

**Table 1.1:** Examples of Text-Hypothesis pairs, taken from the RTE-2 development set, along with the application (Task) from which they were derived and the human annotation of whether the pair satisfies the entailment relation or not (Judgment). SUM=summarization; IR=information retrieval; IE=information extraction; QA=question answering. See Section 1.4 for the data-set generation methodology

ID	Text	Hypothesis	Task	Judgment
77	Google and NASA announced a working agreement, Wednesday, that could result in the Internet giant building a complex of up to 1 million square feet on NASA-owned property, adjacent to Moffett Field, near Mountain View.	Google may build a campus on NASA property.	SUM	YES
110	Drew Walker, NHS Tayside’s public health director, said: “It is important to stress that this is not a confirmed case of rabies.”	A case of rabies was confirmed.	IR	NO
294	Meanwhile, in an exclusive interview with a TIME journalist, the first one-on-one session given to a Western print publication since his election as president of Iran earlier this year, Ahmadinejad attacked the “threat” to bring the issue of Iran’s nuclear activity to the UN Security Council by the US, France, Britain and Germany.	Ahmadinejad is a citizen of Iran.	IE	YES
387	About two weeks before the trial started, I was in Shapiro’s office in Century City.	Shapiro works in Century City.	QA	YES
415	The drugs that slow down or halt Alzheimer’s disease work best the earlier you administer them.	Alzheimer’s disease is treated using drugs.	IR	YES
691	Arabic, for example, is used densely across North Africa and from the Eastern Mediterranean to the Philippines, as the key language of the Arab world and the primary vehicle of Islam.	Arabic is the primary language of the Philippines.	QA	NO

[Dagan *et al.* (2013)]