

Multi-way MT

Firat et al. 2016 (two papers)

Dir			Fr (39m)		Cs (12m)		De (4.2m)		Ru (2.3m)		Fi (2m)	
			→ En	En →	→ En	En →	→ En	En →	→ En	En →	→ En	En →
(a) BLEU	Dev	Single	27.22	26.91	21.24	15.9	24.13	20.49	21.04	18.06	13.15	9.59
		Multi	26.09	25.04	21.23	14.42	23.66	19.17	21.48	17.89	12.97	8.92
	Test	Single	27.94	29.7	20.32	13.84	24	21.75	22.44	19.54	12.24	9.23
		Multi	28.06	27.88	20.57	13.29	24.20	20.59	23.44	19.39	12.61	8.98
(b) LL	Dev	Single	-50.53	-53.38	-60.69	-69.56	-54.76	-61.21	-60.19	-65.81	-88.44	-91.75
		Multi	-50.6	-56.55	-54.46	-70.76	-54.14	-62.34	-54.09	-63.75	-74.84	-88.02
	Test	Single	-43.34	-45.07	-60.03	-64.34	-57.81	-59.55	-60.65	-60.29	-88.66	-94.23
		Multi	-42.22	-46.29	-54.66	-64.80	-53.85	-60.23	-54.49	-58.63	-71.26	-88.09

Table 3: (a) BLEU scores and (b) average log-probabilities for all the five languages from WMT'15.

ok, but what about multi-source?