Problem Statement

- Text summarization is a problem supporting information assurance
- To develop a new algorithm that provides useful, accurate summaries, I plan to ontologically compare the first paragraph of an American news report, which can be viewed as a summary written by journalists, to a computer-generated summary of rest of the article.

Ways to Approach Automatic Summarization

**Previous Attempts – Extractive**

Use a subset of words, phrases, or sentences from the original text

- **FLAWS**
  - Difficult to remove overlapping sentences
  - “Alice ate lunch with Greg at Gino’s.”
  - “Alice ordered a steak, and Greg ordered a hamburger at Gino’s.”
  - “Alice and Greg ordered entrees at Gino’s.”
  - Inaccurate frequency analysis
  - “Allan Smith robbed a 7-11 in Tallahassee, Florida.”
  - “The man entered the convenience store late at night.”

**Ontological Semantics Technology**

- Condensed TMR from Anthrax Article [3]

**My Solution – Abstractive**

Build an internal text meaning representation (TMR) then construct a summary closer to what a human may generate

- Basics of Algorithm
  - Weigh edges by the frequency of their use in text
  - Weigh vertices by the summation of their edge weights
  - For each vertex:
    - Calculate the cost of each path leading to it, where each edge of the path has cost 1/(edge weight)
    - Remove edges corresponding to inefficient paths
  - Remove remaining lowly-weighted vertices with out-degree 0

**Graphical Representation of News Article**

- Vertices represent sentences
- Edge between two vertices means that the two sentences contain the same words
- Algorithm uses high-degree vertices in summary.

---

**Citations**

