CAsT Track, main task results — (HEATWAVE) University of Cambridge Speech Group

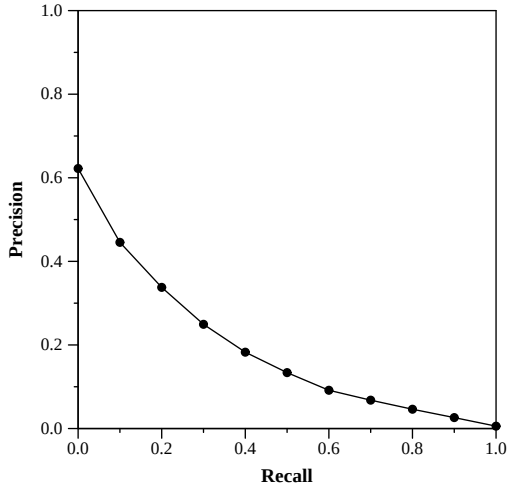
Run Description	
Run ID:	gold
Category:	manually rewritten utterances
Query understanding method:	
Query understanding data:	
Ranking method:	uses traditional unsupervised sparse retrieval (e.g. QL, BM25, etc.); performs re-ranking with a pre-trained neural language model (BERT, Roberta, T5, etc.)
Ranking data:	is trained with TREC Deep Learning Track and/or MS MARCO dataset
Initiative:	does not treat them specially
Response generation:	uses single source (single passage)

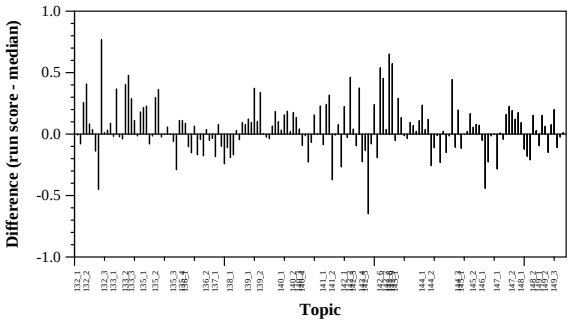
Note: all P/R scores follow the strict criterion $rel >= 2$***. NDCG gains are the rel values.***

Overall measures	
Number of turns	163
Total number retrieved	163000
Total relevant	12318
Total relevant retrieved	3480
MAP	0.1787
Mean NDCG@3	0.4054
Mean NDCG@20	0.3651
Mean Precision@1	0.4479
Mean Precision@3	0.4458

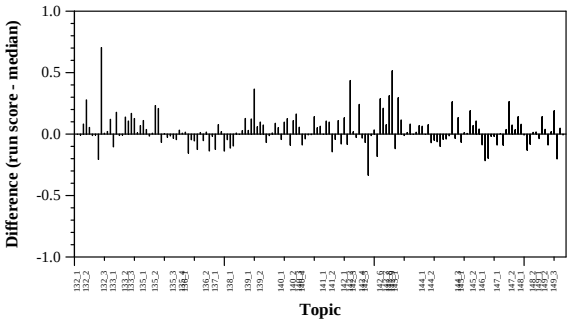
Document Level Averages	
	Precision
At 5 docs	0.4258
At 10 docs	0.3791
At 20 docs	0.3150

R-Precision	
Exact	0.2269





Difference from median NDCG



Difference from median AP