Assignment 3

Team 3

Results

Preprocessing was applied to the bug reports of both systems, and the resulting terms were pos-tagged and counted considering sentence boundaries. The results are presented next:

Derby:

Corpus	Terms		Nouns		Verbs		Adjectives		Adverbs		Sentences	
	Avg	**Med**	**Avg**	**Med**	**Avg**	**Med**	**Avg**	**Med**	**Avg**	**Med**	**Avg**	**Med**
Titles	5.91	5	3.85	3	1.45	1	1.49	1	1.18	1	1.17	1
Descriptions	148.16	44	100.7	25	17.67	8	25.68	9	4.05	2	18.67	6
Both	150.89	47	102.35	28	18.22	9	25.78	9	4.13	2	19.84	7
Total Bugs											6866	

Ofbiz:

Corpus	Terms		Nouns		Verbs		Adjectives		Adverbs		Sentences	
	Avg	**Med**	**Avg**	**Med**	**Avg**	**Med**	**Avg**	**Med**	**Avg**	**Med**	**Avg**	**Med**
Titles	5.59	5	3.81	3	1.38	1	1.47	1	1.05	1	1.05	1
Descriptions	78.51	21	54.14	13	10.14	4	13.64	4	3.59	1	7.93	2
Both	72.72	23	50.02	14	10.01	4	12.84	4	3.57	1	8.99	3
Total Bugs											6912	

Amount of reports without terms after preprocessing by corpus:

- Derby titles: 29
- Derby descriptions: 201
- Derby both: 0
- Ofbiz titles: 17
- Ofbiz descriptions: 547
- Ofbiz both: 0