

Assignment 6: Statistical Parsing and Lexical Semantics

*My posse consists of: null**Salman Ahmad (saahmad@mit.edu)***Part II: Semantic Analysis of “Invest”****Data Collection**

To collect all of the occurrences of “invest” along with their subcategorization frame I used the follow `trexex` query on the supplied subset of the Penn Tree Bank:

```
@VP < ( __< /[Ii]nvest[esi]|^[Ii]nvest$/ )
```

The query returned 150 matches in 149 trees.

Analysis

A summary of the data is shown in the next section.

Data Tabulation

The table below displays the number of times a unique subcategorization frame exists in the Penn Tree Bank.

Table 1: Subcategory counts

Subcategory	Count
[VP (VBG investing) PP-CLR]	13
[VP (VB invest) PP-CLR]	12
[VP (VBG investigating) NP]	11
[VP (VB invest) NP PP-CLR]	8
[VP (VB investigate) NP]	5
[VP (VBG investigating) SBAR]	4
[VP (VBN invested) NP PP-CLR]	4
[VP (VBN invested) PP-CLR]	4
[VP (VBG investing) NP]	3
[VP (VB invest) ADVP-LOC]	2
[VP (VB invest) NP PP-LOC-CLR]	2
[VP (VB invest)]	2
[VP (VB investigate) SBAR]	2
[VP (VBD invested) PP-CLR]	2
[VP (VBD investigated) NP]	2
[VP (VBG investing) NP PP-CLR]	2

Continued on next page

Table 1 – *Continued from previous page*

Subcategory	Count
[VP (VBN invested) NP ADVP-LOC]	2
[VP (VBN invested) NP PP-CLR PP-TMP]	2
[VP (VBN invested) NP PP-CLR-LOC]	2
[VP (VBP invest) PP-CLR]	2
[VP (VBZ investigates) NP]	2
[VP (VB invest) , SBAR-ADV]	1
[VP (VB invest) ADVP-MNR PP-CLR PP-TMP]	1
[VP (VB invest) ADVP-MNR]	1
[VP (VB invest) NP ADVP-LOC PP-MNR]	1
[VP (VB invest) NP ADVP-LOC-CLR]	1
[VP (VB invest) NP PP-CLR , NP-2]	1
[VP (VB invest) NP PP-CLR , PP-2]	1
[VP (VB invest) NP PP-CLR S-PRP]	1
[VP (VB invest) NP PP-CLR SBAR-1]	1
[VP (VB invest) NP PP-TMP PP-LOC-CLR S-CLR]	1
[VP (VB invest) NP S-PRP]	1
[VP (VB invest) NP]	1
[VP (VB invest) PP-CLR PP-PRP]	1
[VP (VB invest) PP-LOC-CLR NP]	1
[VP (VB invest) PP-LOC-CLR]	1
[VP (VB invest) PP-TMP]	1
[VP (VB investigate) CC VB NP]	1
[VP (VB investigate) SBAR-NOM]	1
[VP (VB investigate)]	1
[VP (VBD invested) NP ADVP-TMP PP-CLR]	1
[VP (VBD invested) NP PP-CLR PP-TMP]	1
[VP (VBD invested) NP PP-CLR PP]	1
[VP (VBD invested) NP PP-CLR]	1
[VP (VBD invested) NP PP-LOC-CLR , S-PRP]	1
[VP (VBD invested) NP PP-LOC]	1
[VP (VBD invested) NP PP]	1
[VP (VBD invested) NP S-PRP-CLR]	1
[VP (VBD invested) NP]	1

Continued on next page

Table 1 – *Continued from previous page*

Subcategory	Count
[VP (VBD invested) PP-CLR NP]	1
[VP (VBD investigated) SBAR-NOM]	1
[VP (VBG investigating) NP PP-MNR]	1
[VP (VBG investigating) NP PP-TMP , SBAR-ADV]	1
[VP (VBG investigating) NP-1 PP-CLR]	1
[VP (VBG investigating) PP-MNR]	1
[VP (VBG investing) ADVP PP-CLR]	1
[VP (VBG investing) ADVP-MNR PP-CLR]	1
[VP (VBG investing) ADVP-MNR PP-LOC]	1
[VP (VBG investing) NP PP-LOC-CLR]	1
[VP (VBG investing) NP S-PRP]	1
[VP (VBG investing) PP-CLR ADVP-TMP]	1
[VP (VBG investing) PP-CLR SBAR-1]	1
[VP (VBG investing) PP-LOC-CLR]	1
[VP (VBG investing) PP-PRP]	1
[VP (VBN invested) ADVP-MNR PP-CLR]	1
[VP (VBN invested) ADVP-MNR PP-LOC-CLR , SBAR-TMP]	1
[VP (VBN invested) NP PP-LOC-CLR , PP]	1
[VP (VBN invested) NP PP-LOC-CLR]	1
[VP (VBN invested) NP PP-TMP]	1
[VP (VBN invested) NP PP]	1
[VP (VBN invested) PP-CLR PP-TMP]	1
[VP (VBN invested) PP-LOC-CLR-2 PP-TMP]	1
[VP (VBN investigated) NP PP PP-PRP]	1
[VP (VBN investigated) NP PP]	1
[VP (VBN investigated) NP]	1
[VP (VBP invest) ADVP PP-LOC-CLR ADVP-LOC]	1
[VP (VBP invest) ADVP-LOC]	1
[VP (VBP invest) PP-LOC-CLR]	1
[VP (VBZ invests) ADVP PP-CLR]	1
[VP (VBZ invests) ADVP PP-LOC-CLR]	1
[VP (VBZ invests) ADVP PP-LOC]	1
[VP (VBZ invests) NP NP-TMP PP-CLR SBAR-ADV]	1

Continued on next page

Table 1 – *Continued from previous page*

Subcategory	Count
[VP (VBZ invests) NP PP-CLR]	1