

Run ID:	DFKI1
2-way accuracy:	0.672

		System Response		Total
		Entailment	No Entailment	
Gold Standard	Entailment	333	167	500
	No Entailment	161	339	500
Total		494	506	1000

Table 1: Confusion matrix for the 2-way task for DFKI1

Run ID:	DFKI2
2-way accuracy:	0.699

		System Response		Total
		Entailment	No Entailment	
Gold Standard	Entailment	407	93	500
	No Entailment	208	292	500
Total		615	385	1000

Table 2: Confusion matrix for the 2-way task for DFKI2

Run ID:	DFKI3
2-way accuracy:	0.706

		System Response		Total
		Entailment	No Entailment	
Gold Standard	Entailment	374	126	500
	No Entailment	168	332	500
Total		542	458	1000

Table 3: Confusion matrix for the 2-way task for DFKI3