/NN
  when/WRB
  they/PRP
  gave/VBD
  up/RP
  ./.)
```

Stanford core NLP



Cognitive Computation Tools



INFERENCE:

Spacy: It performs tagging correctly, and finds that there are no named entities in the sentence.

NLTK: It does a good job at shallow parsing and like NLTK finds that there are no named entities.

Stanford core NLP: It correctly tags all the tokens and is the only algorithm to find a named entity “once” as a past reference.

Cognitive computation: It performs better parsing than NLTK parser, and also identifies no named entities in the sentence.

SENTENCE 3: *The Vedic Civilization flourished along the river Saraswati, in a region that now consists of the modern Indian states of Haryana and Punjab.*

Spacy tagging and NER

```
TOKEN, LEMMA, POS, TAG, DEP
The the DET DT det
Vedic Vedic PROPN NNP compound
Civilization Civilization PROPN NNP nsubj
flourished flourish VERB VBD ROOT
along along ADP IN prep
the the DET DT det
river river NOUN NN pobj
Saraswati Saraswati PROPN NNP appos
, , PUNCT , punct
in in ADP IN prep
a a DET DT det
region region NOUN NN pobj
that that DET WDT nsubj
now now ADV RB advmod
consists consist VERB VBZ relcl
of of ADP IN prep
the the DET DT det
modern modern ADJ JJ amod
Indian indian ADJ JJ amod
states state NOUN NNS pobj
of of ADP IN prep
Haryana Haryana PROPN NNP pobj
and and CCONJ CC cc
Punjab Punjab PROPN NNP conj
. . PUNCT . punct

The Vedic Civilization ORG
Saraswati GPE
Indian NORP
Haryana GPE
Punjab GPE
```

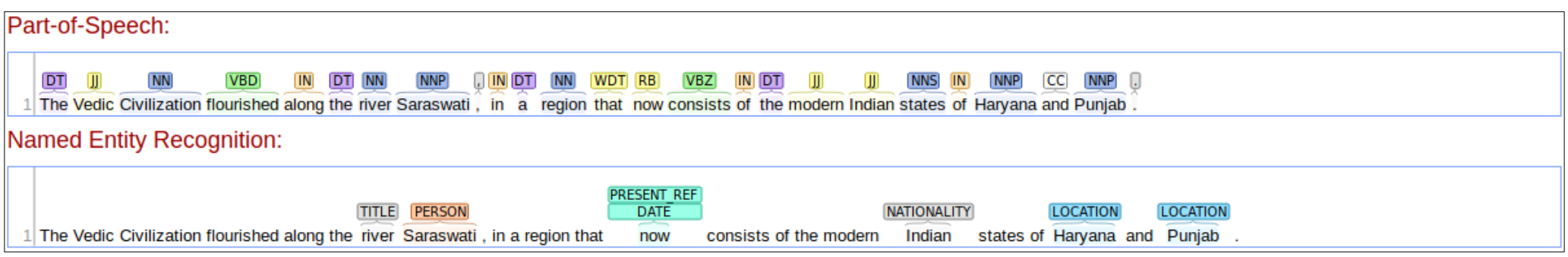
NLTK Parsing

```
(S
  The/DT
  Vedic/NNP
  Civilization/NNP
  flourished/VBD
  along/IN
  (NP the/DT river/NN)
  Saraswati/NNP
  ,/,
  in/IN
  (NP a/DT region/NN)
  that/WDT
  now/RB
  consists/VBZ
  of/IN
  the/DT
  modern/JJ
  Indian/JJ
  states/NNS
  of/IN
  Haryana/NNP
  and/CC
  Punjab/NNP
  ./.)
```

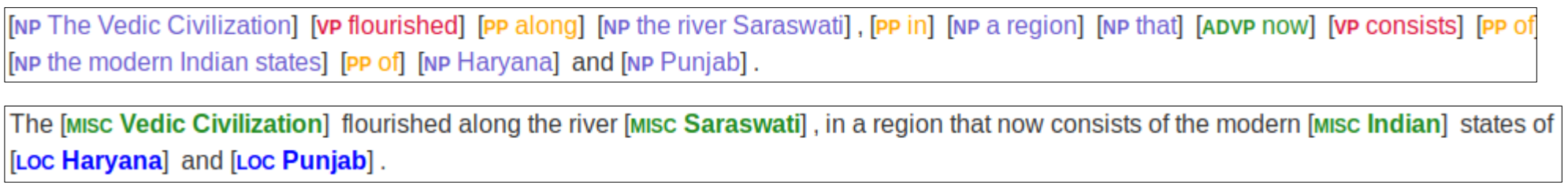
NLTK NER

```
(S
  The/DT
  (ORGANIZATION Vedic/NNP Civilization/NNP)
  flourished/VBD
  along/IN
  the/DT
  river/NN
  (PERSON Saraswati/NNP)
  ,/,
  in/IN
  a/DT
  region/NN
  that/WDT
  now/RB
  consists/VBZ
  of/IN
  the/DT
  modern/JJ
  (GPE Indian/JJ)
  states/NNS
  of/IN
  (ORGANIZATION Haryana/NNP)
  and/CC
  (GPE Punjab/NNP)
  ./.)
```

Stanford core NLP



Cognitive Computation Tools



INFERENCE:

Spacy: Spacy does a good job at determining the tags and also identifies entities like Vedic Civilization.

NLTK: Although it correctly parses the sentence, it incorrectly tags Haryana as an organization and river Saraswati as a person.

Stanford core NLP: Similar to NLTK, although it correctly tags the tokens well, it fails to identify “Vedic Civilization” and incorrectly identifies river Saraswati as a person.

Cognitive computation: It correctly parses and determines the named entities of the entire sentence.

SENTENCE 4: *Dr. APJ Abdul Kalam had an unparalleled career as a defence scientist, culminating the highest civilian award of India, Bharat Ratna.*

Spacy tagging and NER

```
TOKEN, LEMMA, POS, TAG, DEP
Dr. Dr. PROPN NNP compound
APJ APJ PROPN NNP compound
Abdul Abdul PROPN NNP compound
Kalam Kalam PROPN NNP nsubj
had have VERB VBD ROOT
an an DET DT det
unparalleled unparalleled ADJ JJ amod
career career NOUN NN dobj
as as ADP IN prep
a a DET DT det
defence defence NOUN NN compound
scientist scientist NOUN NN pobj
, , PUNCT , punct
culminating culminate VERB VBG advcl
the the DET DT det
highest high ADJ JJS amod
civilian civilian ADJ JJ amod
award award NOUN NN dobj
of of ADP IN prep
India India PROPN NNP pobj
, , PUNCT , punct
Bharat Bharat PROPN NNP compound
Ratna Ratna PROPN NNP appos
. . PUNCT . punct

Abdul Kalam PERSON
India GPE
Bharat NORP
Ratna PERSON
```

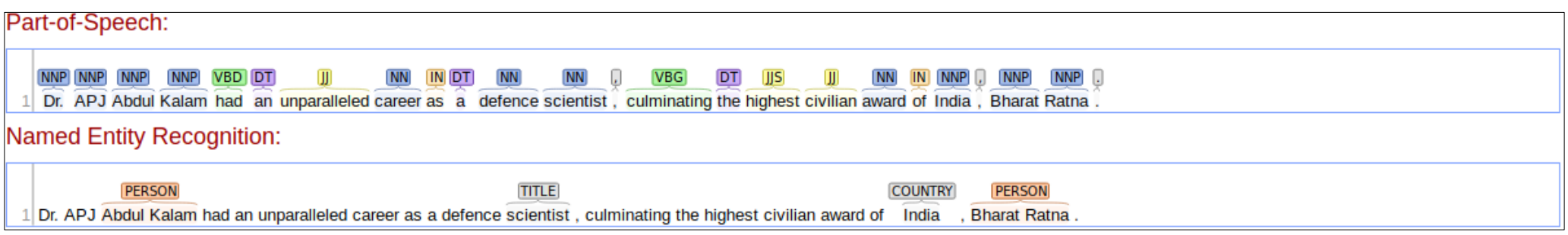
NLTK Parsing

```
(S
  Dr./NNP
  APJ/NNP
  Abdul/NNP
  Kalam/NNP
  had/VBD
  (NP an/DT unparalleled/JJ career/NN)
  as/IN
  (NP a/DT defence/NN)
  (NP scientist/NN)
  ,/,
  culminating/VBG
  the/DT
  highest/JJS
  (NP civilian/JJ award/NN)
  of/IN
  India/NNP
  ,/,
  Bharat/NNP
  Ratna/NNP
  ./.)
```

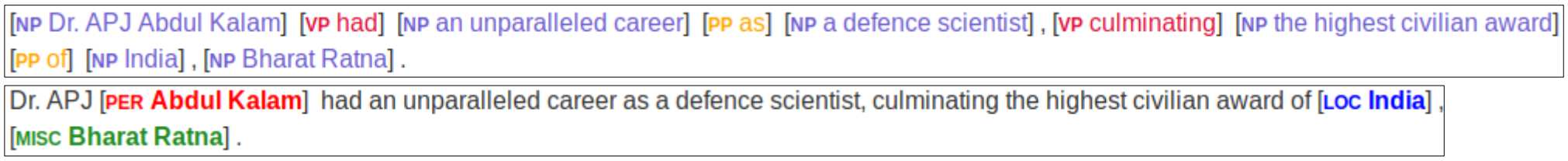
NLTK NER

```
(S
  Dr./NNP
  APJ/NNP
  (PERSON Abdul/NNP Kalam/NNP)
  had/VBD
  an/DT
  unparalleled/JJ
  career/NN
  as/IN
  a/DT
  defence/NN
  scientist/NN
  ,/,
  culminating/VBG
  the/DT
  highest/JJS
  civilian/JJ
  award/NN
  of/IN
  (GPE India/NNP)
  ,/,
  (PERSON Bharat/NNP Ratna/NNP)
  ./.)
```

Stanford core NLP



Cognitive Computation Tools



INFERENCE:

Spacy: Although it tags the tokens of the sentence well and correctly tags most of the entities, it incorrectly tags Bharat Ratna as a person.

NLTK: Similar to Spacy, it also parses the sentence correctly but tags Bharat Ratna as a person.

Stanford core NLP: Similar to the above two, Stanford core NLP incorrectly labels Bharat Ratna as a person.

Cognitive computation: It performs the good parsing and named entity recognition. It categorizes Bharat Ratna in MISC category and not the PER category.

SENTENCE 5: *MS Dhoni's presence in the India team helps skipper Virat Kohli take the right decisions on the field, former India captain and coach Anil Kumble has said.*

Spacy tagging and NER

```
TOKEN, LEMMA, POS, TAG, DEP
MS MS PROPN NNP compound
Dhoni Dhoni PROPN NNP poss
's 's PART POS case
presence presence NOUN NN nsubj
in in ADP IN prep
the the DET DT det
India India PROPN NNP compound
team team NOUN NN pobj
helps help VERB VBZ ccomp
skipper skipper NOUN NN compound
Virat Virat PROPN NNP compound
Kohli Kohli PROPN NNP nsubj
take take VERB VB ccomp
the the DET DT det
right right ADJ JJ amod
decisions decision NOUN NNS dobj
on on ADP IN prep
the the DET DT det
field field NOUN NN pobj
, , PUNCT , punct
former former ADJ JJ amod
India India PROPN NNP compound
captain captain NOUN NN nsubj
and and CCONJ CC cc
coach coach NOUN NN compound
Anil Anil PROPN NNP compound
Kumble Kumble PROPN NNP conj
has have VERB VBZ aux
said say VERB VBN ROOT
. . PUNCT . punct

Dhoni GPE
India GPE
Virat Kohli PERSON
India GPE
Anil Kumble PERSON
```

NLTK Parsing

```
(S
  MS/NNP
  Dhoni/NNP
  's/POS
  (NP presence/NN)
  in/IN
  the/DT
  India/NNP
  (NP team/NN)
  helps/VBZ
  skipper/JJR
  Virat/NNP
  Kohli/NNP
  take/VB
  the/DT
  right/JJ
  decisions/NNS
  on/IN
  (NP the/DT field/NN)
  ,/,
  former/JJ
  India/NNP
  (NP captain/NN)
  and/CC
  (NP coach/NN)
  Anil/NNP
  Kumble/NNP
  has/VBZ
  said/VBD
  ./.)
```

NLTK NER

```
(S
  MS/NNP
  (PERSON Dhoni/NNP)
  's/POS
  presence/NN
  in/IN
  the/DT
  (GPE India/NNP)
  team/NN
  helps/VBZ
  skipper/JJR
  (PERSON Virat/NNP Kohli/NNP)
  take/VB
  the/DT
  right/JJ
  decisions/NNS
  on/IN
  the/DT
  field/NN
  ,/,
  former/JJ
  (GPE India/NNP)
  captain/NN
  and/CC
  coach/NN
  (PERSON Anil/NNP Kumble/NNP)
  has/VBZ
  said/VBD
  ./.)
```

Stanford core NLP

Part-of-Speech:

	NN	NNP	POS	NN	IN	DT	NNP	NN	VBZ	NN	NNP	NNP	VB	DT	NN	NNS	IN	DT	NN	,	JJ	NNP	NN	CC	NN	NNP	NNP	VBZ	VBN	.
1	MS	Dhoni	's	presence	in	the	India	team	helps	skipper	Virat	Kohli	take	the	right	decisions	on	the	field	,	former	India	captain	and	coach	Anil	Kumble	has	said	.

Named Entity Recognition:

	STATE_OR_PROVINCE	PERSON		COUNTRY			PERSON												COUNTRY	TITLE		TITLE	PERSON		
1	MS	Dhoni	's	presence	in	the	India	team	helps	skipper	Virat	Kohli	take	the	right	decisions	on	the	field	,	former	India	captain	and	coach	Anil	Kumble	has	said	.

Cognitive Computation Tools

[NP MS Dhoni] [NP 's presence] [PP in] [NP the India team] [VP helps] [NP skipper Virat Kohli] [VP take] [NP the right decisions] [PP on] [NP the field] , [NP former India captain and coach Anil Kumble] [VP has said] .

[PER MS Dhoni] 's presence in the [LOC India] team helps skipper [PER Virat Kohli] take the right decisions on the field, former [ORG India] captain and coach [PER Anil Kumble] has said.

INFERENCE:

Spacy: It does a good job with tagging the tokens and identifying/labelling most named entities, except that Dhoni is labelled as a country.

NLTK: It does a good job with parsing the sentence and labelling all the named entities correctly.

Stanford core NLP: It correctly labels all the tokens, and is mostly correct in the named entities and titles identified, except that it takes the “MS” from “MS Dhoni” and categorizes it as a state or province.

Cognitive computation: It correctly parses the sentence and identifies all the named entities.

SENTENCE 6: *A member of the Democratic Party, Barack Obama was the first African American to be elected to the presidency.*

Spacy tagging and NER

```
TOKEN, LEMMA, POS, TAG, DEP
A a DET DT det
member member NOUN NN nsubj
of of ADP IN prep
the the DET DT det
Democratic Democratic PROPN NNP compound
Party Party PROPN NNP pobj
, , PUNCT , punct
Barack Barack PROPN NNP compound
Obama Obama PROPN NNP nsubj
was be VERB VBD ROOT
the the DET DT det
first first ADJ JJ amod
African african ADJ JJ amod
American American PROPN NNP attr
to to PART TO aux
be be VERB VB auxpass
elected elect VERB VBN relcl
to to ADP IN prep
the the DET DT det
presidency presidency NOUN NN pobj
. . PUNCT . punct

the Democratic Party ORG
Barack Obama PERSON
first ORDINAL
African NORP
American NORP
```

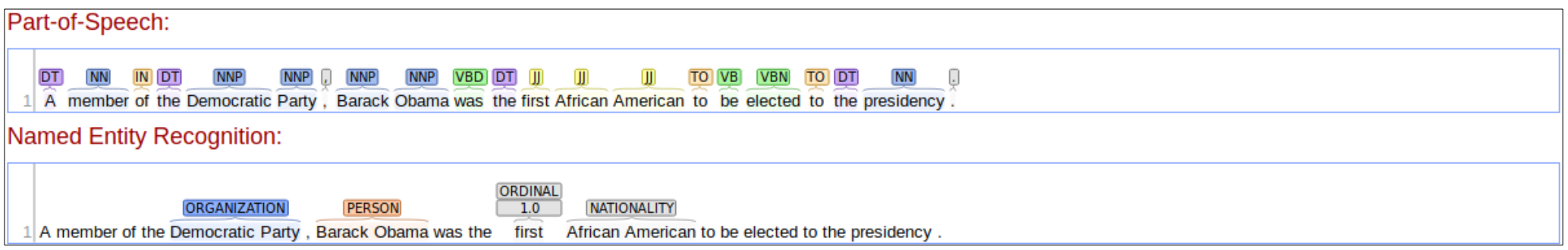
NLTK Parsing

```
(S
  (NP A/DT member/NN)
  of/IN
  the/DT
  Democratic/NNP
  Party/NNP
  ,/,
  Barack/NNP
  Obama/NNP
  was/VBD
  the/DT
  first/JJ
  African/JJ
  American/NNP
  to/TO
  be/VB
  elected/VBN
  to/TO
  (NP the/DT presidency/NN)
  ./.)
```

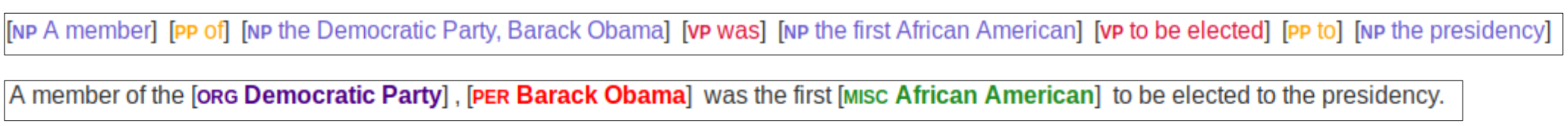
NLTK NER

```
(S
  A/DT
  member/NN
  of/IN
  the/DT
  (ORGANIZATION Democratic/NNP Party/NNP)
  ,/,
  (PERSON Barack/NNP Obama/NNP)
  was/VBD
  the/DT
  first/JJ
  (ORGANIZATION African/JJ American/NNP)
  to/TO
  be/VB
  elected/VBN
  to/TO
  the/DT
  presidency/NN
  ./.)
```

Stanford core NLP



Cognitive Computation Tools



INFERENCE:

Spacy: Spacy does a good job at tagging and finding the named entities in the sentence.

NLTK: Although it parses the sentence well and also tags most of the named entities correctly, it categorizes African American as an organization, which is not correct.

Stanford core NLP: It tags all the tokens correctly and determines the named entities too, including "first" which it puts under the ordinal category.

Cognitive computation: It correctly performs shallow parsing and named entity recognition of the sentence.

SENTENCE 7: *Napoleon Bonaparte had famously said that impossible is a word to be found only in the dictionary of fools.*

Spacy tagging and NER

```
TOKEN, LEMMA, POS, TAG, DEP
Napoleon Napoleon PROPN NNP compound
Bonaparte Bonaparte PROPN NNP nsubj
had have VERB VBD aux
famously famously ADV RB advmod
said say VERB VBN ROOT
that that ADP IN mark
impossible impossible ADJ JJ nsubj
is be VERB VBZ ccomp
a a DET DT det
word word NOUN NN attr
to to PART TO aux
be be VERB VB auxpass
found find VERB VBN relcl
only only ADV RB advmod
in in ADP IN prep
the the DET DT det
dictionary dictionary NOUN NN pobj
of of ADP IN prep
fools fool NOUN NNS pobj
. . PUNCT . punct

Napoleon Bonaparte PERSON
```

NLTK Parsing

```
(S
  Napoleon/NNP
  Bonaparte/NNP
  had/VBD
  famously/RB
  said/VBD
  that/IN
  impossible/JJ
  is/VBZ
  (NP a/DT word/NN)
  to/TO
  be/VB
  found/VBN
  only/RB
  in/IN
  (NP the/DT dictionary/NN)
  of/IN
  fools/NNS
  ./.)
```

NLTK NER

```
(S
  (PERSON Napoleon/NNP)
  (PERSON Bonaparte/NNP)
  had/VBD
  famously/RB
  said/VBD
  that/IN
  impossible/JJ
  is/VBZ
  a/DT
  word/NN
  to/TO
  be/VB
  found/VBN
  only/RB
  in/IN
  the/DT
  dictionary/NN
  of/IN
  fools/NNS
  ./.)
```

Stanford core NLP

Part-of-Speech:

	NNP	NNP	VBD	RB	VBD	IN	JJ	VBZ	DT	NN	TO	VB	VBN	RB	IN	DT	NN	IN	NNS	.
1	Napoleon	Bonaparte	had	famously	said	that	impossible	is	a	word	to	be	found	only	in	the	dictionary	of	fools	.

Named Entity Recognition:

	PERSON																			
1	Napoleon	Bonaparte	had	famously	said	that	impossible	is	a	word	to	be	found	only	in	the	dictionary	of	fools	.

Cognitive Computation Tools

[NP Napoleon Bonaparte] [VP had famously said] [SBAR that] [NP impossible] [VP is] [NP a word] [VP to be found] [ADVP only] [PP in] [NP the dictionary] [PP of] [NP fools] .

[PER **Napoleon Bonaparte**] had famously said that impossible is a word to be found only in the dictionary of fools.

INFERENCE:

Spacy: It does a good job at tagging the sentence well and also identifies Napoleon Bonaparte as a person.

NLTK: Although it correctly tags the sentence, it identifies Napoleon and Bonaparte as two different people.

Stanford core NLP: It does a good job at tagging the sentence well and also identifies the named entity correctly.

Cognitive computation: It does a good job at parsing the sentence well and also identifies the named entity correctly.

SENTENCE 8: *Some current major algorithms for part-of-speech tagging include the Viterbi algorithm, Brill tagger, Constraint Grammar, and the Baum-Welch algorithm also known as the forward-backward algorithm.*

Spacy tagging and NER

```
TOKEN, LEMMA, POS, TAG, DEP
Some some DET DT det
current current ADJ JJ amod
major major ADJ JJ amod
algorithms algorithm NOUN NNS nsubj
for for ADP IN prep
part part NOUN NN nmod
- - PUNCT HYPH punct
of of ADP IN prep
- - PUNCT HYPH punct
speech speech NOUN NN pobj
tagging tagging NOUN NN pobj
include include VERB VBP ROOT
the the DET DT det
Viterbi Viterbi PROPN NNP compound
algorithm algorithm NOUN NN dobj
, , PUNCT , punct
Brill Brill PROPN NNP compound
tagger tagger NOUN NN appos
, , PUNCT , punct
Constraint Constraint PROPN NNP compound
Grammar Grammar PROPN NNP conj
, , PUNCT , punct
and and CCONJ CC cc
the the DET DT det
Baum Baum PROPN NNP compound
- - PUNCT HYPH punct
Welch Welch PROPN NNP compound
algorithm algorithm NOUN NN conj
also also ADV RB advmod
known know VERB VBN acl
as as ADP IN prep
the the DET DT det
forward forward ADJ JJ advmod
- - PUNCT HYPH punct
backward backward NOUN NN amod
algorithm algorithm NOUN NN pobj
. . PUNCT . punct

Viterbi ORG
Brill ORG
Constraint Grammar PERSON
Baum-Welch ORG
```

NLTK Parsing

```
(S
  (NP Some/DT current/JJ major/JJ algorithms/NN)
  for/IN
  part-of-speech/JJ
  tagging/VBG
  include/VBP
  the/DT
  Viterbi/NNP
  (NP algorithm/NN)
  ,/,
  Brill/NNP
  (NP tagger/NN)
  ,/,
  Constraint/NNP
  Grammar/NNP
  ,/,
  and/CC
  the/DT
  Baum-Welch/NNP
  (NP algorithm/NN)
  also/RB
  known/VBN
  as/IN
  (NP the/DT forward-backward/JJ algorithm/NN)
  ./.)
```

NLTK NER

```
(S
  Some/DT
  current/JJ
  major/JJ
  algorithms/NN
  for/IN
  part-of-speech/JJ
  tagging/VBG
  include/VBP
  the/DT
  (ORGANIZATION Viterbi/NNP)
  algorithm/NN
  ,/,
  (GPE Brill/NNP)
  tagger/NN
  ,/,
  (ORGANIZATION Constraint/NNP Grammar/NNP)
  ,/,
  and/CC
  the/DT
  Baum-Welch/NNP
  algorithm/NN
  also/RB
  known/VBN
  as/IN
  the/DT
  forward-backward/JJ
  algorithm/NN
  ./.)
```

Stanford core NLP

Part-of-Speech:

1	Some	current	major	algorithms	for	part-of-speech	tagging	include	the	Viterbi	algorithm	,	Brill	tagger	,	Constraint	Grammar	,	and	the	Baum-Welch	algorithm	also	known	as	the	forward-backward	algorithm	.			
	DT	JJ	JJ	NNS	IN	JJ	NN	VBP	DT	NNP	NN	,	NNP	NN	,	NNP	NNP	,	CC	DT	NNP	NN	RB	VBN	IN	DT	JJ	NN	.			

Named Entity Recognition:

	PRESENT_REF DATE							PERSON		PERSON						LOCATION															
1	Some	current	major	algorithms	for	part-of-speech	tagging	include	the	Viterbi	algorithm	,	Brill	tagger	,	Constraint	Grammar	,	and	the	Baum-Welch	algorithm	also	known	as	the	forward-backward	algorithm	.			

Cognitive Computation Tools

[NP Some current major algorithms] [PP for] [NP part-of-speech] [VP tagging include] [NP the Viterbi algorithm] , [NP Brill tagger] , [NP Constraint Grammar] , and [NP the Baum-Welch algorithm] [ADVP also] [VP known] [PP as] [NP the forward-backward algorithm] .

Some current major algorithms for part-of-speech tagging include the [MISC Viterbi] algorithm, [PER Brill] tagger, Constraint Grammar, and the [MISC Baum-Welch] algorithm also known as the forward-backward algorithm.

INFERENCE:

Spacy: Although Spacy tags the tokens well, it is incorrect in all the named entity categorizations.

NLTK: Similar to Spacy, although the parsing is correct, it incorrectly categorizes all the named entities.

Stanford core NLP: It correctly tags the tokens and also labels Viterbi and Brill as persons, it miscategorises Baum-Welch as a location.

Cognitive computation: It does a good job at parsing the sentence, and decently well at named entity, however categorises Viterbi in the MISC category.

SENTENCE 9: *Apple's revolutionary products, which include the iPod, iPhone and iPad, are now seen as dictating the evolution of modern technology.*

Spacy tagging and NER

```
TOKEN, LEMMA, POS, TAG, DEP
Apple Apple PROPN NNP poss
's 's PART POS case
revolutionary revolutionary ADJ JJ amod
products product NOUN NNS nsubjpass
, , PUNCT , punct
which which DET WDT nsubj
include include VERB VBP relcl
the the DET DT det
iPod iPod PROPN NNP dobj
, , PUNCT , punct
iPhone iPhone PROPN NNP conj
and and CCONJ CC cc
iPad iPad PROPN NNP conj
, , PUNCT , punct
are be VERB VBP auxpass
now now ADV RB advmod
seen see VERB VBN ROOT
as as ADP IN prep
dictating dictate VERB VBG pcomp
the the DET DT det
evolution evolution NOUN NN dobj
of of ADP IN prep
modern modern ADJ JJ amod
technology technology NOUN NN pobj
. . PUNCT . punct

Apple ORG
iPhone ORG
```

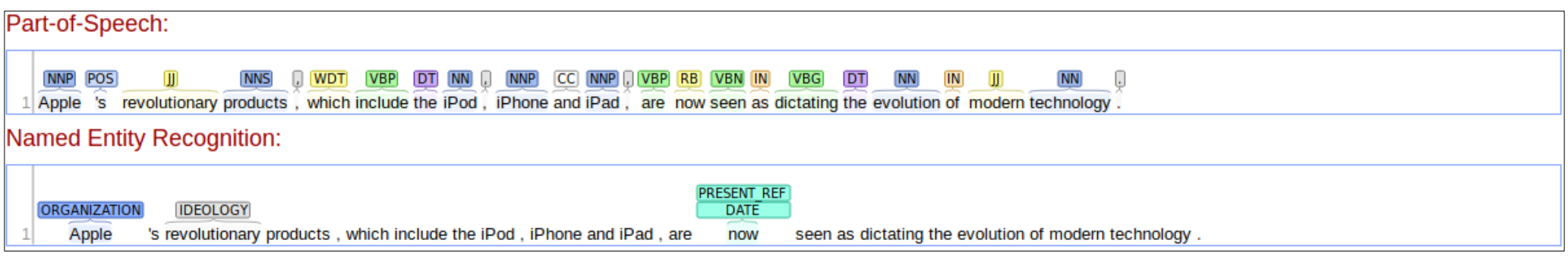
NLTK Parsing

```
(S
  Apple/NNP
  's/POS
  revolutionary/JJ
  products/NNS
  ,/,
  which/WDT
  include/VBP
  (NP the/DT iPod/NN)
  ,/,
  (NP iPhone/NN)
  and/CC
  (NP iPad/NN)
  ,/,
  are/VBP
  now/RB
  seen/VBN
  as/IN
  dictating/VBG
  (NP the/DT evolution/NN)
  of/IN
  (NP modern/JJ technology/NN)
  ./.)
```

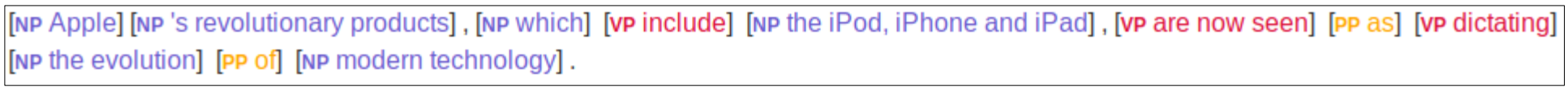
NLTK NER

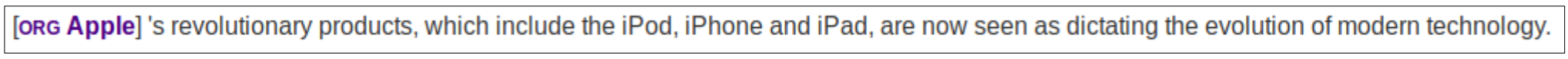
```
(S
  (GPE Apple/NNP)
  's/POS
  revolutionary/JJ
  products/NNS
  ,/,
  which/WDT
  include/VBP
  the/DT
  (ORGANIZATION iPod/NN)
  ,/,
  (ORGANIZATION iPhone/NN)
  and/CC
  (ORGANIZATION iPad/NN)
  ,/,
  are/VBP
  now/RB
  seen/VBN
  as/IN
  dictating/VBG
  the/DT
  evolution/NN
  of/IN
  modern/JJ
  technology/NN
  ./.)
```

Stanford core NLP



Cognitive Computation Tools





INFERENCE:

Spacy: It correctly tags the tokens and Apple as an organization, but incorrectly tags iPhone as an organization too.

NLTK: While it parses the sentences correctly, it incorrectly categorizes Apple as a country and iPhone, iPad and iPod as organizations.

Stanford core NLP: It performs a good job at tagging and named entity recognition, but it does not identify iPhone, iPad and iPod as named entities.

Cognitive computation: It correctly parses the sentence and recognizes Apple as an organization, but fails to recognize iPod, iPhone and iPad as named entities.

SENTENCE 10: *Chennai, originally known as Mudhiras, was located in the province of Tondaimandalam, an area lying between Penna River of Nellore and the Ponnaiyar river of Cuddalore.*

Spacy tagging and NER

```
TOKEN, LEMMA, POS, TAG, DEP
Chennai Chennai PROPN NNP nsubjpass
, , PUNCT , punct
originally originally ADV RB advmod
known know VERB VBN acl
as as ADP IN prep
Mudhiras Mudhiras PROPN NNP pobj
, , PUNCT , punct
was be VERB VBD auxpass
located locate VERB VBN ROOT
in in ADP IN prep
the the DET DT det
province province NOUN NN pobj
of of ADP IN prep
Tondaimandalam Tondaimandalam PROPN NNP pobj
, , PUNCT , punct
an an DET DT det
area area NOUN NN appos
lying lie VERB VBG acl
between between ADP IN prep
Penna Penna PROPN NNP compound
River River PROPN NNP pobj
of of ADP IN prep
Nellore Nellore PROPN NNP pobj
and and CCONJ CC cc
the the DET DT det
Ponnaiyar Ponnaiyar PROPN NNP compound
river river NOUN NN conj
of of ADP IN prep
Cuddalore Cuddalore PROPN NNP pobj
. . PUNCT . punct

Mudhiras ORG
Tondaimandalam GPE
Penna River of Nellore LOC
Ponnaiyar GPE
```

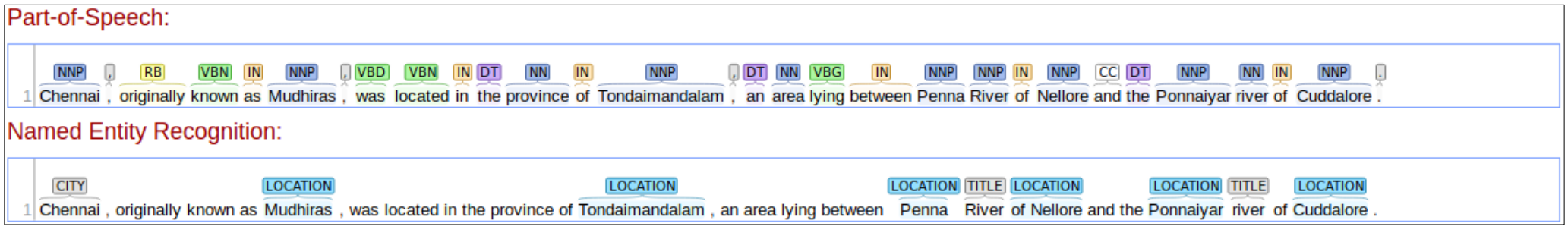
NLTK Parsing

```
(S
  Chennai/NNP
  ,/,
  originally/RB
  known/VBN
  as/IN
  Mudhiras/NNP
  ,/,
  was/VBD
  located/VBN
  in/IN
  (NP the/DT province/NN)
  of/IN
  Tondaimandalam/NNP
  ,/,
  (NP an/DT area/NN)
  lying/VBG
  between/IN
  Penna/NNP
  River/NNP
  of/IN
  Nellore/NNP
  and/CC
  the/DT
  Ponnaiyar/NNP
  (NP river/NN)
  of/IN
  Cuddalore/NNP
  ./.)
```

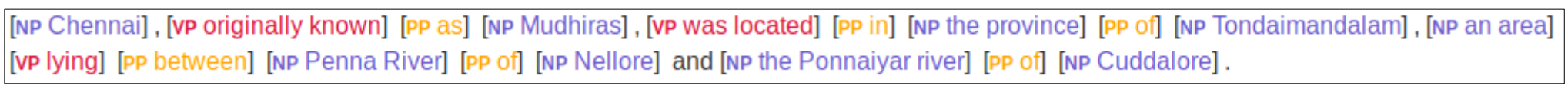
NLTK NER

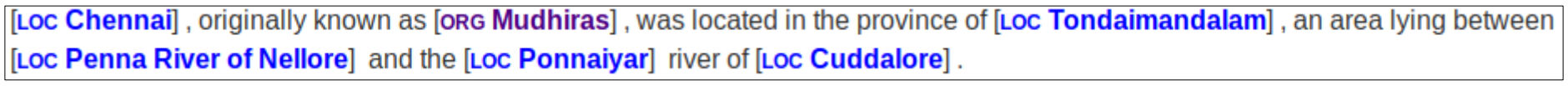
```
(S
  (GPE Chennai/NNP)
  ,/,
  originally/RB
  known/VBN
  as/IN
  (GPE Mudhiras/NNP)
  ,/,
  was/VBD
  located/VBN
  in/IN
  the/DT
  province/NN
  of/IN
  (GPE Tondaimandalam/NNP)
  ,/,
  an/DT
  area/NN
  lying/VBG
  between/IN
  (PERSON Penna/NNP River/NNP)
  of/IN
  (GPE Nellore/NNP)
  and/CC
  the/DT
  (ORGANIZATION Ponnaiyar/NNP)
  river/NN
  of/IN
  (GPE Cuddalore/NNP)
  ./.)
```

Stanford core NLP



Cognitive Computation Tools





INFERENCE:

Spacy: It tags the tokens and identifies most named entities correctly, but it fails to recognize Chennai as a named entity and labels Mudhiras as an organization.

NLTK: It does a good job at parsing the sentence and most named entities too, but it labels Penna river as a person and Ponnaiyar as an organization.

Stanford core NLP: It correctly tags all the tokens and also correctly categorizes all the named entities.

Cognitive computation: It correctly parses the sentence and also correctly categorizes all the named entities.