NLP
Introduction to NLP

Summarization Techniques 2/3
Mani/Bloedorn (1997, 1999)

- Graph-based method for identifying similarities and differences between documents
- Single event or sequence of events
- Content entities (nodes) and relations (edges)
- Relation types
  - SAME, ADJACENT, ALPHA (WordNet, NetOwl), COREF
- Spreading activation
Mr. Kenny is the person that invented the anesthetic machine which uses micro-computers to control the rate at which an anesthetic is pumped into the blood. Such machines are nothing new. But his device uses two micro-computers to achieve much closer monitoring of the pump feeding the anesthetic into the patient.
Barzilay and M. Elhadad (1997)

- WordNet-based
- Three types of relations:
  - extra-strong (repetitions)
  - strong (WordNet relations)
  - medium-strong (link between synsets is longer than one + some additional constraints)
Scoring chains:

- Length
- Homogeneity index:
  \[= 1 - \left(\frac{\text{# distinct words in chain}}{\text{length}}\right)\]

Score = Length * Homogeneity

Score > Average + 2 * st.dev.

Marcu (1997-1999)

• Focuses on text coherence
• Based on Rhetorical Structure Theory (Mann and Thompson 1988)
  – Nucleus+Satellite

[The truth is that the pressure to smoke in junior high is greater than it will be any other time of one’s life:]_N
[we know that 3,000 teens start smoking each day.]_S

• Example: evidence
  – N+S combination increases reader’s belief in N
| With its distant orbit (50 percent farther from the sun than Earth) and slim atmospheric blanket, | Mars experiences frigid weather conditions (2) | Surface temperatures typically average about -60 degrees Celsius (-76 degrees Fahrenheit) at the equator and can dip to -123 degrees C near the poles (3) | Only the midday sun at tropical latitudes is warm enough to thaw ice on occasion, (4) but any liquid water formed in this way would evaporate almost instantly (5) | because of the low atmospheric pressure (6) | Although the atmosphere holds a small amount of water, and water-ice clouds sometimes develop, (7) Most Martian weather involves blowing dust and carbon monoxide. (8) | Each winter, for example, a blizzard of frozen carbon dioxide rages over one pole, and a few meters of this dry-ice snow accumulate as previously frozen carbon dioxide evaporates from the opposite polar cap. (9) | Yet even on the summer pole, where the sun remains in the sky all day long, temperatures never warm enough to melt frozen water. (10) |
Mars experiences frigid weather conditions (2).

Surface temperatures typically average about -60 degrees Celsius (-76 degrees Fahrenheit) at the equator and can dip to -123 degrees C near the poles (3).

Although the atmosphere holds a small amount of water, and water-ice clouds sometimes develop, (4) Most Martian weather involves blowing dust and carbon monoxide. (8)

Each winter, for example, a blizzard of frozen carbon dioxide rages over one pole, and a few meters of this dry-ice snow accumulate as previously frozen carbon dioxide evaporates from the opposite polar cap. (9)

Yet even on the summer pole, where the sun remains in the sky all day long, temperatures never warm enough to melt frozen water. (10)
Mars experiences frigid weather conditions (2). Surface temperatures typically average about -60 degrees Celsius (-76 degrees Fahrenheit) at the equator and can dip to -123 degrees C near the poles (3).

Although the atmosphere holds a small amount of water, and water-ice clouds sometimes develop, Most Martian weather involves blowing dust and carbon monoxide. (7) Each winter, for example, a blizzard of frozen carbon dioxide rages over one pole, and a few meters of this dry-ice snow accumulate as previously frozen carbon dioxide evaporates from the opposite polar cap. (8) Yet even on the summer pole, where the sun remains in the sky all day long, temperatures never warm enough to melt frozen water. (10)

Only the midday sun at tropical latitudes is warm enough to thaw ice on occasion, (4) but any liquid water formed in this way would evaporate almost instantly because of the low atmospheric pressure (6).
Mars experiences frigid weather conditions (2).

Surface temperatures typically average about -60 degrees Celsius (-76 degrees Fahrenheit) at the equator and can dip to -123 degrees C near the poles (3).

Although the atmosphere holds a small amount of water, and water-ice clouds sometimes develop, (7)

Most Martian weather involves blowing dust and carbon monoxide. (8)

Each winter, for example, a blizzard of frozen carbon dioxide rages over one pole, and a few meters of this dry-ice snow accumulate as previously frozen carbon dioxide evaporates from the opposite polar cap. (9)

Yet even on the summer pole, where the sun remains in the sky all day long, temperatures never warm enough to melt frozen water. (10)

With its distant orbit (50 percent farther from the sun than Earth) and slim atmospheric blanket, (1)

Only the midday sun at tropical latitudes is warm enough to thaw ice on occasion, (4)

but any liquid water formed in this way would evaporate almost instantly (5)

because of the low atmospheric pressure (6)

Evidence Cause

Contrast

Elaboration

Example

Concession

Antithesis
Mars experiences frigid weather conditions (2).

Surface temperatures typically average about -60 degrees Celsius (-76 degrees Fahrenheit) at the equator and can dip to -123 degrees C near the poles (3).

Contrast

Although the atmosphere holds a small amount of water, and water-ice clouds sometimes develop, (7)

Most Martian weather involves blowing dust and carbon monoxide. (8)

Evidence Cause

Every winter, for example, a blizzard of frozen carbon dioxide rages over one pole, and a few meters of this dry-ice snow accumulate as previously frozen carbon dioxide evaporates from the opposite polar cap. (9)

Yet even on the summer pole, where the sun remains in the sky all day long, temperatures never warm enough to melt frozen water. (10)

With its distant orbit (50 percent farther from the sun than Earth) and slim atmospheric blanket, (1)

Only the midday sun at tropical latitudes is warm enough to thaw ice on occasion, (4)

but any liquid water formed in this way would evaporate almost instantly (5)

because of the low atmospheric pressure (6)
Noisy Channel Models

- Source/target language
- Coding process
Berger and Mittal 2000

• Source language
  – full document

• Target language
  – Summary

• Gisting (OCELOT)

\[ g^* = \arg\max_g p(g|d) = \arg\max_g p(g) \cdot p(d|g) \]
Berger & Mittal 2000

- Training on 100K summary+document pairs
- Testing on 1046 pairs
- Use Viterbi–type search
- Evaluation: word overlap (0.2–0.4)
- No word ordering
Audubon society atlanta area savannah georgia chatham and local birding savannah keepers chapter of the audubon georgia and leasing
Carbonell and Goldstein (1998)

• Maximal marginal relevance
• Greedy selection method
• Query-based summaries
• Law of diminishing returns

\[
\text{MMR} = \arg\max_{D_i \in R \setminus S} \left[ \lambda \left( \text{Sim}_1(D_i, Q) \right) - (1-\lambda) \max_{D_j \in S} \text{Sim}_2(D_i, D_j) \right]
\]

\[C = \text{doc collection}\]
\[Q = \text{user query}\]
\[R = \text{IR}(C, Q, \theta)\]
\[S = \text{already retrieved documents}\]
\[\text{Sim} = \text{similarity metric used}\]
Mead (Radev et al. 2000)

• Salience-based extractive summarization
• Centroid-based method
• Vector space model
• Features: position, length, centroid
• Reranker – similar to MMR
• Open source library–
  www.summarization.com/mead
\[ \cos(\bar{x}, \bar{y}) = \frac{\bar{x} \cdot \bar{y}}{|\bar{x}| |\bar{y}|} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2}} \]
Mead

• **Input**
  – Cluster of $d$ documents with $n$ sentences (compression rate = $r$)

• **Output**
  – $(n \times r)$ sentences from the cluster with the highest scores

$$SCORE (s) = \sum_i (w_c C_i + w_p P_i + w_f F_i)$$
NewsInEssence (Radev et al. 2001)

• Web-based multi-document news summarization system
A man arrested in Sweden on suspicion of attempting to hijack a plane for British has denied any knowledge of a gun in his luggage. The man was arrested at Vasteras airport, 100km (62 miles) north-west of Stockholm, after allegedly being caught trying to board the plane with a gun. **Summary**

A Swedish airport authority has praised for their security measures after a man was arrested on suspicion of attempting to hijack a plane, when trying to board a plane to London carrying a handbag. A 29-year-old Tunisian terrorist is being held by Swedish police after he allegedly attempted to board the plane at Vasteras Airport (100 kilometers (62 miles) northwest of the capital, Stockholm, with a gun in his handbag. **Police** are looking into a preliminary charge of planning to hijack a plane.

**Swede charged with planes to hijack plane**

Produced on 08/30, 01:52 PM

89% Summary

An organizer of an Islamic conference in Birmingham says it is "highly unlikely" that the man charged with trying to hijack a Ryanair flight in Sweden, was due to attend the conference. [2] Security officers at Stockholm's Vasteras Airport say they found a handbag in a toilet bag when they scanned the 30-year-old man's hand luggage. [2] The man - who was born in Sweden to Tunisian parents - was trying to board a Ryanair flight to London's Stansted airport. [2] Swedish police say he was with a party traveling to an Islamic conference in Birmingham and believe the man was going to hijack the plane. [2] A man arrested in Sweden on suspicion of attempting to hijack a plane for Britain has denied any knowledge of a gun in his luggage. The man was arrested at Vasteras airport, 100km (62 miles) north-west of Stockholm, after allegedly being caught trying to board the plane with a gun. **Summary**

A Swedish airport authority has praised for their security measures after a man was arrested on suspicion of attempting to hijack a plane, when trying to board a plane to London carrying a handbag. A 29-year-old Tunisian terrorist is being held by Swedish police after he allegedly attempted to board the plane at Vasteras Airport (100 kilometers (62 miles) northwest of the capital, Stockholm, with a gun in his handbag. **Police** are looking into a preliminary charge of planning to hijack a plane.

**Swede charged with planes to hijack plane**

Produced on 08/30, 01:52 PM

89% Summary

An organizer of an Islamic conference in Birmingham says it is "highly unlikely" that the man charged with trying to hijack a Ryanair flight in Sweden, was due to attend the conference. [2] Security officers at Stockholm's Vasteras Airport say they found a handbag in a toilet bag when they scanned the 30-year-old man's hand luggage. [2] The man - who was born in Sweden to Tunisian parents - was trying to board a Ryanair flight to London's Stansted airport. [2] Swedish police say he was with a party traveling to an Islamic conference in Birmingham and believe the man was going to hijack the plane. [2] A man arrested in Sweden on suspicion of attempting to hijack a plane for Britain has denied any knowledge of a gun in his luggage. The man was arrested at Vasteras airport, 100km (62 miles) north-west of Stockholm, after allegedly being caught trying to board the plane with a gun. **Summary**

A Swedish airport authority has praised for their security measures after a man was arrested on suspicion of attempting to hijack a plane, when trying to board a plane to London carrying a handbag. A 29-year-old Tunisian terrorist is being held by Swedish police after he allegedly attempted to board the plane at Vasteras Airport (100 kilometers (62 miles) northwest of the capital, Stockholm, with a gun in his handbag. **Police** are looking into a preliminary charge of planning to hijack a plane.
Other News Summarization Systems

• Newsblaster (McKeown et al. 2002)
• Google News
NLP