

Factbank tags in Hubdub

Chris Potts, Ling 236: Seminar in Lexical & Constructional Pragmatics, Fall 2009

Nov 4

fb_factValue.txt

Compile a dictionary

```
... 'said' ||| 'AUTHOR' ||| 'CT+'
... 'raised' ||| 'CORP._AUTHOR' ||| 'CT+'
... 'raised' ||| 'AUTHOR' ||| 'Uu'
... 'offering' ||| 'CORP._AUTHOR' ||| 'CT+'
... 'offering' ||| 'AUTHOR' ||| 'Uu'
... 'reduce' ||| 'company_AUTHOR' ||| 'CT+'
... 'reduce' ||| 'AUTHOR' ||| 'Uu'
... 'said' ||| 'AUTHOR' ||| 'CT+'
... 'due' ||| 'company_AUTHOR' ||| 'CT+'
...
```

```
said:      Uu: 5
           CT-: 1
           NA: 4
           CT+: 751
raised:    Uu: 3
           CT-: 3
           CT+: 8
offering:  Uu: 8
           CT+: 24
...
```

Hubdub 'annotations'

Go through the Hubdub comments with associated predictions, mapping words occurring in fb_factValue.txt to their most frequent tags.

Word	MedianWager	MeanWager	MedianPercentWagered	MeanPercentWagered	MedianLevel	MeanLevel	MedianAgreement	MeanAgreement
PR-	500	1443.18	0.00155	0.02902	15.50	14.50	46	47.73
Uu	300	789.03	0.00141	0.02887	15	13.92	35	39.96
CT+	350	789.00	0.00133	0.02713	16	13.92	34	40.48
CT-	400	763.43	0.00128	0.04788	15	13.92	31	38.27
PR+	200	744.36	0.00118	0.03573	15	14.05	29	35.77
PS+	200	645.61	0.00128	0.07579	16	13.80	30	36.22
PS-	300	491.22	0.00155	0.00799	14	13.57	33.5	37.38

From Sauri & Pustejovsky 2009, 'FactBank: a corpus annotated with event factuality', p. 257 (subpart of table 5):

Table 5 Factuality values and their use

VALUE	USE
-------	-----

Committed values:

- CT+ According to the source, it is **certainly** the case that X
- PR+ According to the source, it is **probably** the case that X
- PS+ According to the source, it is **possibly** the case that X
- CT- According to the source, it is **certainly not** the case that X
- PR- According to the source it is **probably not** the case that X
- PS- According to the source it is **possibly not** the case that X