The Computer-based Analysis of Narrative and Multimodality

The Media School Research Seminar Series
Bournemouth University

26th March 2007

Dr Andrew Salway

www.bbrel.co.uk
Why a computer-based analysis?

- The need for metadata that describes 'semantic multimedia content'
- A new way to investigate old questions about how stories are told and understood using different modes of communication
Approach

Working within the constraints of what computers can do...

- Concentrate on repeating patterns in the surface forms of texts, images, videos, etc.
- Avoid encoding world knowledge / experience
Overview

- Analysis of narrative in film
- Analysis of multimodality in webpages
- Some questions for discussion
Questions for Discussion

- To what extent do interesting narrative and multimodal phenomena manifest as patterns of media features that can be detected by a computer-based analysis?

- What new applications will be enabled by giving computers some 'understanding' of narrative and multimodality?

- How can computer-based analysis feed into the development of theories about narrative and multimodality?
Narrative comprises *story* and *discourse*: the same story can be told in different discourses (Chatman 1978). A story is a chain of *events*, in *cause-effect* relationships, organised in *space* and *time*. The agents of cause-effect are *characters* with *goals*, *beliefs* and *emotions* (Bordwell & Thompson 1997).

For eight criteria to assess a text's degree of narrativity, see (Ryan 2006). For a richer account of storyworlds, see (Herman 2002)
The limits of automatic video analysis

- Computer vision is an unsolved problem
- Automatic video analysis can do:
  - shot segmentation and classification;
  - some scene segmentation;
  - detect changes in 'Tempo'
Collateral Texts for Films...

INT. CASINO FLOOR

Tess and Rusty hold a stare as Rusty holds the phone.

TESS
Where's Danny?

RUSTY
He's fine. He wants you to go upstairs, and watch TV.

TESS
(a little pissed)
He does?

BENEDICT (V.O.)
You have a deal.

RUSTY
It's alright, Tess. I promise.
(back on phone)
Good. Here's what you do. Five minutes from now, the men in the vault are going to deposit six bags in the vault elevator.

Tess isn't sure what to do. As Rusty continues on the phone, she backs off, debating: can she blow the whistle on her ex?

INT. VAULT ELEVATOR - CLOSE ON SIX CANVAS BAGS

each sealed tight, each marked with an X, loaded onto the vault elevator.

RUSTY (V.O.)
INT. CASINO FLOOR

Tess and Rusty hold a stare as Rusty holds the phone.

TESS
Where's Danny?

RUSTY
He's fine. He wants you to go upstairs, and watch TV.

TESS
(a little pissed)
He does?

BENEDICT (V.O.)
You have a deal.

RUSTY
It's alright, Tess. I promise.
(back on phone)
Good. Here's what you do. Five minutes from now, the men in the vault are going to deposit six bags in the vault elevator.

Tess isn't sure what to do. As Rusty continues on the phone, she backs off, debating: can she blow the whistle on her ex?

INT. VAULT ELEVATOR - CLOSE ON SIX CANVAS BAGS

each sealed tight, each marked with an X, loaded onto the vault elevator.

RUSTY (V.O.)
Where's Danny? - He's fine. He's in good form.

He asks that you go up and watch TV. - He does?

We have three men with explosives who have taken control of our vault.

Okay. - It's okay.

You have a deal. - I promise.

Fantastic. Here's what you do:

The guys in the vault will put six bags in the elevator.
Where's Danny? - He's fine. He's in good form.

He asks that you go up and watch TV. - He does?

We have three men with explosives who have taken control of our vault.

Okay. - It's okay.

You have a deal. - I promise.

Fantastic. Here's what you do:
The guys in the vault will put six bags in the elevator.
When Daniel Ocean is released from prison in New Jersey, his next heist is already planned. Danny's target are three Las Vegas casinos: The Bellagio, the Mirage and the MGM Grand. They all belong to ruthless entrepreneur Terry Benedict, who, by the way, also shows a certain interest in Danny's beautiful ex-wife Tess. During a much-anticipated boxing event (Lennox Lewis vs. Wladimir Klitschko), there will be $150 million in the safe, 70 yards below the strip.

So, Danny starts to hire professionals from all over the country: There's the card magician Rusty Ryan, the perfect pickpocket Linus Caldwell and the ingenious pyrotechnician Basher Tarr. Reuben Tishkoff, who lost a casino to Benedict, provides funding, the brothers Virgil and Turk Malloy will drive and help, and Frank Catton, a professional card dealer, gets a job at the casino to watch the routines. Saul Bloom, already retired, will play the weaponry dealer and live in the hotel, while Livingston Dell bugs the place to look over the shoulders of the security personnel.
When Daniel Ocean is released from prison in New Jersey, his next heist is already planned. Danny's target are three Las Vegas casinos: The Bellagio, the Mirage and the MGM Grand. They all belong to ruthless entrepreneur Terry Benedict, who, by the way, also shows a certain interest in Danny's beautiful ex-wife Tess. During a much-anticipated boxing event (Lennox Lewis vs. Wladimir Klitschko), there will be $150 million in the safe, 70 yards below the strip.

So, Danny starts to hire professionals from all over the country: There's the card magician Rusty Ryan, the perfect pickpocket Linus Caldwell and the ingenious pyrotechnician Basher Tarr. Reuben Tishkoff, who lost a casino to Benedict, provides funding, the brothers Virgil and Turk Malloy will drive and help, and Frank Catton, a professional card dealer, gets a job at the casino to watch the routines. Saul Bloom, already retired, will play the weaponry dealer and live in the hotel, while Livingston Dell bugs the place to look over the shoulders of the security personnel.
1:32:25 Tess steps up to Russ in the casino. He holds her steely gaze.

1:32:34 Russ covers his phone’s mouthpiece. Grim-faced, Benedict stares at the cellphone in his hand for a long while, mulling over his options, then turns to Walsh.

1:32:45 Walsh fishes a cellphone out of his trouser pocket, dials, and raises it to his ear.

1:32:59 In the casino.

1:33:02 Walsh nods to Benedict, who speaks into his phone.
Generating film metadata from collateral texts

- What kinds of information about a film's story can be extracted automatically?
  - Characters emotions
  - Frequently depicted events
  - Frequent speech acts

- What can this be used for?
  - Film retrieval by story similarity
  - Hypervideo browsing of films by story structure
Emotions in *Captain Correlli’s Mandolin* (Salway and Graham 2003)
Emotions in *Captain Correlli’s Mandolin* (Salway and Graham 2003)

15-20 minutes: Pelagria’s betrothal to Madras
20-30 minutes: invasion of the island
68-74 minutes: Pelagria and Correlli’s growing relationship
92-95 minutes: German soldiers disarm Italians
Emotions in *The Postman* (Salway and Graham 2003)
Frequently depicted events
(Salway, Vassilou and Ahmad 2005; Vassilou 2006)

Some common phrases in audio description and film scripts...

- “X looks at Y”
- “X goes to the Y”
- “X opens the door and leaves”
- “X nods her head”
- “INT, daytime, an office”

>>> a local grammar of audio description and film scripts
An Automatically-extracted Database of Film Events

<table>
<thead>
<tr>
<th>CFOA ID</th>
<th>Focus Type</th>
<th>Character Focussing</th>
<th>Character Focussed On</th>
<th>Object Focussed On</th>
<th>Occurrence Time</th>
<th>Text String</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOAAD5</td>
<td>ACTIVE</td>
<td>Jim</td>
<td>(their) car</td>
<td>00:15:22:24</td>
<td>00:15:22:24 looking at</td>
<td></td>
</tr>
<tr>
<td>FOAAD6</td>
<td>ACTIVE</td>
<td>Carl (He)</td>
<td>Jim</td>
<td>00:23:32:07</td>
<td>00:23:32:07 turns to see</td>
<td></td>
</tr>
<tr>
<td>FOAAD7</td>
<td>PASSIVE</td>
<td>Jim</td>
<td>the gun</td>
<td>00:25:11:36</td>
<td>00:25:11:36 takes</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ScCh ID</th>
<th>INT\EXT</th>
<th>Location</th>
<th>Time of Day</th>
<th>Line No.</th>
<th>% Film Time</th>
<th>Text String</th>
</tr>
</thead>
<tbody>
<tr>
<td>ScChAD1</td>
<td>Interior</td>
<td>FITTS HOUSE - RICKY'S BEDROOM</td>
<td>Night</td>
<td>2</td>
<td>0.07</td>
<td>INT. FITTS HOUSE - RICKY'S BEDROOM- NIGHT</td>
</tr>
<tr>
<td>ScChAD10</td>
<td>Exterior</td>
<td>SALE HOUSE</td>
<td>Day</td>
<td>322</td>
<td>11.25</td>
<td>EXT. SALE HOUSE- DAY</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NVC ID</th>
<th>Body Part</th>
<th>Character Communicating</th>
<th>Action Involved</th>
<th>Occurrence Time</th>
<th>Text String</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVCAD9</td>
<td>Head</td>
<td>Madox</td>
<td>shakes</td>
<td>00:25:09:16</td>
<td>00:25:09:16 shakes his head</td>
</tr>
<tr>
<td>NVCAD10</td>
<td>Eyes</td>
<td>Caravaggio</td>
<td>closes</td>
<td>00:25:34:20</td>
<td>00:25:34:20 closes his eyes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COL ID</th>
<th>Character State</th>
<th>Character Changing Location</th>
<th>Occurrence Time</th>
<th>Text String</th>
</tr>
</thead>
<tbody>
<tr>
<td>COLAD2</td>
<td>ENTERING</td>
<td>Donkey</td>
<td>00:06:26:17</td>
<td>00:06:26:17 runs into</td>
</tr>
<tr>
<td>COLAD3</td>
<td>LEAVING</td>
<td>Donkey</td>
<td>00:10:51:06</td>
<td>00:10:51:06 walks out</td>
</tr>
<tr>
<td>COLAD4</td>
<td>WITHINNON</td>
<td>Donkey</td>
<td>00:11:55:04</td>
<td>00:11:55:04 walking over to</td>
</tr>
</tbody>
</table>
What do people say in films?
(Lingabavan and Salway 2006)

Some common phrases subtitles...

- “I don't know”
- “I'm gonna X”; “I'm going to X”
- “I wanna X”; “I want to X”
- “I gotta X”; “I'm going to X”

cf. Kozloff's (2000) functions of film dialogue:
Anchorage of the diegesis and characters;
Communication of narrative causality; Enactment of
narrative events; Character revelation; Adherence to
the code of realism; Control of viewer evaluation
NAFI: NAvigating Films (Salway and Xu 2005)
Kress and Van Leeuwen (2006)

“... any text whose meanings are realized through more than one semiotic code is **multimodal**...

the question arises **whether the products of the various modes should be analysed separately or in an integrated way**;

whether the meanings of the whole should be treated as **the sum of the meanings of the parts, or whether the parts should be looked upon as interacting with and affecting one another**.

It is the latter path that we will pursue...” (p.177)
DNA Molecule - Two Views

Legend:

The double helix of the DNA is shown along with details of how the bases, sugars and phosphates connect to form the structure of the molecule.

DNA is a double-stranded molecule twisted into a helix (think of a spiral staircase). Each spiraling strand, comprised of a sugar-phosphate backbone and attached bases, is connected to a complementary strand by non-covalent hydrogen bonding between paired bases. The bases are adenine (A), thymine (T), cytosine (C) and guanine (G).

A and T are connected by two hydrogen bonds. G and C are connected by three hydrogen bonds.
Queen in Liverpool for ceremony

The Queen is visiting Liverpool for the ancient Maundy Thursday ceremony.

During the service at the city's Anglican cathedral the Queen will hand out money to 79 men and 79 women.

Maundy Thursday marks the washing of the disciple's feet by Jesus. British monarchs used to wash subjects' feet, but James II ended the practice.

It is the first time the ceremony has been performed at the cathedral. The locals were selected for their contributions to their communities.

The ceremony is due to take place at 1100 BST.

Organisers have selected 78 men and women as a way of marking the Queen's 78th birthday celebrations which take place on 21 April.

The Queen and the Duke of Edinburgh, will arrive at Liverpool's Lime Street station before being taken the short distance to the cathedral where they will receive the traditional posies of strong-smelling flowers.

After the service they will attend a civic reception at Liverpool Hope University.

Lord Mayor of Liverpool, Councillor Ron Gould, said: "It is a wonderful honour that the Queen has chosen to visit Liverpool this year for the Maundy Thursday celebrations.

'I am delighted that some of the people who make such a great contribution to our city are receiving this prestigious award from Her Majesty.'

The Queen has only missed four Maundy Thursday ceremonies since coming to the throne in 1952.
The Goddess of Discord Choosing the Apple of Contention in the Garden of the Hesperides

Oil on canvas
Support: 1563 x 2184 mm
Painting

Bequeathed by the artist 1856

N00477

The Hesperides were the three daughters of Hesperus, the evening star. They kept watch over a tree of golden apples in a garden on the slopes of Mount Atlas. Here, the goddess Discord chooses the apple that will eventually lead to the Trojan war. Paris offered the apple to the goddess Aphrodite who, in return, offered him the most beautiful woman in the world: the Greek Queen, Helen.

The classical grandeur of this landscape is based on the work of the seventeenth-century French artist, Nicolas Poussin. Its background reflects Turner’s experience of the Alps in 1802

(From the display caption February 2004)
In what ways are images and texts used together to convey meanings? How do their meanings relate in particular combinations?

How do we know when they are being used in one way or another?

==> a functional-systemic account of image-text relations
The relation between an image and a text is **equal** when:
- both the image and the text are required for successful communication, in which case they are **equal-complementary**; **OR**
- both the image and the text can be understood individually, in which case they are **equal-independent**.

The relation between an image and a text is **unequal** when either the image or the text can be understood individually - that which cannot be understood individually is **subordinate**.
DNA Molecule - Two Views

Legend:

The double helix of the DNA is shown along with details of how the bases, sugars and phosphates connect to form the structure of the molecule.

DNA is a double-stranded molecule twisted into a helix (think of a spiral staircase). Each spiraling strand, comprised of a sugar-phosphate backbone and attached bases, is connected to a complementary strand by non-covalent hydrogen bonding between paired bases. The bases are adenine (A), thymine (T), cytosine (C) and guanine (G).

A and T are connected by two hydrogen bonds. G and C are connected by three hydrogen bonds.
Queen in Liverpool for ceremony

The Queen is visiting Liverpool for the ancient Maundy Thursday ceremony.

During the service at the city's Anglican cathedral the Queen will hand out money to 78 men and 79 women.

Maundy Thursday marks the washing of the disciple's feet by Jesus. British monarchs used to wash subjects' feet, but James II ended the practice.

It is the first time the ceremony has been performed at the cathedral. The locals were selected for their contributions to their communities.

The ceremony is due to take place at 1100 BST.

Organisers have selected 78 men and women as a way of marking the Queen's 78th birthday celebrations which take place on 21 April.

The Queen and the Duke of Edinburgh will arrive at Liverpool's Lime Street station before being taken the short distance to the cathedral where they will receive the traditional nosegays - posies of strong-smelling flowers.

After the service they will attend a civic reception at Liverpool Hope University.

Lord Mayor of Liverpool, Councillor Ron Gould, said: "It is a wonderful honour that the Queen has chosen to visit Liverpool this year for the Maundy Thursday celebrations."

'I am delighted that some of the people who make such a great contribution to our city are receiving this prestigious award from Her Majesty.'

The Queen has only missed four Maundy Thursday ceremonies since coming to the throne in 1952.
The Goddess of Discord Choosing the Apple of Contention in the Garden of the Hesperides exhibited 1805

Oil on canvas
support: 1583 x 2184 mm
painting

Bequeathed by the artist 1856

N00477

The Hesperides were the three daughters of Hesperus, the evening star. They kept watch over a tree of golden apples in a garden on the slopes of Mount Atlas. Here, the goddess Discord chooses the apple that will eventually lead to the Trojan war. Paris offered the apple to the goddess Aphrodite who, in return, offered him the most beautiful woman in the world: the Greek Queen, Helen.

The classical grandeur of this landscape is based on the work of the seventeenth-century French artist, Nicolas Poussin. Its background reflects Turner’s experience of the Alps in 1802

(From the display caption February 2004)
Logico-Semantic Relations
(Martinec and Salway 2005)

- A text **elaborates** the meaning of an image, and vice versa, by further specifying or describing it.

- A text **extends** the meaning of an image, and vice versa, by adding new information.

- A text **enhances** the meaning of an image, and vice versa, by qualifying it with reference to time, place and/or cause-effect.
Queen in Liverpool for ceremony

The Queen is visiting Liverpool for the ancient Maundy Thursday ceremony.

During the service at the city's Anglican cathedral the Queen will hand out money to 78 men and 78 women.

Maundy Thursday marks the washing of the disciple's feet by Jesus. British monarchs used to wash subjects' feet, but James II ended the practice.

It is the first time the ceremony has been performed at the cathedral. The locals were selected for their contributions to their communities.

The ceremony is due to take place at 1100 BST.

Organisers have selected 78 men and women as a way of marking the Queen's 78th birthday celebrations which take place on 21 April.

The Queen and the Duke of Edinburgh, will arrive at Liverpool's Lime Street station before being taken the short distance to the cathedral where they will receive the traditional nosegays - posies of strong-smelling flowers.

After the service they will attend a civic reception at Liverpool Hope University.

Lord Mayor of Liverpool, Councillor Ron Gould, said: "It is a wonderful honour that the Queen has chosen to visit Liverpool this year for the Maundy Thursday celebrations.

'I am delighted that some of the people who make such a great contribution to our city are receiving this prestigious award from Her Majesty.'

The Queen has only missed four Maundy Thursday ceremonies since coming to the throne in 1952.
The Goddess of Discord Choosing the Apple of Contention in the Garden of the Hesperides 1805

Oil on canvas
support: 1563 x 2134 mm
painting

Bequeathed by the artist 1856

N00477
The Hesperides were the three daughters of Hesperus, the evening star. They kept watch over a tree of golden apples in a garden on the slopes of Mount Atlas. Here, the goddess Discord chooses the apple that will eventually lead to the Trojan war. Paris offered the apple to the goddess Aphrodite who, in return, offered him the most beautiful woman in the world: the Greek queen, Helen.

The classical grandeur of this landscape is based on the work of the seventeenth-century French artist, Nicolas Poussin. Its background reflects Turner’s experience of the Alps in 1802.
(From the display caption February 2004)
Baghdad stunned by hotel attack

Rescuers in Baghdad have given up the search for survivors in the rubble left by a massive blast which killed at least 17 people, most of them Iraqis.

The explosion on Wednesday night wrecked the small Mount Lebanon hotel and nearby houses in the city centre.

The US military says a 450kg car bomb packed with artillery ordinance killed at least 17 people and injured 45.

On Thursday, three people also died in an attack on staff at a US-funded TV station in Baghda, north of Baghdad.

US officials say the hotel attack, in the busy Kamada district, was intended to cause maximum casualties and bore the hallmarks of an al-Qaeda-related group.

One Briton was killed, and another injured, the UK government said.

US officials initially put the death toll as high as 27, but later revised it downwards. Iraqi officials said the toll was as low as six.

Hours after the blast, a mortar attack on a US logistics base north of Baghdad killed two soldiers and injured six.

'Friendly' hotel

Most of the victims of the hotel bombing were Iraqis living...
Media features to recognise image-text relations?

**Page Layout and Formatting**
- Relative size and position of image and text; font type and size; use of borders

**Text**
- Lexical references in text: ‘This picture shows...’; ‘See Figure 1’; ‘on the left’; ‘is shown by’
- Grammatical characteristics of the text: tense – past / present; quantification – single / many; full sentences or short phrases

**Image**
- Modality of images: a scale from realistic to abstract, or from photographic to graphic
- Framing of images: for example, one centred subject, or no particular subject
Our system of image-text relations captures our intuitions about how the visual and verbal combine in print and on web-pages, but the ‘realisations’ of the relations have yet to be tested experimentally.

Does our system generalise to ‘multimedia relations’? Could it be applied to the relationship between the moving image and film dialogue?
Questions for Discussion

- To what extent do interesting narrative and multimodal phenomena manifest as patterns of media features that can be detected by a computer-based analysis?

- What new applications will be enabled by giving computers some 'understanding' of narrative and multimodality?

- How can a computer-based analysis feed into the development of theories about narrative and multimodality?


Herman, David (2002) *Story Logic: problems and possibilities of narrative*, University of Nebraska Press.


