KBM – Video Application

Frank Nack

Outline

- Organisation

- Last lecture

- AUTEUR – automatic generation of slapstick video sequences

- Additional applications

Video – summary

Investigated
- Temporal visual medium
- Effect of time on interaction and creativity
- AI techniques for the representation of time, action, plan

Example

Video, Frames, Scripts, Conceptional Dependencies, Plans, Goals,
Story Grammar, Semantic Links

Findings

- 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).
- 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 workable for automatic video generation.
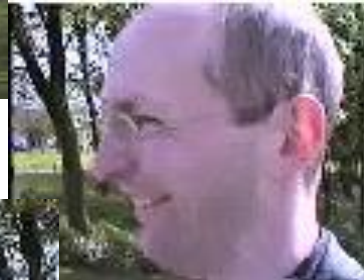
AUTEUR
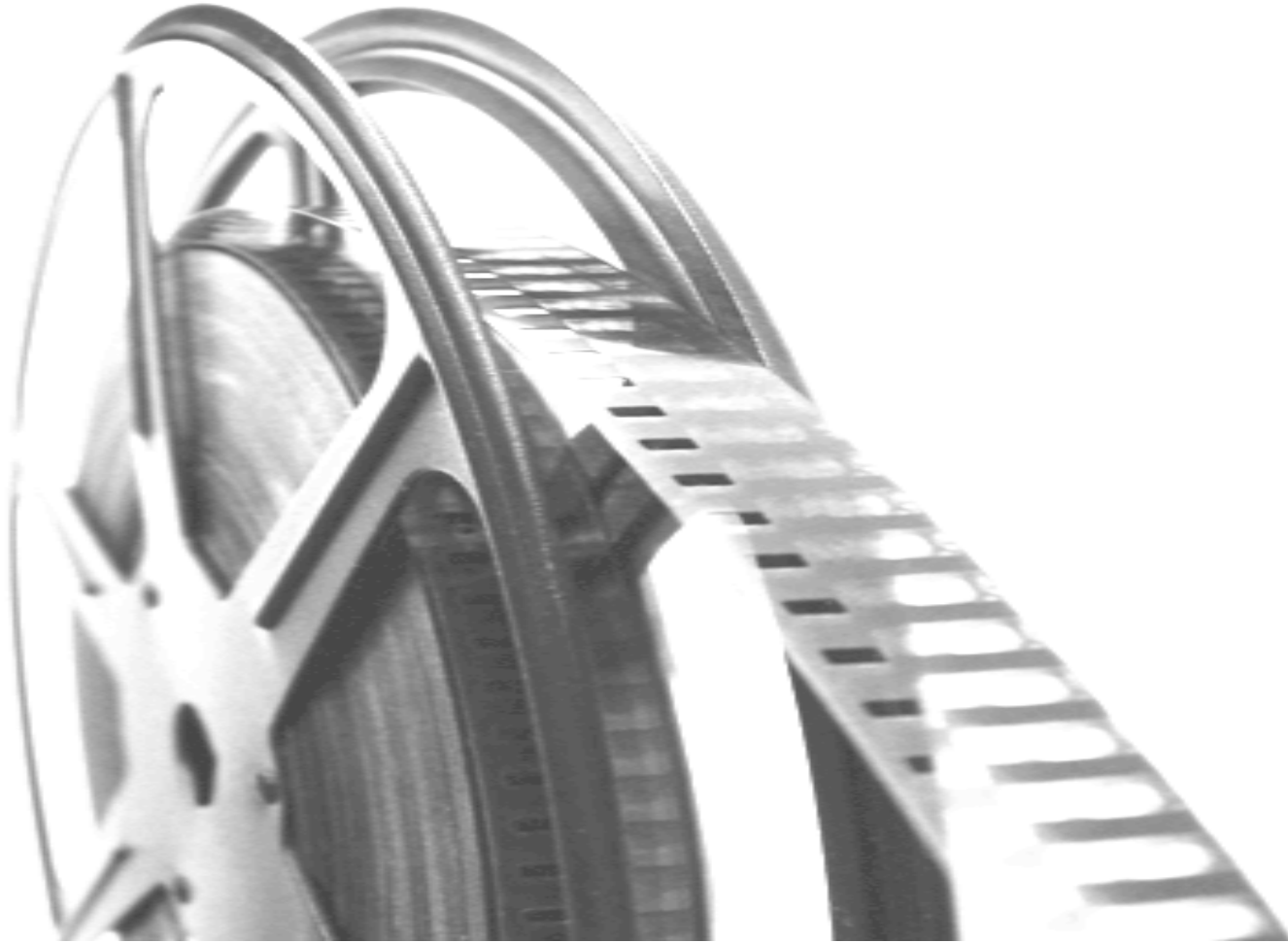
AUTEUR - Aim

- Automatic composition of visual slapstick sequences.

- Use an existing, arbitrary data base.

- Investigate machine creativity.
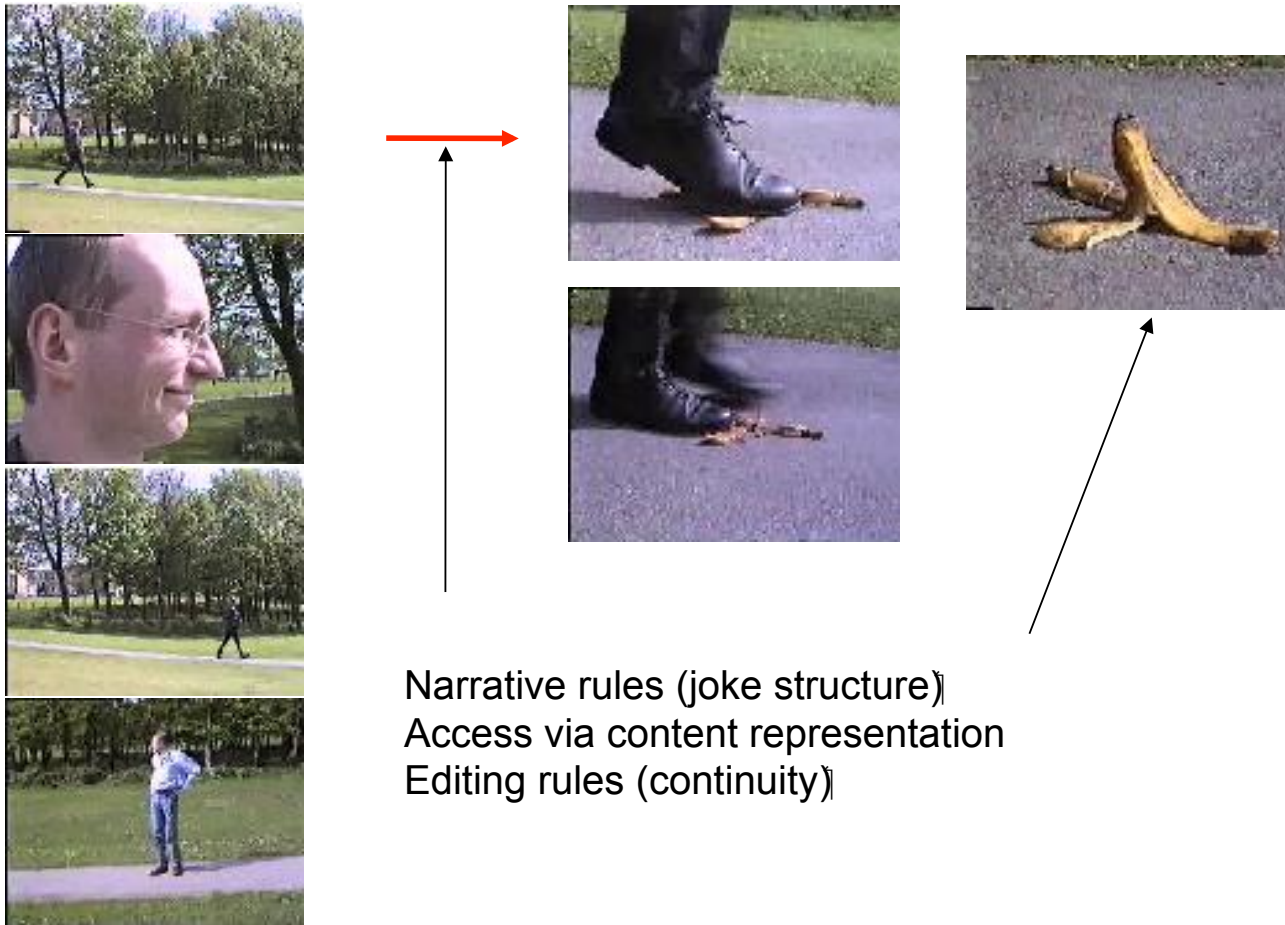
AUTEUR – Generation Example I

Motivation

Narrative rules (joke structure)
Access via content representation
Editing rules (continuity)

Narrative rules (continuity)
Access via content representation
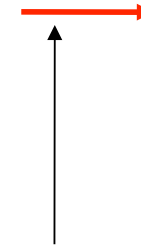Editing rules (continuity)

Frank Nack   Kennisgebaseerde media

AUTEUR – Generation Example II

Realisation

Narrative rules (joke structure)
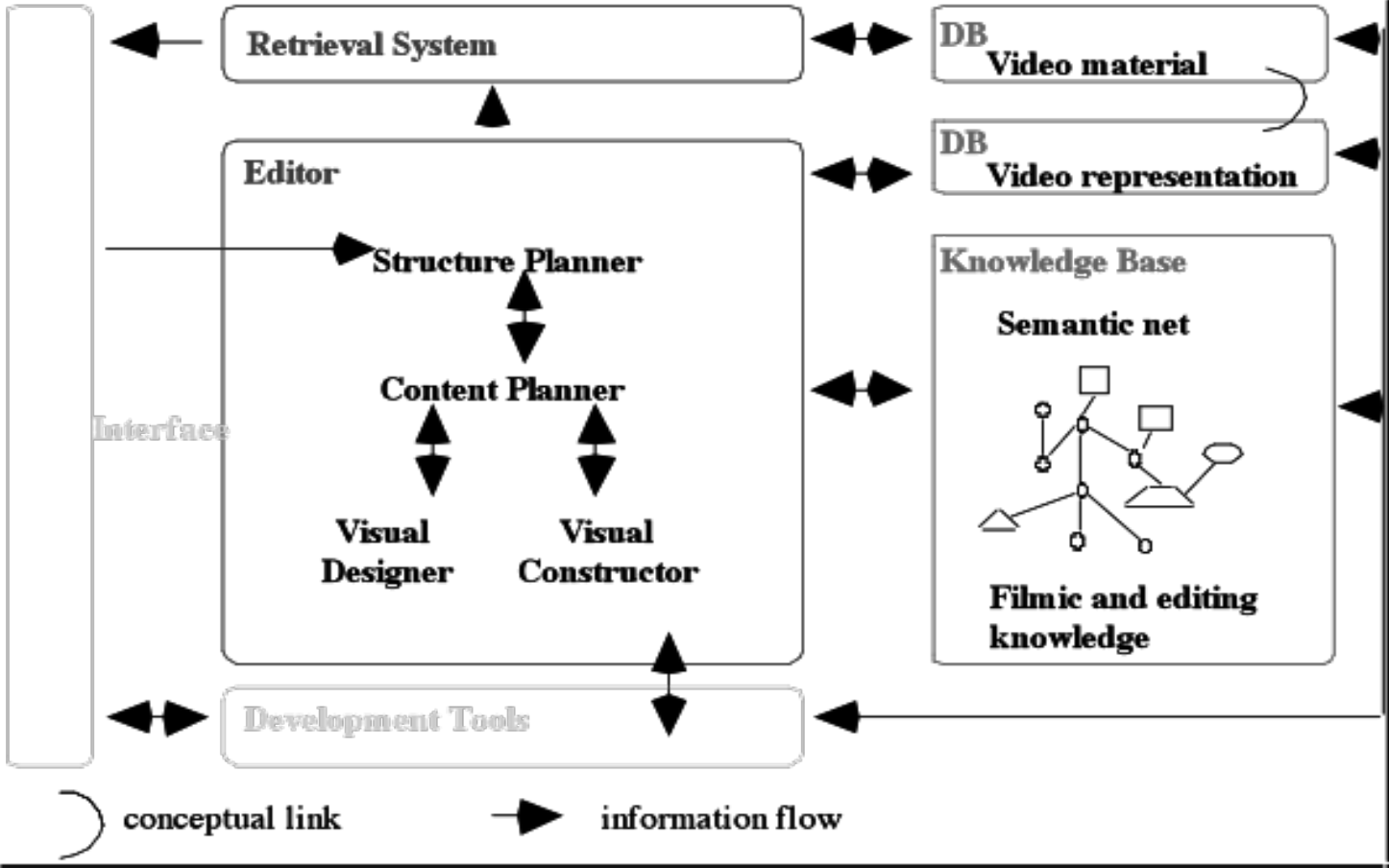Access via content representation
Editing rules (continuity)

Resolution

Narrative rules (joke structure)
Access via content representation
Editing rules (continuity)

AUTEUR – Generation Example IV

Result
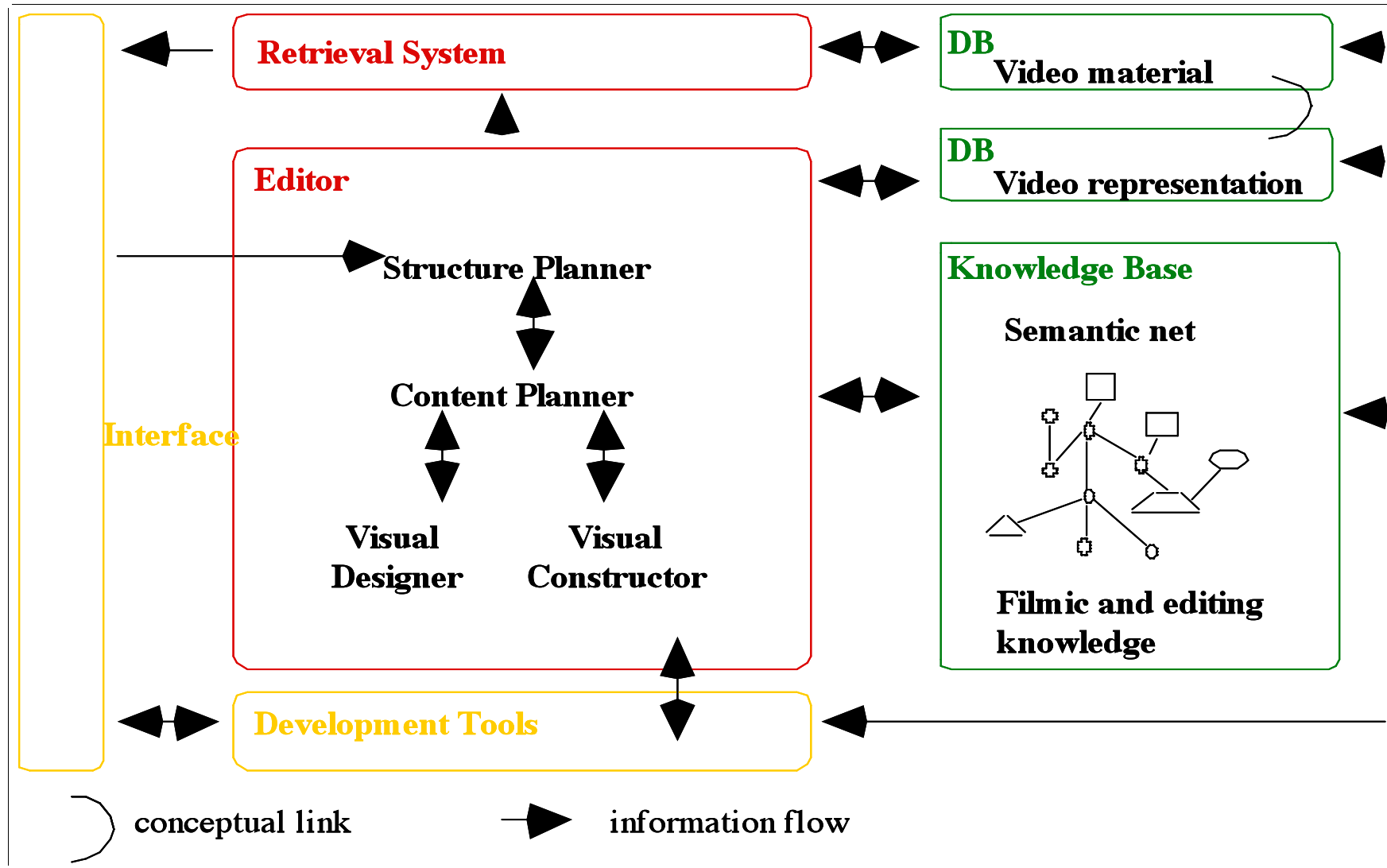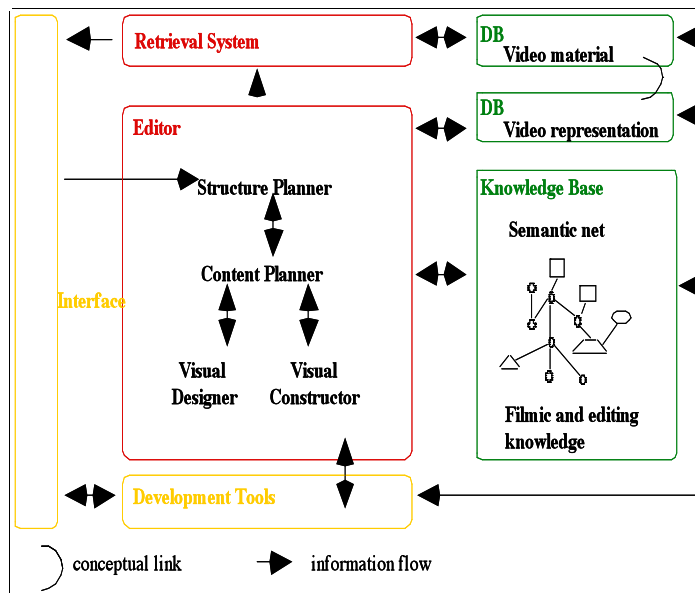
Frank Nack   Kennisgebaseerde media

AUTEUR - Architecture

Retrieval System

DB Video material

Editor

DB Video representation

Structure Planner

Knowledge Base

Semantic net

Content Planner

Interface

Visual Designer

Visual Constructor

Filmic and editing knowledge

Development Tools

⌒ conceptual link   → information flow

Frank Nack   Kennisgebaseerde media

AUTEUR - Architecture

Retrieval System

DB Video material

DB Video representation

Editor

Structure Planner

Content Planner

Visual Designer

Visual Constructor

Interface

Knowledge Base

Semantic net

Filmic and editing knowledge

Development Tools

conceptual link   information flow

AUTEUR - Architecture

- 145 Conceptual Dependencies

- 26 humour rules

- 37 editing rules

- 85 clips (5 sec – 9 min)

Frank Nack   Kennisgebaseerde media

AUTEUR – Structure Planner

conceptual link → information flow

AUTEUR – Plot Model I

Motivation	Realisation	Resolution
	Derision	
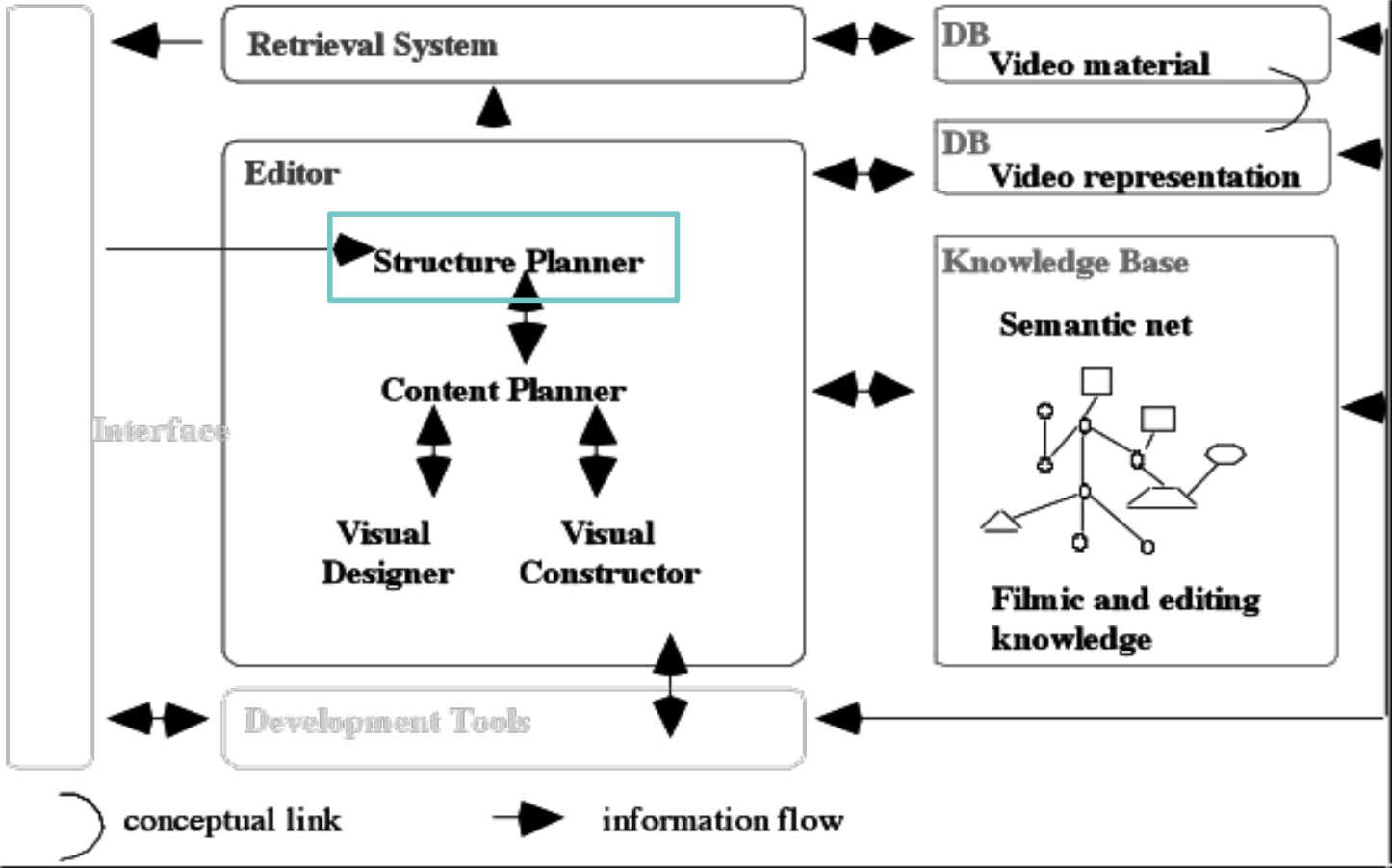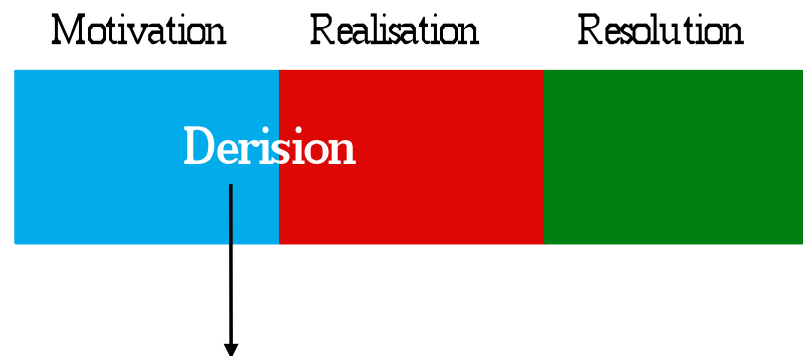
The explicit transformation of moods or states of the portrayed character.

The implicit upgrading of the mood of the viewer.

Motivation	Realisation	Resolution
	Incongruity	
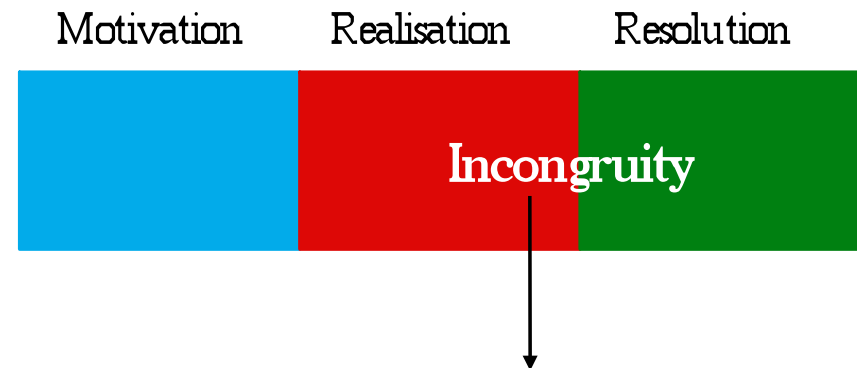
The conflict between the expected and what actually occurs.

AUTEUR – Plot Model II

AUTEUR – Humour Planner

conceptual link — information flow
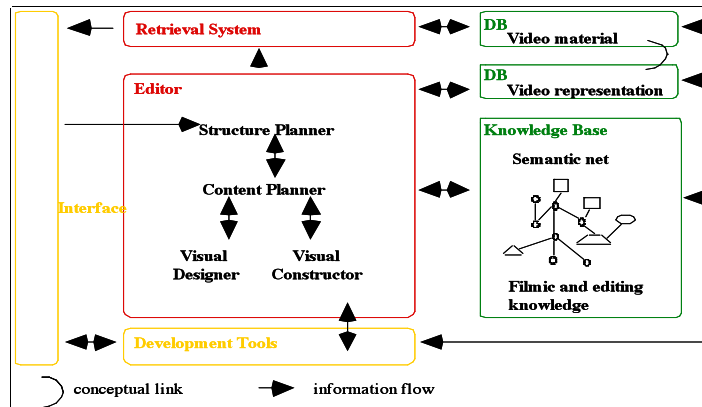
H-Strategy 1
An action forms the most suitable subject for a joke, then an actor, then an object, and finally a location.

H-Strategy 2
If the action portrays an intention [goal], interrupt the action in a way that is unexpected by the character, so that the goal cannot be fulfilled and the character's mood is downgraded or he or she suffers in some way.
(Mischief + Schadenfreude)

H-Strategy 16
A sequence of actions that is meaningful is more preferable for the construction of jokes than a sequence of unrelated actions.

H-Strategy 24
A relationship between two oppositional characters should be established in such a way that the goal of one character is to interrupt the goal of the other in such a way, that is unexpected by the second character. The reaction of the second character must then be influenced by the first so that the second character's mood is downgraded or he suffers in some way. (Mischief + Schadenfreude).

Frank Nack   Kennisgebaseerde media

AUTEUR – Content Planner
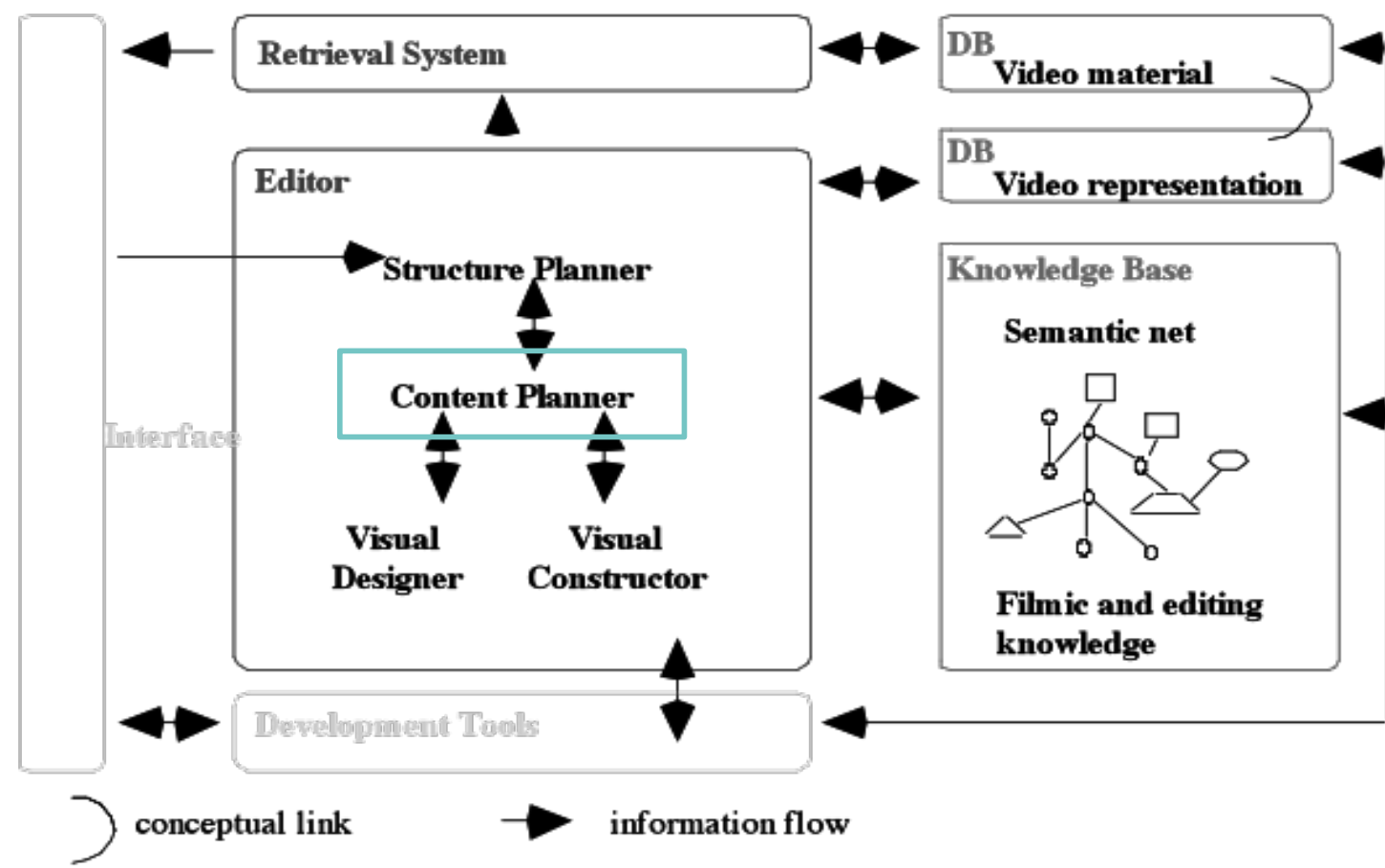
AUTEUR – Content Planner
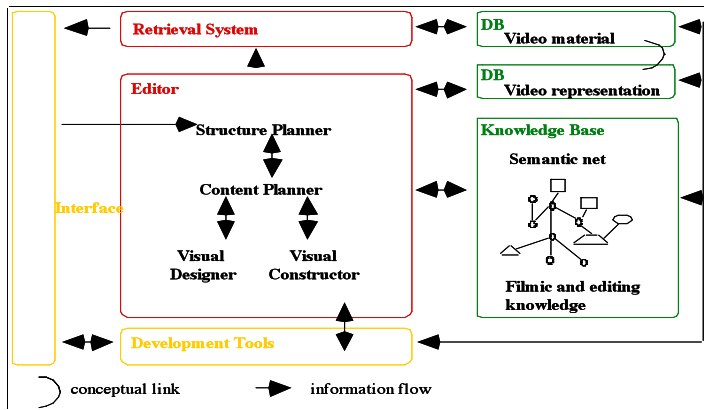
Motivation => Realisation => Resolution
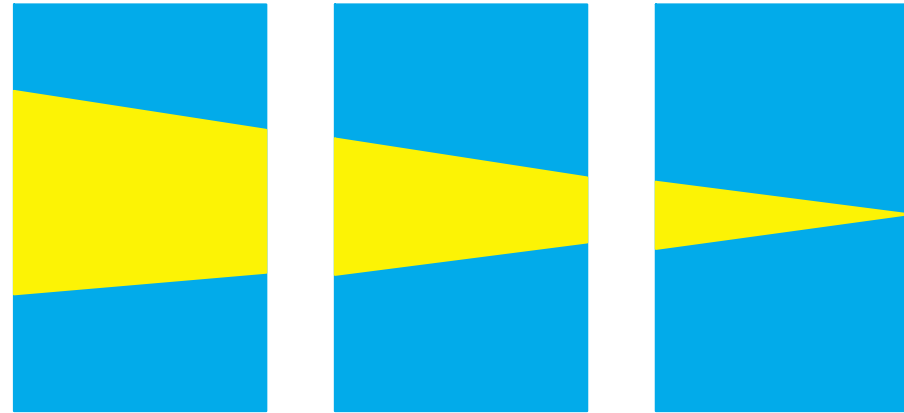
available video material
scene relevant material

Motivation	Realisation	Resolution
[approach]	[[search_money, insert_money+], [wait]]	[look_change, take_cup+, leave]
[]	[[process],[provide_cup+]	[]

Frank Nack   Kennisgebaseerde media

AUTEUR – Semantic Descriptions - Action

Frame

Name	slip
Domain	motion
Nature of location	outdoors
Set of objects	[banana_peel, dog_shit, soap, ice]
Body part / related object	[shoe]
Location	[road]
Relation Location -> Object	under
Relation Object -> Body part	under
Intention	[unintentional]
Result actions	[sit, lie, kneel, shake, look_back]
Result mood	[anger, rage, astonishment]

AUTEUR – Semantic Description – Event and Concept

Script (to meet)

Name	Actor number	Gender	Intention	Motivation	Realisation	Resolution	Episode
meeting	2	any any	meet	[walk] [wait]	[look at] [look at]	[shake_hand] [shake_hand]	date

Conceptual dependency

Emotional class name	Body part	Action
pleasure	Head	[lip, up]
	--	whistle

Abstract concept name	Representation structure [character/object, action]
time	[[shadows], [passing]]

Frank Nack   Kennisgebaseerde media

AUTEUR – Semantic Description – Semantic Graph
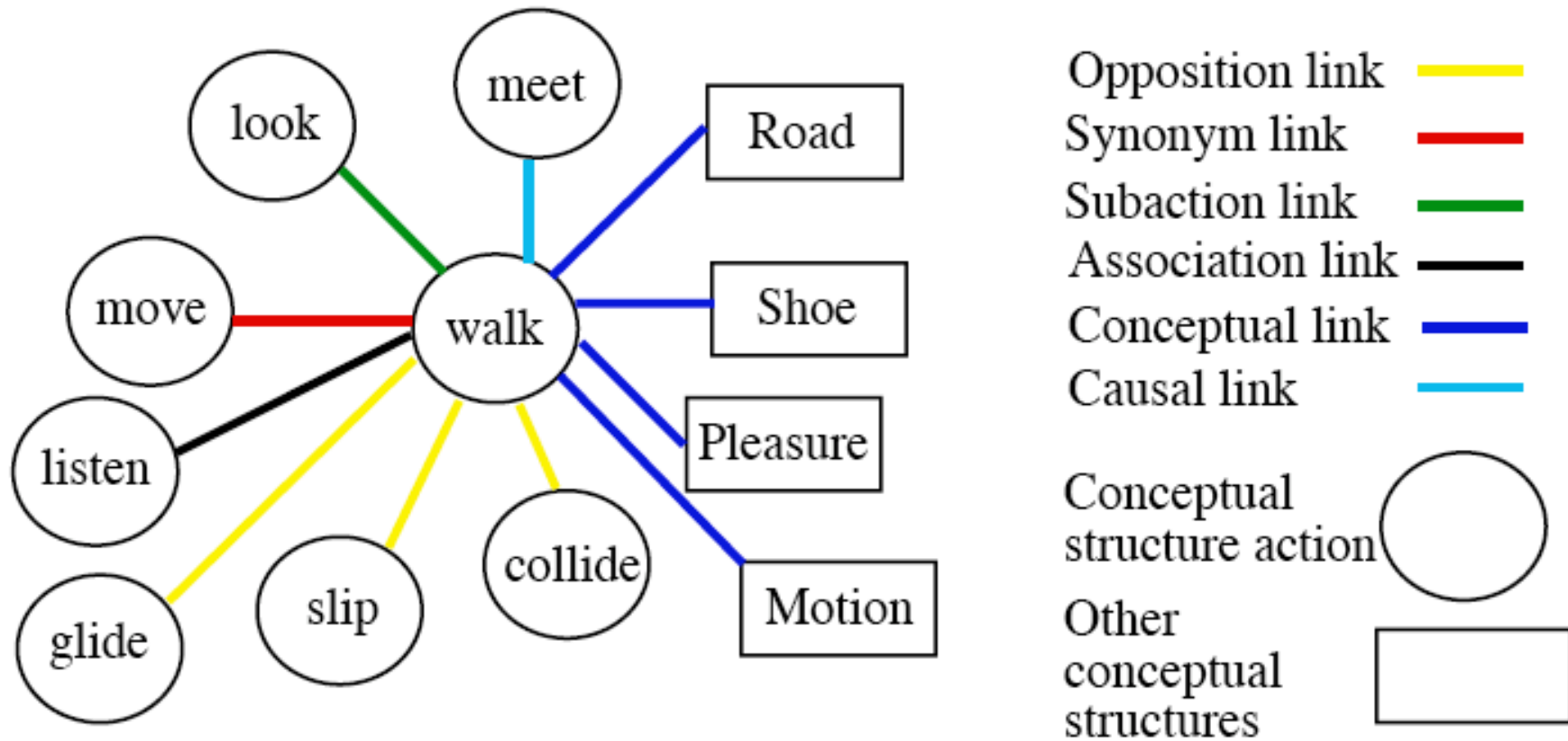
AUTEUR – Visual Designer
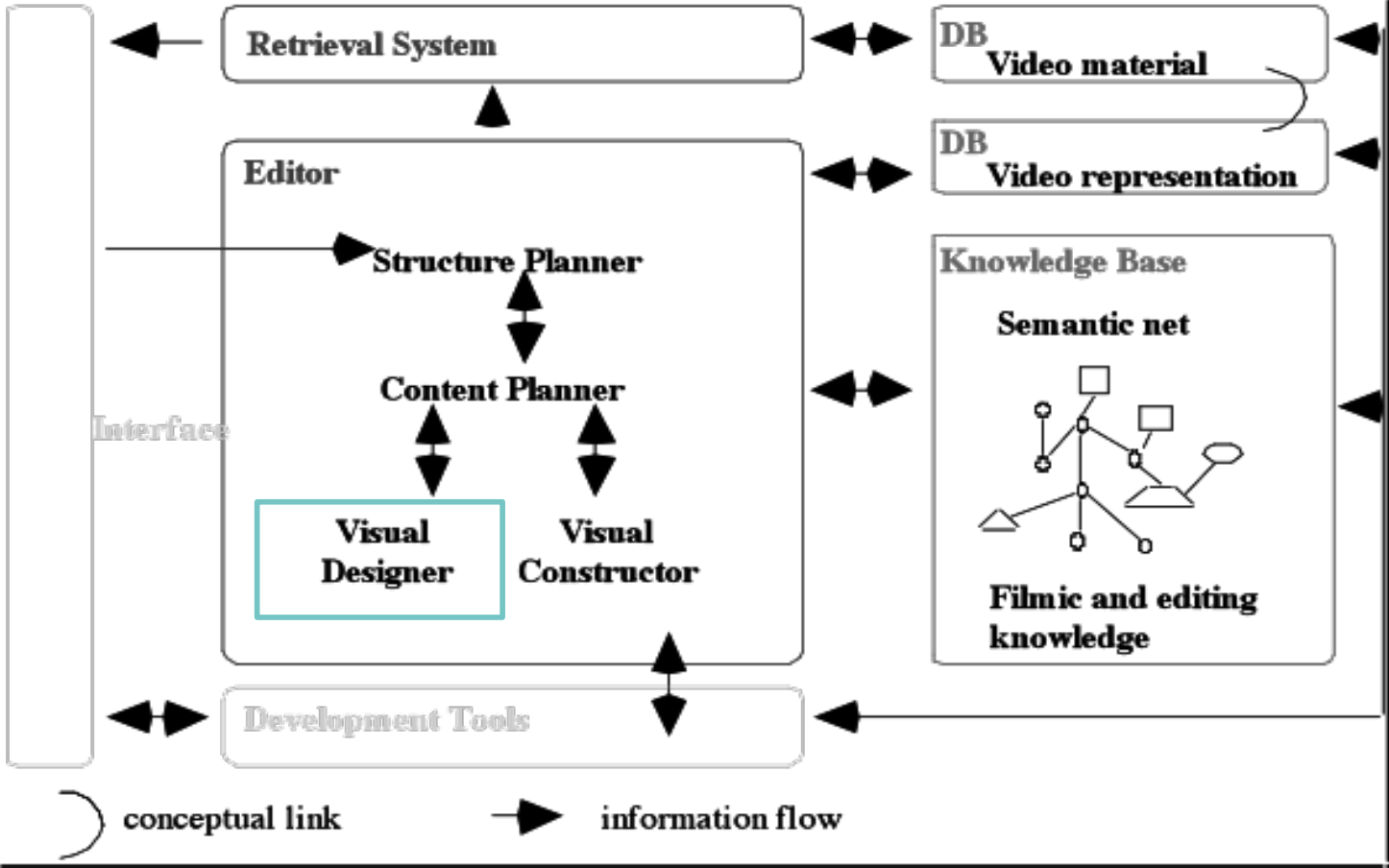
AUTEUR – Video Content Representation I

Description Scheme (Character)

Name	Description
Shot ID	
Startframe	
Endframe	
Identifier	Identifier for a character, e.g. a name or a number
Gender	male, female, hermaphrodite, artificial
Age	e.g. young, old, 25, etc.
Race	e.g. black, white, Asian, etc.
Appearance	a structure including: **role** e.g. lawyer, plumber, stewardess, etc. **Costume** ….**kind** e.g. business suit, apron dress, overall, etc. ….**colour** a doublet list providing the major colour for the top and bottom part e.g. [black, white] **appeal** e.g. casual, formal, etc.

Frank Nack   Kennisgebaseerde media

AUTEUR – Video Content Representation II

Description Scheme (Denotative Video Aspects)

Name	Description
Shot ID	
Startframe	
Endframe	
Identifier	Identifier for a character, e.g. name or number
Relative Position	(Screen position first frame, screen position last frame), e.g. (left, right), (left, middle), (right, right), etc.
Action	e.g. eat, drink, walk, read, etc.
Speed of action	e.g. slow, medium, fast
Direction of action	left, up-left, up, up-right, right, down-right, down, down-left, front, back, circular

Bodygesture	a structure containing:
	full body horizontal, vertical, left-diagonal, right-diagonal
	Head
	  profile right, left, half-left, half-right
	  movement up-down, left-right, up, down, left right, circle
	  eyebrows up, down, straight, etc.
	  line of sight left, right, straight, up, down, etc.
	  mouth up, down, straight, open
	Hand
	  left action/related object e.g.(tap/table)
	  right action/related object e.g.(holding/head)
	Foot
	  left e.g. tap, lift, etc.
	  right e.g. tap, lift, etc.

AUTEUR – Video Content Representation III

Description Scheme (Media Dependent Aspects)

Name	Description
Shot ID	Identifier
Shotlength	in frames (25 frames for a second)
Startframe	
Endframe	
Shot kind	a structure including: **lens movement** 　　**zoom-in**　[start camera dist., end camera dist.] 　　**zoom-out**　[start camera dist., end camera dist.] 　　**masking**　　left, middle, right **lens state**　(deep focus, foreground-focus, 　　　　　　　background-focus) **camera distance**　(extreme close-up, close-up, medium, 　　　　　　　medium long, long, extreme long) **camera movement**　(pan_left,pan_right, tilt_up, 　　　　　　　tilt_down,roll_left, roll_right) **camera position**　(left, midle, right) **camera angle**　(overhead, high-angle, eye-level, 　　　　　　　low-angle) **film speed**　(slow motion, normal, fast motion)
Shot colour	**colour**　　　　list of the dominant colours **black & white**
Shot granularity	fine, medium, strong
Shot contrast	high, medium, low

AUTEUR – Visual Designer

E-Strategy 1
If sequence.kind = Motivation
then

      Camera distance of Shot to be chosen is long
            => location.function = outdoor
medium long or medium
=> location.function = indoor

AUTEUR – Video spatial relations between frames

Shot B / Shot A	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) extreme close-up	X	X					
(2) close-up	X	X	X	X		O	
(3) medium close-up		X	X	X			
(4) medium		X	X	X	X	X	O
(5) medium long			X	X	X	X	
(6) long		O			X	X	X
(7) extreme long				O		X	X

Frank Nack   Kennisgebaseerde media

AUTEUR – Visual Designer
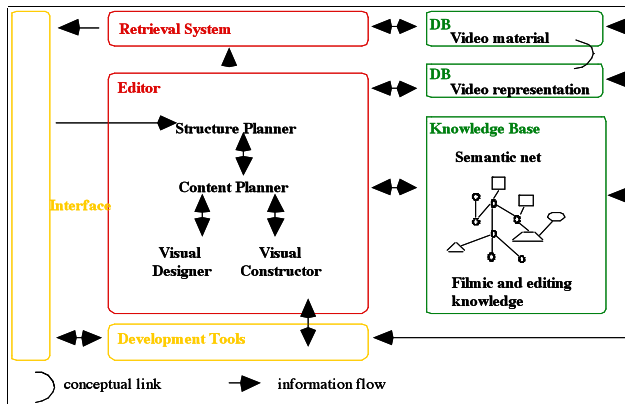
Value of camera device	Level of content detailness
extreme close-up	object:subparts[form,colour] character: subparts one detail of either head, hand, feet
close-up	object:Instance shape character: Instance or parts of either Head+Id, Hand or Foot
medium close-up medium	object:Instance shape character: Instance Appearance, Head
medium long long	object:Instance shape character: Instance Appearance setting: Time, Location, Lighting
extreme long	object: Instance shape setting: Time,Location, Lighting

Frank Nack   Kennisgebaseerde media

Legend:

Symbols: ☐ shot, | cut, ∘ subject

Type of spatial relations:

— line,

△ triangle or half circle,

☐ square or circle,

✕ over cross

Spatial relationships in the shot and created via a cut

based on the decomposition (1):
hierarchical knowledge
representation of subjects, the
decomposition starts
on parts level :

Class
|
Subclass
|
Instance
|
Parts
|
Subparts

decomposition (2 - n):
the content of each shot
should present the same
hierarchical level within
the knowledge
representation, e.g
Instance - Instance, or
Parts - Parts.
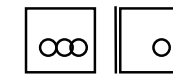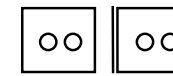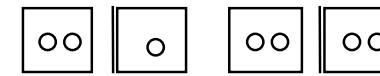
break up into
sub-groups
of the kind
described
to the left and
then use
their
configuration

1 subject   **2 subjects**   **3 subjects**   **4 subjects**   **n subject (where n > 4)**

number and size of
characters to be
portrayed and their
spatial relationships
involved.

Frank Nack   Kennisgebaseerde media

AUTEUR – Visual Designer
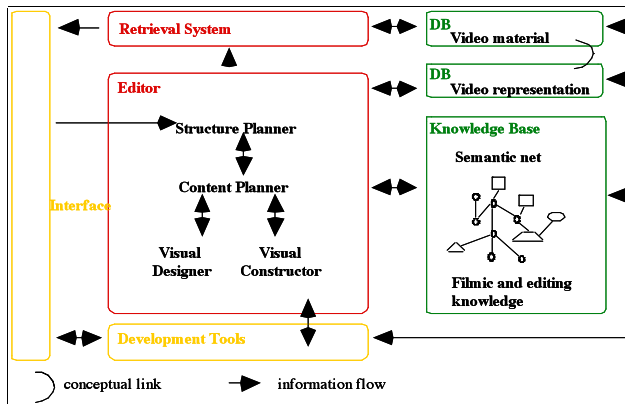
E-Strategy 6

If a sequence is to be established
> *where location of shot A ≠ location of shot B*
> *or the sequence is the first sequence to be*
> *established*

then

> create a memory structure of the spatial relations
> between all characters of Shot B

Location-Memory-Structure

Start	Shot-id
End	Shot-id

List of structures
> List_of_content_relations
> List_of_used_shots

E-Strategy 9

If number of character = 3 and
  camera distance of both shots ≥ medium long
> then
> shot A ([left | right]) with shot B ([middle]) => circle / triangle
> shot A ([left | middle]) with shot B ([right]) => circle / triangle
> shot A ([middle | right]) with shot B ([left]) => circle / triangle
> shot A ([left]) with shot B ([middle]) with shot C ([left]) =>
>> line or its permutations.

AUTEUR – Visual Designer - Action
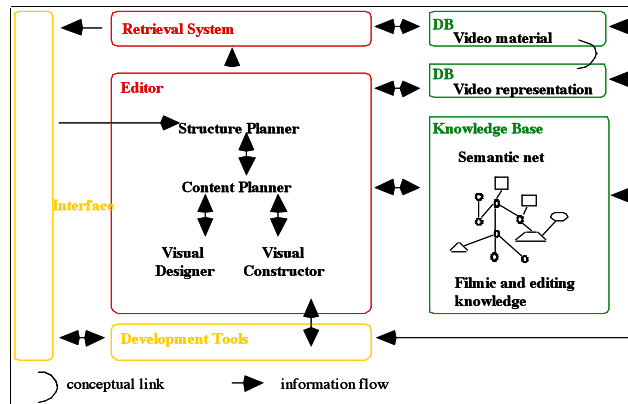
E-Strategy 12

If an action for a character is required and
there is no shot available to portray that action
then
isolate the character in a shot
retrieve the body part related to the action
retrieve a suitable shot where a body part
performs the required action
build a bridge into or out of this sequence if
necessary
continue with the sequence which was
interrupted by *this subsequence.*

AUTEUR – Visual Constructor

conceptual link   information flow

AUTEUR – Video shaping

Frames

TIME (H:Min:Sec:Frames)

Stratamodel

0

140

eat

walk

sit

talk

Shot

Annotations

Frank Nack   Kennisgebaseerde media
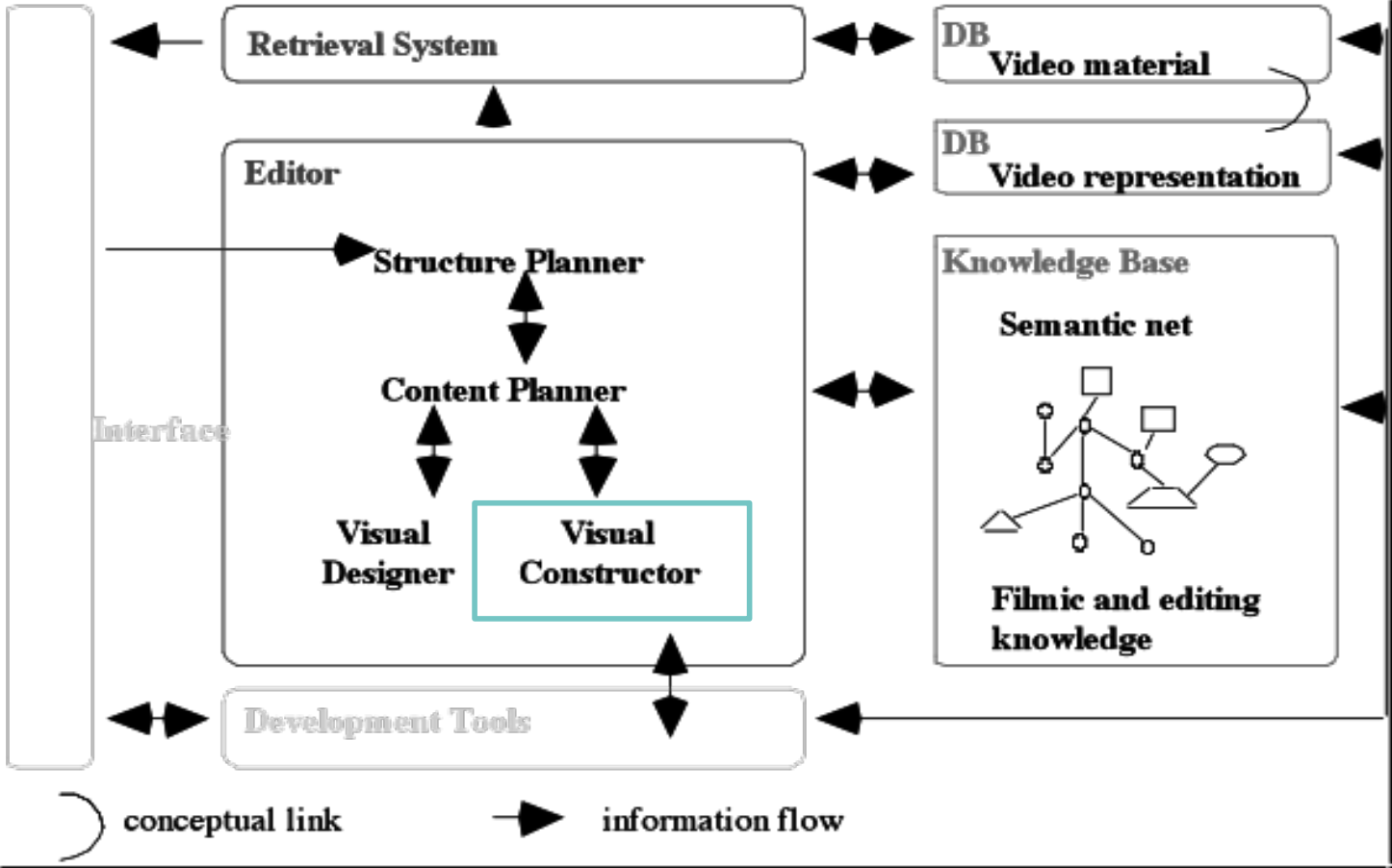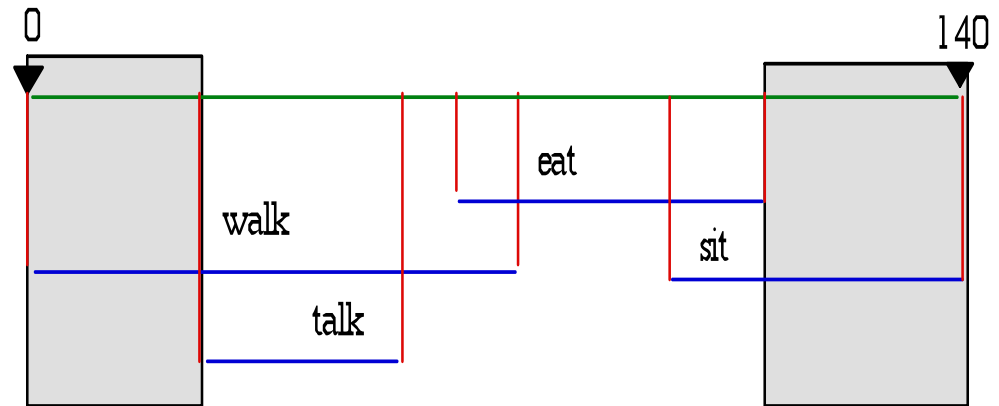
AUTEUR – Visual Constructor
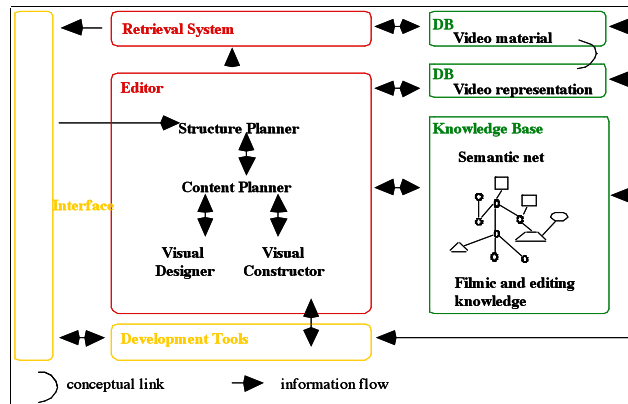
E-Strategy 29
If sequence.action.tempform = contraction and
   action is a single action
then
   favour decomposed forms of presentation
   where the camera distance of shot A ≥ camera
distance shot B

E-Strategy 31
If camera distance of a shot ≤ close-up and
then
         clip it to a length ≤ 60 Frames.

E-Strategy 32
If close-up < camera distance of a shot < long and
   sequence. kind = motivation
then
   clip it to a length ≤ 108 Frames.

AUTEUR – Summary

- Planner approach (backward chaining – goal driven)

- Addresses continuity problem BUT incompletely

- Closed system

- Requires lots of knowledge

- Editing strategies are explicit BUT not final

- Story generation is flexible BUT presentation should be generated rather than using existing material

Video – Additional Applications

Video – Additional Application

Video – Additional Application II

Video – Additional Application III

Video – Additional Application VI

Summary

- **The content determines the application**
- **Content description**
  - **Application dependent**
  - **Complex**
  - **Recourse demanding**
  - **Time critical**
  - **Incomplete**
- **Modular Schemata**
- **Description environment**
  - **Production supportive**
  - **Archive supportive**