

Run ID:	AUEBNLP1
3-way accuracy:	0.546
2-way accuracy:	0.58
Average precision:	0.5654

		System Response			Total
		Entailment	Unknown	Contradiction	
Gold Standard	Entailment	382	109	9	500
	Unknown	187	160	3	350
	Contradiction	115	31	4	150
Total		684	300	16	1000

Table 1: Confusion matrix for the 3-way task for AUEBNLP1

Run ID:	AUEBNLP2
3-way accuracy:	0.547
2-way accuracy:	0.579
Average precision:	0.562

		System Response			Total
		Entailment	Unknown	Contradiction	
Gold Standard	Entailment	385	115	0	500
	Unknown	188	162	0	350
	Contradiction	118	32	0	150
Total		691	309	0	1000

Table 2: Confusion matrix for the 3-way task for AUEBNLP2

Run ID:	AUEBNLP3
3-way accuracy:	0.554
2-way accuracy:	0.584
Average precision:	0.522

		System Response			Total
		Entailment	Unknown	Contradiction	
Gold Standard	Entailment	399	101	0	500
	Unknown	195	155	0	350
	Contradiction	120	30	0	150
Total		714	286	0	1000

Table 3: Confusion matrix for the 3-way task for AUEBNLP3