

Run ID:	CERES1
3-way accuracy:	0.405
2-way accuracy:	0.526

		System Response			Total
		Entailment	Unknown	Contradiction	
Gold Standard	Entailment	133	353	14	500
	Unknown	65	269	16	350
	Contradiction	42	105	3	150
Total		240	727	33	1000

Table 1: Confusion matrix for the 3-way task for CERES1

Run ID:	CERES2
3-way accuracy:	0.416
2-way accuracy:	0.526

		System Response			Total
		Entailment	Unknown	Contradiction	
Gold Standard	Entailment	133	363	4	500
	Unknown	65	282	3	350
	Contradiction	42	107	1	150
Total		240	752	8	1000

Table 2: Confusion matrix for the 3-way task for CERES2