

NLP

Introduction to NLP

Summarization Evaluation

Evaluation Criteria

- Appropriate length
- Fidelity
- Salience
- Grammaticality
- Non-redundancy
- Referential well-formedness
- Structure and coherence

Ideal Evaluation

Information content

$$C = \text{Compression Ratio} = \frac{|S|}{|D|}$$

$$R = \text{Retention Ratio} = \frac{i(S)}{i(D)}$$

$$R > C$$

Types Of Evaluation Methods

- **Extrinsic techniques (task-based)**
 - Can you make the same decision using the summary as with the full text, but in less time
- **Intrinsic techniques**
 - Comparing summaries against gold standards

Precision and Recall

	Relevant	Non-relevant
System: relevant	A	B
System: non-relevant	C	D

$$\text{Precision: } P = \frac{A}{A + B}$$

$$\text{Recall: } R = \frac{A}{A + C}$$

$$F = \frac{2PR}{(P + R)}$$

Rouge (Lin and Hovy 2003)

- Very convenient for prototyping
- Based on Papineni et al's BLEU (used for MT)
- R stands for Recall
- ROUGE-n is a measure of n-gram overlap between a summary and a set of reference summaries
- ROUGE-L uses longest common subsequence instead of n-gram overlap
- Correlated with manual evaluations when averaged over many examples
- It can be easily gamed

Relative Utility (Radev et al. 2000)

	Ideal	System 1	System 2
S1	+	+	-
S2	+	+	+
S3	-	-	-
S4	-	-	+
S5	-	-	-
S6	-	-	-
S7	-	-	-
S8	-	-	-
S9	-	-	-
S10	-	-	-

	Ideal	System 1	System 2
S1	10(+)	10(+)	5
S2	8(+)	9(+)	8(+)
S3	2	3	4
S4	7	6	9(+)

$U(\text{system}, \text{ideal}) = \text{\% of ideal utility covered by system summary}$

Relative Utility

$$RU = \overline{17}$$

	<i>Judge1</i>	<i>Judge2</i>	<i>Judge3</i>
Sentence 1	10	10	5
Sentence 2	8	9	8
Sentence 3	2	3	4
Sentence 4	5	6	9

Relative Utility

$$RU = \frac{13}{17} = 0.765$$

	<i>Judge1</i>	<i>Judge2</i>	<i>Judge3</i>
Sentence 1	10	10	5
Sentence 2	8	9	8
Sentence 3	2	3	4
Sentence 4	5	6	9

Normalized System Performance

	Judge 1	Judge 2	Judge 3	Average
Judge 1	1.000	1.000	0.765	0.883
Judge 2	1.000	1.000	0.765	0.883
Judge 3	0.722	0.789	1.000	0.756

Normalized system performance

$$D = \frac{(S-R)}{(J-R)}$$

System performance

Random performance

Interjudge agreement

Random Performance

average of all $\frac{n!}{(n(1-r))! (r*n)!}$ systems

$$D = \frac{(S-R)}{(J-R)}$$

↑

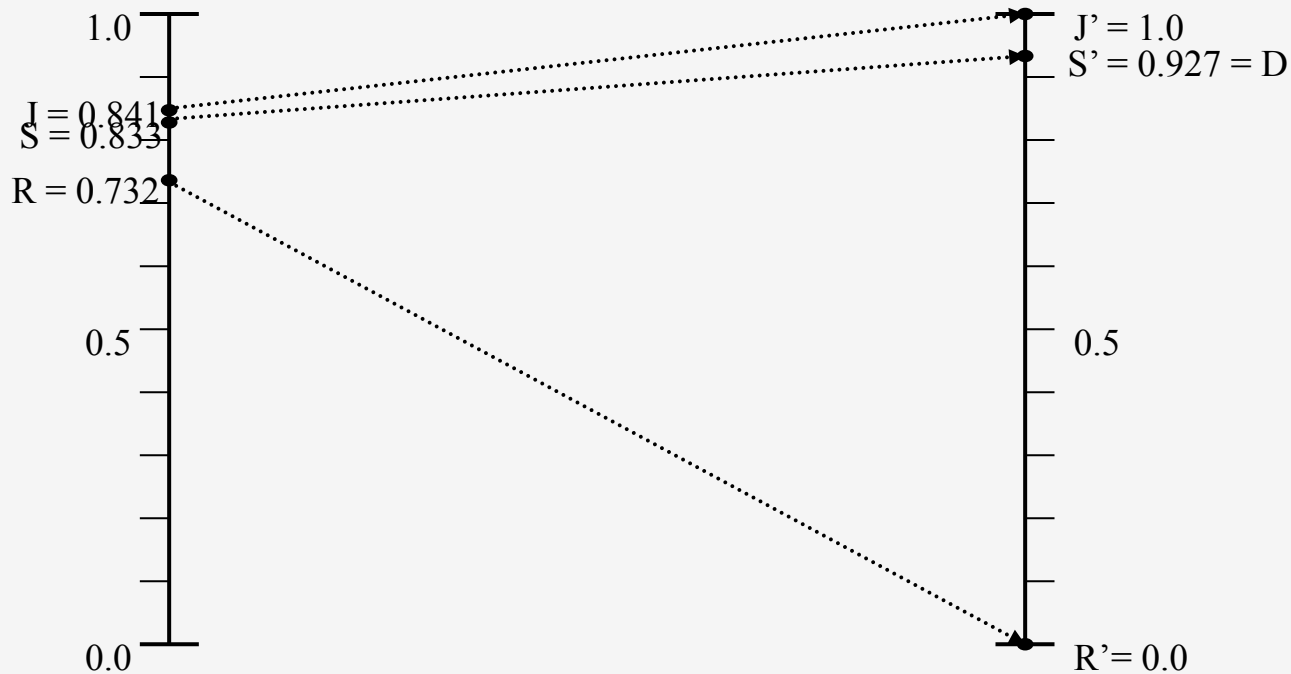
{12}
{13}
{14}
{23}
{24}
{34}

Example

$$D_{\{14\}} = \frac{(S-R)}{(J-R)} = \frac{0.833 - 0.732}{0.841 - 0.732} = 0.927$$

$$D_{\{24\}} = 0.963$$

Normalized Evaluation of {14}



ARTICLE 18853: ALGIERS, May 20 (AFP)

1. Eighteen decapitated bodies have been found in a mass grave in northern Algeria, press reports said Thursday, adding that two shepherds were murdered earlier this week.
2. Security forces found the mass grave on Wednesday at Chbika, near Djelfa, 275 kilometers (170 miles) south of the capital.
3. It contained the bodies of people killed last year during a wedding ceremony, according to Le Quotidien Liberte.
4. The victims included women, children and old men.
5. Most of them had been decapitated and their heads thrown on a road, reported the Es Sahafa.
6. Another mass grave containing the bodies of around 10 people was discovered recently near Algiers, in the Eucalyptus district.
7. The two shepherds were killed Monday evening by a group of nine armed Islamists near the Moulay Slissen forest.
8. After being injured in a hail of automatic weapons fire, the pair were finished off with machete blows before being decapitated, Le Quotidien d'Oran reported.
9. Seven people, six of them children, were killed and two injured Wednesday by armed Islamists near Medea, 120 kilometers (75 miles) south of Algiers, security forces said.
10. The same day a parcel bomb explosion injured 17 people in Algiers itself.
11. Since early March, violence linked to armed Islamists has claimed more than 500 lives, according to press tallies.

ARTICLE 18854: ALGIERS, May 20 (UPI)

1. Algerian newspapers have reported that 18 decapitated bodies have been found by authorities in the south of the country.
2. Police found the "decapitated bodies of women, children and old men, with their heads thrown on a road" near the town of Jelfa, 275 kilometers (170 miles) south of the capital Algiers.
3. In another incident on Wednesday, seven people -- including six children -- were killed by terrorists, Algerian security forces said.
4. Extremist Muslim militants were responsible for the slaughter of the seven people in the province of Medea, 120 kilometers (74 miles) south of Algiers.
5. The killers also kidnapped three girls during the same attack, authorities said, and one of the girls was found wounded on a nearby road.
6. Meanwhile, the Algerian daily Le Matin today quoted Interior Minister Abdul Malik Silal as saying that "terrorism has not been eradicated, but the movement of the terrorists has significantly declined."
7. Algerian violence has claimed the lives of more than 70,000 people since the army cancelled the 1992 general elections that Islamic parties were likely to win.
8. Mainstream Islamic groups, most of which are banned in the country, insist their members are not responsible for the violence against civilians.
9. Some Muslim groups have blamed the army, while others accuse "foreign elements conspiring against Algeria."

ARTICLE 18853: ALGIERS, May 20 (AFP)

1. Eighteen decapitated bodies have been found in a mass grave in northern Algeria, press reports said Thursday, adding that two shepherds were murdered earlier this week.

2. Security forces found the mass grave on Wednesday at Chbika, near Djelfa, 275 kilometers (170 miles) south of the capital.

3. It contained the bodies of people killed last year during a wedding ceremony, according to Le Quotidien Liberte.

4. The victims included women, children and old men.

5. Most of them had been decapitated and their heads thrown on a road, reported the Es Sahafa.

6. Another mass grave containing the bodies of around 10 people was discovered recently near Algiers, in the Eucalyptus district.

7. The two shepherds were killed Monday evening by a group of nine armed Islamists near the Moulay Slissen forest.

8. After being injured in a hail of automatic weapons fire, the pair were finished off with machete blows before being decapitated, Le Quotidien d'Oran reported.

9. Seven people, six of them children, were killed and two injured Wednesday by armed Islamists near Medea, 120 kilometers (75 miles) south of Algiers, security forces said.

10. The same day a parcel bomb explosion injured 17 people in Algiers itself.

11. Since early March, violence linked to armed Islamists has claimed more than 500 lives, according to press tallies.

ARTICLE 18854: ALGIERS, May 20 (UPI)

1. Algerian newspapers have reported that 18 decapitated bodies have been found by authorities in the south of the country.

2. Police found the "decapitated bodies of women, children and old men, with their heads thrown on a road" near the town of Jelfa, 275 kilometers (170 miles) south of the capital Algiers.

3. In another incident on Wednesday, seven people -- including six children -- were killed by terrorists, Algerian security forces said.

4. Extremist Muslim militants were responsible for the slaughter of the seven people in the province of Medea, 120 kilometers (74 miles) south of Algiers.

5. The killers also kidnapped three girls during the same attack, authorities said, and one of the girls was found wounded on a nearby road.

6. Meanwhile, the Algerian daily Le Matin today quoted Interior Minister Abdul Malik Silal as saying that "terrorism has not been eradicated, but the movement of the terrorists has significantly declined."

7. Algerian violence has claimed the lives of more than 70,000 people since the army cancelled the 1992 general elections that Islamic parties were likely to win.

8. Mainstream Islamic groups, most of which are banned in the country, insist their members are not responsible for the violence against civilians.

9. Some Muslim groups have blamed the army, while others accuse "foreign elements conspiring against Algeria."

Subsumption and Equivalence

- *Subsumption*: If the information content of sentence a (denoted as $I(a)$) is contained within sentence b, then a becomes informationally redundant and the content of b is said to subsume that of a:

$$I(a) \subset I(b)$$

- *Equivalence*: If $I(a) \subset I(b) \wedge I(b) \subset I(a)$

Subsumption

- (1) John Doe was found guilty of the murder.
- (2) The court found John Doe guilty of the murder of Jane Doe last August and sentenced him to life.

Subsumption

	Article 1	Article 2	Article 3
S1	10 —→ 10		5
S2	8	9	8
S3	2	3	4
S4	7	6 —→ 9	

Subsumption (Cont'd)

$$SCORE(s) = \sum_i (w_c C_i + w_p P_i + w_f F_i) - w_R R_s$$

R_s = *cross-sentence word overlap*

$$R_s = 2 * (\# \text{ overlapping words}) / (\# \text{ words in sentence 1} + \# \text{ words in sentence 2})$$

$$w_R = \text{Max}_s (SCORE(s))$$

Pyramid

- Nenkova and Passonneau 2004
- Based on Semantic Content Units (SCU)
- Used for multi-document summarization
- Different surface realizations with equivalent meanings

Pyramid

A1 In 1998 two Libyans indicted in 1991 for the Lockerbie bombing were still in Libya.

B1 Two Libyans were indicted in 1991 for blowing up a Pan Am jumbo jet over Lockerbie, Scotland in 1988.

C1 Two Libyans, accused by the United States and Britain of bombing a New York bound Pan Am jet over Lockerbie, Scotland in 1988, killing 270 people, for 10 years were harbored by Libya who claimed the suspects could not get a fair trial in America or Britain.

D2 Two Libyan suspects were indicted in 1991.

SCU1 ($w=4$): two Libyans were officially accused of the Lockerbie bombing

A1 [two Libyans]1 [indicted]1

B1 [Two Libyans were indicted]1

C1 [Two Libyans,]1 [accused]1

D2 [Two Libyan suspects were indicted]1

SCU2 ($w=3$): the indictment of the two Lockerbie suspects was in 1991

A1 [in 1991]2

B1 [in 1991]2

D2 [in 1991.]2

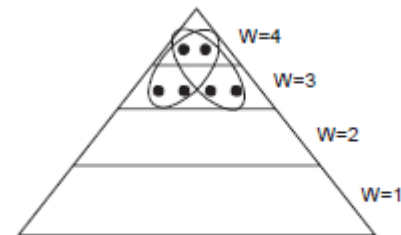


Figure 2: Two of six optimal summaries with 4 SCUs

Available Corpora

- DUC and TAC corpora
 - <http://duc.nist.gov>
 - <http://tac.nist.gov>
- SummBank corpus
 - <http://www.summarization.com/summbank>
- SUMMAC corpus
- NY Times corpus (from LDC)

NLP