Screenplay to Plot Summary Alignment

Jennifer Gillenwater Ben Taskar

Computer & Information Science
University of Pennsylvania

May 11, 2009

Motivation

- Searching through film is difficult
- Screenplays are more easily searchable media
- But queries are usually more similar to a summary
- Ability to align screenplays to plot summaries is a first step towards tractable film searching

Fantastic Four Example

He is about to deal the finishing strike, when a recovered Reed and Susan arrive to save Ben.

EXT. NEW YORK CITY STREET - SAME TIME

On the street, cars SCREAM to stops, people GASP. Doom strides up to Ben, and raises the ELECTRICITY POLE for the death-blow. Currents surge. Right before he swings down –

VOICE (O.S.): I can't let you do that.

Simple, strong. Doom turns to see ... REED. Alone. Bruised and battered. Slowly stepping toward him. Doom smiles.

...

Data

Movie	**Scenes**	**Sentences**
Batman	**374**	**23**
Fargo	126	**55**
Good Will Hunting	124	35
Jerry Maguire	226	37
O Brother Where Art Thou	**69**	44
One Flew Over the Cuckoo's Nest	139	36
The Fantastic Four	180	50
You've Got Mail	209	34
Average	181	39

Gold Alignments - The Fantastic Four

Simple alignment - each summary sentence is assigned one scene

Figure 8: fantastic four

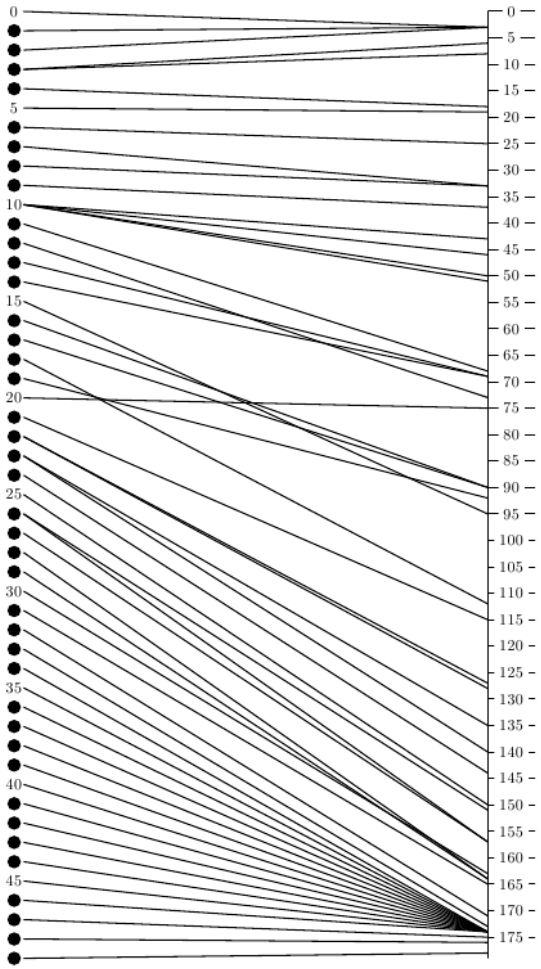


Figure 7: fantastic four

Gold Alignments - Batman

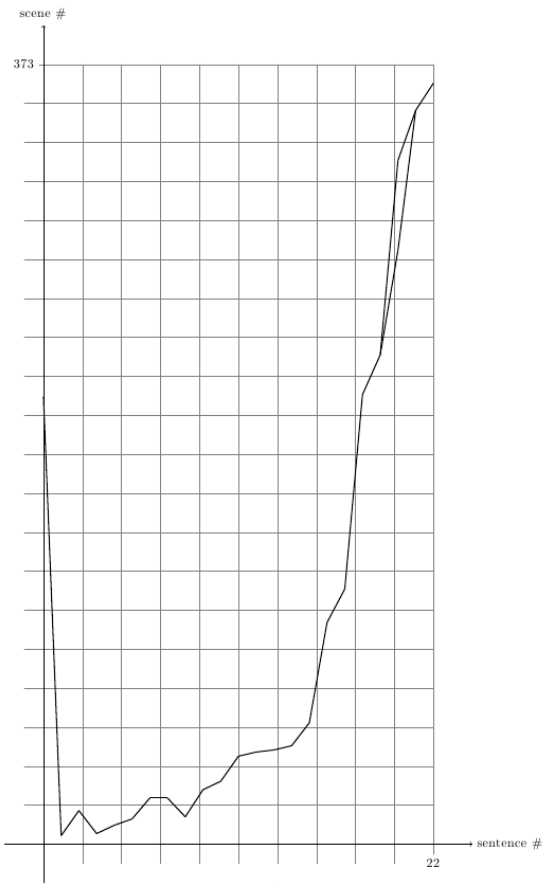


Figure 2: batman

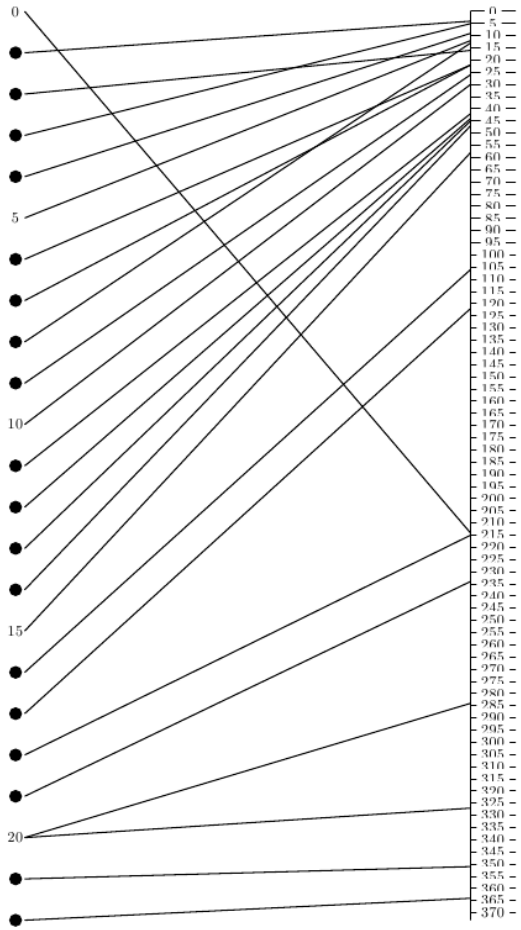
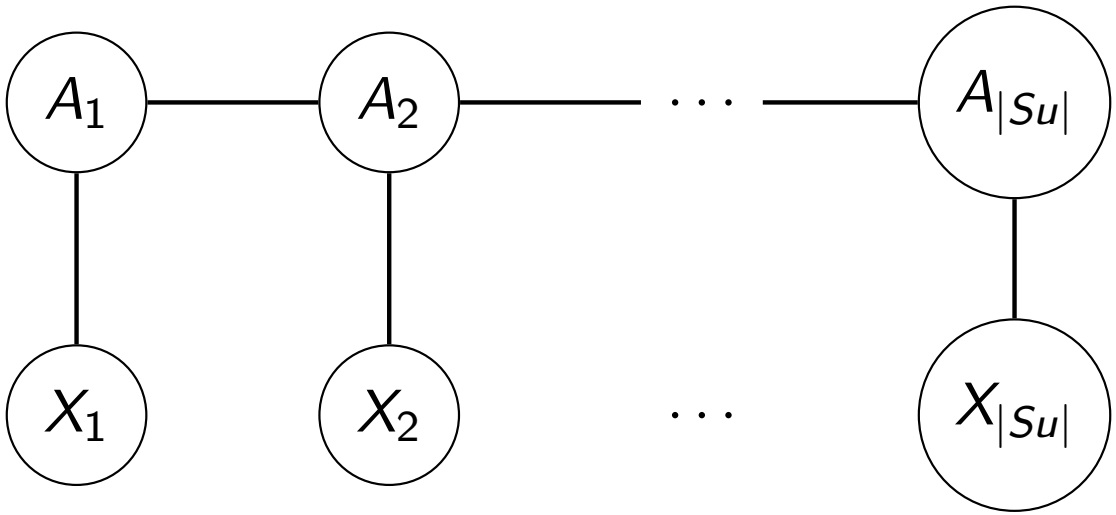


Figure 1: batman

Conditional Random Field



$$A_i = \text{scene aligned to } i^{th} \text{ summary sentence}$$

$$X_i = \{Su_i, Sc\} = \{i^{th} \text{ summary sentence, entire screenplay}\}$$

$$P(A \mid X) = \phi_{01}(A_1)\phi_1(A_1, X_1) \prod_{i=2}^{|Su|} \phi_{(i-1)i}(A_{i-1}, A_i)\phi_i(A_i, X_i)$$

Methods Overview

1. Supervised training - Learn weights α, β for

$$\phi_i(A_i, X_i) = e^{\alpha \cdot f(Su_i, Sc_{a_i})}$$
$$\phi_{(i-1)i}(A_{i-1}, A_i) = e^{\beta \cdot b(A_{i-1}, A_i)}$$

   where f, b are features and Sc_{a_i} is the scene aligned to summary sentence i

2. Testing
   1. Compute f for all $|Sc||Su|$ possible (A_i, X_i) pairs
   2. Compute b for all $|Sc|^2$ possible (A_{i-1}, A_i) pairs
   3. Use Viterbi algorithm to find best alignment
   4. Compare to gold alignment

Methods Specifics

- Weight learning - gradient descent on $\log P(A|X)$, using all movies except one heldout for testing
- Features
  - $\phi_i(A_i, X_i)$: noun overlap, verb overlap, adjective overlap, adverb overlap
  - $\phi_{(i-1)i}(A_{i-1}, A_i)$: temporal difference $A_{i-1} - A_i$
- Evaluation
  - Baselines - random, most common scene
  - Metrics - exact scene match, ± 5 scenes slackness

Evaluation Results

		Baselines	
Movie	CRF	Random	MCS
Batman	$5/23 = 22\%$	0.28	8.7
Fargo	$26/55 = 47\%$	0.85	7.3
Good Will Hunting	$9/35 = 26\%$	0.81	8.6
Jerry Maguire	$7/37 = 19\%$	0.45	5.4
Brother Where Art Thou	$20/44 = 45\%$	1.52	16
Cuckoo's Nest	$5/36 = 14\%$	0.78	11
The Fantastic Four	$17/50 = 34\%$	0.63	26
You've Got Mail	$7/34 = 26\%$	0.48	8.8

Predicted Alignments - The Fantastic Four

Gold

Figure 8: fantastic four

Predicted

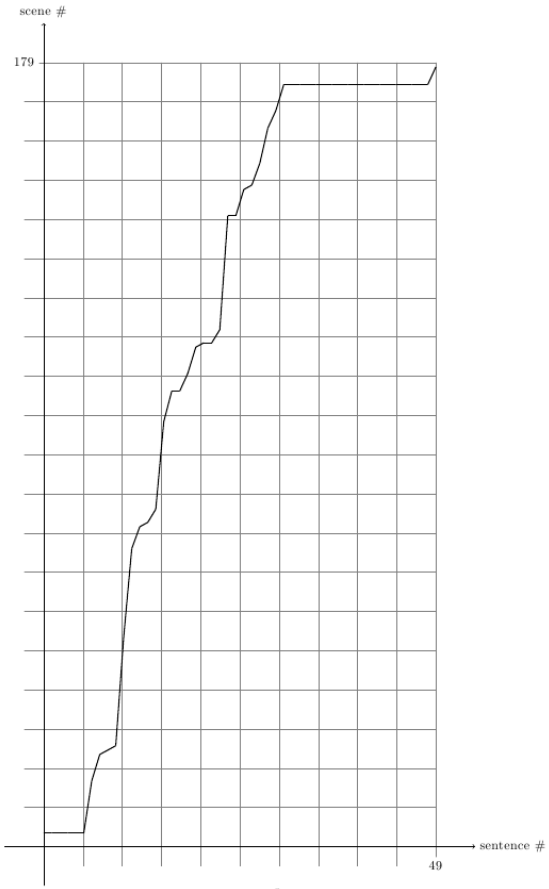


Figure 8: fantastic four

Predicted Alignments - Batman

Gold

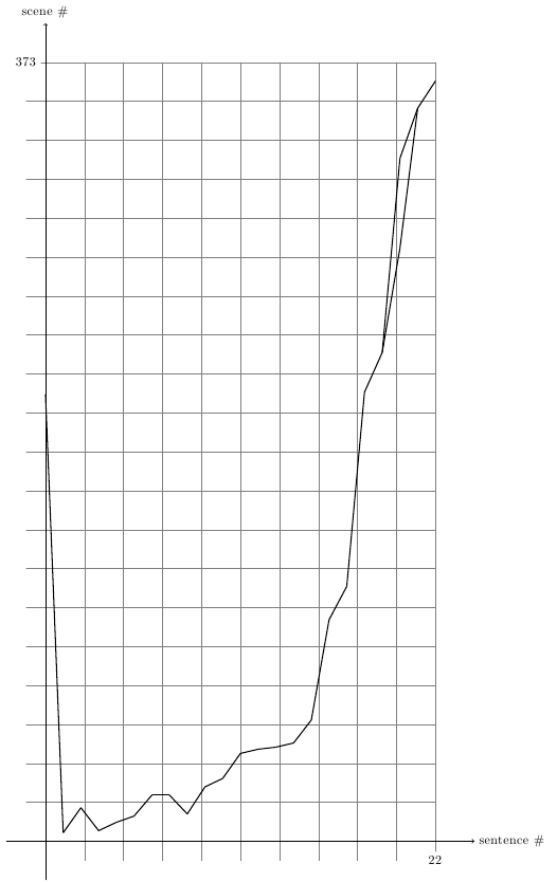


Figure 2: batman

Predicted

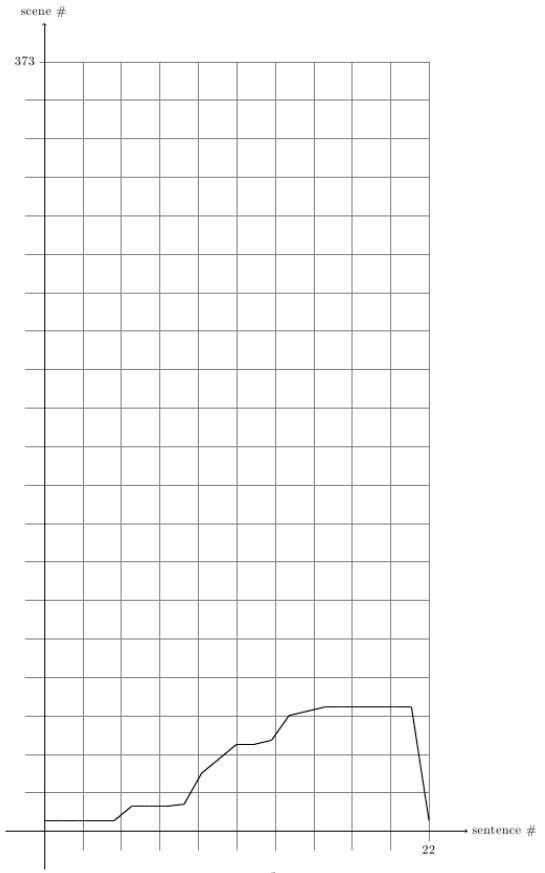


Figure 2: batman

Feature Ranking

Which features are the main contributors?

Word Overlaps (gradient descent weights)

1. Nouns (0.58)
2. Verbs (0.17)
3. Adjectives (0.03)
4. Adverbs (0.05)

Future Work - Scene Clustering

Scene 84
ANGLE ON BRUCE
BRUCE is still frozen to the spot and exposed to the "mimes."

Scene 85
A FEW FEET AWAY
Vicki crouches behind the car and beckons to him.
VIKI: Bruce! Get DOWN!
He totally ignores her and BEGINS TO WALK. At first slowly, then faster, right toward the Joker.

Scene 86
ON JOKER
He LAUGHS at the DIN and walks as if immortal through the mayhem. (He never sees Bruce.)

Scene 87
ON BRUCE
He's walking the same way, still at a distance from Joker. But his recognition of him is growing.

Future Work - Pre-Proccessing

- Pronoun resolution - Eleni Miltsakaki's Antelogue
- Synonym identification
  - WordNet
  - Chris Callison-Burch's work
  - Distributional clustering with data from Google n-gram

Future Work - Additional Features

$\phi_i(A_i, X_i)$

- Noun-noun bigrams
- LDA with scenes as training "documents"
- Weighting by salience (more weight for subj. than obj.)

$\phi_{(i-1)i}(A_{i-1}, A_i)$

- Entity grid, centering relations
- Scene-scene comparison for scenes aligned to adjacent sentences
- Take advantage of timing words such as "later", "while", etc.
- Account for paragraph breaks