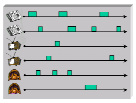
National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce

TDT2000 Link Detection Evaluation Results

Jonathan Fiscus and George Doddington

Spoken Natural Language Processing Group

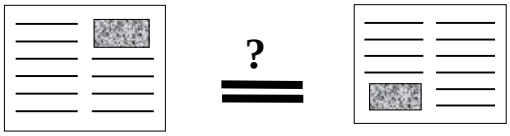


TDT Link Detection Task

One Participant: University of Massachusetts

System Goal:

- **To detect whether a pair of stories discuss the same topic.**
  (Can be thought of as a "primitive operator" to build a variety of applications)

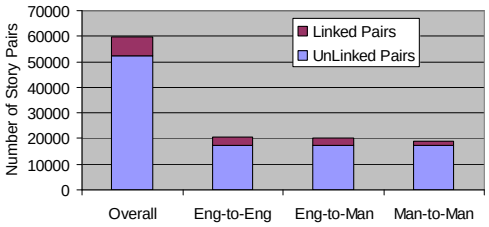


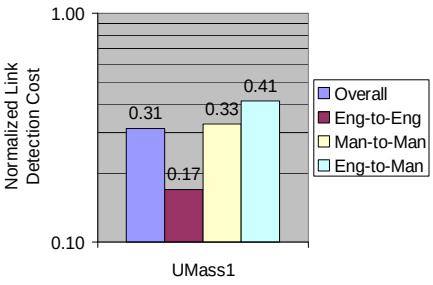
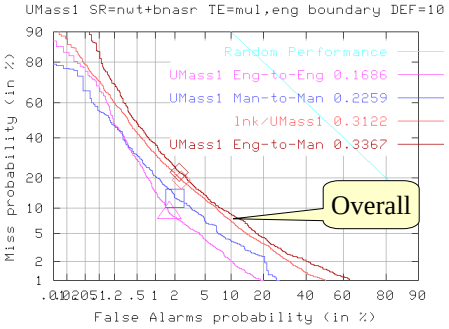
NIST Speech Group

2000 Link Detection Results

- A lot was learned last year:
  - The test set must be properly sampled
    - "Linked" story pairs were selected by randomly sampling all possible on-topic story pairs
    - "Unlinked" pairs were selected using all on-topic stories as one of the pair, and a randomly chosen story was chosen as the second
  - This year, the task was made multilingual
  - More story pairs were used

Link Detection Test Set Composition



Link Detection Results

- Required Condition
  - Multilingual texts
  - Newswire + Broadcast News ASR,
  - Reference story boundaries
  - 10 file decision deferral