

Run ID:	Boeing1
3-way accuracy:	0.438
2-way accuracy:	0.553

		System Response			Total
		Entailment	Unknown	Contradiction	
Gold Standard	Entailment	77	221	2	300
	Unknown	27	181	2	210
	Contradiction	18	67	5	90
Total		122	469	9	600

Table 1: Confusion matrix for the 3-way task for Boeing1

Run ID:	Boeing2
3-way accuracy:	0.463
2-way accuracy:	0.567

		System Response			Total
		Entailment	Unknown	Contradiction	
Gold Standard	Entailment	101	196	3	300
	Unknown	34	171	5	210
	Contradiction	27	57	6	90
Total		162	424	14	600

Table 1: Confusion matrix for the 3-way task for Boeing2

Run ID:	Boeing3
3-way accuracy:	0.547
2-way accuracy:	0.615

		System Response			Total
		Entailment	Unknown	Contradiction	
Gold Standard	Entailment	177	120	3	300
	Unknown	60	145	5	210
	Contradiction	48	36	6	90
Total		285	301	14	600

Table 1: Confusion matrix for the 3-way task for Boeing3