

SRL Data Analysis

We will perform some data analysis on the SRL data. With iterations, We will continue to add more information which will be for analysis, organizing and improving SRL and other associated tasks.

SRL Data Sources

Here we will provide information on the SRL data.

Total sentences : 12271

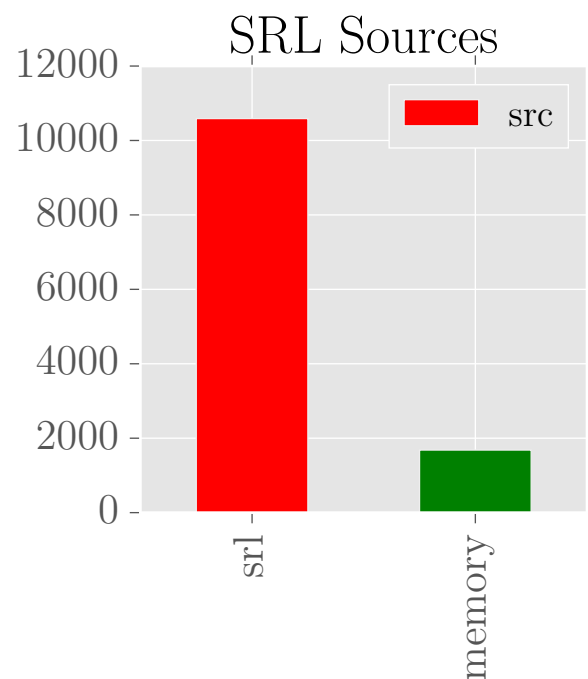
Sentences with multiple predicates : 6269

```
[5]: #df['sentence']
```

Composition of the SRL data

	src
srl	10589
memory	1682

Plot of the Composition of SRL data



Plot of counts of the SRL Relations.

Each relation is counted only once per sentence.

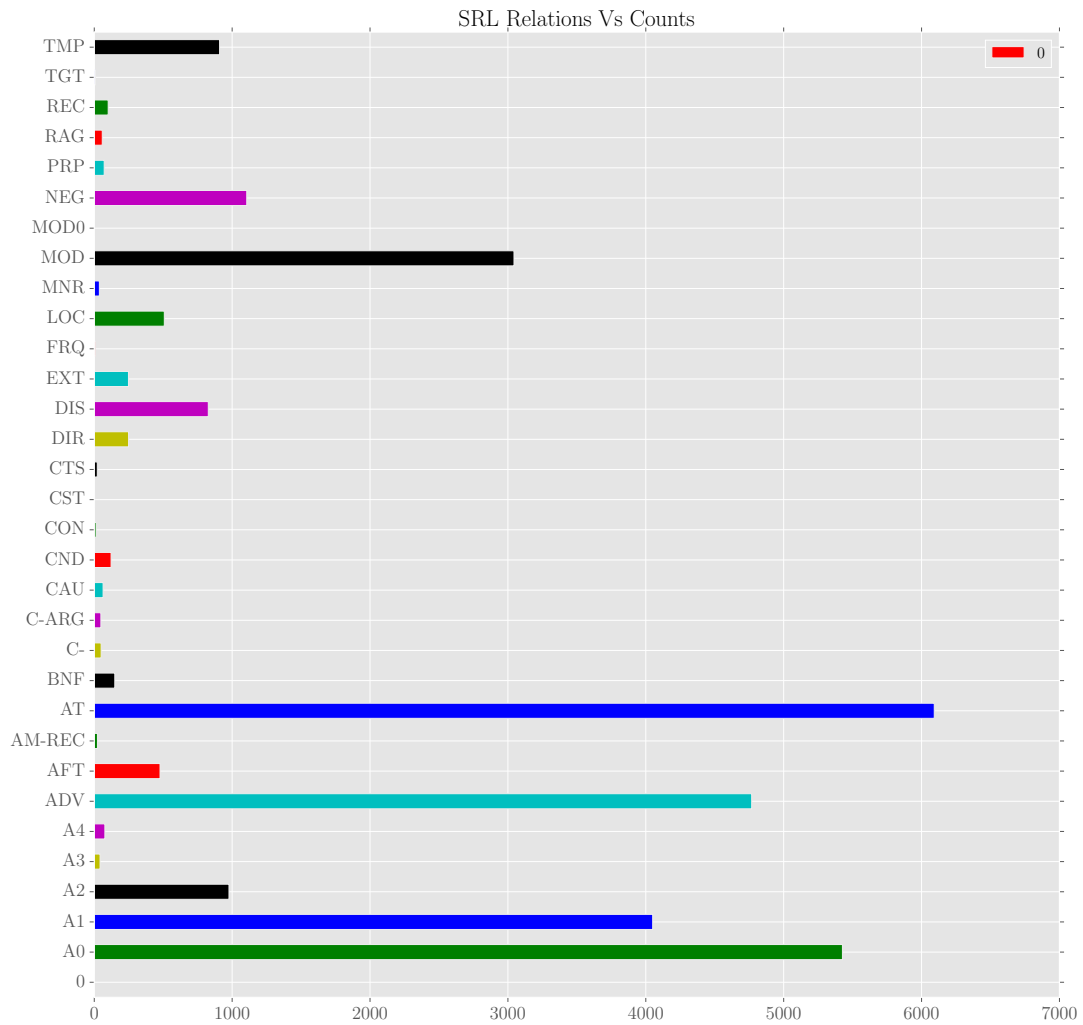


Table for the Count of the SRL relations.

Here again each relation is counted only once per sentence.

	0
0	1
A0	5426
A1	4050
A2	977
A3	42
A4	77
ADV	4768
AFT	477
AM-REC	23
AT	6092
BNF	149
C-	50
C-ARG	48
CAU	64

CND	122
CON	13
CST	3
CTS	22
DIR	249
DIS	827
EXT	248
FRQ	8
LOC	508
MNR	37
MOD	3044
MOD0	1
NEG	1106
PRP	72
RAG	58
REC	101
TGT	9
TMP	909

Plot of sentence token length in the data.

Average token length of sentence in SRL: 7.208295982397523

