Centrum voor Wiskunde en Informatica

Creation and use of video annotations for presentation generation

Lynda Hardman
Frank Nack

http://www.cwi.nl/~lynda

Outline

- Re-use of video
  - Documentary Generation
    Vox Populi (Stefano Bocconi, CWI & Uni. Turin)
- Associating video data and metadata
  - Canonical processes of media production
  - IPTC news codes, NewsML2
- Generating news presentations

Generating video documentaries from annotated media repositories

Stefano Bocconi
CWI Amsterdam
The Netherlands
Contact: Stefano.Bocconi@cwi.nl

Talk Outline

- ☐ Motivation
- ☐ Example
- ☐ Scenarios
- ☐ Technical details
  - ■ Annotations
  - ■ Editing Process
- ☐ Conclusions

Video Documentaries on the Web

- ☐ Traditional video authoring: there is only one final version, what is shown is the choice of the author/editor
- ☐ Proposed video authoring:
  - ■ Annotate the video material semantics
  - ■ Show automatically what the user asks to see, using presentation forms a film editor would use

Video material

- ☐ Focus on video interviews about controversial issues
- ☐ **Interview with America** video footage with interviews and background material about the opinion of American people after 9-11
  www.interviewwithamerica.com

Example: *What do you think of the war in Afghanistan?*

"*I am never a fan of military action, in the big picture I don't think it is ever a good thing, but I think there are circumstances in which I certainly can't think of a more effective way to counter this sort of thing…*"

What do you think of the war in Afghanistan?

I am not a fan of military actions

I cannot think of a more effective solution

War has never solved anything

Two billions dollar bombs on tents

Scenarios

- ☐ Augmenting one interview with man-on-the-street opinion ("Vox Populi" documentary)
- ☐ Overview of the content of video footage:
  - ■ Example: trailers ("Voices of Iraq")
  - ■ Browse the content by opinion

The annotations

- ☐ Rhetorical
  - ■ Rhetorical Statement (mostly verbal, but visual also possible)
  - ■ Argumentation model: Toulmin model
- ☐ Descriptive
  - ■ Question asked
 - ■ Interviewee (social)
  - ■ Filmic (e.g. location/time/framing/gaze)

Encode statements

☐ Statement formally annotated:
  - \<subject\> \<modifier\> \<predicate\>
  - E.g. "**war best solution**"

☐ A thesaurus containing:
  - Terms (155)
  - Relations between terms: *similar* (72), *opposite* (108), *generalization* (10), *specialization* (10)
  - E.g. **war** *opposite* **diplomacy**

Connect statements

❑ Using the thesaurus, generate related statements and query the repository
  "*war best solution*",
  "*diplomacy best solution*",
  "*war not solution*"

☐ Create a graph of related statements
  - nodes are the statements (corresponding to video segments)
  - edges are either *support* or *contradict*

Toulmin model

| Data | → | Claim |

Qualifier

Warrant

Condition

Backing
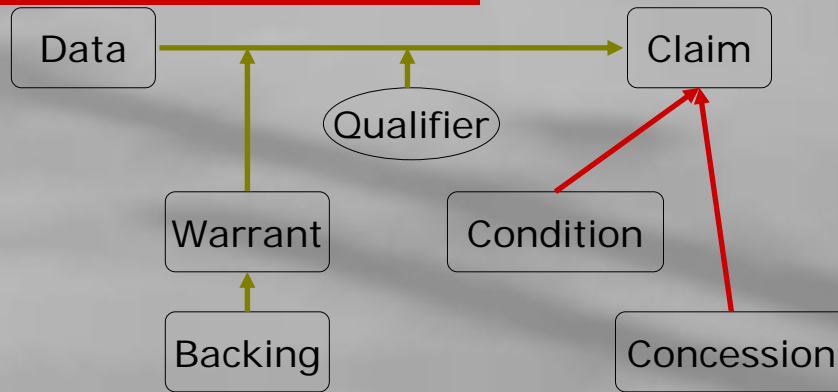
Concession

57 Claims, 16 Data, 4 Concessions,
3 Warrants, 1 Condition

Analysis of the Example

Two billions dollar bombs on tents

Claim — *contradict* →

Claim

**I cannot think of a more effective solution**

weaken

Concession — **I am not a fan of military actions**

Claim — *support*

War has never solved anything

Facts and features

- ☐ Annotations: 1 hour annotated, 15 interviews, 60 interview segments, 120 statements
- ☐ Partially tunable: examining the Segment graph gives feedback on the quality of the annotations and the thesaurus

←——→ = support

←——→ = contradict
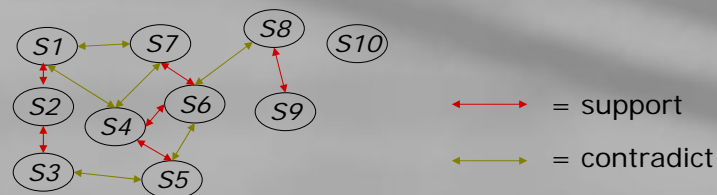
Controlling the Bias

- ☐ Video documentaries are not neutral account of reality: the selection and editing of the footage expresses a point of view
- ☐ Editing strategy:
  - ■ Balanced
  - ■ Pro opinion X
  - ■ Against opinion X
- ☐ We use:
  - ■ Logos (the statements)
  - ■ Ethos (based on user profile)
  - ■ Film editing (e.g. framing, gaze)

Vox Populi interface

Question	Interviewee	Opinion	
Why did they do what they did? What do you think of the casualties among civilians? What do you think of the Afghanistan war? What are the consequences of the war? What are the roots of the problem? What do you think about the Anthrax?	Cameroun Parking Guard at Stamford Lawyer in Harward	War in Afghanistan - Pro	**Position**

Age	Education	Employment	GeoLocation	Race	Religion	Sex	
Middleage Old Teenager Young	HighEducated LowEducated MediumEducated	HighIncomeJob LowIncomeJob MiddleIncomeJob Retired Student	NotUSA USA	AmericanIndian Asian Black Hispanic White	Atheist Christian Muslim	Female Male	**First Character**

Age	Education	Employment	GeoLocation	Race	Religion	Sex	
Middleage Old Teenager Young	HighEducated LowEducated MediumEducated	HighIncomeJob LowIncomeJob MiddleIncomeJob Retired Student	NotUSA USA	AmericanIndian Asian Black Hispanic White	Atheist Christian Muslim	Female Male	**Second Character**

Strategy
- ○ None
- ● Create Clash
- ○ Create Support
- ○ Vox Populi

Bandwidth
- ○ Low Bandwidth
- ● Medium Bandwidth
- ○ High Bandwidth

Intercut
- ● True
- ○ False

Caption
- ○ On (can cause problems)
- ● Off

Done   Reset

Conclusions

☐ Automatic generation of video interviews augmented with supporting and/or contradicting material

☐ The **user** can determine the subject and the bias of the presentation

☐ The **documentarist** can add material and let the system generate new documentaries
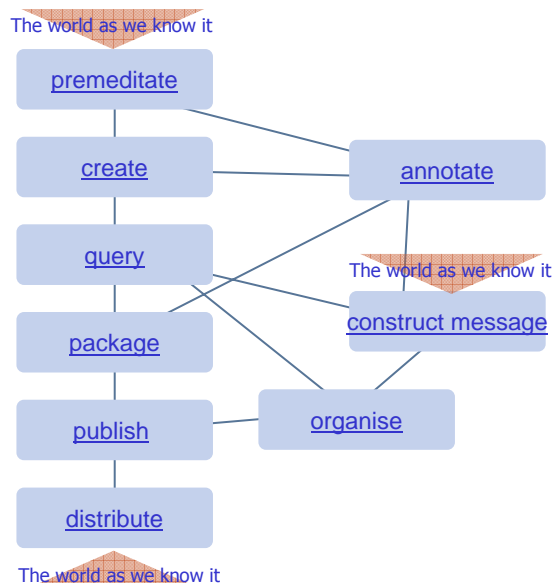
Pointers & Acknowledgments

- ☐ This presentation and a Demo available at:

  http://www.cwi.nl/~media/demo/IWA/

- ☐ This research was funded by the Dutch national ToKeN I^{2}RP and CHIME projects.

Canonical processes of media production

http://www.cwi.nl/~media/projects/canonical/

The world as we know it

premeditate

create — annotate

query

The world as we know it

package — construct message

publish — organise

distribute

The world as we know it

Video data + metadata

- High quality *semantic* multimedia metadata enables:
  - Easy exchange of news items
  - Semantic search of particular news items
  - Delivery of personalized news content to customers
    - Interactive browsing in a news archive
    - Cross-modality: packaging the news stories, photos, graphics, audio, videos
    - For different end-user platforms (mobiles, PC, handhelds, etc.)

IPTC Metadata Standards

- IPTC has defined 28 sets of multilingual News Codes (de en-GB es fr it ja-JP)
  - NewsCodes use numeric strings = language agnostic
  - Subject $\approx$ 1300 terms, 3 levels hierarchy
  - NewsCodes Viewer application View
- XML Wrapper
  - Metadata embedded in a photo: XMP
  - Metadata stored in a separate file: NewsML

Role of the Semantic Web

- "Oh no! Not yet another metadata standard!"
  Like we don't have enough of them already:
  - MPEG-7, EXIF, Dublin Core, VRA Core, IPTC Core, XMP, Creative Commons, ... ?
- But again: No single standard can cover all metadata needs
- SW is a framework that could make existing metadata standards and tools interoperable ... and make them interoperable with the rest of the Web!

NewsML2 and the SW

- Common basis
  - Distributed resources (news item) globally and uniquely identified => URI
  - Use of shared and controlled vocabularies
- Natural switch and numerous benefits
  - Better control of NewsML2 descriptions (logical consistency check)
  - Enhanced search of News topic (logical inferences)
  - Intelligent presentation – Semantic interfaces
  - Unified news management – Semantic CMS

What we have done so far?

- Creation of a News domain ontology in OWL
  - Based on the UML model specifications of NewsML2
- Online conversion service
  - Mapping of the IPTC NewsCodes into various SKOS thesaurus
  - Transforming dynamically the NewsML2 (XML) descriptions in its equivalent RDF counterpart
    - ‣ Using to the NewsML ontology
    - ‣ Linking to the SKOS IPTC NewsCodes

http://newsml.cwi.nl/

What is the added value?

- Example: A "normal" day in AFP
- Dataset
  - 200 NewsML2 stories, 35 photos (original size + thumbnails) + 35 NewsML2 descriptions
  - Covering various subjects e.g.:
    - ‣ A military drill for dealing with contaminations (toxic, nuclear or biological) - Photo
    - ‣ A protest made on the Arch of Triumph in Paris, related to the Iran nuclear crisis - Photo
    - ‣ A meeting between the French president and Israeli prime minister - Photo

Example 1: reasoning on the content

■ Find all related news about "Nuclear"

Nucléaire ──────▶ Military drill (NBC)

Arc de Triomphe protest

Nuclear ──────▶ Iran nuclear crisis

Chirac – Elmer summit
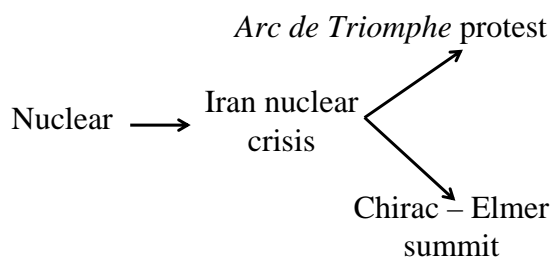
W3C Multimedia Semantics Incubator Group

W3C Multimedia Semantics Incubator Group

- Light-weight, one year (May 2007) group looking at image and other multimedia metadata on the Web
- Focus on interoperability with existing standards
- You can help and shape the future of multimedia metadata on the Web!
- Need input from real media users:
  - Provide use cases & examples
  - State which standards are most important to you
  - Review the notes written by the group

MM Sem XG

Athens face2face December 2006

Conclusions

- Re-using video for end users
  - Vox Populi
  - News quiz (Masanori Sano, NHK/NII)
- Look at use cases in MM Sem XG
  http://www.w3.org/2005/Incubator/mmsem/wiki/News_Use_Case
  http://www.w3.org/2005/Incubator/mmsem/wiki/
- Discussion on scenarios for international news in Japanese, French, (maybe also English, Dutch and Irish) repositories

NewsCodeViewer

back