

Relationship Pilot

- Summary from analyst breakout (kickoff)
 - relationship ~ ability of one object to influence another
 - determine how the influence occurs
 - determine why the influence occurs (motivation)
 - multiple spheres of influence
 - financial movement of goods family
 - communication organizational location
 - temporal common interests
 - need to recognize when support is lacking
 - doesn't exist? misinformation? missing?

Summary, cont'd

- Relationships of interest highly dependent on analyst, situation, and purpose
- Some existing technology addresses problem
 - link analysis tools
 - information extraction for known relationship types

Pilot task

- Use topics such as those used in TREC-1 to set context. Topic explicitly says which relation type is of interest with respect to the broad categories defined before. Question is either a yes/no question with the expectation that system will support answer, or an explicit request for evidence.
 - What financial dealings has John Smith had with XYZ Corporation?
 - Did Jane Doe meet with Tom Buck in August 2000?
 - What were the movements of John Doe in the month before the ZXY bombing?

Evaluation

- Analogous to Definition Pilot
 - human assessor creates list of known support
 - assessor marks presence of support in each system's response
 - calculate concept recall over vital concepts
 - compute concept precision as inverse function of response length
 - combine recall and precision using F with emphasis on recall
- As always, all judgments are in the opinion of the assessor, who has final say

Schedule

- Analysts created topics in June
 - 13 analysts from military
 - 1 NSA/NIST analyst
- 5 sample topics released July 2
 - samples consisted of statement & analyst prose answer (i.e., not nuggets)
- 50 test topics released mid-July
- Result returned to NIST by Aug 31
- Prose answers returned early September; nuggets and scored results returned 9/25

Sample Topics

- The analyst is interested in information regarding the Nobel Prize winners from previous years. Records indicate that David Trimble and John Hume shared the Nobel Peace Prize in 1998. Who are Trimble and Hume, and what was their relationship?
- The analyst is interested in proliferation of various weapons systems on a global scale. Specifically, the analyst is looking for proliferation of diesel submarines and what countries are obtaining the capabilities of extremely quiet submarine operation. What countries have imported the Russian produced Kilo SS submarine?

Sample Topic with Nuggets

- The analyst is interested in the status of Gibraltar. Specifically, the analyst wants to know who currently controls Gibraltar and whether there is any on-going controversy or or disagreement over its status.
 - ∇ controlled by Britian
 - ∇ Spain wants to rule Gibraltar
 - Gibraltar formerly belonged to Spain
 - Spain and Britian vie for control

Participation

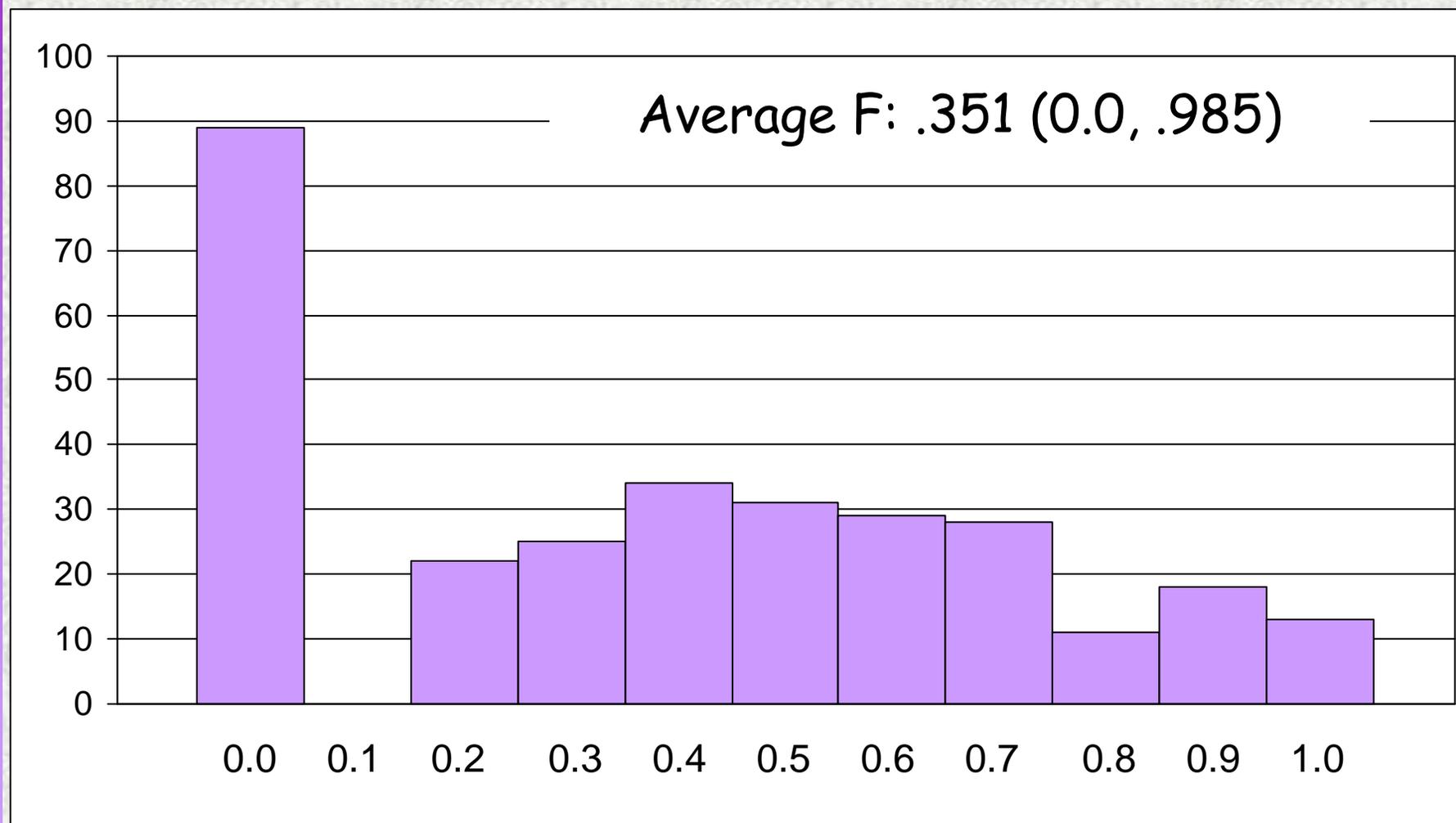
- 6 runs from 4 contractors
 - BBN
 - LCC
 - CMU
 - MIT

Special thanks to BBN who was "volunteered" at the last minute and produced runs in slightly more than a day

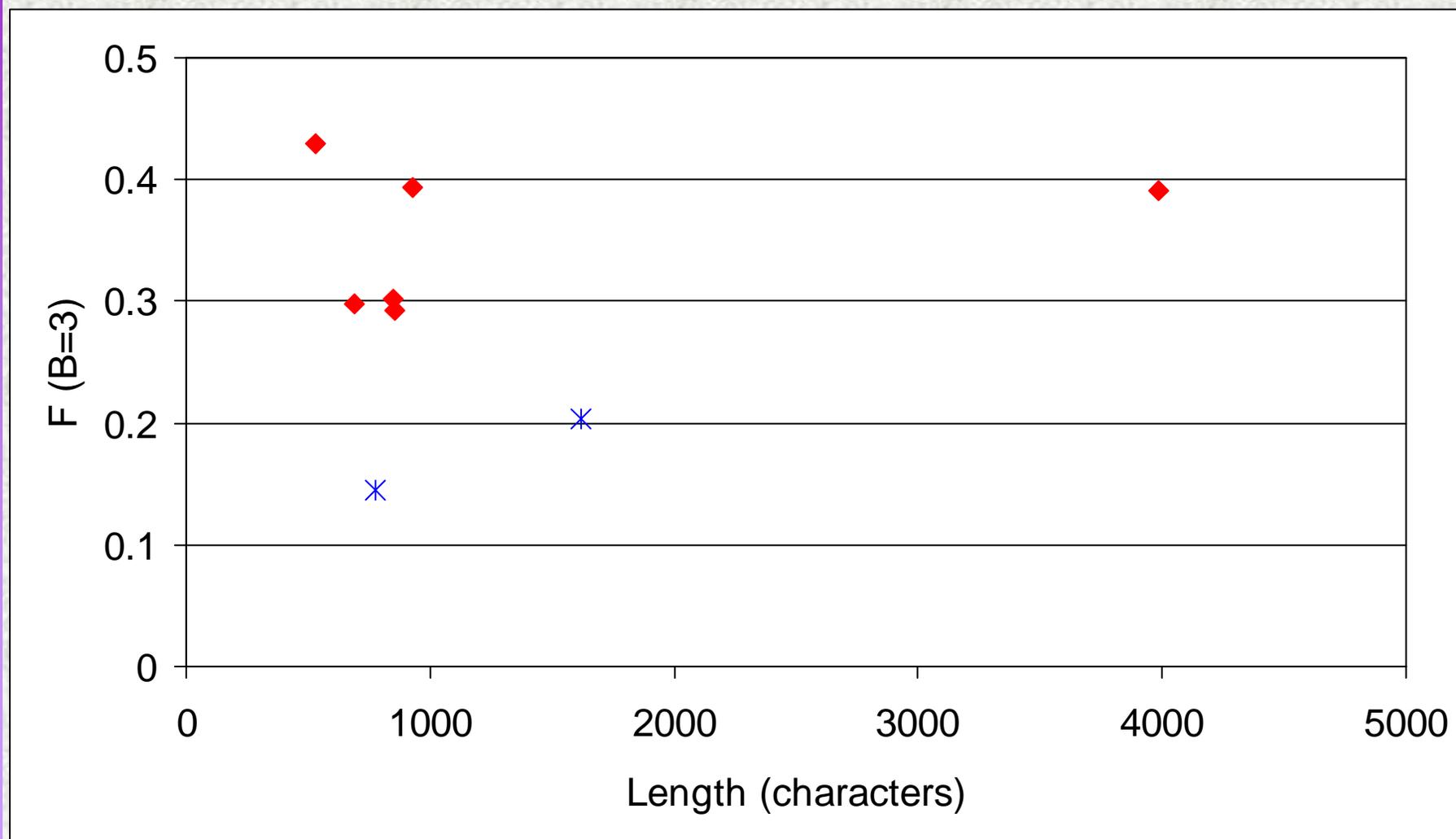
Nugget Assignment

- Responses judged by "nugget" approach
 - 6 analysts did judging
- Distinction between vital and okay nuggets
 - vital nuggets only participate in recall
 - okay nuggets don't harm precision
- Averages across 50 topics
 - vital: 3.0 (1, 10)
 - total: 6.7 (1,17)

Histogram of F Scores



F Scores vs. Length



Now What?

- Were the questions what were expected?
- Do they represent "relationships"?
- Did the scores correlate (tightly) with quality?
- What's missing in the systems?