

Run ID:	Emory1
2-way accuracy:	0.588
Average precision:	0.5998

		System Response		Total
		Entailment	No Entailment	
Gold Standard	Entailment	340	160	500
	No Entailment	252	248	500
Total		592	408	1000

Table 1: Confusion matrix for the 2-way task for Emory1

Run ID:	Emory2
2-way accuracy:	0.57
Average precision:	0.6012

		System Response		Total
		Entailment	No Entailment	
Gold Standard	Entailment	323	177	500
	No Entailment	253	247	500
Total		576	424	1000

Table 2: Confusion matrix for the 2-way task for Emory2

Run ID:	Emory3
2-way accuracy:	0.511
Average precision:	0.5115

		System Response		Total
		Entailment	No Entailment	
Gold Standard	Entailment	500	0	500
	No Entailment	489	11	500
Total		989	11	1000

Table 3: Confusion matrix for the 2-way task for Emory3