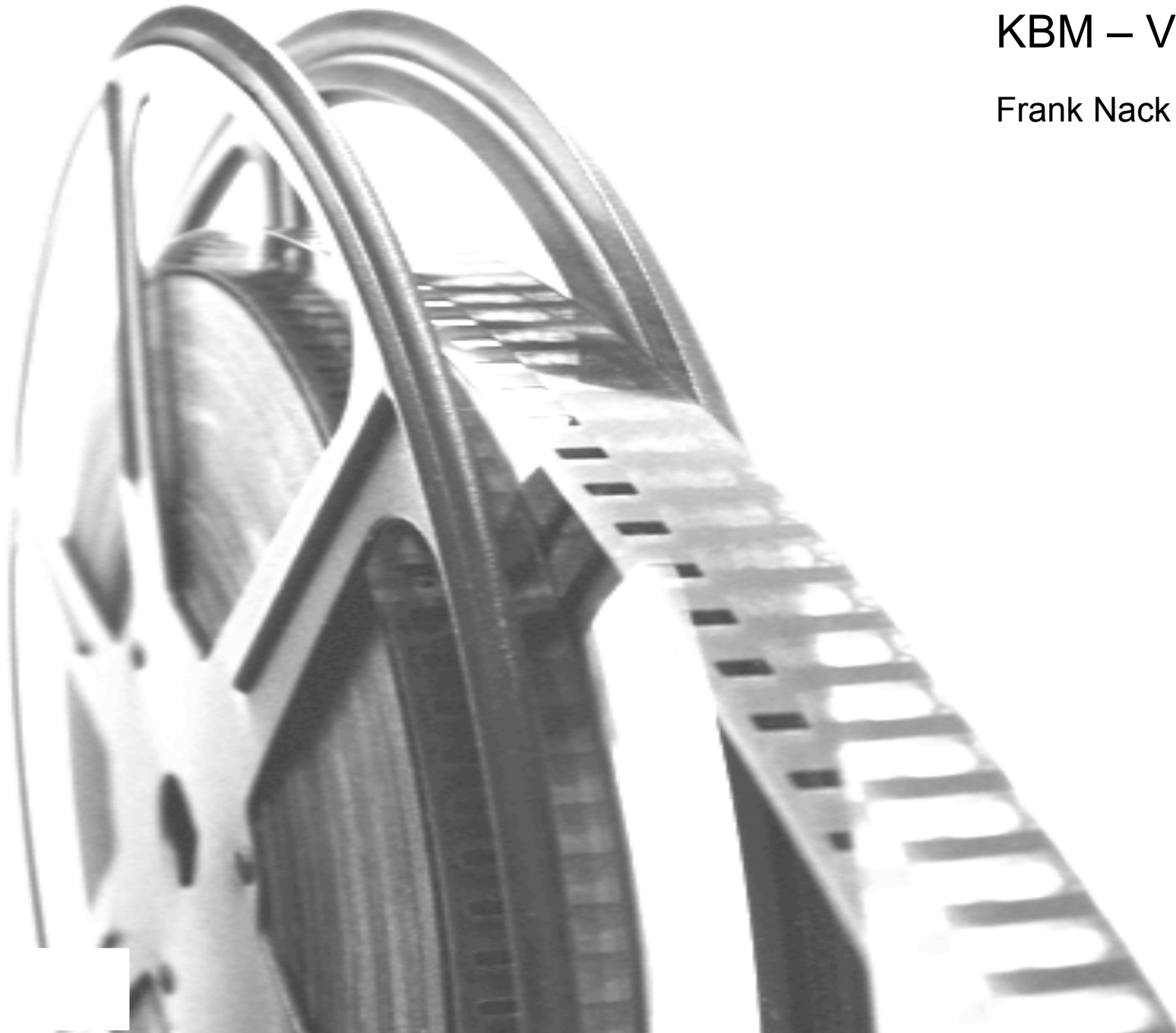


KBM – Video

Frank Nack



# Outline

---

- Video – another visual sign system
- Media-dependent context
- Meta-semantic context
- AI techniques

# Text and Image Application – summary

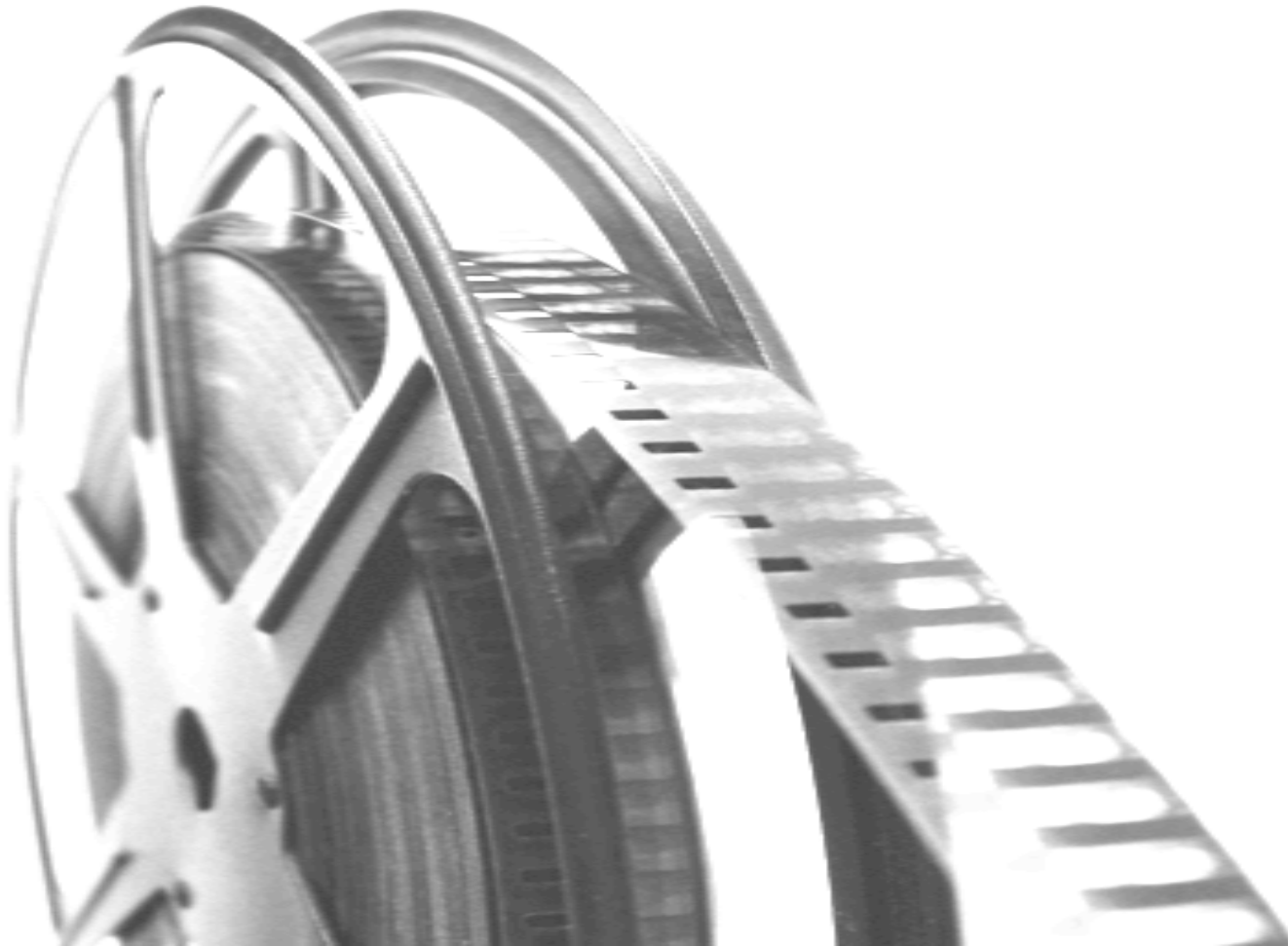
## Investigated

- Word of bag in Image Analysis
- The MultimediaN E-Culture Demonstrator

## Findings

- word of bags work on the denotative level of the image (1<sup>st</sup> step of signification)
- a method to statistically establish the objects in an image
- image descriptors are not given with the image: one needs to *extract* or *interpret* them
- Web2.0 facilities fit well:
  - involving community experts in annotation
  - personalization, myArt
- more links, more data, more metadata
- social barriers have to be overcome

## Video – a temporal visual sign system



## Video – Examples

---



Titanic – The sequel



The shining – a romantic comedy



Schindler's list – the romantic comedy



Mary Poppins – the scary original

All images from YouTube ([www.youtube.com](http://www.youtube.com))

## Video – Internal and external context

---



**An image is an index to a story**



**A video is an iconic representation of a story**

Images from Stanley Kubrick's 'Shining'

# Video – Internal and external context

---

## Media-dependent context

- Spatial
- Temporal



Images from Stanley Kubrick's 'Shining'

## Meta-semantic context

- Plot structure
- Genre
- References
- Reviews
- Personal preferences
- ....

# Video – Media-dependent context - temporal



Episode context: temporal relationship between scenes

Scene context: temporal relationship between shots, e.g. insert

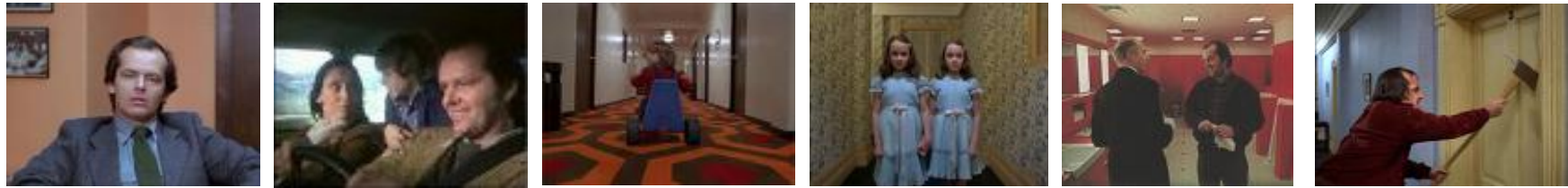
Shot context: temporal relationship between frames

## Montage

- Metric (absolute length)
- Rhythmic
- Tonal
- Compression
- Expansion
- Insertion
- Deletion



# Video – Media-dependent context - summary



All frames from "The Shining" (1980)

# Video – Internal and external context

---

## Media-dependent context

- Spatial
- Temporal



Images from Stanley Kubrick's 'Shining'

## Meta-semantic context

- Plot structure
- Genre
- References
- Reviews
- Personal preferences
- ....

# Video – Meta-Semantic context – Narration



**Fabula**

A set of of logically and chronologically related events that are caused or experienced by actors.

**Theme**

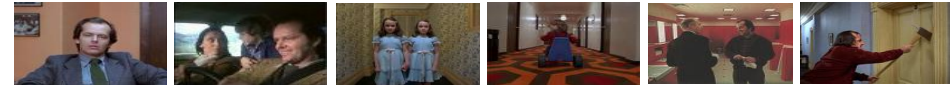
a high level concept that is part of the subtext of a narrative, not directly present within the narrative but is connoted through the presence of motifs.

Horror: supernatural, macabre, fear, terror,

**Genre**

a loose set of criteria for a category of composition  
epic, tragedy, comedy, novel, short story, drama, documentary, etc.

**Plot**



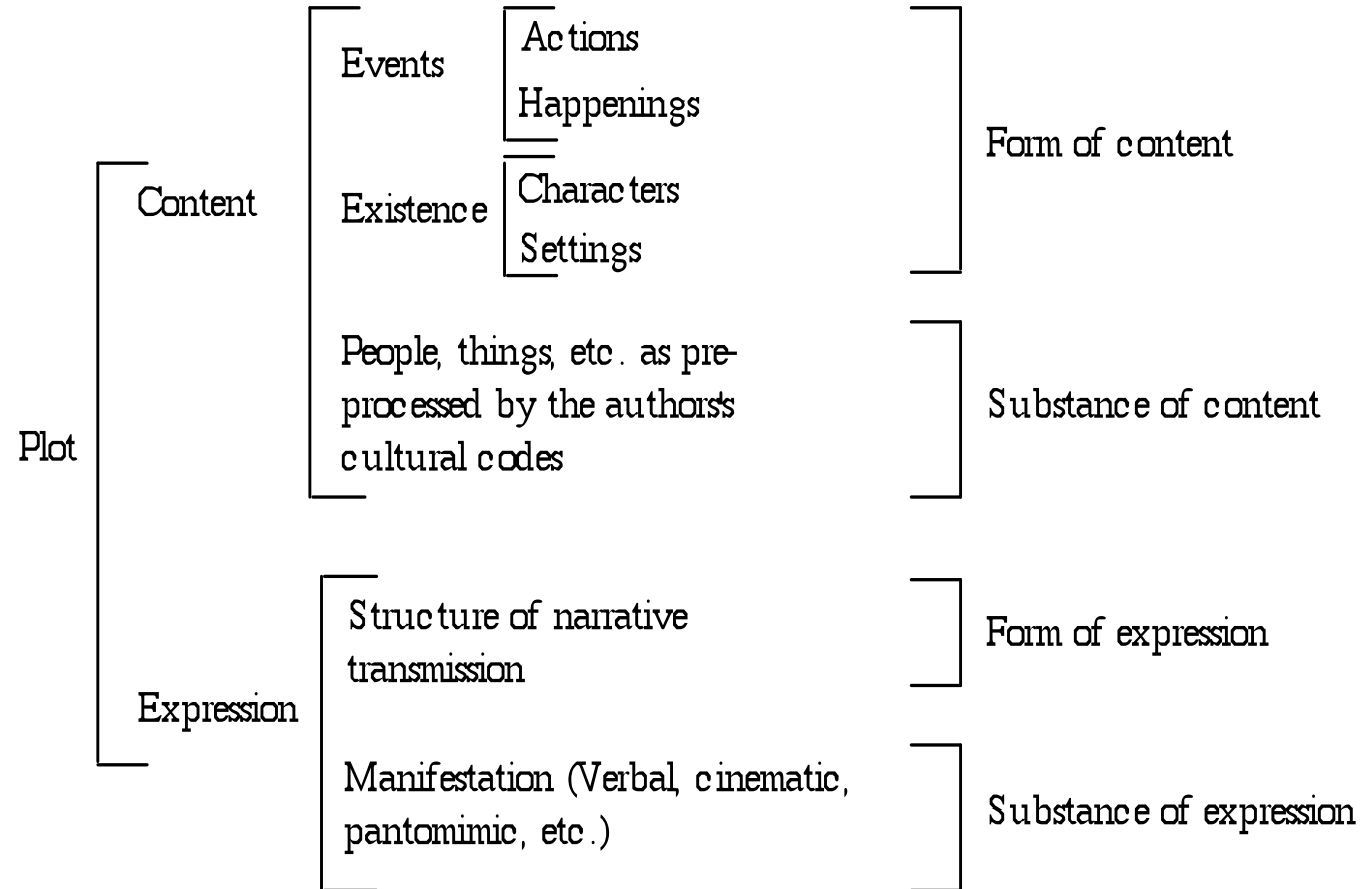
**Structure**

Catalyst – Conflict - Consequence

**Style**

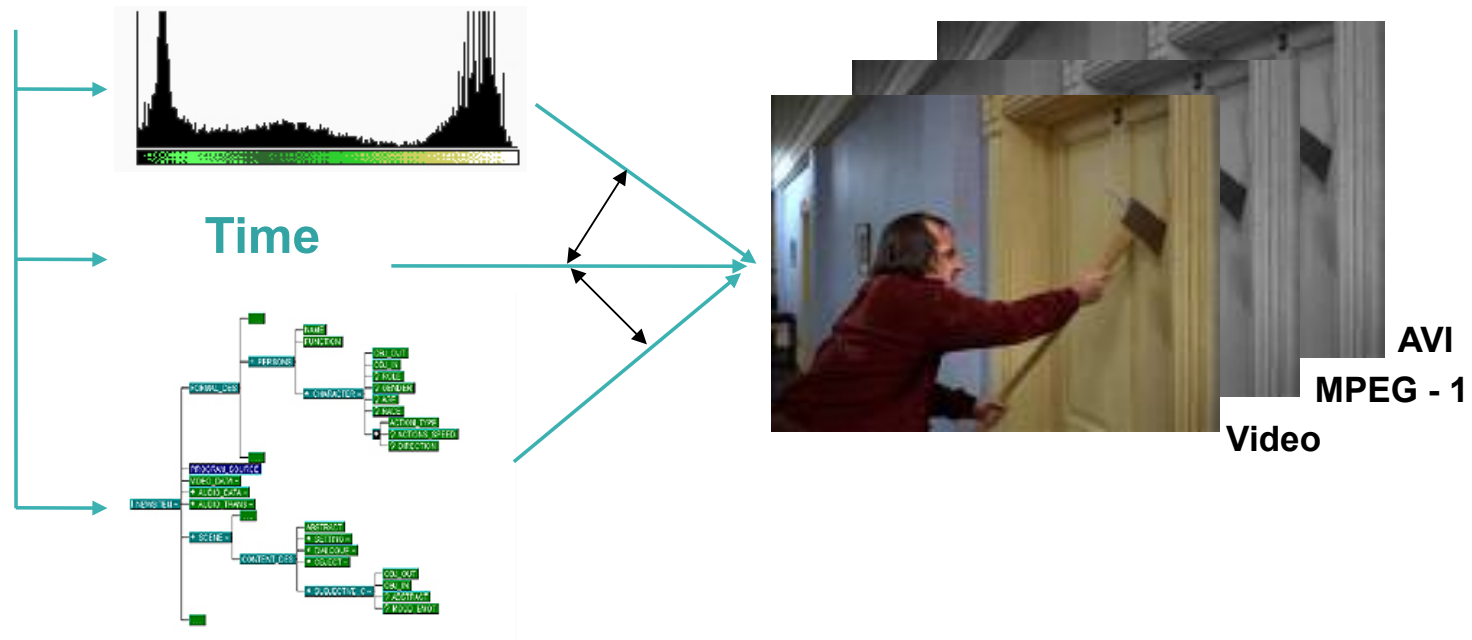
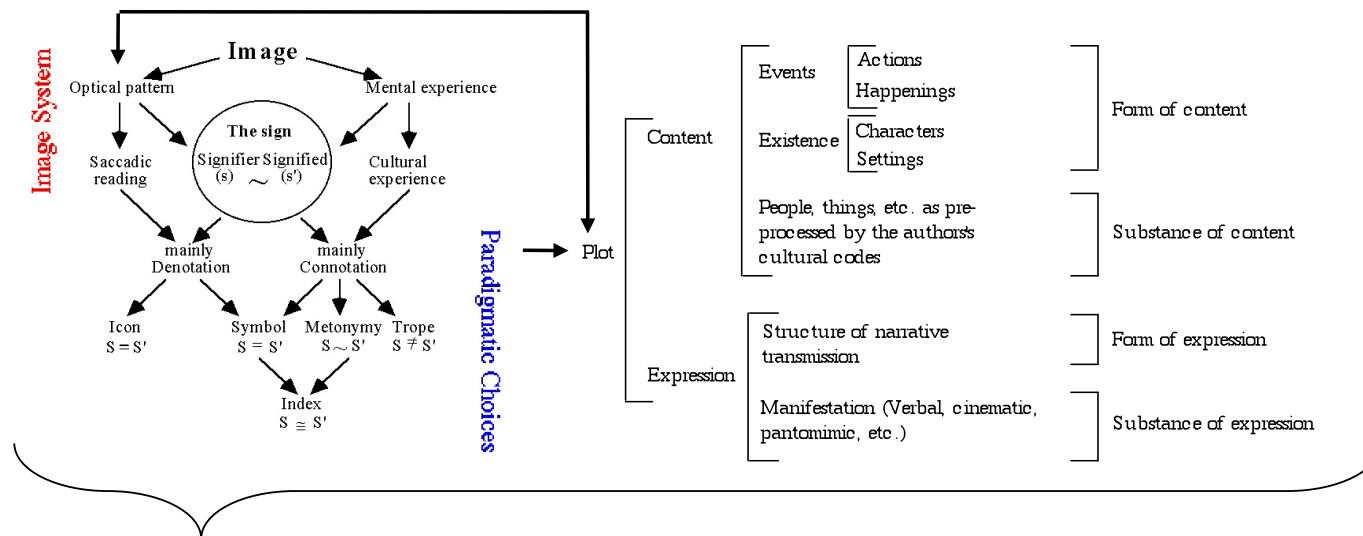
Characteristic objects and actions from the real world:  
ghosts, aliens, vampires, werewolves, curses, satanism, torture, vicious animals, zombies, cannibals, and serial killers

# Video – Meta-Semantic context - Narration II

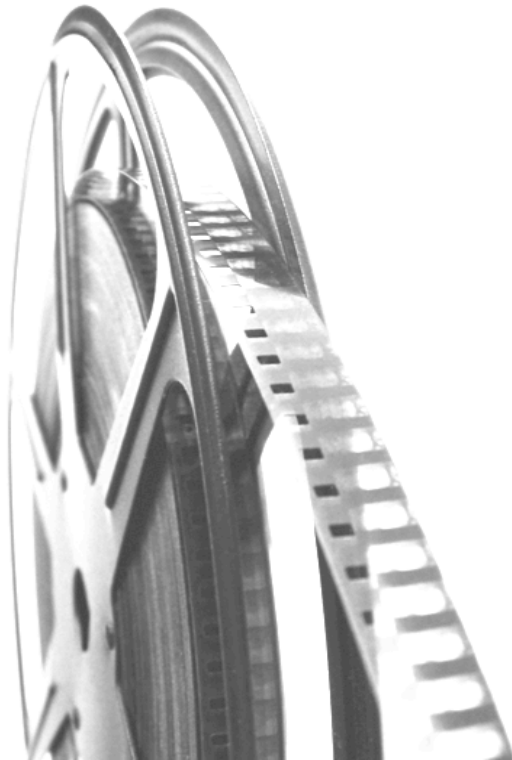


S. Chatman (1978, p. 26).

# Video – Context - summary



## Video – a temporal visual sign system - summary

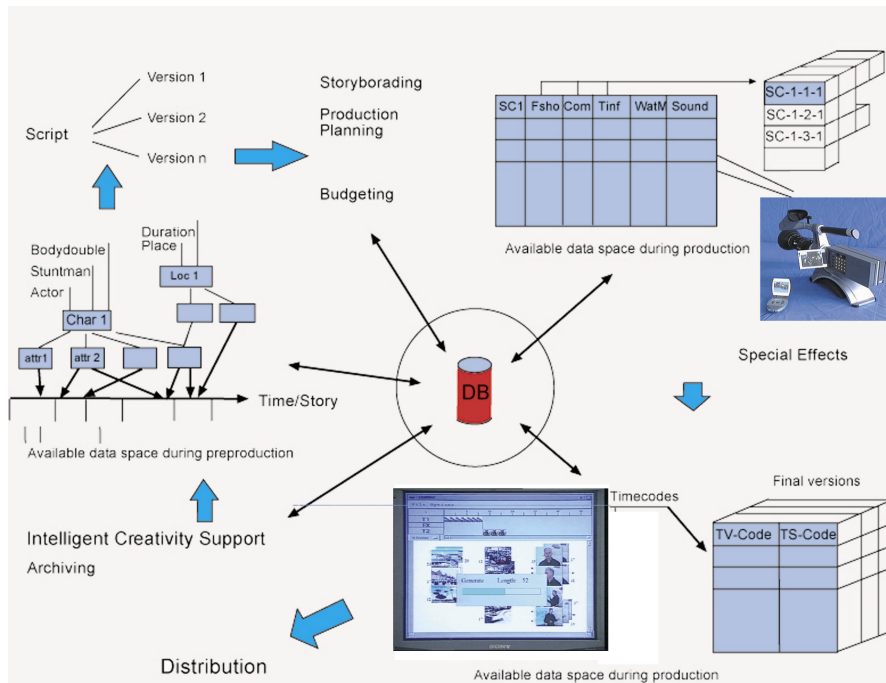


- A video is a compositional unit with individualised semantics.
- The semantics may change if a shot is juxtaposed with another shot.
- A distinction between filmic (codify the relation to reality) and cinematic codes (codify narrative communication) must be made.
- Video, though based on common human content and thematic structures, provides its own realities of time and space which are interwoven in the narrative structure.
- A story is a representational system based on two main layers, structure and content, each serving two distinct purposes (form and substance) simultaneously.

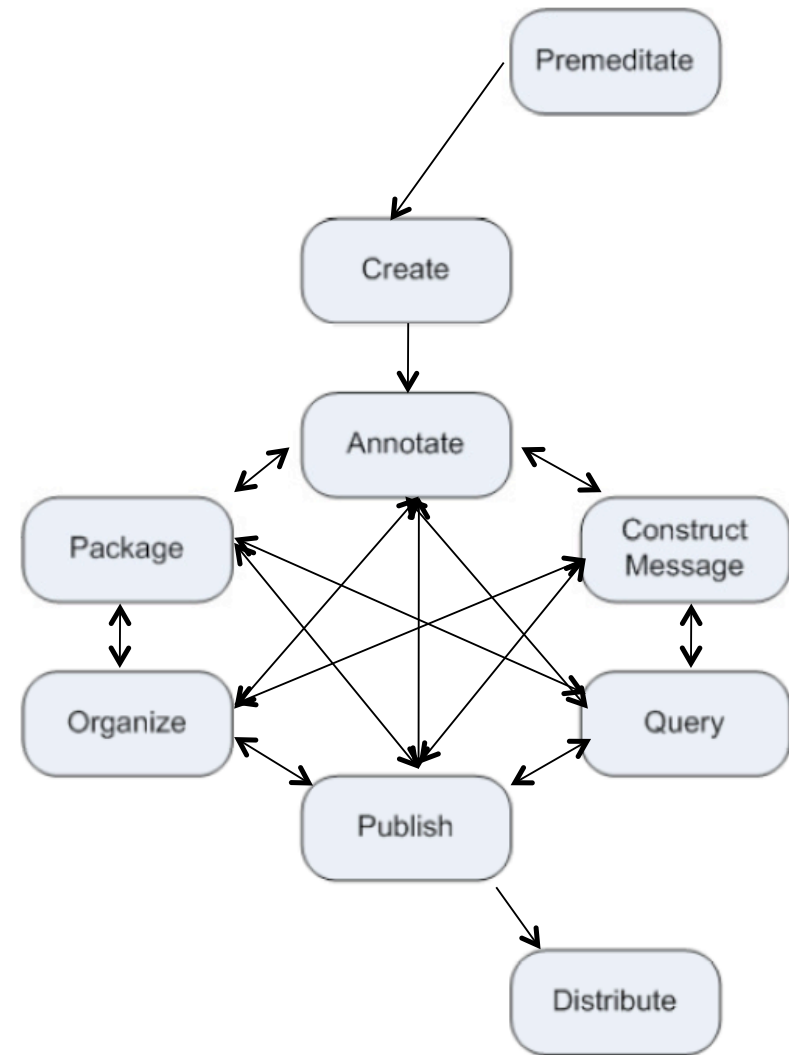
# Video – Processes



# About media processes



Details in Nack & Putz (2004)



Details in Hardman et al. (2008)

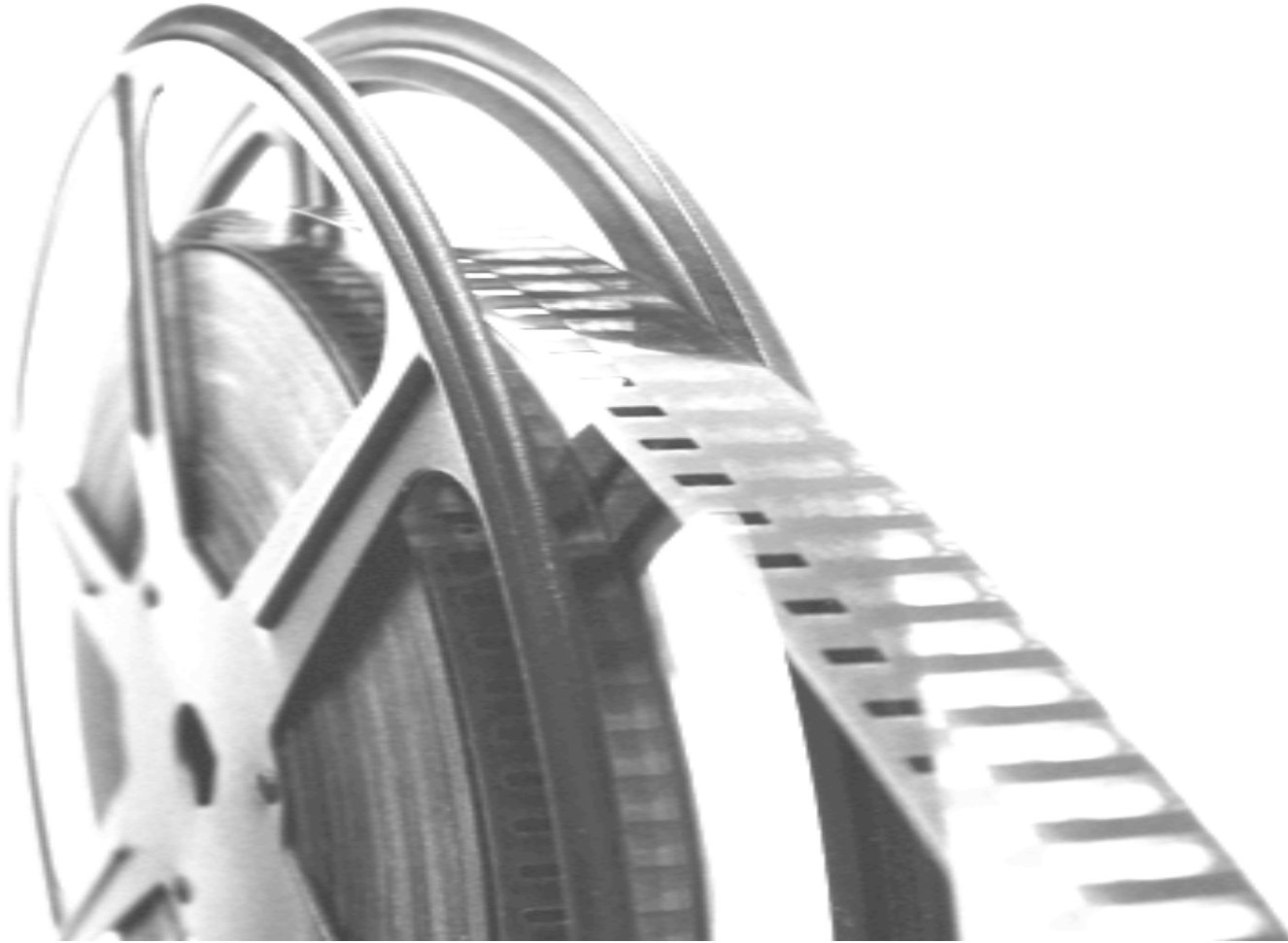


# Video – Processes



- **Access**
  - Search
  - Filter
  - Sharing
- **Creativity**
  - Manipulation / Authoring
  - Generate
- **Interaction**
  - Stimulus - response
  - Jump
  - Browse
  - Question answering

## Video – AI techniques



## Video – AI techniques

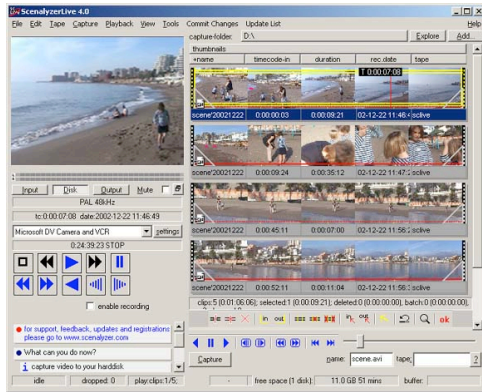


A video is spatial, temporal, situated and active item

We need

- temporal/spatial representations
- context representations
- action representations

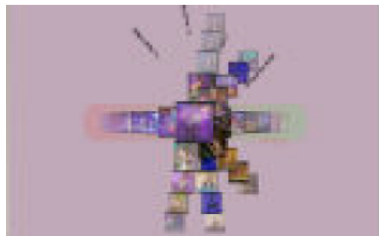
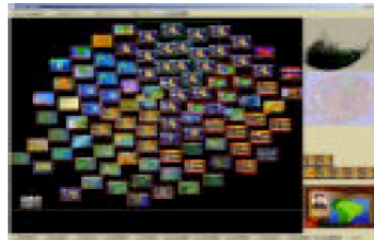
# Video – AI techniques



Make use of denotative attributes

Cut detection (change of motion, change of colour, change of objects)

Order of shots based on keyframe organised by some thread



Definition

A **thread** is a linked sequence of shots in a specified order, based upon an aspect of their content.

Static threads: pre-computed

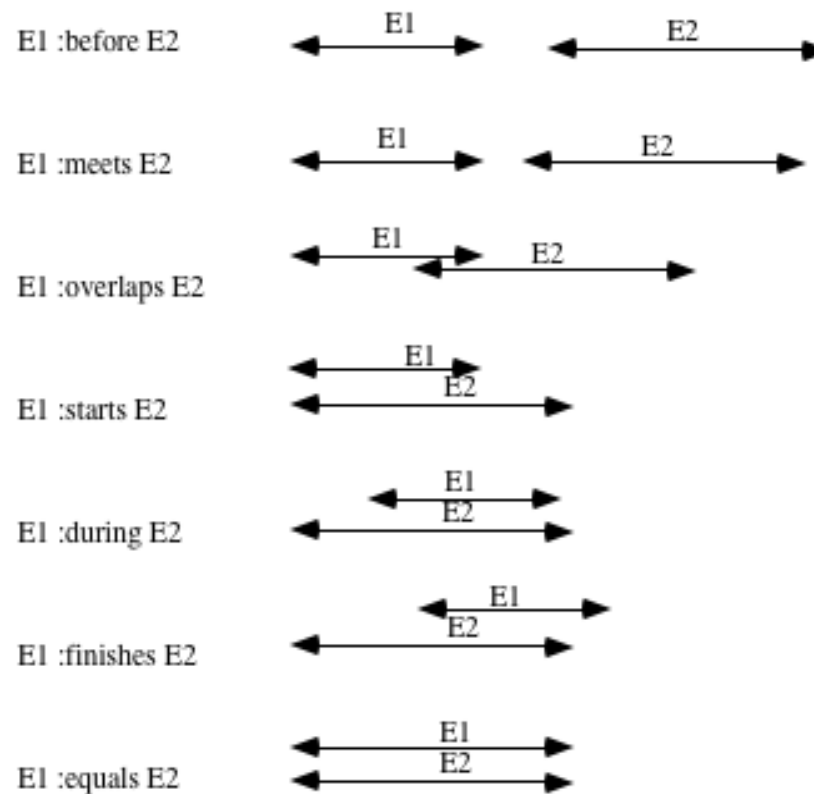
Dynamic threads: created on the fly

# Video – Representation of time I

## Time

a measuring system used to sequence events, to compare the durations of events and the intervals between them, and to quantify the motions of objects

### James Allen's classification of time

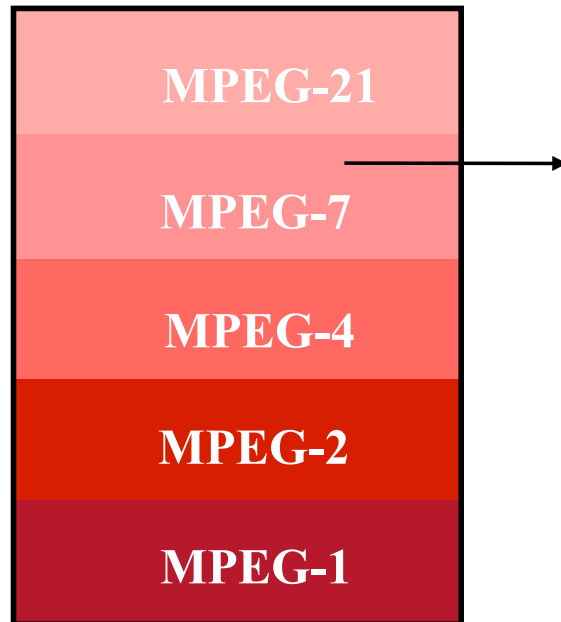


# Video – Representation of time II

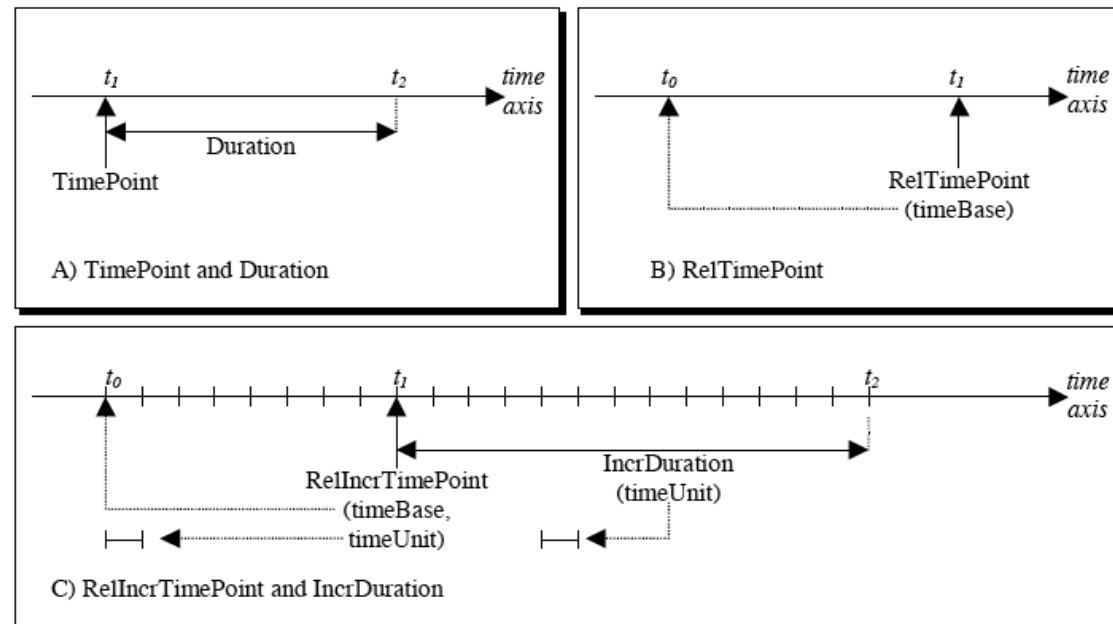
The time description in XML-S is based on a restricted set of lexical expressions of the ISO 8601 standard.

The lexical time expressions can represent

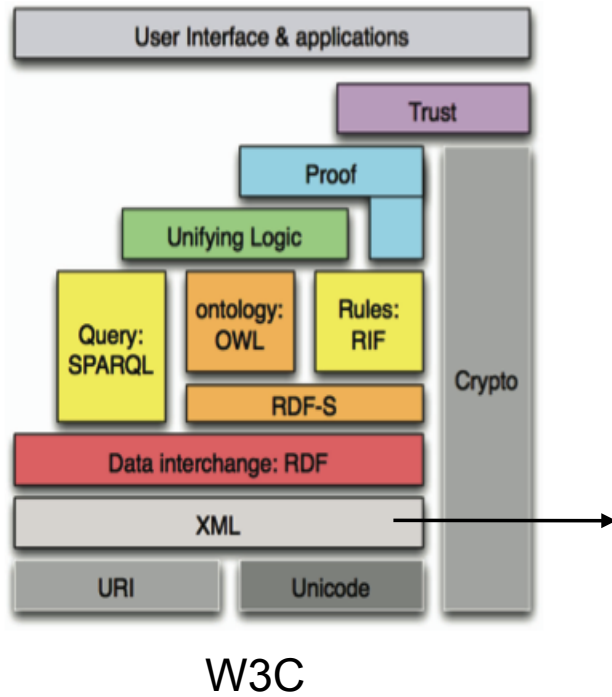
- the real world time (**Time data type**)
- the time as used in the AV data (**mediaTime data type**).



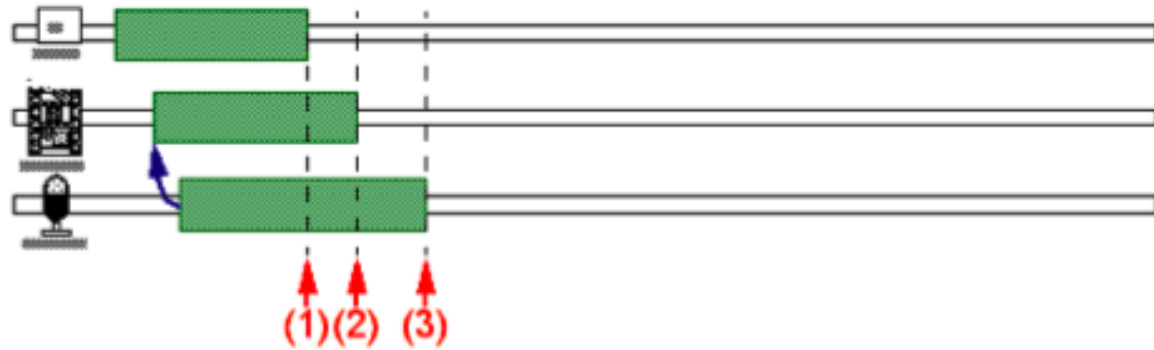
ISO



# Video – Representation of time III



## Synchronized Multimedia Integration Language (SMIL)



# Video – Representation of situation I

---



Marvin Minsky

## Frames,

a data structure used to divide knowledge into substructures by representing "stereotyped situations."

A frame contains information on

- how to use the frame,
- what to expect next
- what to do when these expectations are not met.

Some of the slots are generally unchanged.

Some slots (terminals) can change => value range

Different frames may share the same terminals.



# Video – Representation of situation I a

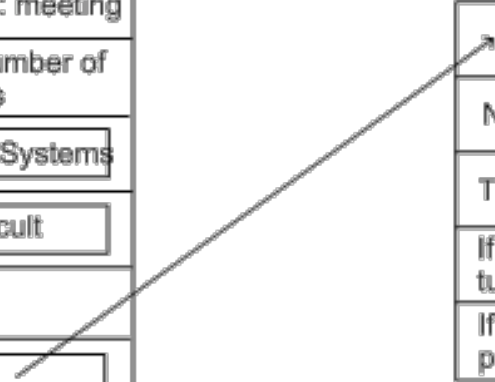
## Frames



Marvin Minsky

| Lecture                          |  |
|----------------------------------|--|
| Specialisation of:               | meeting                                  |
| Context:                         | large number of students                 |
| Course:                          | <input type="text" value="Op. Systems"/> |
| Level:                           | <input type="text" value="Difficult"/>   |
| If difficult, then pay attention |  |
| Lecturer:                        | <input type="text"/>                     |
| Room <sup>^</sup> :              | <input type="text"/>                     |

| Lecturer                                  |   |
|---|---|
| Name:                                     | <input type="text" value="Prof Jones"/> |
| Tolerance:                                | <input type="text" value="Intolerant"/> |
| If intolerant, then turn off mobile phone |   |
| If intolerant, then pay attention         |   |



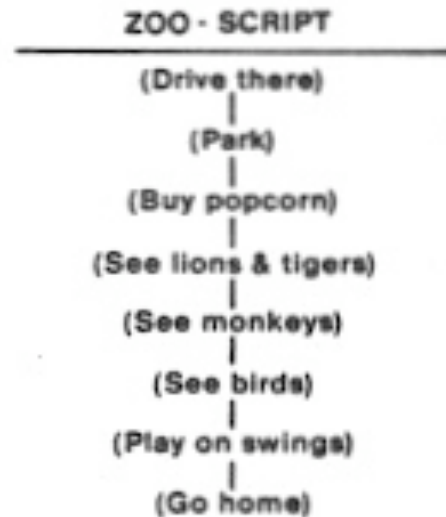
# Video – Representation of situation II

## Scripts

describe stereotyped sequences of events in a particular context (situational, personal or instrumental) Schank and Abelson

- (1) Graham went to see the tigers last night.
- (2) Paul went to see the Cubs last night.
- (3) Howard went to see the expos last night.

(1a) He saw Eric in the monkey house.



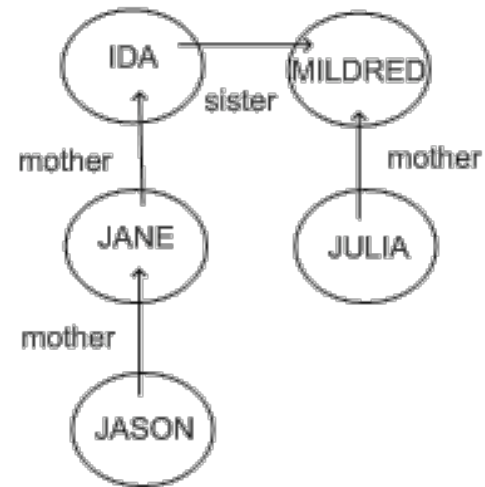
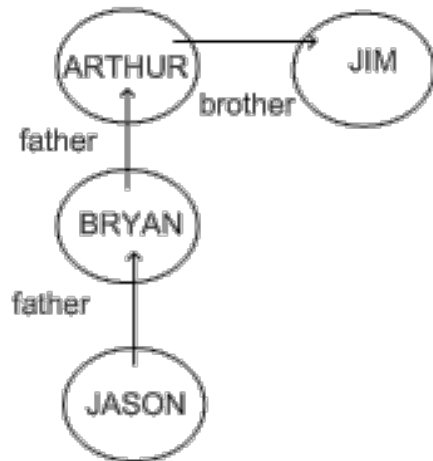
(1b) He saw Eric during the National Anthem



## Video – Representation of situation III

### Semantic graph

is a network which represents semantic relations among concepts.

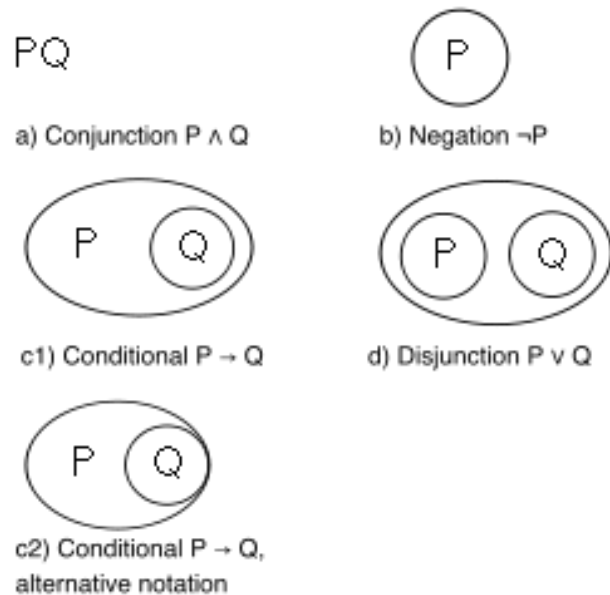


# Video – Representation of situation III

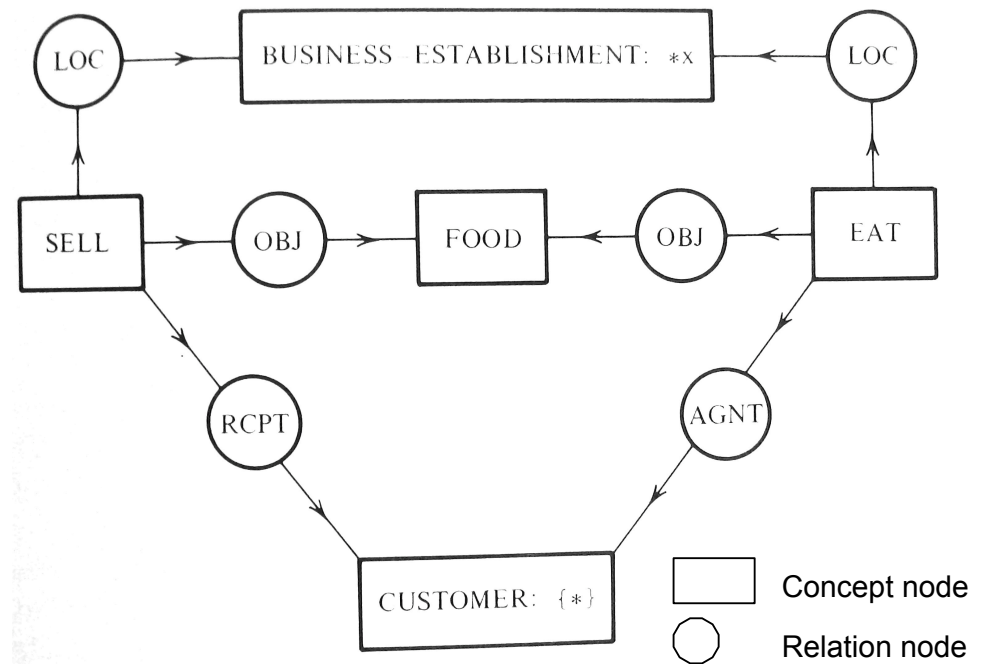
## Conceptual graph

A conceptual graph (CG) is a notation for logic based on the existential graphs of Charles Sanders Peirce and the concept of a semantic network. First used to represent the conceptual schema in databases Sowa

Existential graph (Peirce)



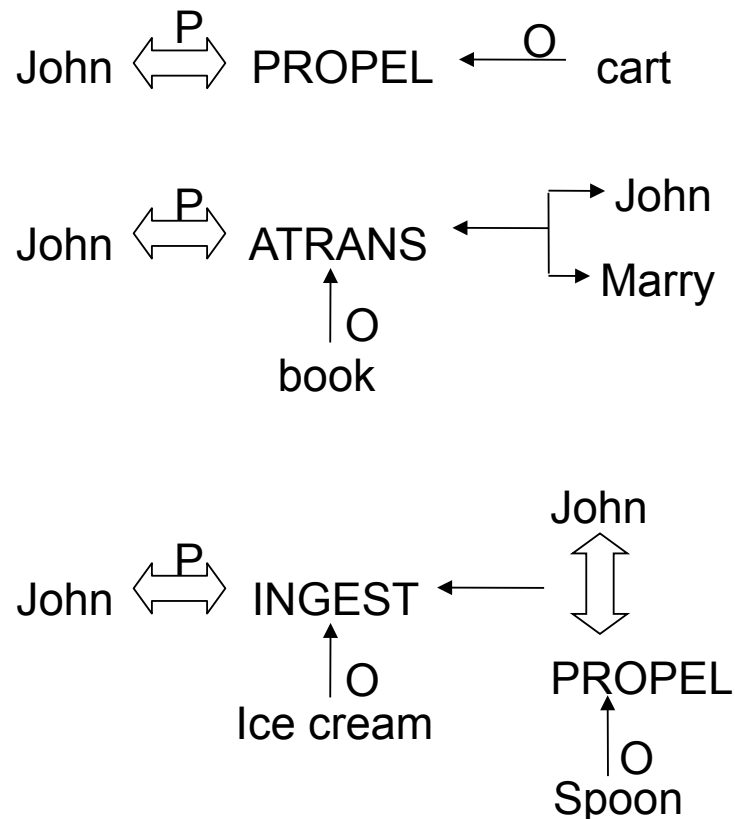
Conceptual graph (Sowa)



# Video – Representation of action

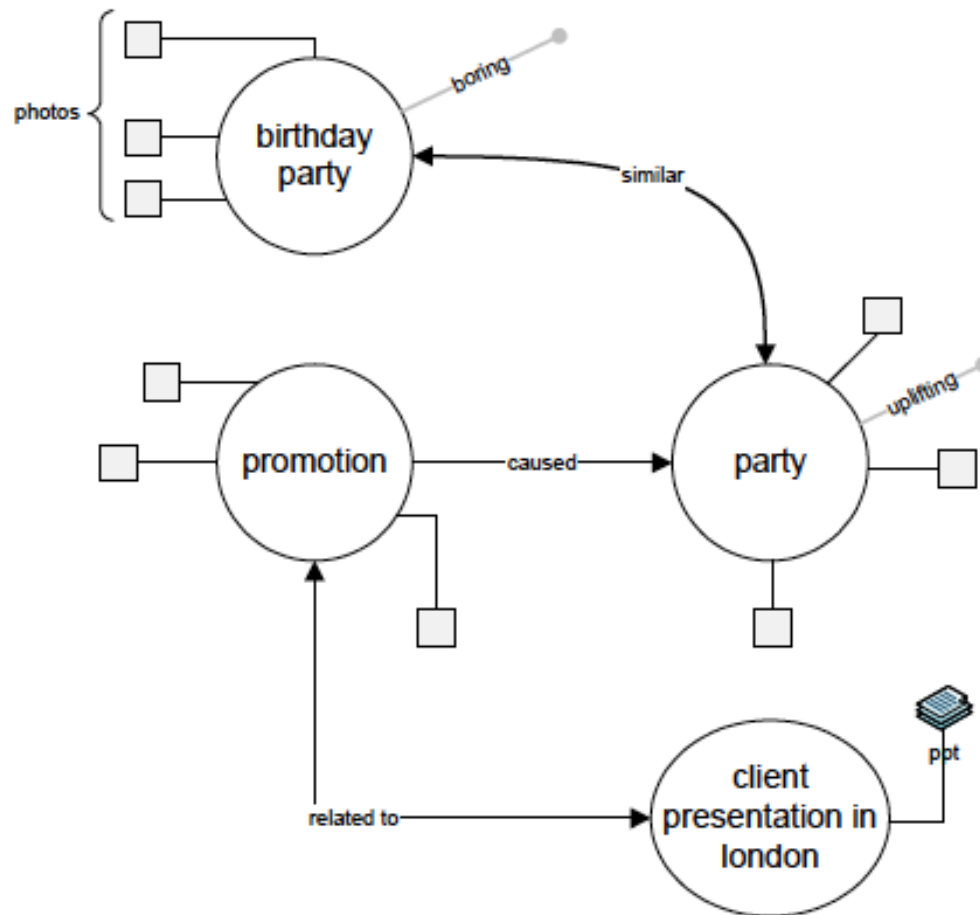
## Conceptual dependencies (Schank)

which describe human action through a small set of 11 composable primitives



- ATRANS - transfer of an abstract relationship
- PTRANS - physical transition [of things]
- PROPEL - apply physical force to an object
- MTRANS - transfer of mental information
- MBUILD - construct new information
- SPEAK - Utter a sound
- ATTEND - Focus a sense on a stimulus
- MOVE - Movement of a body part by owner
- GRASP- Actor grasping an object. e.g. clutch.
- INGEST - Actor ingesting an object. e.g. eat.
- EXPEL - Actor getting rid of an object from body

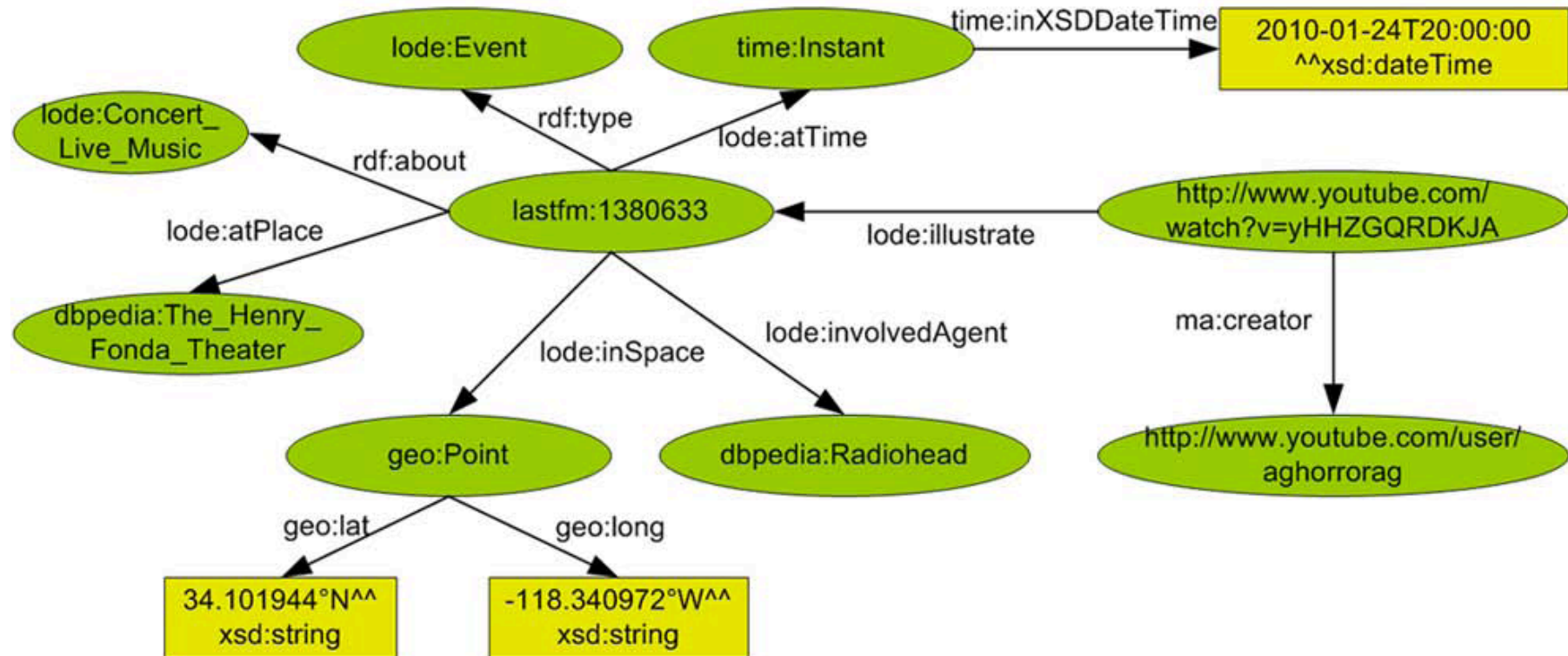
# Event - Eventory



An Event Based Media Repository with event description in MySQL

X. Wang, S. Mamadgi, A. Thekdi, A. Kelliher, and H. Sundaram. Eventory – an event based media repository. In *Semantic Computing. IEEE, 2007*

# Event - LOD



Raphaël Troncy, Bartosz Malocha and André Fialho. Linking Events with Media. In the Open Track of the Linked Data Triplification Challenge, colocated with the 6th International Conference on Semantic Systems (I-SEMANTICS'10), Graz, Austria, September 1-3, 2010

# Video – Interaction



A

B



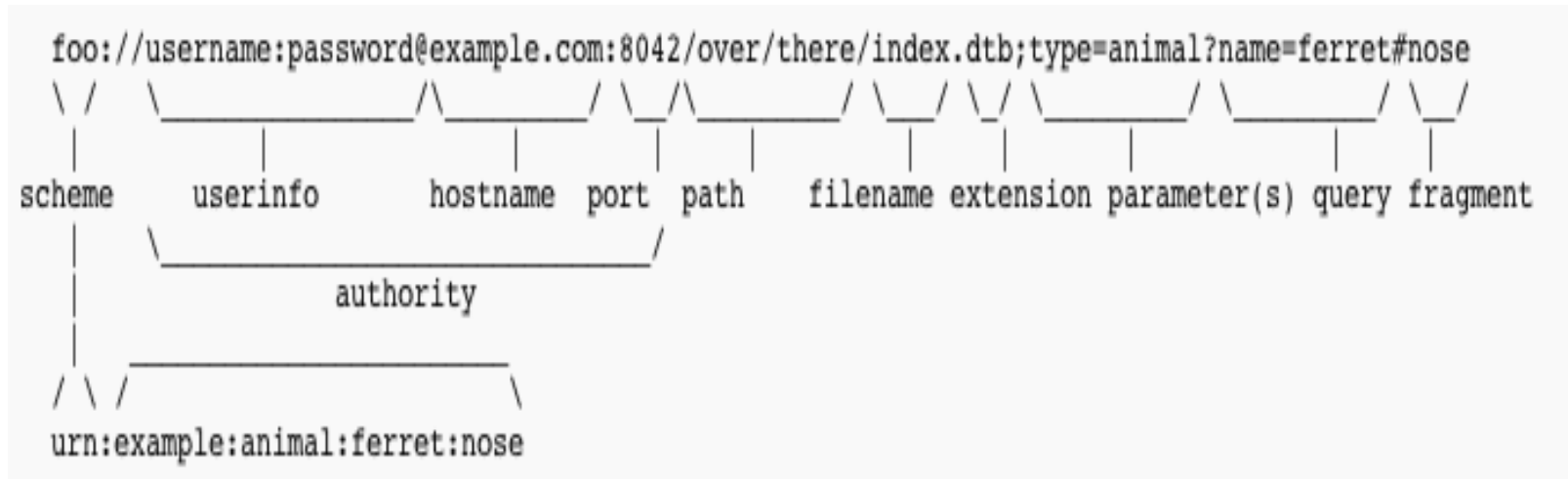
The degree of **responsiveness**, examined as a **communication process** in which each message is related to the previous messages exchanged, and to the relation of those messages to the messages preceding them.



**Semantic LINKS**



## Video – Representation of story flow (links)



But Hypervideo offers more:

- The link is the mechanism to represent the dynamics and rhetoric of the video [Sawhney et al.1996]
- Work that describes in some detail the relation between the structure of hypertext and film, and links to edits [Miles 2000, Mancini 2000]

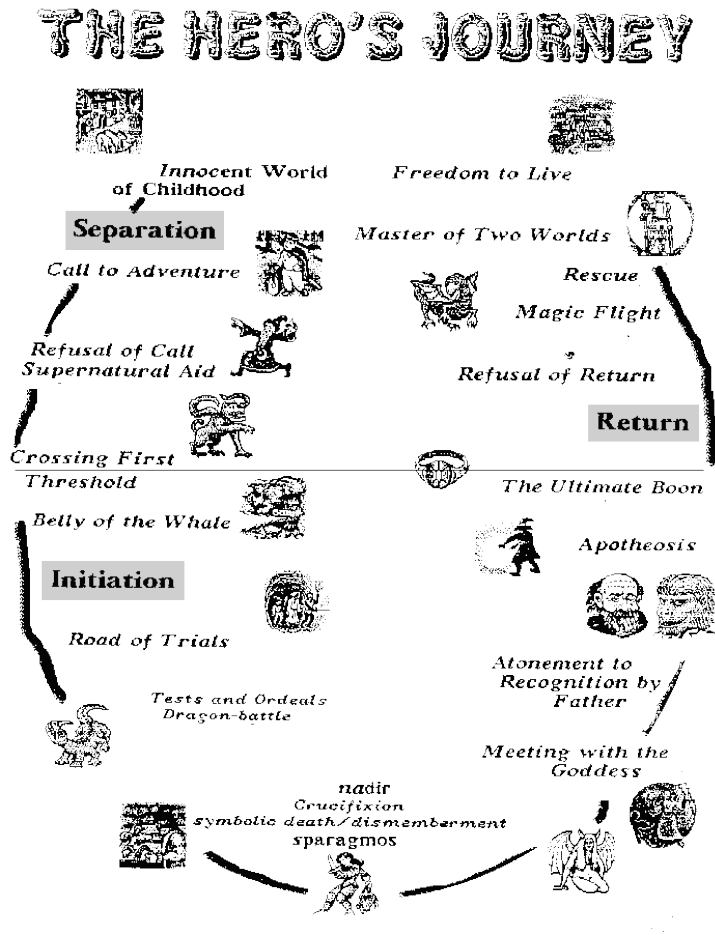
# Video – Representation of story flow (grammar)

## Story Grammar

describes the structure of a story including story elements, such as setting, characters, events, resolution, etc., and the relationships between them.

Main influenced by Propp's work on Russian folktales and Chomsky's transformational grammar.

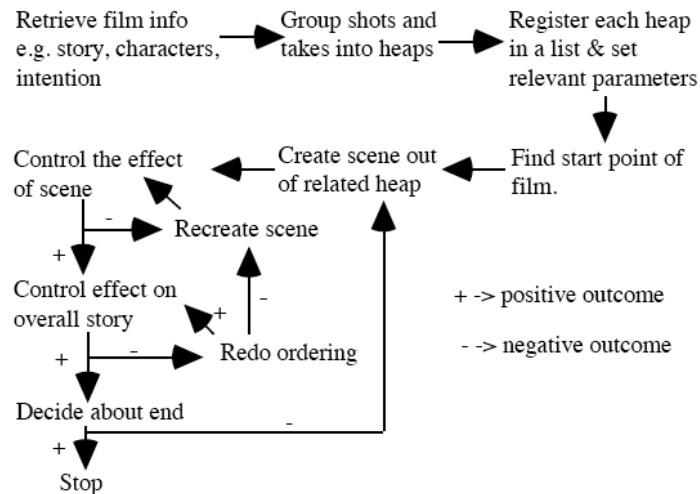
[Colby (1973); Knitsch & van Dijk (1978); Lakoff (1972); Mandler (1977); Rumelhart (1975, 1977); Thorndyke (1977)].



# Video – Representation of story flow (planner)

## Goals / Plans (Schank & Abelson)

represent high-level structures that control processes.

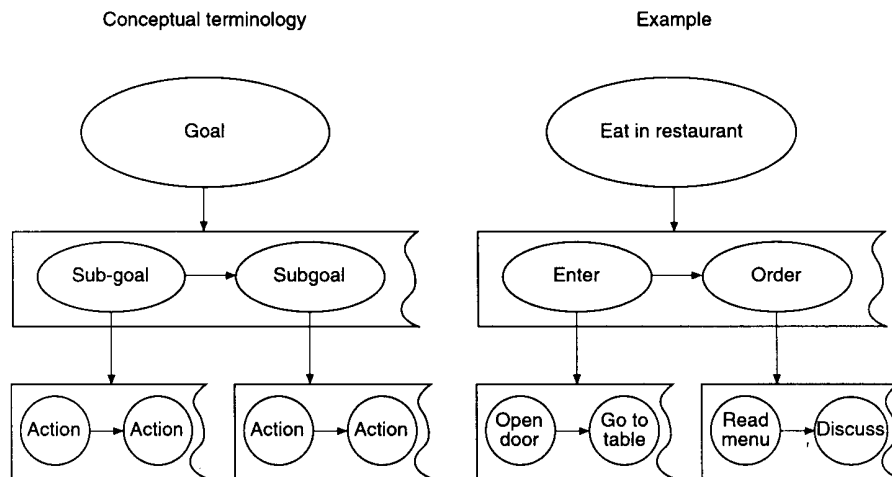


## Pro Plan Arguments

The main arguments against this approach are advanced by Black & Wilensky (1979)

- There is always a story more complex than the latest grammar
- Grammars are static, storytelling as a process is fluid

# Video – Representation of story flow (planner)



A typical planner takes three inputs:

- a description of the initial state of the world,
- a description of the desired goal
- a set of possible actions.

The planner produces a sequence of actions that lead from the initial state to a state meeting the goal

**forward** versus **back-ward chaining**.

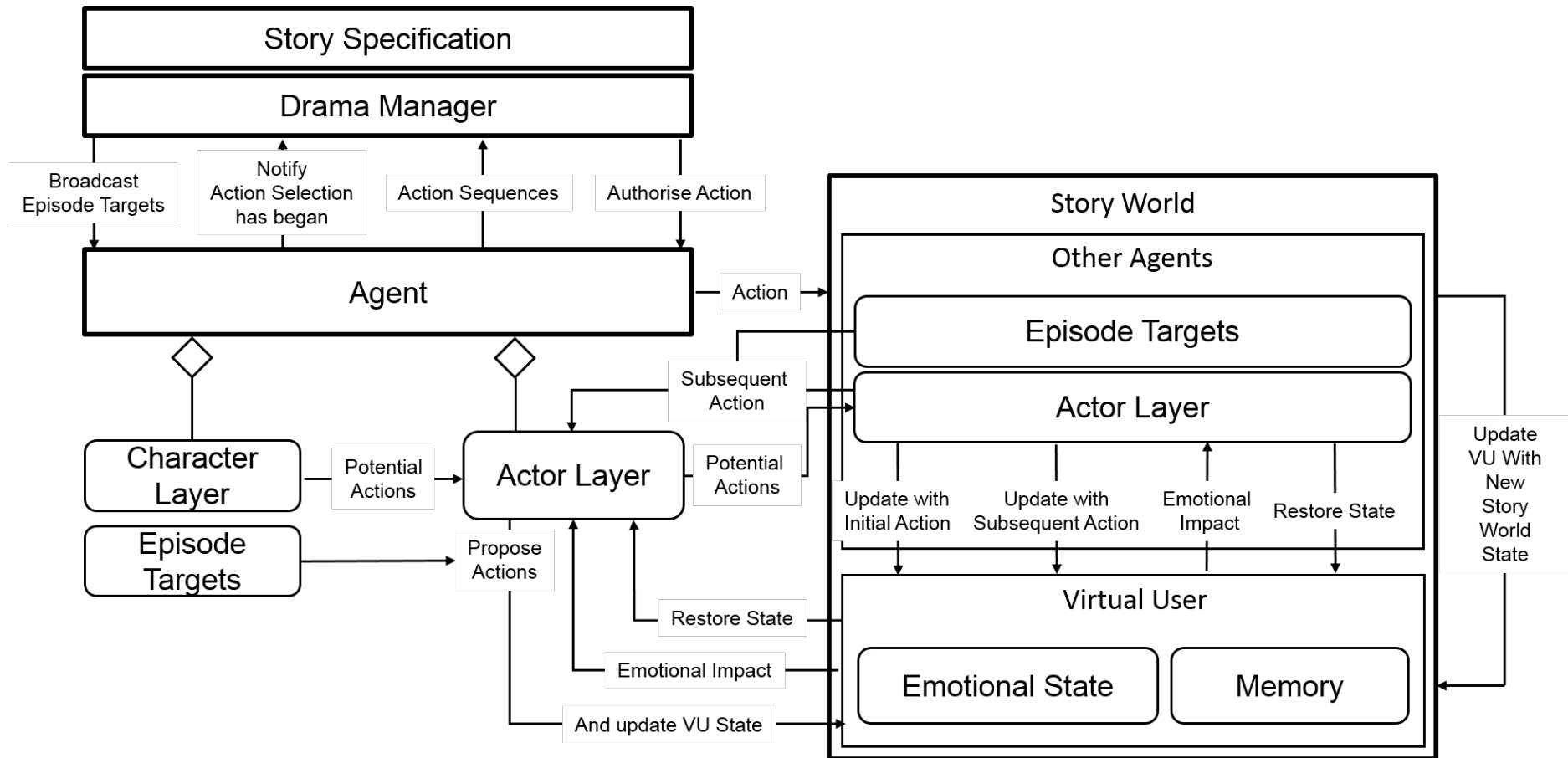


data-driven



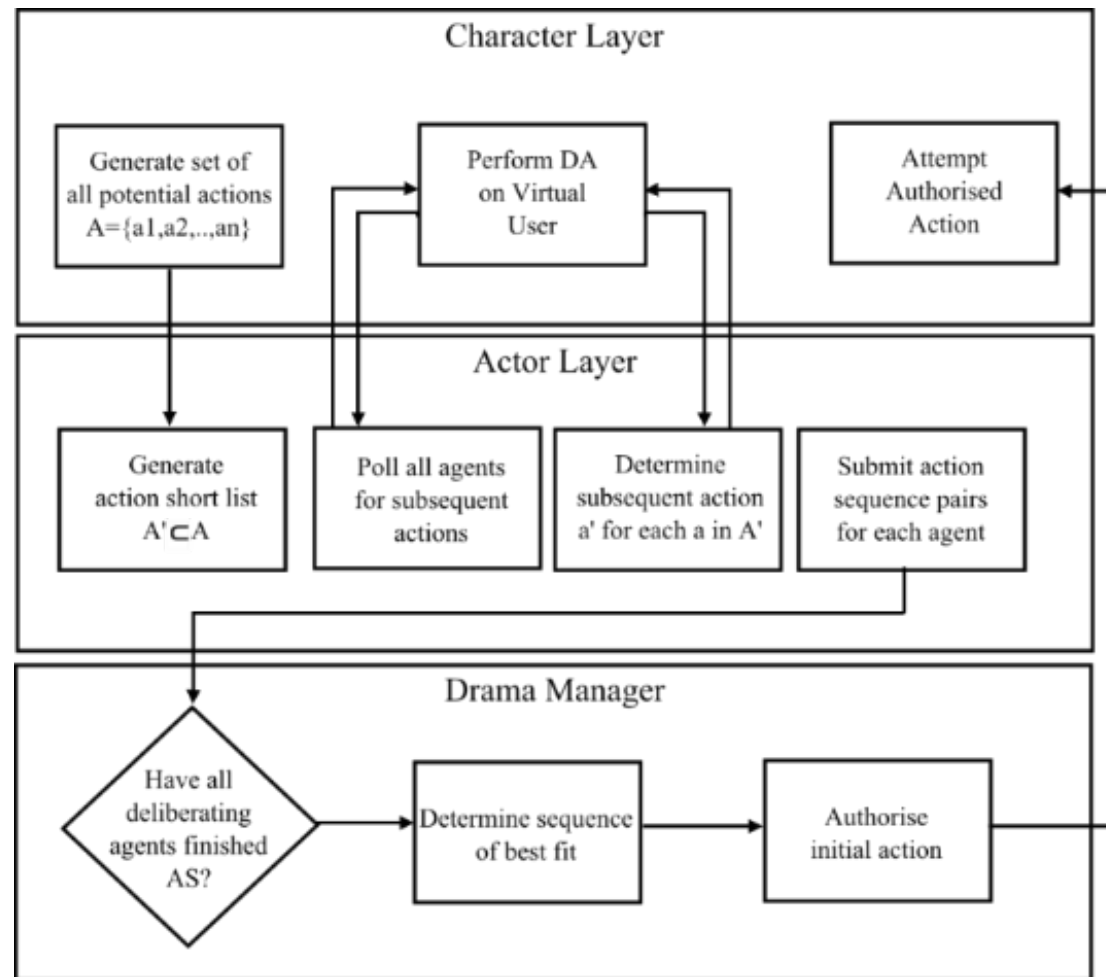
goal-driven

# Video – Representation of story flow (agent)



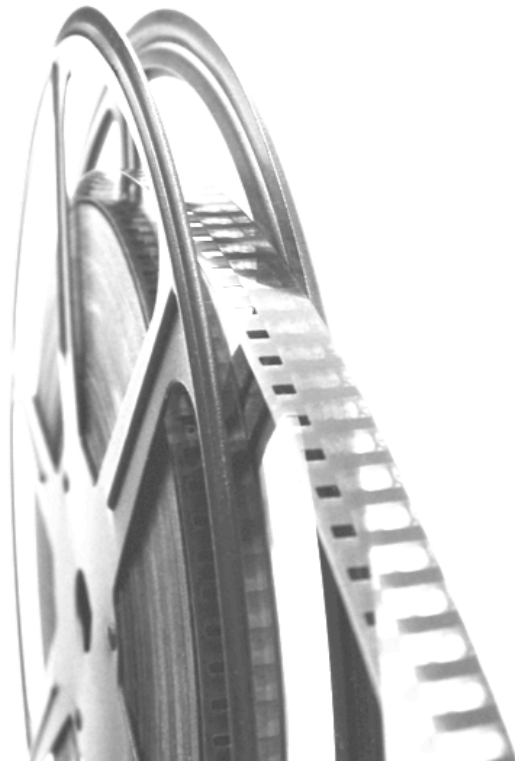
Weallans, Allan, Sandy Louchart, and Ruth Aylett. "Distributed drama management: beyond double appraisal in emergent narrative." *Interactive Storytelling*. Springer Berlin Heidelberg, 2012. 132

## Video – Representation of story flow (agent)



Weallans, Allan, Sandy Louchart, and Ruth Aylett. "Distributed drama management: beyond double appraisal in emergent narrative." *Interactive Storytelling*. Springer Berlin Heidelberg, 2012. 132

# Video – AI techniques - summary



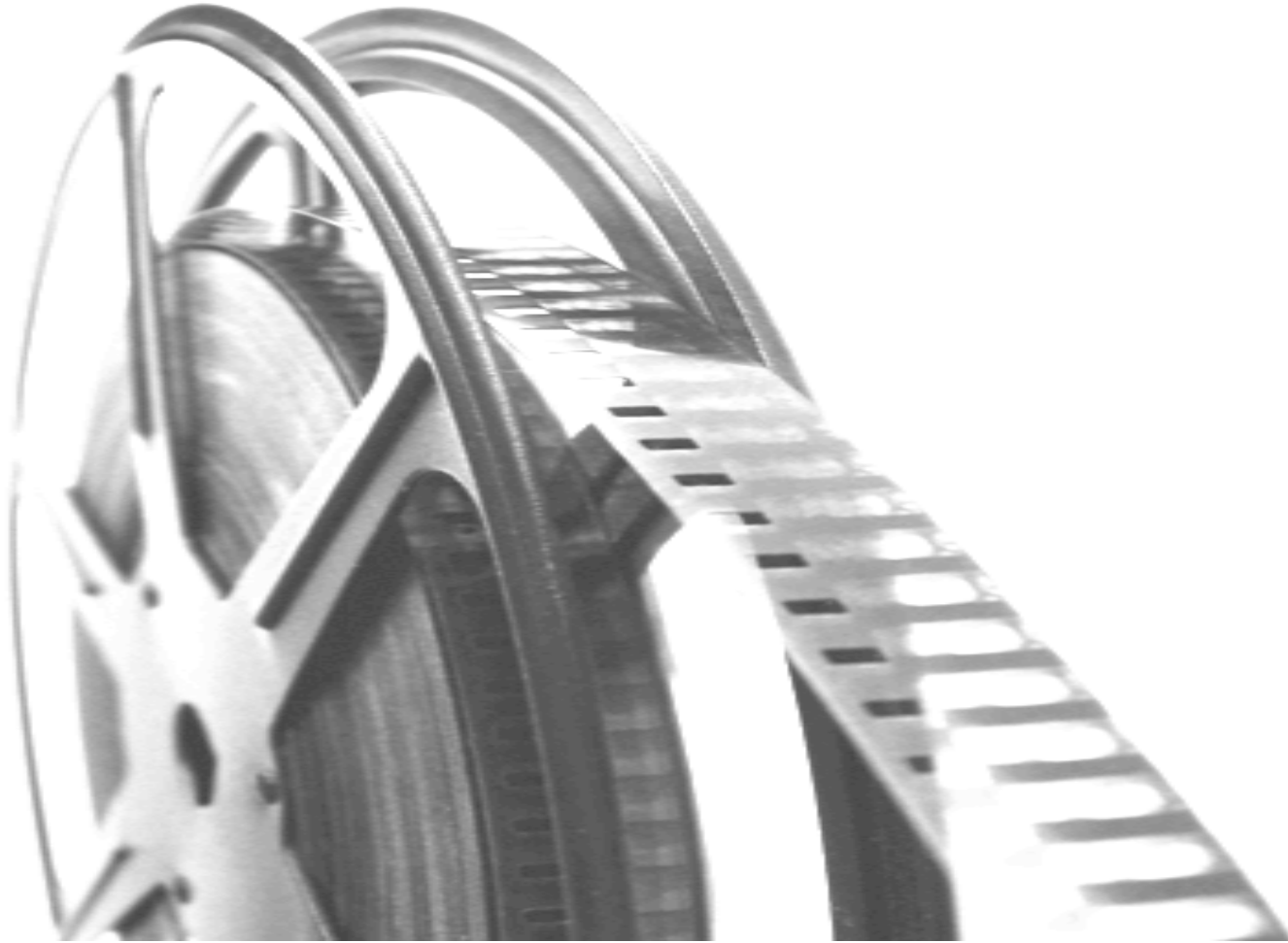
Relevant content representation mechanisms for video are:

- Scripts
- Conceptual Dependencies
- Semantic networks
- Semantic links

A primarily structure-oriented approach (grammar) to the temporal aspects of video with respect to generation and interaction is not appropriate.

A planning approach (planner or agent) seems more promising, as then the different levels can be separated, while maintaining the interaction between the structure and content layers.

# Video – Applications





# Video – Applications I

The screenshot shows the YouTube homepage in a browser window titled "YouTube - Broadcast Yourself." The address bar shows "http://www.youtube.com/". The browser's taskbar includes various open tabs like "delicio.us", "tag this", "Manuscript Central", "Google Documenten", "LEO Deutsch-Englisc...", "ACM Digital Library", "caisME", "Van Dale Taalweb | ...", and "SWI-Prolog's Home".

The YouTube logo is at the top left with the tagline "Broadcast Yourself™". Navigation tabs for "Home", "Videos", "Channels", and "Community" are visible. A search bar is located below the tabs, and an "Upload" button is on the right.

Under "Videos being watched right now...", there are five video thumbnails with durations: 00:04, 00:46, 00:25, 05:14, and 00:46.

The "Promoted Videos" section features four thumbnails with titles: "Midnight Vignette" (SecretyJag), "Donkere dagen voor" (BeeldenGeluid), "Hoe staat jouw Kers..." (sterrennl), and "MAX & Loretta - Naa..." (omroepmax).

The "Featured Videos" section has three items:

- Pirates VS Ninjas**: From booshoe37, 15,019 views, 5 stars, 02:21. More in Comedy.
- Star Wars according to a 3 year ...**: From fistofblog, 1,274,242 views, 5 stars, 01:30. More in People & Blogs.
- Do you like movies? Do you make...**: From fromheretoawesome, 244,123 views, 5 stars, 04:59. More in Film & Animation.

On the right side, there is a "Login" form with fields for "Username: fnack" and "Password: \*\*\*\*\*", a "Login" button, and links for "Forgot Username", "Forgot Password", and "Login with your Google account".

The "What's New" section includes:

- YouTube Mobile**: New! Watch ALL YouTube videos on your mobile device.
- Warp!**: Visually fly through YouTube videos in the Fullscreen player.
- RSS Feeds**: Click on the "RSS this page" link to get fresh videos delivered.

A "Taking Films 'From Here To Awesome'" section mentions that technology has democratized film production and distribution, listing success stories like "Four Eyed Monsters", "M dot Strange", "Lance Weiler" and "Robert...". A link "Read more in our Blog" is provided.

At the bottom left, it says "Read i.ytimg.com". At the bottom right, there is a "Tor Disabled" notification.

# Video – Applications II

Storyspace: HyperCafe

http://www.eastgate.com/storyspace/film/HyperCafe.html

serious hypertext Eastgate Systems Inc.

hypertext tools fiction nonfiction poetry books Hypertext NOW

## Storyspace

Download your FREE Storyspace demo now!

for Macintosh

Storyspace 2.5.0 for MacOS X: \$295  
[Add to shopping cart](#)  
You can always remove it later.

for Windows

Storyspace 2.0 for Windows: \$295  
[Add to shopping cart](#)  
You can always remove it later.

Need to upgrade?

Storyspace Macintosh Upgrade: \$95  
[Add to shopping cart](#)  
You can always remove it later.

Need to upgrade?

Storyspace Windows Upgrade: \$95  
[Add to shopping cart](#)  
You can always remove it later.

### USER STORIES

[Berreby Ucerler](#)

### ABOUT STORYSPACE

[Overview](#)  
[Storyspace maps](#)  
[Writing spaces](#)

### STORYSPACE AND HYPERVIDEO SCREENWRITING

HyperCafe, an award-winning interactive film, places the viewer inside a virtual cafe. It's a strangely real yet cinematic experience, a video environment where stories unfold around the viewer.

Select a conversation, and the navigation pan fades into a closeup, two men talking, and one's saying to the other, "in fact, our words over here don't affect their words over there."

Another table comes into view, another possible conversation, next to the first. After a few moments, the second table fades away if not selected. The link unrealized (which is a choice in itself), the story continues: "I find that highly questionable," the other man says. "What if I yell 'fire'? What then?" [1]

Written, directed, and produced by the team of Nitin Sawhney, David Balcom, and Ian Smith, a report on HyperCafe won the prestigious Engelbart award at Hypertext 96. A fuller discussion of HyperCafe has just appeared in the *IEEE Multimedia Journal*, and the work itself has been featured at conferences and festivals throughout the U.S. and Europe.

A screen shot of the HyperCafe shooting script. In Storyspace. Copyright 1996 by Nitin Sawhney, David Balcom, and Ian Smith, used by permission.

The best HyperCafe success in making out cinema and interactivity.

Done Tor Disabled

# Video – Applications III

Lev Manovich: Soft Cinema

http://softcinema.net/?reload

Welcome to the HCS ... UvA - Medewerkers ... Blackboard Learning ... Manuscript Central Google Documenten LEO Deutsch-Englisc... ACM Digital Library caisME Van Dale Taalweb | ... SWI-Prolog's Home

Lev Manovich Lev Manovich: Soft Cinema

## SOFT CINEMA : AMBIENT NARRATIVE

- + VISUALS
- + CONCEPTS
- + MOVIE EDITIONS
- + PUBLICATIONS
- + ARCHITECTURE
- + EXHIBITIONS
- + CREDITS
- + LINKS

### PROJECT SUMMARY

Soft Cinema project mines the creative possibilities at the intersection of software culture, cinema, and architecture. Its manifestations include **films**, dynamic **visualizations**, computer-driven **installations**, architectural **designs**, print **catalogs**, and DVDs. In parallel, the project investigates how the new representational techniques of soft(ware) cinema can be deployed to address the new dimensions of our time, such as the rise of mega-cities, the "new" Europe, and the effects of information technologies on subjectivity.

At the heart of the project is custom **software** and media databases. The software edits movies in real time by choosing the elements from the **database** using the systems of rules defined by the authors.

contact: manovich [at] ucsd [dot] edu

### SOFT CINEMA: NAVIGATING THE DATABASE DVD now available

Lev Manovich and Andreas Kratky

SOFT CINEMA: Navigating the Database

DVD-video with 40 page color booklet  
The MIT Press, 2005  
ISBN 0-262-13456-X

SOFT CINEMA: NAVIGATING THE DATABASE is available through The MIT Press <mitpress.mit.edu>, online resellers <www.amazon.com, www.barnesandnoble.com, etc.> and selected bookstores.

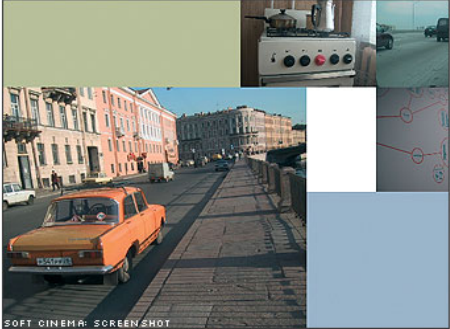
What kind of cinema is appropriate for the age of Google and blogging? Automatic surveillance and self-guided missiles? Consumer profiling and CNN? To investigate answers to this question Lev Manovich - one of today's most influential thinkers in the fields of media arts and digital culture - has paired with award-winning new media artist and designer Andreas Kratky to create the Soft Cinema project. They have also invited contributions from leaders in other cultural fields: DJ Spooky, Scanner, George Lewis and Jóhann Jóhannsson (music), servo and Andreas Angelidakis (architecture), Schoenerwissen/Office for Computational Design (data visualization), and Ross Cooper Studios (media design).

SOFT CINEMA: Navigating the Database is the Soft Cinema project's first DVD published and distributed by The MIT Press (2005). Although the three films presented on the DVD reference the familiar genres of cinema, the process by which they were created and the resulting aesthetics fully belong to the software age. They demonstrate the possibilities of soft(ware) cinema - a 'cinema' in which human subjectivity and the variable choices made by custom software combine to create films that can run infinitely without ever exactly repeating the same image sequences, screen layouts and narratives.

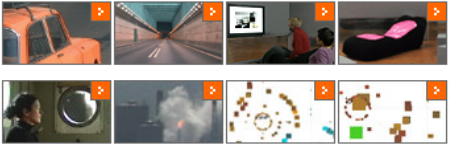
'Mission to Earth' is a science fiction allegory of the immigrant experience. It adopts the variable choices and multi-frame layout of the Soft Cinema system to represent 'variable identity'. 'Absences' is a lyrical black and white narrative that relies on algorithms normally deployed in military and civilian surveillance applications to determine the editing of video and audio. 'Texas' is a 'database narrative', which assembles its visuals, sounds, narratives, and even the identities of its characters from multiple databases.

The DVD was designed and programmed so that there is no single version of any of the films. All the elements - including screen layout, the visuals and their combination, the music, the narrative, and the length - are subject to change every time the film is viewed.

The development of Soft Cinema project was made possible by the commissions from ZKM Center for Art and Media and the BALTIC, The Centre for Contemporary Art. The resulting computer-driven installations and films have been exhibited in museums, galleries, media and film festivals around the world.



SOFT CINEMA: SCREENSHOT



Done

Tor Disabled

# Video – Applications IV

Advene project

http://liris.cnrs.fr/advene/#anglais

Most Visited SAP NetWeaver Portal Second-hand Amste... Blackboard Learning... WordPress.com » Yo... Vertalen .nu : vertali... LEO Deutsch-Englis... ACM Digital Library Online Mind Mappin... iipcreate.com Twitter / Home

## ADVENE

Annotate Digital Video, Exchange on the Net

Project Application Examples Support Research

Presentation News Team Contact Screenshots Download

### Project presentation

Advene (*Annotate Digital Video, Exchange on the NET*) is an ongoing project in the **LIRIS laboratory** (UMR 5205 CNRS) at **University Claude Bernard Lyon 1**. It aims at providing a model and a format to share annotations about digital video documents (movies, courses, conferences...), as well as tools to edit and visualize the hypervideos generated from both the annotations and the audiovisual documents. Teachers, moviegoers, etc. can use them to exchange multimedia comments and analyses about video documents. The project also aims at studying the way that communities of users (teachers, moviegoers, students...) will use these self-publishing tools to share their audiovisual "readings", and to envision new editing and viewing interfaces for interactive comment and analysis of audiovisual content.

Package creation (definition of annotations and views)

Package utilisation (visualisation, enrichment)

Package sharing:  
- copy via webservice, email, etc.  
- reference through imports

As seen on the figure above, Advene prototype features:

1. At the package creation level : annotation of audiovisual documents (association of typed information to temporal fragments), creation of visualisation means (views).
2. Exchange of annotations and visualization modes in packages independently from the audiovisual material (images and sounds). If needed for the visualization of the data, pictures and sound clips can be extracted from the digital video support (e.g. file, DVD). The user of the data is then required to possess the video to take full advantage of the analysis and comments.
3. At the package use level : visualisation of augmented movie (the annotations are used to display supplementary information on the video, to control the playing of the video, to navigate the video), visualisation of hypertext documents constructed from annotation and AV material, use of ad-hoc views (e.g. timeline view).

The Advene project is supported by the French **National Research Agency** for the **Cinelab project**.

### Présentation du projet

Le projet Advene (*Annotations de vidéo numérique (Digital Video) échangées sur le Net*) est mené depuis plusieurs années au **laboratoire LIRIS** (UMR 5205 CNRS) de **Université Claude Bernard Lyon 1**. L'objectif général du projet est de fournir un modèle et un format permettant de partager des annotations de documents audiovisuels numériques (films, cours, conférence, etc. ainsi que des outils permettant d'éditer et de visualiser/jouer des hypervidéos générées à partir des informations liées aux annotations et aux documents audiovisuels. Les enseignants, les cinéphiles peuvent utiliser les hypervidéos pour échanger des commentaires et des analyses multimédiatiques à propos des documents vidéos. Le projet a aussi pour objectif d'étudier la manière dont les communautés d'utilisateurs (étudiants, cinéphiles, étudiants...) s'emparent de ces outils d'auto-publication pour partager leurs "lectures audiovisuelles", ainsi que de participer à la définition de nouveaux modes d'édition et de visualisation interactive de contenus audiovisuels.

Comme décrit dans le schéma ci-dessus, le prototype Advene permet :

1. au niveau de la création de recueil (à gauche du schéma) : d'annoter des documents audiovisuels en associant des informations à des

Find: Fear Next Previous Highlight all Match case

Done Tor Disabled

# Video – Applications V

Welcome to Terminal Time!

http://www.terminaltime.com/

Welcome to the HCS ... UvA – Medewerkers ... Blackboard Learning ... Manuscript Central Google Documenten LEO Deutsch-Englisc... ACM Digital Library caisME Van Dale Taalweb | ... SWI-Prolog's Home

TERMINAL TIME Introduction Description Support Technical Screenings Products Contacts

**Project Informa**

History is in your hands! Through an audience response-measuring device (applause-meter) connected to a computer, viewing audiences respond to periodic questions reminiscent of marketing polls. These questions occur every 6 minutes during the story. The loudest applause determines the winning answer.

1. What is the most pressing issue facing the world today?

20 A. Men are becoming too feminine and women too masculine.  
231 B. People are forgetting their cultural heritage.  
C. Machines are becoming smarter than people.  
D. It's getting harder to earn a living and support a family.

C. Machines are becoming smarter than people.

Your answers to these questions allow the computer program to create historical narratives that mirror and even exaggerate your biases and desires. Just clap, watch and enjoy. At long last, Terminal Time gives you the history you deserve!

The Terminal Time engine uses the past 1,000 years of world history as "fuel" for creating these custom-made historical documentaries. Each program generated by the machine can be either projected on a screen or broadcast on television monitors. (Although the video and sound are constructed in the computer, the signal is compatible with standard video technology.) Each program lasts approximately 30 minutes.

Audiences experience the project in a similar manner to a movie. A short breakdown of how the audience experiences the work follows:

first questions posed to audience → 2 min. → Q1 → 6 min. → Q2 → 6 min. → Q3 → 6 min. → ending credits → Discussion with Audience

Intro: Similar to Masterpiece Theater Introduction  
Section 1: 1000 to 1750 A.D.  
Section 2: roughly 1750-1950  
Section 3: roughly 1950-present

1. Audience members find their seat and the "movie" begins with a short title sequence and introduction.
2. A series of three questions are asked -- each is shown on screen and read aloud by the digital voiced narrator. The audience answers the questions as a group by applauding. (The answers to these questions changes the story that will follow.)
3. The history lesson begins with a 6 minute section of the movie covering the years 1000-1750.

Done Tor Disabled

# Video – Application Via

Media Streams: Representing Video for Retrieval and Repurposing

http://acg.media.mit.edu/people/golan/mediastreams/

## Media Streams

Video annotation and editing system  
By Marc Davis, Brian Williams, and Golan Levin

*[Personal note: Media Streams was conceived by Marc Davis and developed between 1991 and 1997 at the Machine Understanding Group of the MIT Media Laboratory and at Interval Research Corporation by Marc Davis, Brian Williams, and Golan Levin [myself]. I was involved from the project's inception as the principal designer of its iconic visual language and as a contributing designer of its interface and interaction.]*



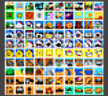
**Media Streams** is a system for annotating, retrieving, repurposing, and automatically assembling digital video. It uses a stream-based, semantic representation of video content with an iconic visual language interface of hierarchically structured, composable, and searchable primitives. Media Streams addresses problems of annotation convergence and human-computer communication by creating a standardized, computationally readable and writable visual language for representing consensual interpretations of video content.

The Media Streams video content annotation system uses a vocabulary of more than 6000 icons to represent the characters, objects, behaviors and settings of the broadcast universe. Because they are combined in a grammar with a syntax and semantics—permitting meaningful combinations numbering in the hundreds of millions—the Media Streams icons are not merely an iconography but a true visual language. This generative and searchable language supports gestalt information visualization, quick recognition and browsing of annotations, the potential for global use, and the representation of semantic, relational and temporal video content. Creating the Media Streams lexicon involved knowledge-engineering a sensible relational hierarchy of thousands of concepts, and inventing a consistent and readable set of recombinable sub-iconic graphic elements.

The Media Streams iconic annotation language has numerous advantages over traditional keyword annotation systems, including its ability to describe relations between descriptions, its ability to clearly render overlapping and contained actions, its ability to refer more directly to the intrinsic visual qualities of the video medium, and its ability to serve as a "consensus" language for multimedia professionals.

[More information on Media Streams can be found here.](#)

Detailed images and descriptions of the Media Streams user interface and visual language can be accessed from the links below:

-  [\[Link\]](#) The Media Streams *Media Timeline*, on which iconic annotations of video are temporally indexed. Each stream in the Media Timeline contains annotations about a unique aspect of video content, such as settings, characters, objects, actions, camera motions, etcetera.
-  [\[Link\]](#) The Media Streams *Icon Space*, an atemporal, hierarchically-indexed "dictionary" of iconic descriptors. The Icon Space incorporates utilities for icon construction and search.
-  [\[Link\]](#) A small but representative selection of icons from the Media Streams visual annotation lexicon. In the interest of space, the icons displayed here have not been arranged to reflect the grammars or knowledge hierarchies into which they are ordinarily structured.

A *Media Streams* bibliography:

Davis, Marc. "Media Streams: An Iconic Visual Language for Video Annotation." *Elektronikk* 4.93 (1993a): 59-71.

Davis, Marc. "Media Streams: An Iconic Visual Language for Video Annotation." In: *Proceedings of 1993 IEEE Symposium on Visual Languages in Bergen, Norway*, IEEE Computer Society Press, 196-202, 1993b.

Davis, Marc. "Knowledge Representation for Video." In: *Proceedings of Twelfth National Conference on Artificial Intelligence (AAAI-94) in Seattle, Washington*, AAAI Press, 120-127, 1994.

Davis, Marc. "Media Streams: An Iconic Visual Language for Video Representation." In: *Readings in Human-Computer Interaction: Toward the Year 2000*, ed. Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg, 854-866, 2nd ed.

# Video – Applications V1b



<http://acg.media.mit.edu/people/golan/mediastreams/>

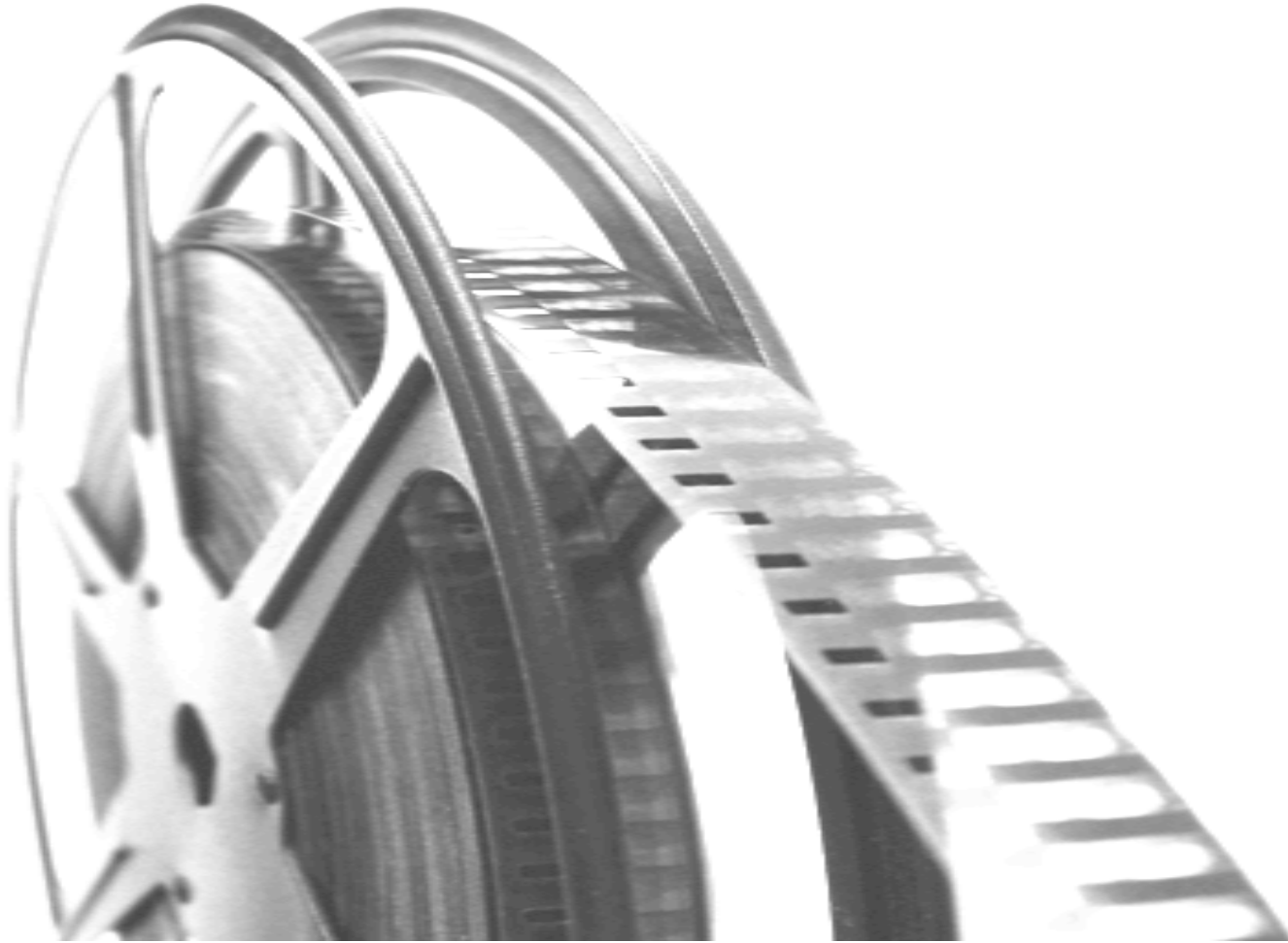
## Video – summary



- A video is a compositional unit with individualised semantics that may change if a shot is juxtaposed with another shot.
- A distinction between filmic (codify the relation to reality) and cinematic codes (codify narrative communication) must be made.
- Video provides its own realities of time and space which are interwoven in the narrative structure.
- A story is a representational system based on two main layers, structure and content, each serving two distinct purposes (form and substance) simultaneously.
- A primarily structure-oriented approach (grammar) to the temporal aspects of video with respect to generation and interaction is not appropriate.
- A planning approach (planner or agent) seems more promising.



## Video – References



# Video – References I

---

- Bernstein, M. (1998). Patterns of Hypertext. Proceedings of the 9th ACM conference on Hypertext and Hypermedia, pp. 21 - 29, June 20-24, Pittsburgh, USA.
- Black, J. B., & Wilensky, R. (1979). An evaluation of story grammars. *Cognitive Science*, 3, 213 - 230.
- Buckingham Shum S., Domingue J., & Motta E. (2000). Scholarly Discourse as Computable Structure, Technical Report KMI-TR-93. The Knowledge Media Institute, Open University, UK, April 25, 2000.
- Chatman, S.B. *Story and Discourse: Narrative Structure in Fiction and Film*. Cornell University Press, Ithaca, N.Y., 1978.
- Colby, B. N. (1973). A partial grammar of Eskimo folktales. *American Anthropologist*, 75, 645 – 62.
- Hardman, L., Obrenovic, Z., Nack, F., Kerherve, B., and Piersol, K. (2008) Canonical Processes of Semantically Annotated Media Production. *Special Issue on 'Canonical Process of Media Production', Multimedia Systems Journal, 14(6), pp. 327 – 340.*
- Minsky, M. L. (1988). *The Society of mind*. London: Picador.
- Liestøl, G. (1994). Aesthetic and Rhetorical Aspects of Linking Video in Hypermedia. Proceedings of the ACM European Conference Hypermedia Tchnology 94 (ECHT '94), pp. 217 – 223.
- Mancini, C. (2000). From Cinematographic to Hypertext Narrative. Proceedings of the 11th ACM conference on Hypertext and Hypermedia, pp. 236 -237, May 30 – June 4, San Antonio, Texas, USA.
- Mandler, J. M., & Johnson, N. S. (1977). Remembrance of Things Parsed: Story Structure and Recall. *Cognitive Psychology*, 9, 111 – 151.

## Video – References II

---

- Miles A. (2000). Hypertext in the Dark: cinematic narration with links / Hypertext Syntagmas: Cinematic Narration with Links. *Journal of Digital Information*, Vol. 1, No. 7, December 2000 <http://jodi.ecs.soton.ac.uk/Articles/v01/i07/Miles/>
- Moulthrop, S. (1991). Beyond the Electronic Book: A Critique of Hypertext Rhetoric. Proceedings of the 4th ACM conference on Hypertext, pp. 291 - 298, December 15-18, San Antonio, USA.
- Nack, F. & Putz, W. (2004) Saying What It means: Semi-automated (News) Media Annotation. *Multimedia Tools and Applications*, 22, pp. 263 – 302,
- Rumelhart, D. E. (1977). Understanding and summarizing brief stories. In D. Laberge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension* (pp. 265 - 303). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Sawhney, N., Balcom, D., and Smith, I. (1996). *HyperCafe: Narrative and Aesthetic Properties of Hypervideo*. Proceedings of the 7th ACM conference on Hypertext, pp. 1 - 10, March 16 – 20, Washington DC, USA.
- Schank, R. C. (1972). Conceptual Dependency: A theory of natural language understanding. *Cognitive Psychology*, 3, 552 – 631.
- Schank, R. C., & Abelson, R. (1977). *Scripts, Plans, Goals And Understanding*. Hillsdale, New Jersey: Lawrence Earlbaum Associates.
- Schank, R. C. (1982). *Dynamic memory*. New York: Cambridge University Press.
- Sowa, J. F. (1984). *Conceptual Structures – Information processing in Mind and Machine*. Addison-Wesley Publishing Company.
- Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9, 77 – 100.
- Walker, J. (1999). Piercing together and tearing apart: finding the story in afternoon. Proceedings of the 10th ACM conference on Hypertext and Hypermedia, pp. 111-117, February 21-25, Darmstadt, Germany.