

1 Grammatical Relations

The simplest classification approach considered the relative frequency of different grammatical relations. For this approach, the governor and the dependent of the dependencies were ignored, with only the relation itself being used.

Each data set instance contained 52 numerical attributes, one for each relation in the Stanford Dependency system. For each attribute A_r corresponding to the relation r , the corresponding value was n_r/n_t , where n_r and n_t were the number of occurrences of the relation r and the total number of relations in the text, respectively.

A C4.5 decision tree classifier trained on these instances produces the decision tree shown in Algorithm 1.1. The full names for the seven relations are shown in Table 1.1. The following subsections explore the linguistic reasons why these particular relations should be so useful in classifying the texts.

Algorithm 1.1 C4.5 decision tree classifier

```
if complm ≤ 0.011635 then
  if purpcl ≤ 0.000856 then
    if rcmod ≤ 0.012254 then
      en(34.0)
    else
      if pvt ≤ 0.002113 then
        es(4.0/1.0)
      else
        en(6.0)
    else
      if purpcl ≤ 0.001191 then
        if advmod ≤ 0.045825 then
          es(6.0)
        else
          en(2.0)
      else
        en(7.0)
  else
    if mark ≤ 0.00808 then
      en(6.0)
    else
      if aux ≤ 0.044037 then
        en(6.0/1.0)
      else
        es(60.0)
```

Table 1.1: Relation abbreviations

advmod	adverbial modifier
aux	auxiliary
complm	complementizer
mark	marker
prt	phrasal verb particle
purpcl	purpose clause modifier
rcmod	relative clause modifier

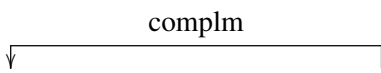
1.1 Adverbial Modifier

1.2 Auxiliary

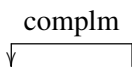
1.3 Complementizer

A complementizer is a word that signals the beginning of a clausal complement. The Stanford Parser recognizes the complementizers *that* and *whether* as shown in example 1.1 (Sulec, Micus). The governor of a complementizer dependency is the root of the clause: a verb generally or, in the case of copular clauses, the complement.


(1.1) a. ... I will consider ... whether the world is a safe place



b. At least you choose whether to go to a pub or not.



c. They state that climate generally predicts that temperatures should rise ...



Whitley [1986] points out that while English tends to allow complementizers introducing clausal complements in the object position to be deleted, Spanish is much more restrictive in this regard (see examples 1.2 and 1.3). [Whitley, 1986, p.278].

(1.2) a. I say that he'll do it.

b. I say he'll do it.

(1.3) a. Digo que lo har .


b. *Digo lo har .

1.4 Marker

1.5 Phrasal Verb Particle

The phrasal verb particle relation ties the head word of a phrasal verb to its particle as shown in Example 1.4.

(1.4) ...the reduction of superfluous proteins will free up resources ...



1.6 Purpose Clause Modifier

1.7 Relative Clause Modifier

References

M. Stanley Whitley. *Spanish/English Contrasts: A Course in Spanish Linguistics*. Georgetown University Press, 1986.