

Run ID:	OAQA1
3-way accuracy:	0.616
2-way accuracy:	0.688
Average precision:	0.5811

		System Response			Total
		Entailment	Unknown	Contradiction	
Gold Standard	Entailment	390	81	29	500
	Unknown	148	163	39	350
	Contradiction	54	33	63	150
Total		592	277	131	1000

Table 1: Confusion matrix for the 3-way task for OAQA1

Run ID:	OAQA2
3-way accuracy:	0.52
2-way accuracy:	0.54
Average precision:	0.5811

		System Response			Total
		Entailment	Unknown	Contradiction	
Gold Standard	Entailment	443	34	23	500
	Unknown	268	75	7	350
	Contradiction	135	13	2	150
Total		846	122	32	1000

Table 2: Confusion matrix for the 3-way task for OAQA2

Run ID:	OAQA3
3-way accuracy:	0.432
2-way accuracy:	0.547
Average precision:	0.581

		System Response			Total
		Entailment	Unknown	Contradiction	
Gold Standard	Entailment	178	272	50	500
	Unknown	80	242	28	350
	Contradiction	51	87	12	150
Total		309	601	90	1000

Table 3: Confusion matrix for the 3-way task for OAQA3