Part II

Ontology for Multimedia Applications

Hiranmay Ghosh
TCS Innovation Labs, Delhi

Part II

Agenda

Part I

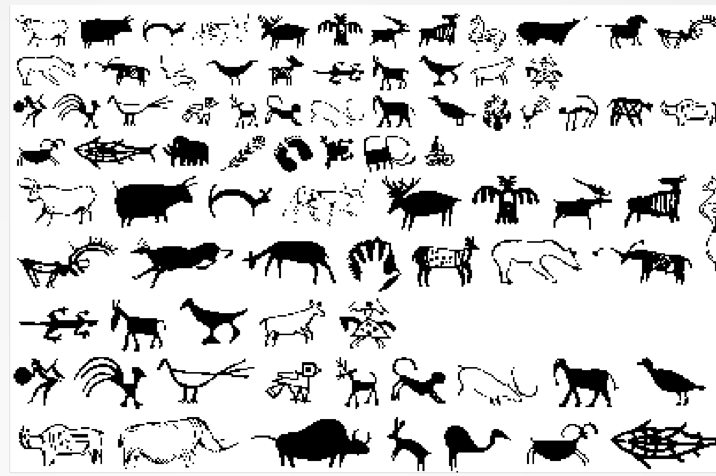
- **Introduction**
- Semantic Web and Ontology
- Semantic Multimedia Content Processing
- Ontology for Multimedia Data Interpretation

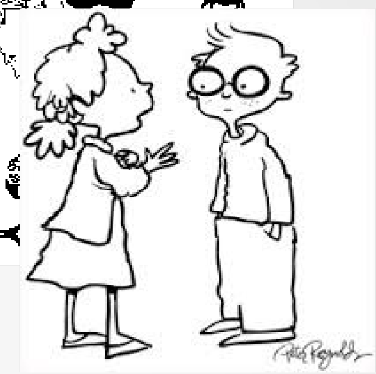
Part II

- **Multimedia Web Ontology Language**
- Application Examples
- Distributed Multimedia Applications
- Conclusion

Media for expression and communication

Cave Paintings





Towards symbolic representation: Icons and Speech

Hunting is the practice of pursuing any living organism, usually
trade. Animals may also hunt other animal species, but this is u
distinguished from poaching, which is the killing, trapping or cap
species that are hunted are referred to as game and are usually

Hunting can also involve the elimination of vermin, as a means
overpopulation. Hunting advocates state that hunting can be a
example, to help maintain a population of healthy animals withi
checks such as predators are absent.[2] In the United States, wil
licensing bodies, where they help to set rules on the number, m

Text: A symbolic form of expression

Patna railway police seems to have committed gr
filed after the first of the serial blasts in Patna. Go
Police (Patna) station house officer stated in the F
"could not produce valid documents for keeping
it".

To multimedia: symbolic and perceptual

Text *vs.* Media

- Text is a conceptual description
- Human created
  - *Reflects author's viewpoint*
  - *Abstract: Leaves out “unnecessary” details

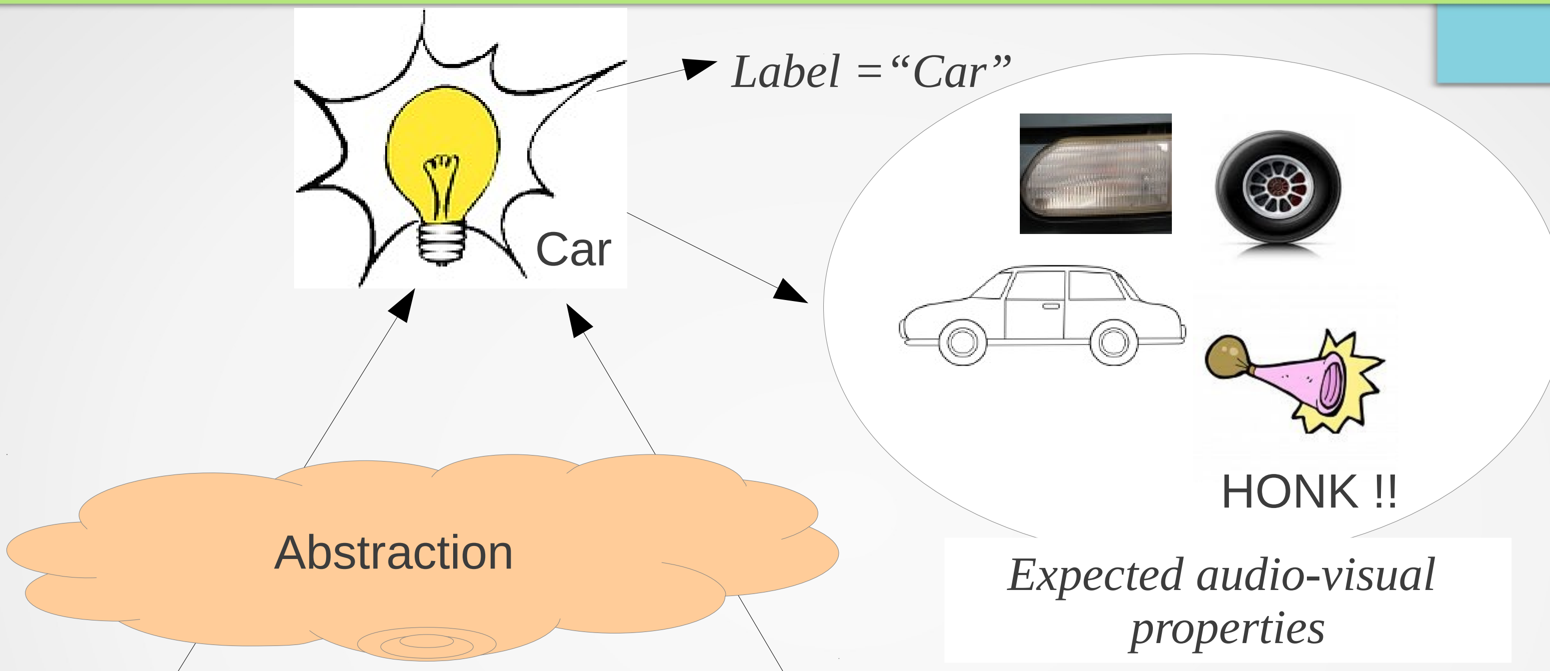
- Media provides a perceptual description
- Machine recorded
  - *Matter of fact recording of an event*
  - *Contains all details, e.g. “background”*
  - *But there is a person behind the machine telling a story*

Shortcoming of existing approaches

- Ontology is a conceptual model of a domain
  - *It is an abstraction over all forms of expression*
  - *But, it needs a body for expression and communication*
  - *Expressed with textual artefacts (closest to the abstract world)*
- Existing methods use conceptual domain ontology *on top of* intelligent media processing
  - *Media interpretation does not benefit from domain knowledge*
- Specific issues of modeling of multimedia knowledge have not been addressed
  - *Uses crisp reasoning rules – inflexible for media data interpretation*
  - *Cannot reason with media properties of concepts*

Are concepts really separated from the media world ?

Perceptual model of a concept

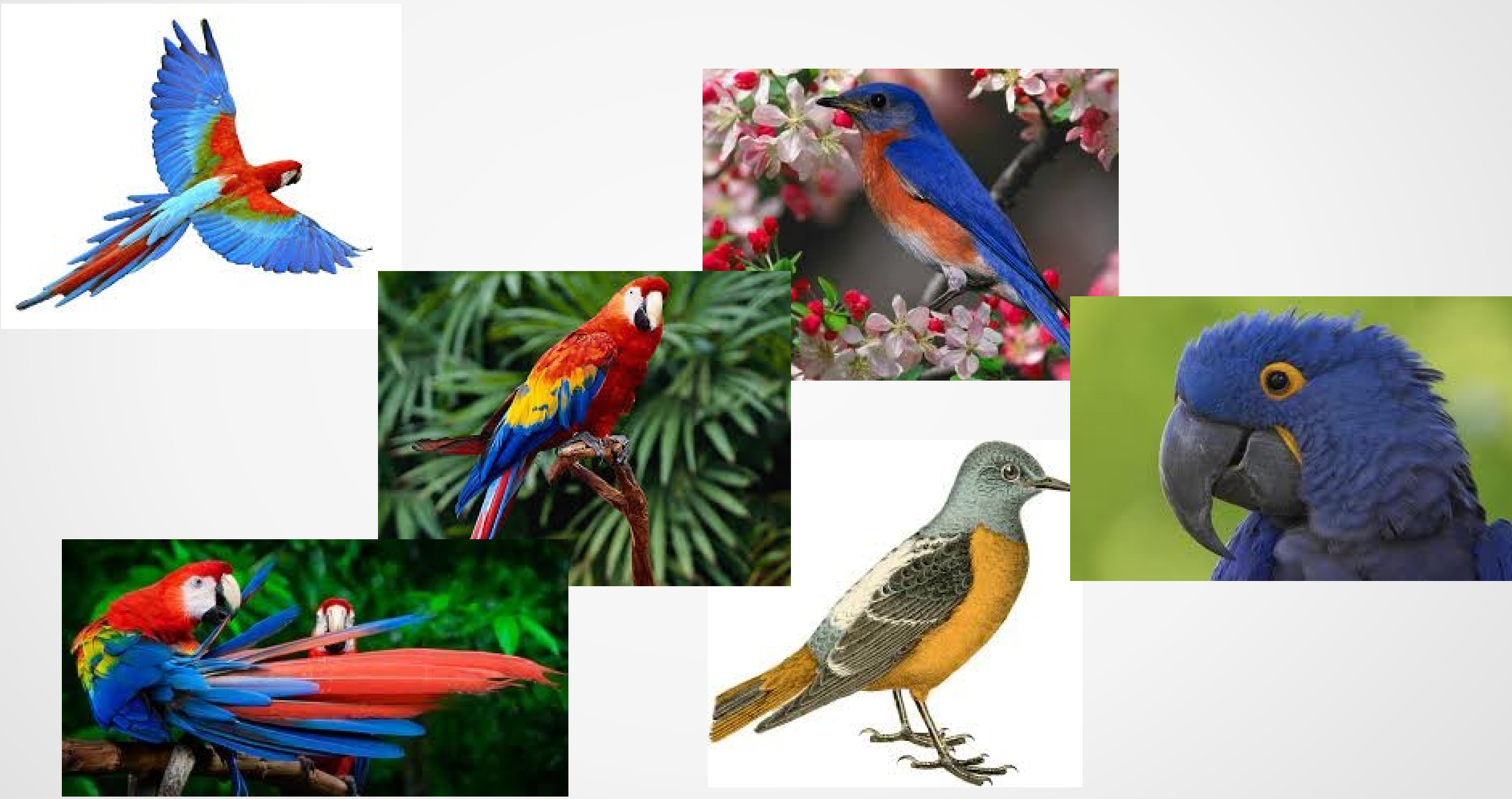


An alternate representation for concepts?

A causal model of the world

- Concepts *causes* some media forms to appear in a multimedia instance
  - Media manifestations are *evidences* for those concepts
- There are uncertainties in media manifestations
  - Variations in realizations of concepts
    - *Intraclass variations*
  - Due to changes in viewpoint, occlusion, environmental noise, etc.

Illustrations



Is it possible to abstract ?

What is common to all (most) birds

- *Body shape*
- *Chirping*
- *Ability to fly*

With lots of variations and uncertainty

But, if we observe any of them (in a given context), it is an evidence towards sighting a bird

Formalizing the model

$$\text{If } C_X \rightarrow M$$

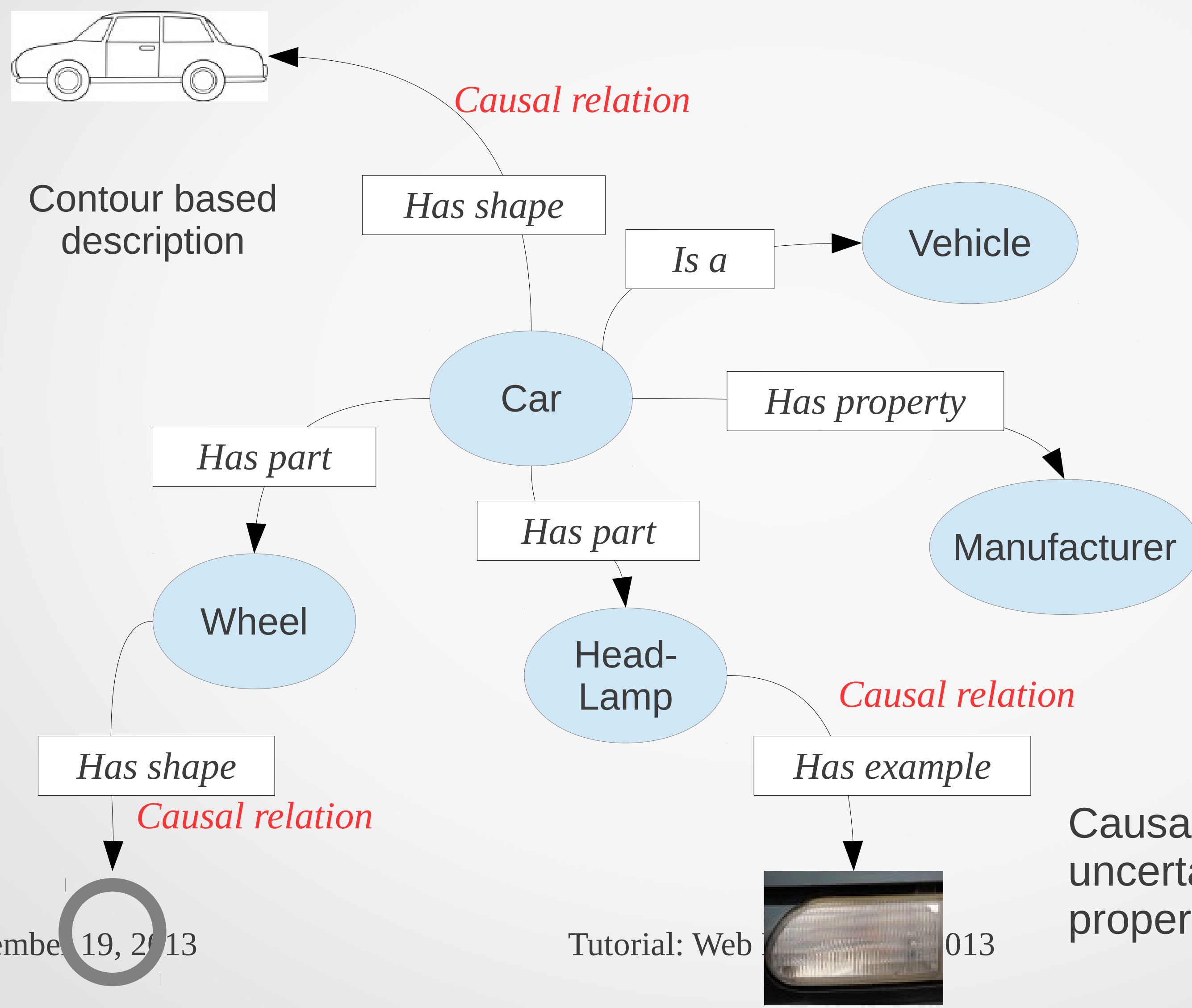
$$\text{Then } m \in M_X \Rightarrow C$$

C: Concept
M: Media patterns
X: Domain Context

Abductive reasoning

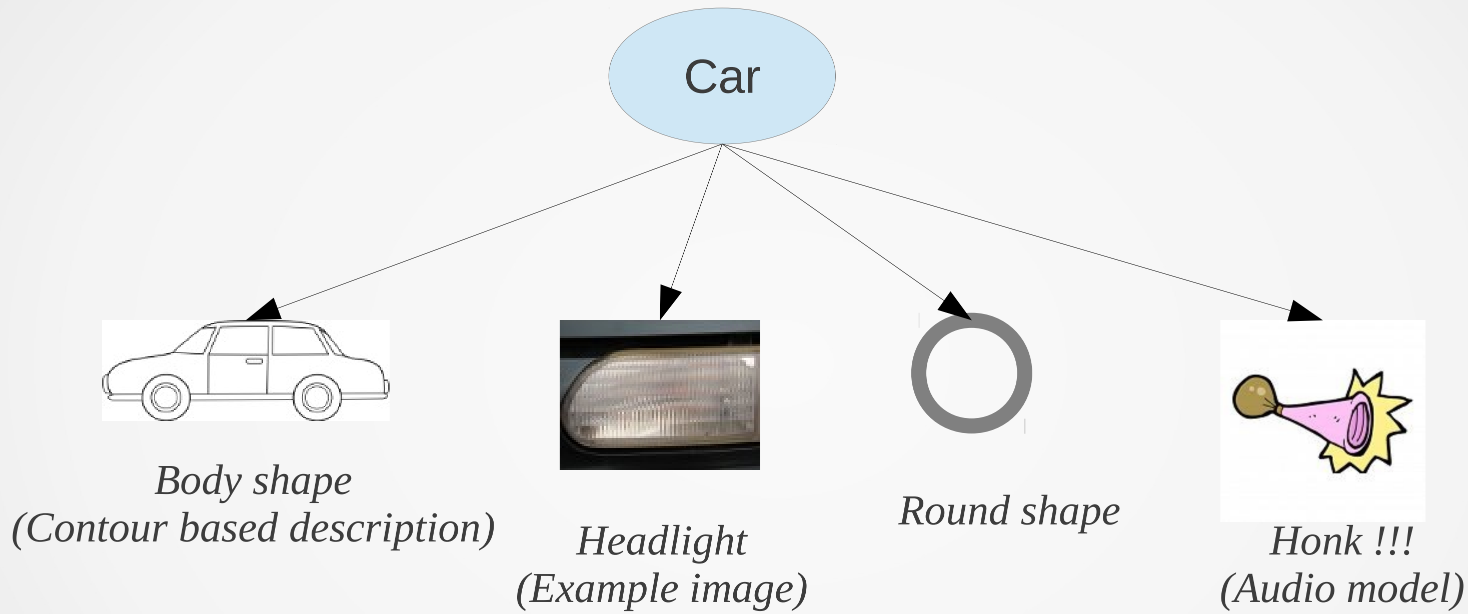
- Reasoning with evidence
- Weak but robust
- Inferencing by best explanation
- Closed world assumption

Ontology for perceptual modeling



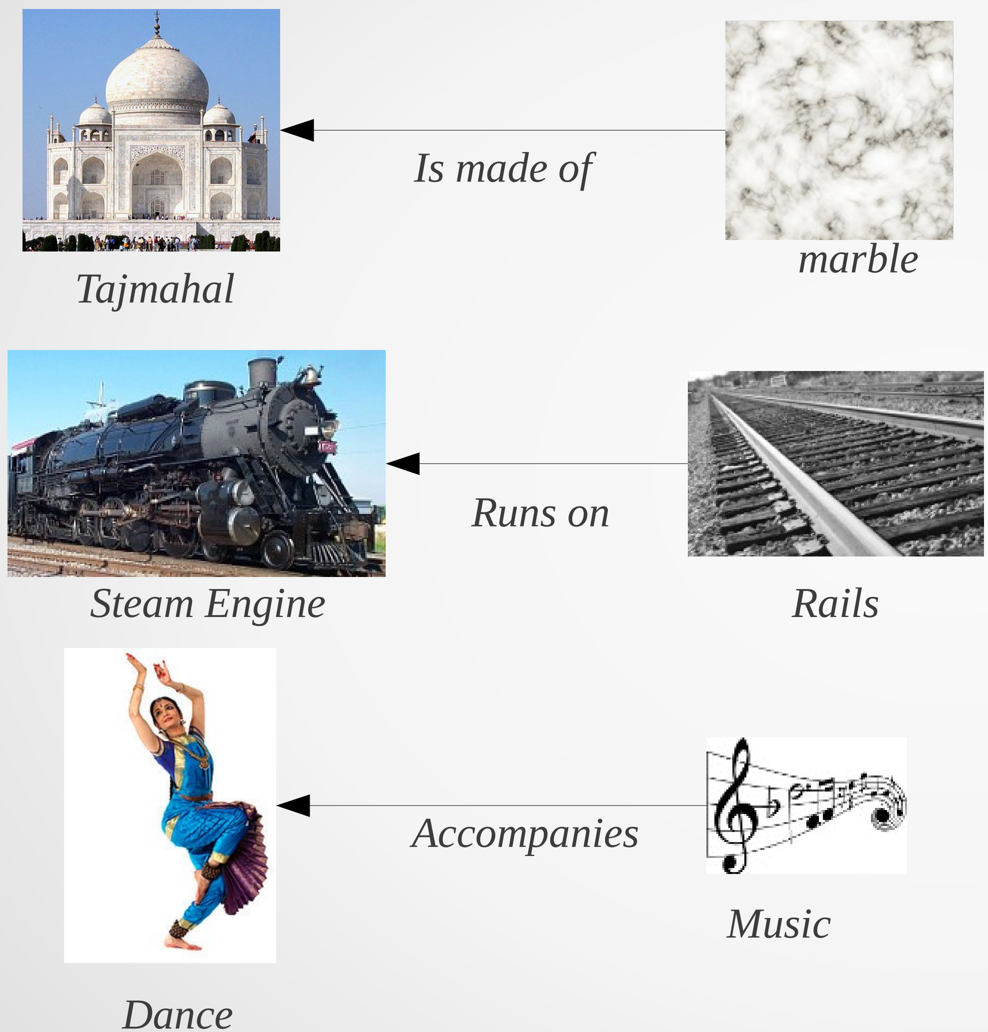
Causal relations with uncertainties bind media properties with concepts

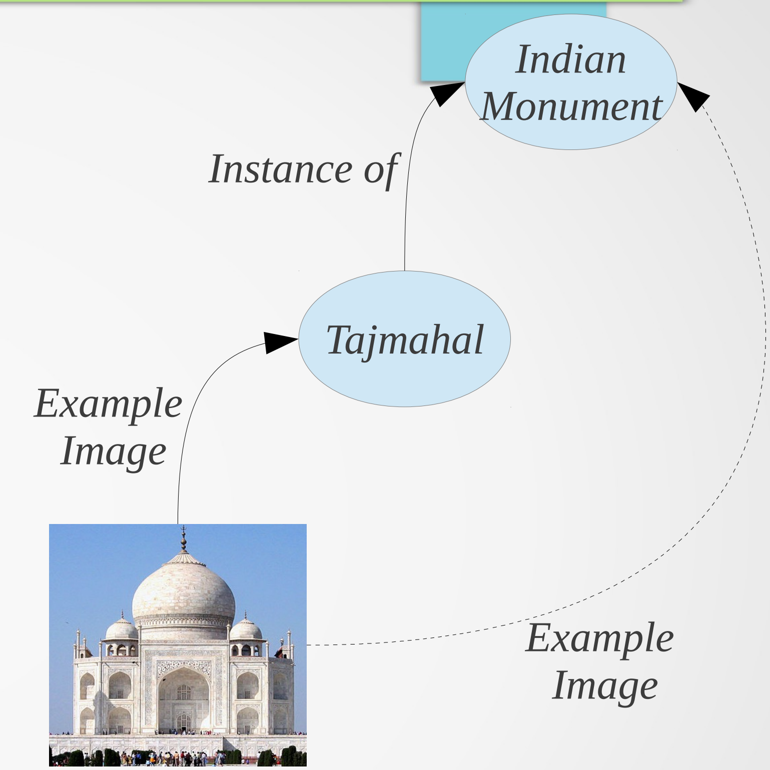
Observation Model for concept recognition



- *The observable media properties are evidences for the concept*
- *Concept can be recognized using Bayesian reasoning*

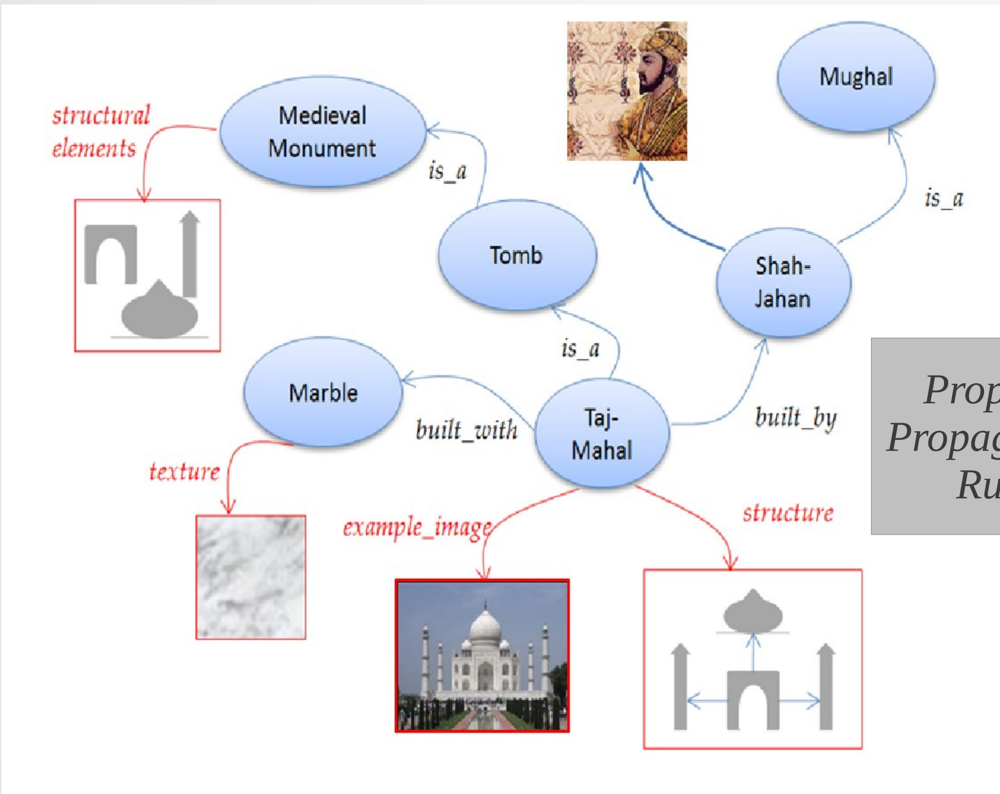
Media property propagation





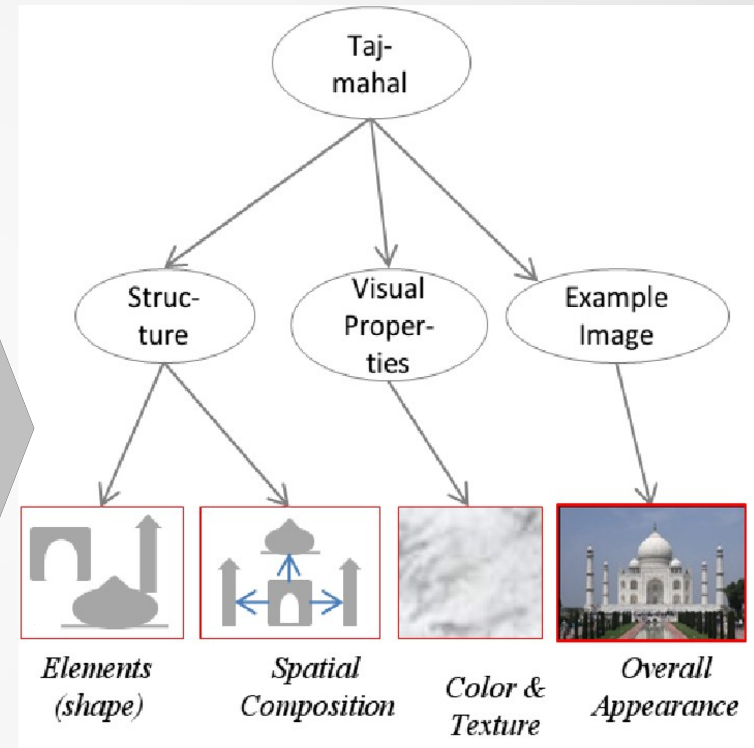
- *Media properties flow across domain-specific relations*
- *Super-classes "inherit" media examples*

Observation Model



Ontology

Property Propagatopn Rules



Observation Model

Incorporates domain-specific contextual Information

Multimedia Web Ontology Language

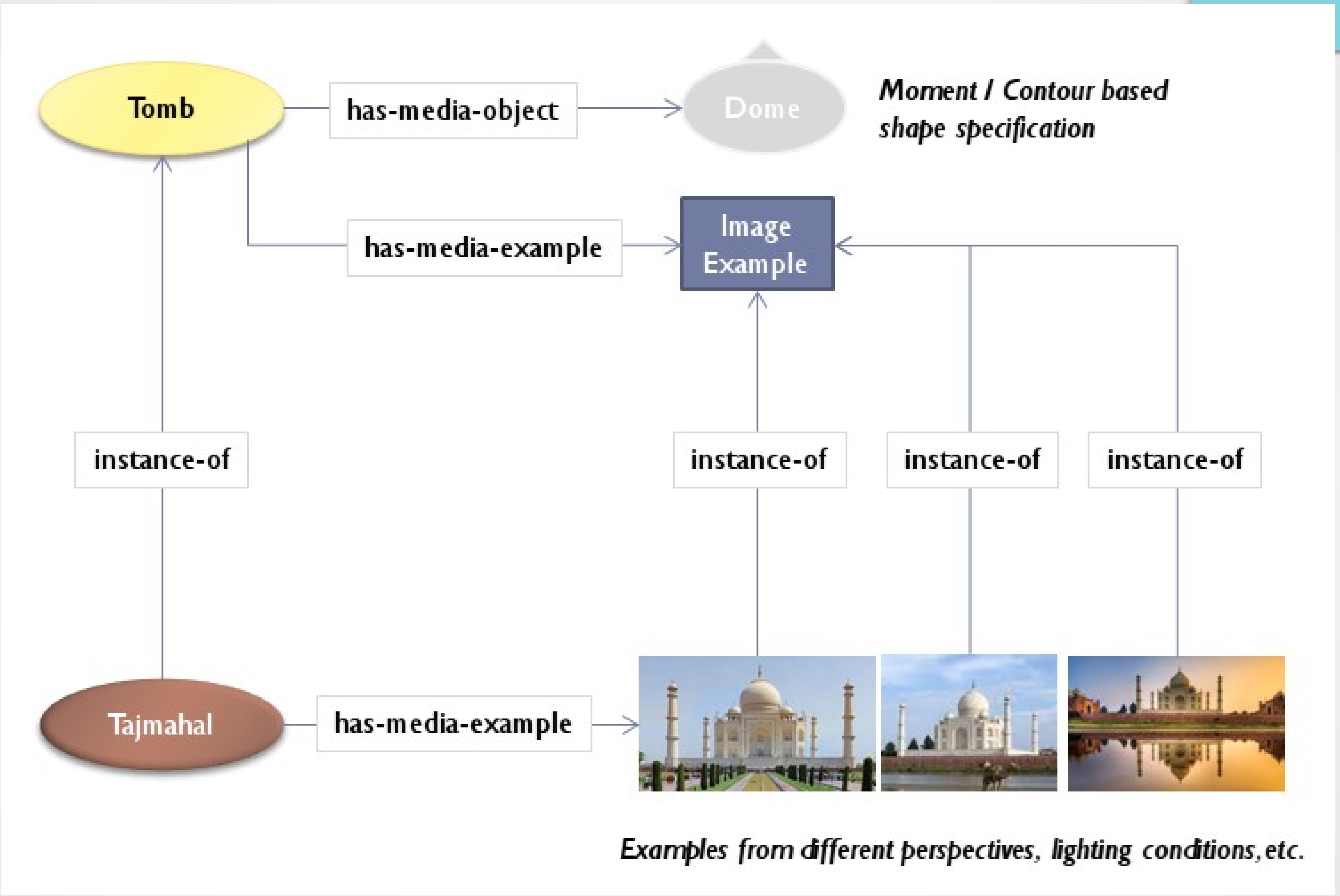
MOWL

- Supports perceptual modeling
- Causal model of the domain
  - Concepts manifests in media patterns
- Different ways of media property specifications
  - Formal specification of spatio-temporal constructs
- Reasoning with media properies
- Abductive reasoning for concept recognition
  - Probabilistic framework with Bayesian network (OM)
  - Dynamically derived from the ontology
- *Formalism with flexibility*

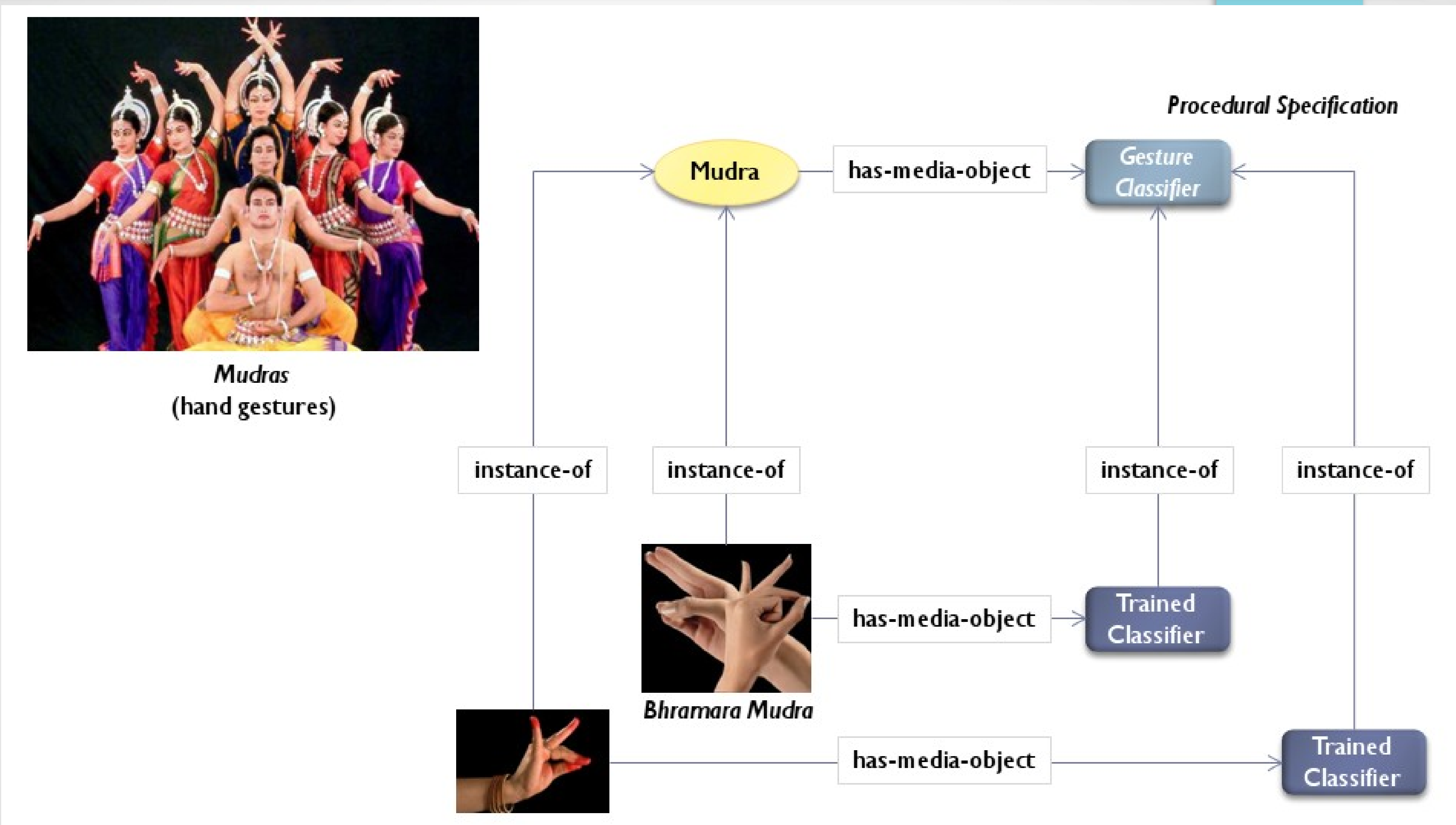
Media properties

- ***Declarative***
  - At different level of complexity, e.g. dominant colors, shape contour, pitch, timbre, MPEG-7 descriptors, …
- ***By Example***
  - Several illustrative media instances
- ***Procedural***
  - Specification of an algorithm / Pointer to a piece of code that can recognize a concept, e.g. a dance posture
- ***Structural Composition***
  - Spatial or temporal combinations
  - Hierarchical structure

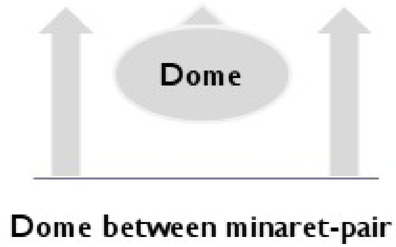
Declarative specs & Media Examples



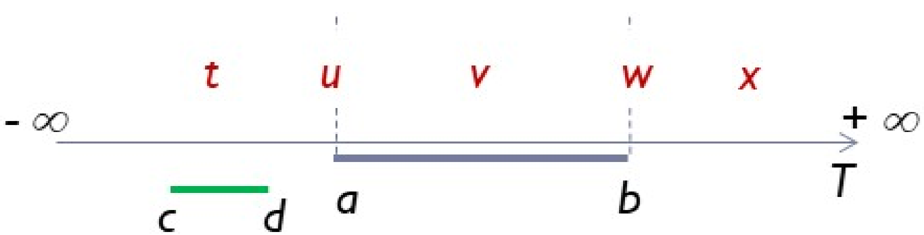
Procedural specifications



Structural specification of media properties



- Define the relations in a formal way
- Allow for uncertainty



- **Interval algebra**

<cd> ***before*** <ab> iff

- <cd> ∩ <t> ≠ ∅
- <cd> ∩ <u> = ∅
- <cd> ∩ <v> = ∅
- <cd> ∩ <w> = ∅
- <cd> ∩ <x> = ∅

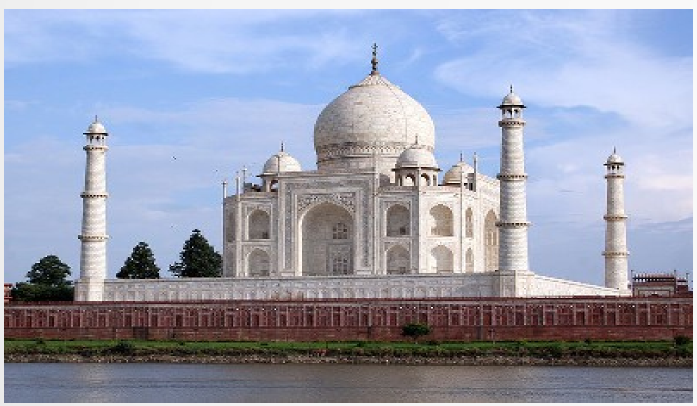
R_10000

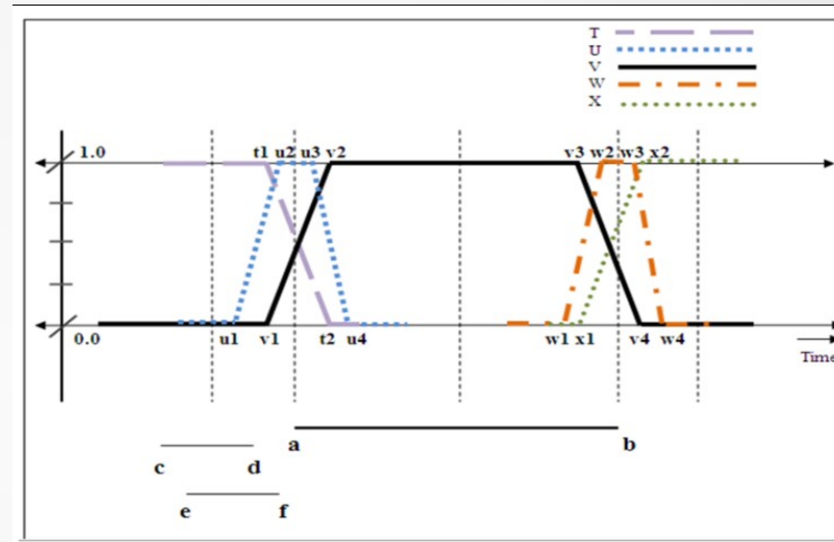
<cd> ***meets*** <ab> iff

- <cd> ∩ <t> ≠ ∅
- <cd> ∩ <u> ≠ ∅
- <cd> ∩ <v> = ∅
- <cd> ∩ <w> = ∅
- <cd> ∩ <x> = ∅

R_11000

Fuzzy Interval algebra

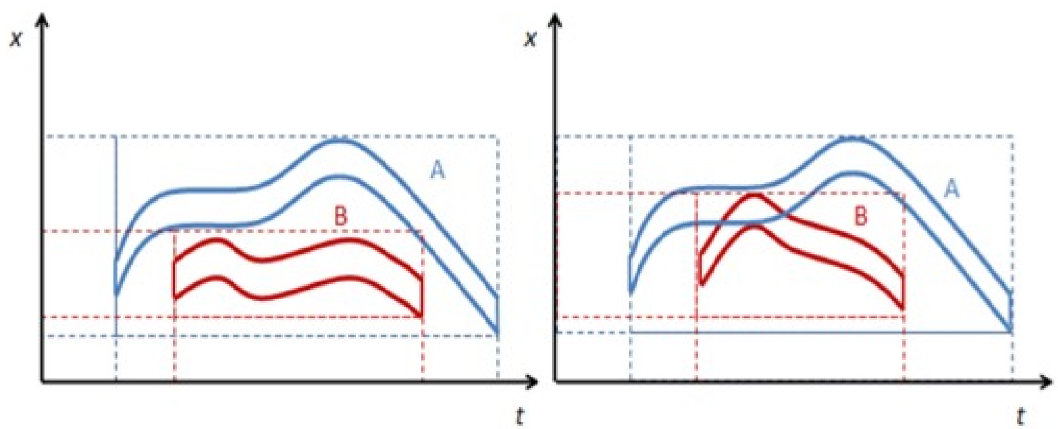




- Replace crisp intersections with fuzzy membership functions
- Represent each relation with a 5-tuple
  - $<\tau.\upsilon,\varpi,\omega,\xi> \in$ [0.1]

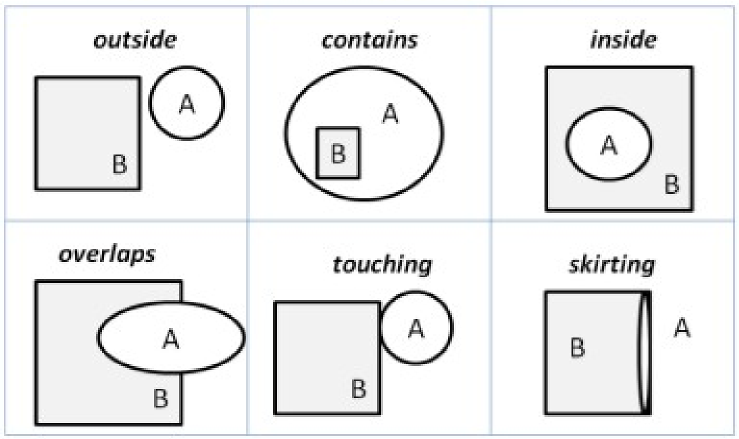
Containment relation

- Consider projections of an event in 4 dimensions <x,y,z,t>
- Define a spatio-temporal relation *formally* with 4 such 5-tuples <τ.υ,ϖ,ω,ξ>



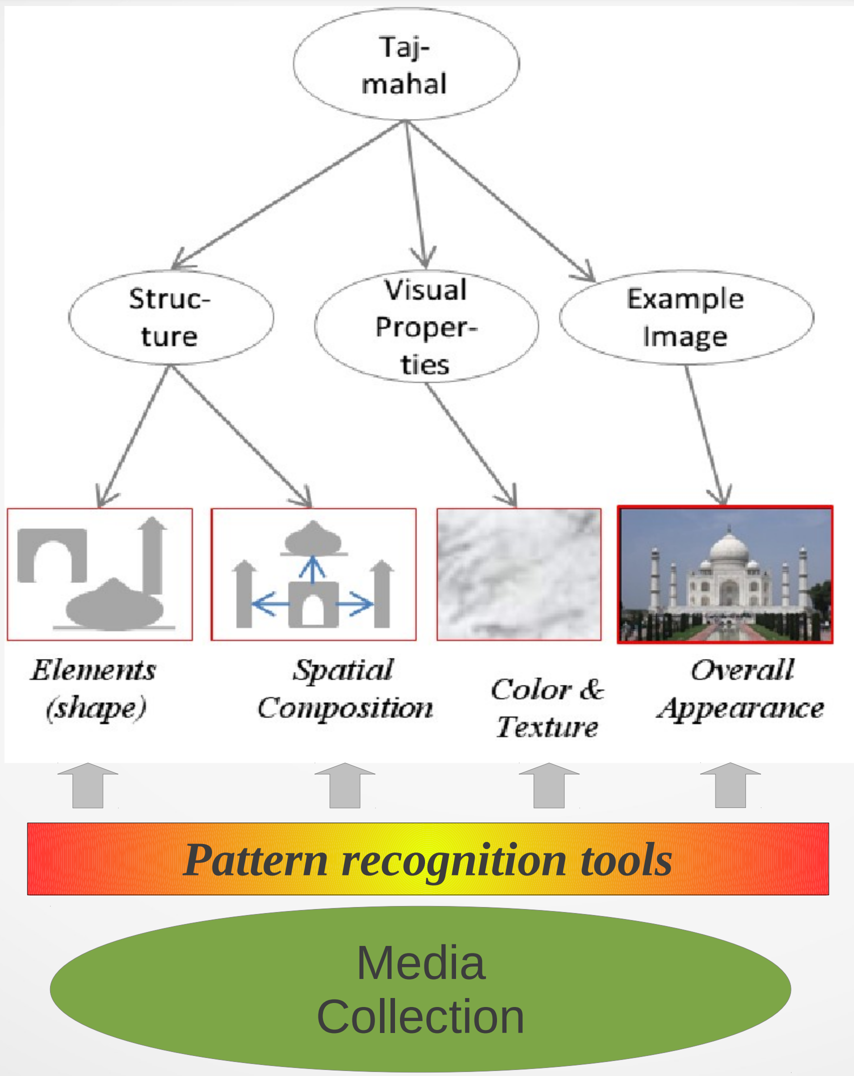
(a) non-intersecting trajectory (b) intersecting trajectory

Distinct containment relations



- Add1for containment relation (4-tuple)
- *Not all the relations are always required in practice*

Concept recognition: Abductive reasoning



Learning the ontology

- Supervised learning: Based on labeled data-set
- Approach
  - Create Obs Model (Bayesian Network)
  - Learn (refine) the network based on labelled data
  - Reverse Engineer ontology from Obs Model
- Two types of learning in Bayesian Network
  - Assume topology, learn probabilities only
  - Learn the network topology

See Heckerman's tutorial

Agenda

Part I

- **Introduction**
- Semantic Web and Ontology
- Semantic Multimedia Content Processing
- Ontology for Multimedia Data Interpretation

Part II

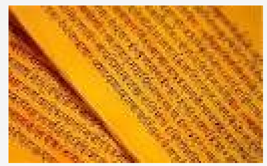
- Multimedia Web Ontology Language
- **Application Examples**
- Distributed Multimedia Applications
- Conclusion

Nrityakosha

An archive of Indian Classical Dance and Music

Indian Classical Dance & Music

Genres
Bharatnatyam, Kuchipudi, Kathakali, Odissi ...



Scriptures

Sculptures

Myths & Mythical characters

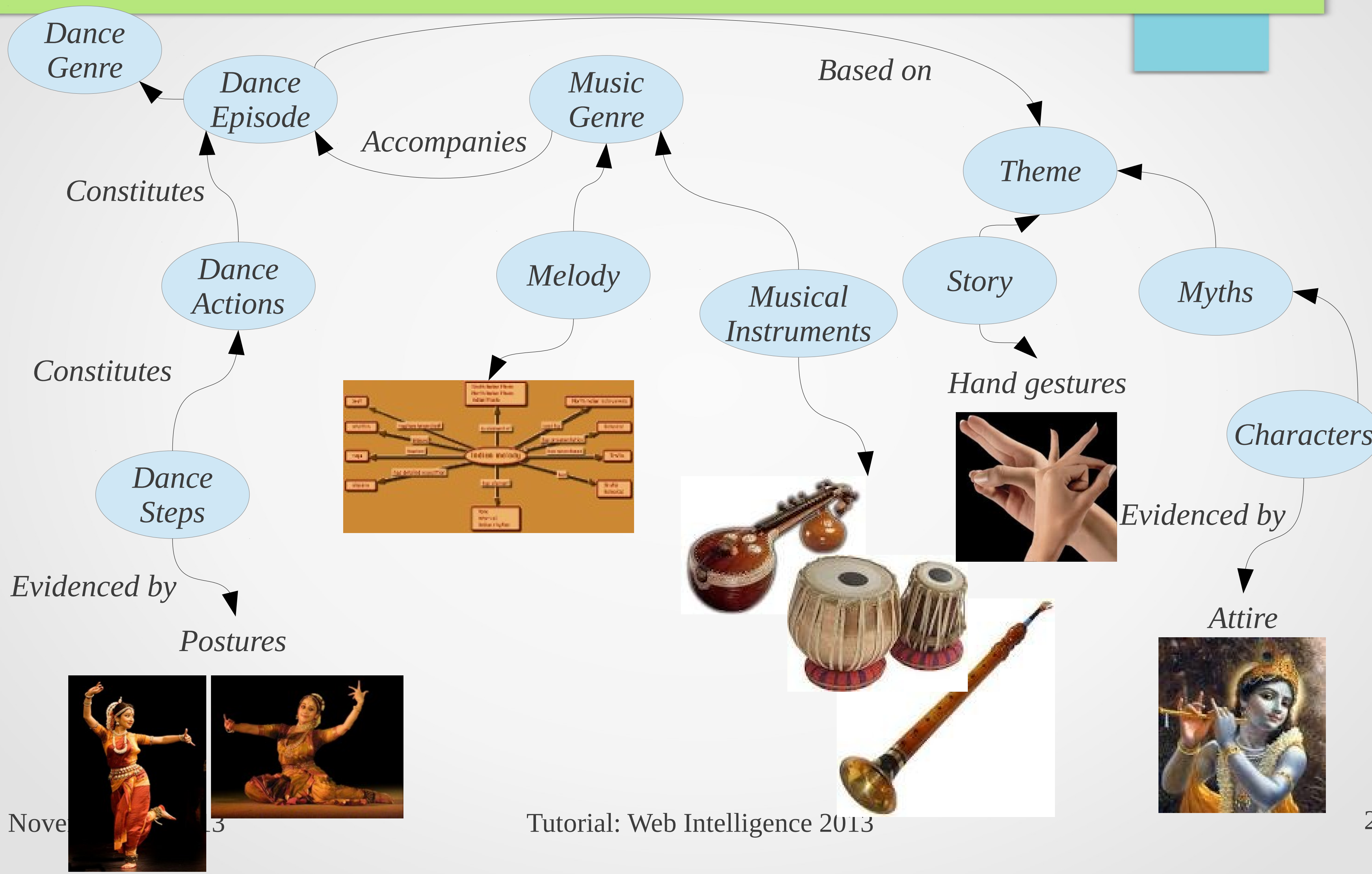
Exponents

Gharanas (The schools)

Musical Instruments

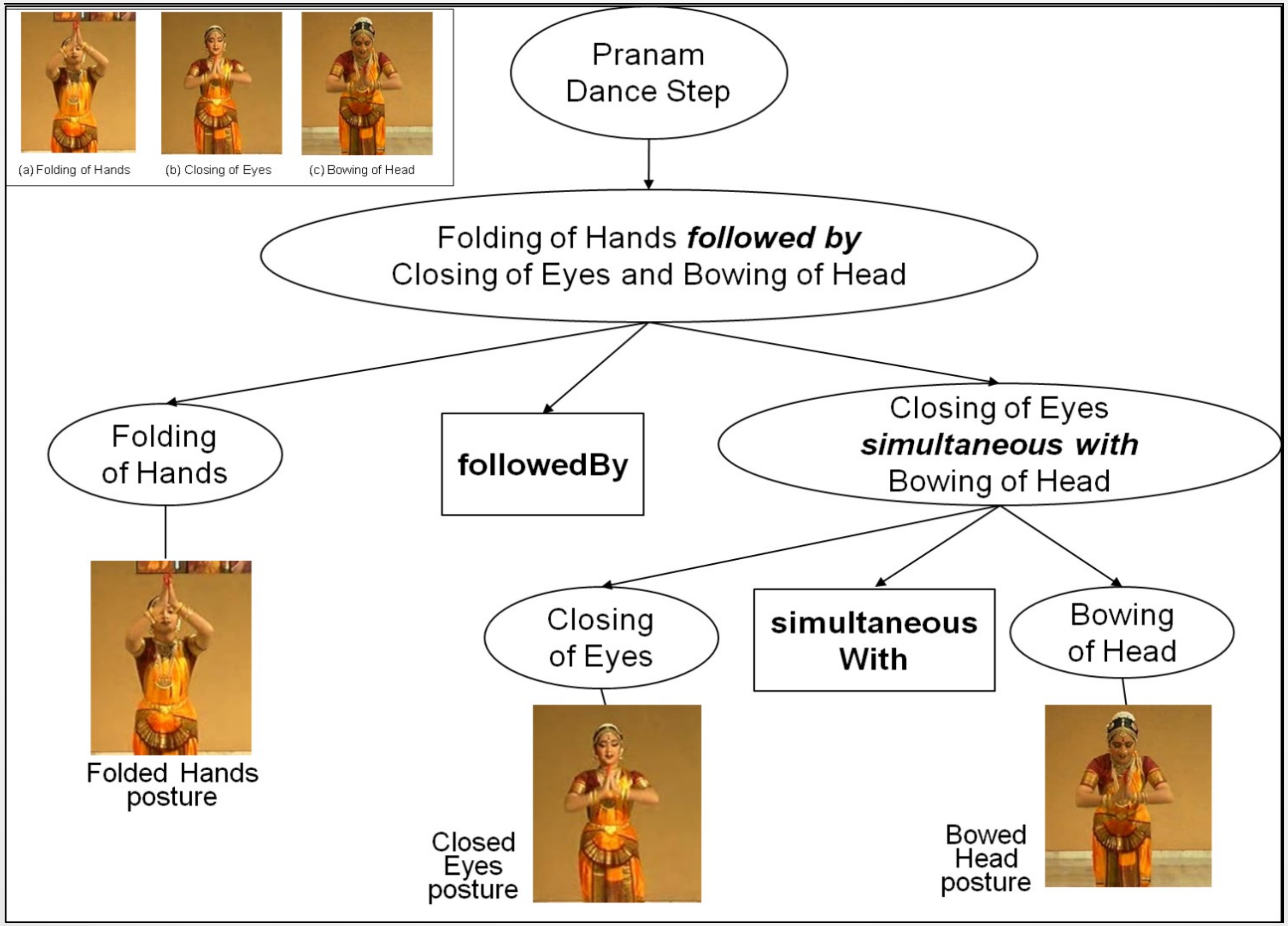
Regions (History, Culture)
Tamilnadu, Odissa, Kerala

Semantic Modeling of the domain



Dance steps and postures

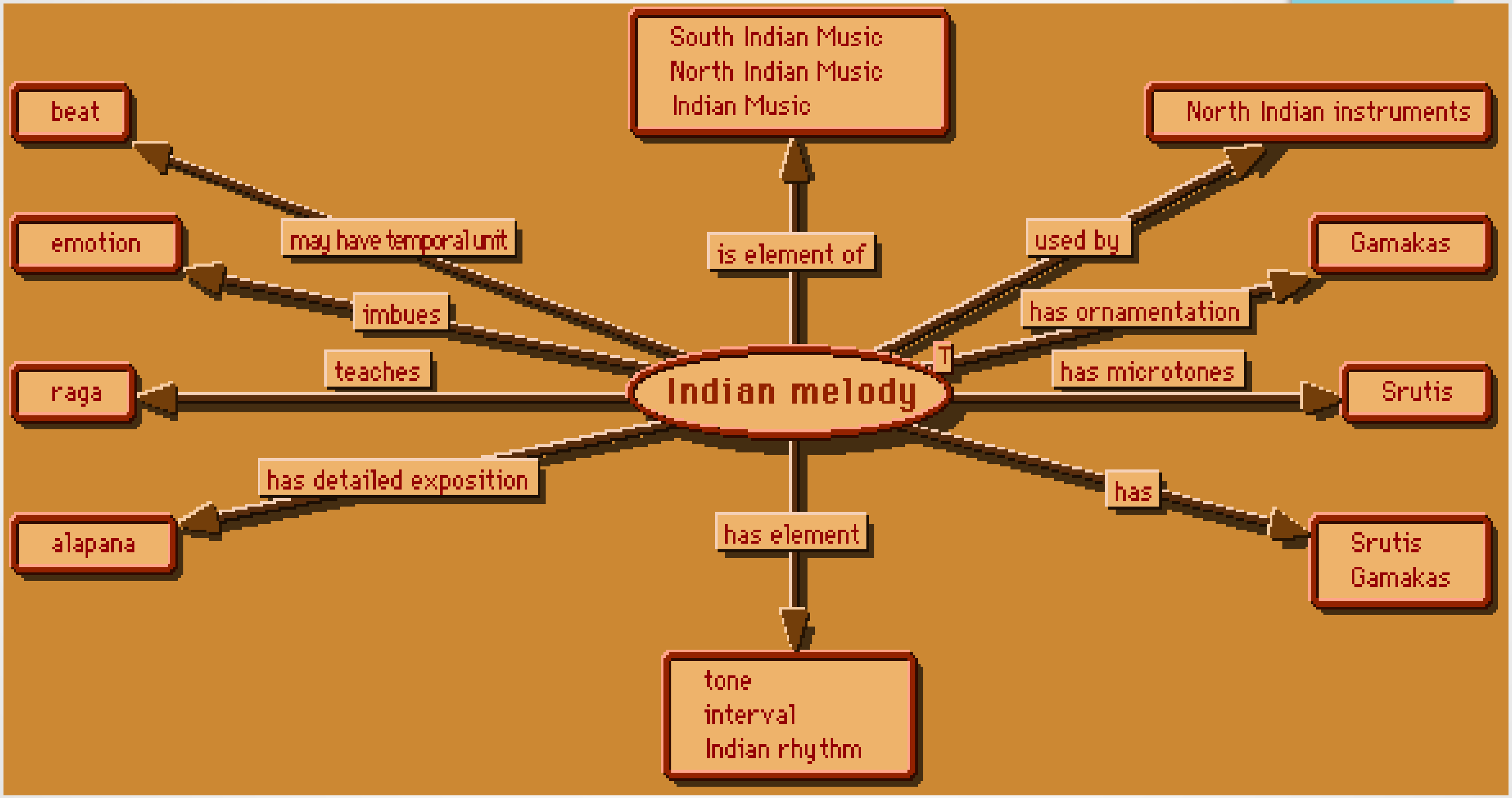
A Dance Step is a temporal composition of elementary Postures



Trained classifiers for posture recognition

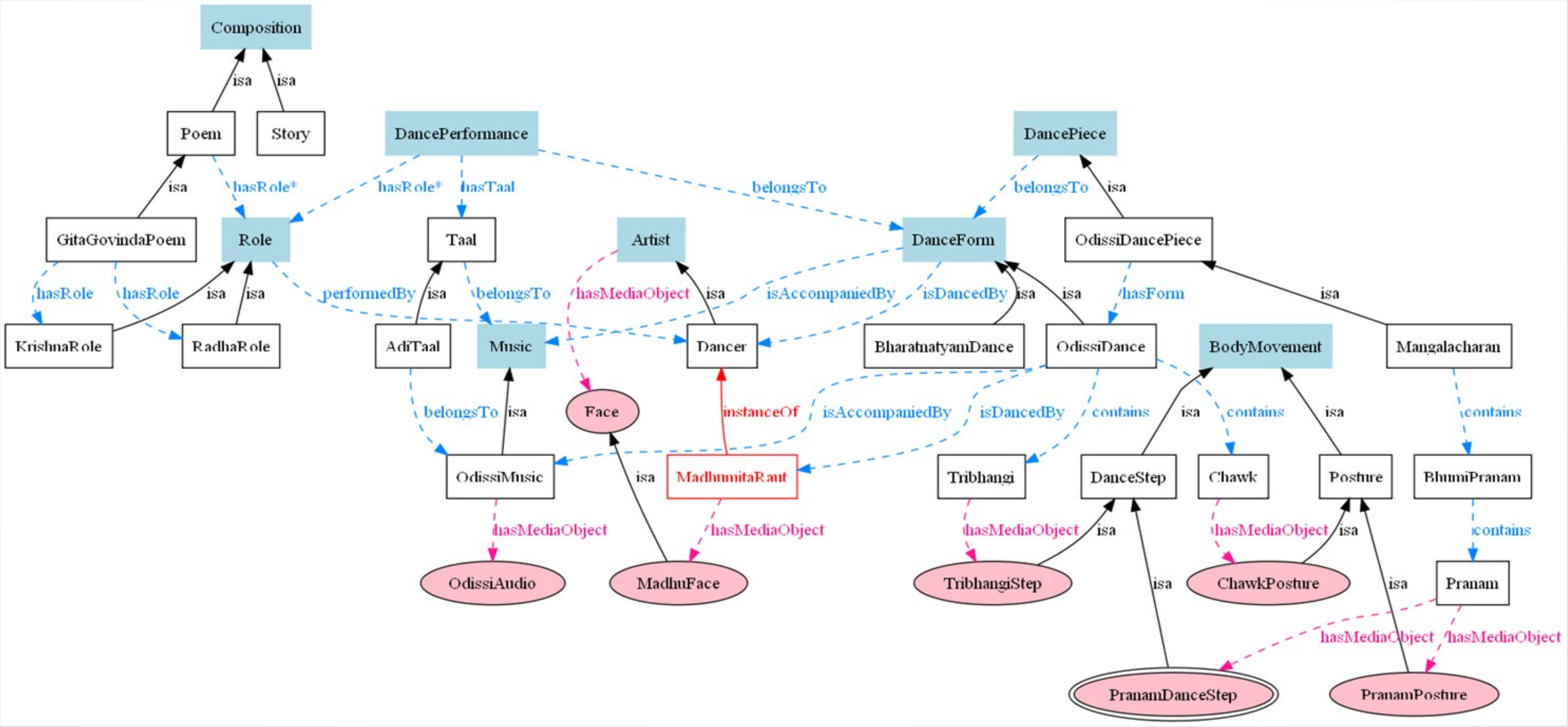
SIFT, SVM

Indian Melody



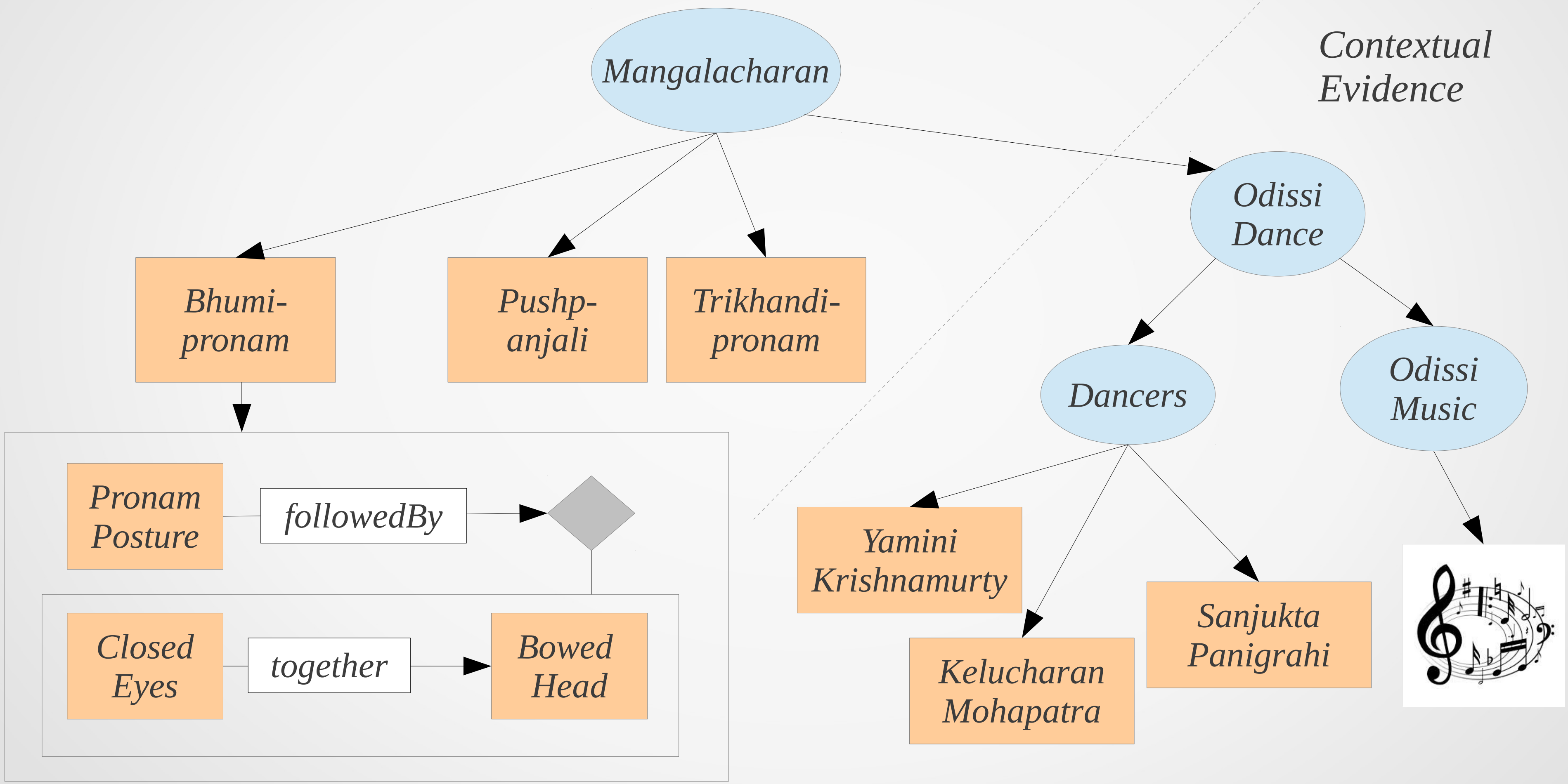
Source:
http://trumpet.sdsu.edu/M151/Knowledge_Webs/7Indian_MusicY/Indian_melody.htm

An ontology snippet



Observation Model for a dance action

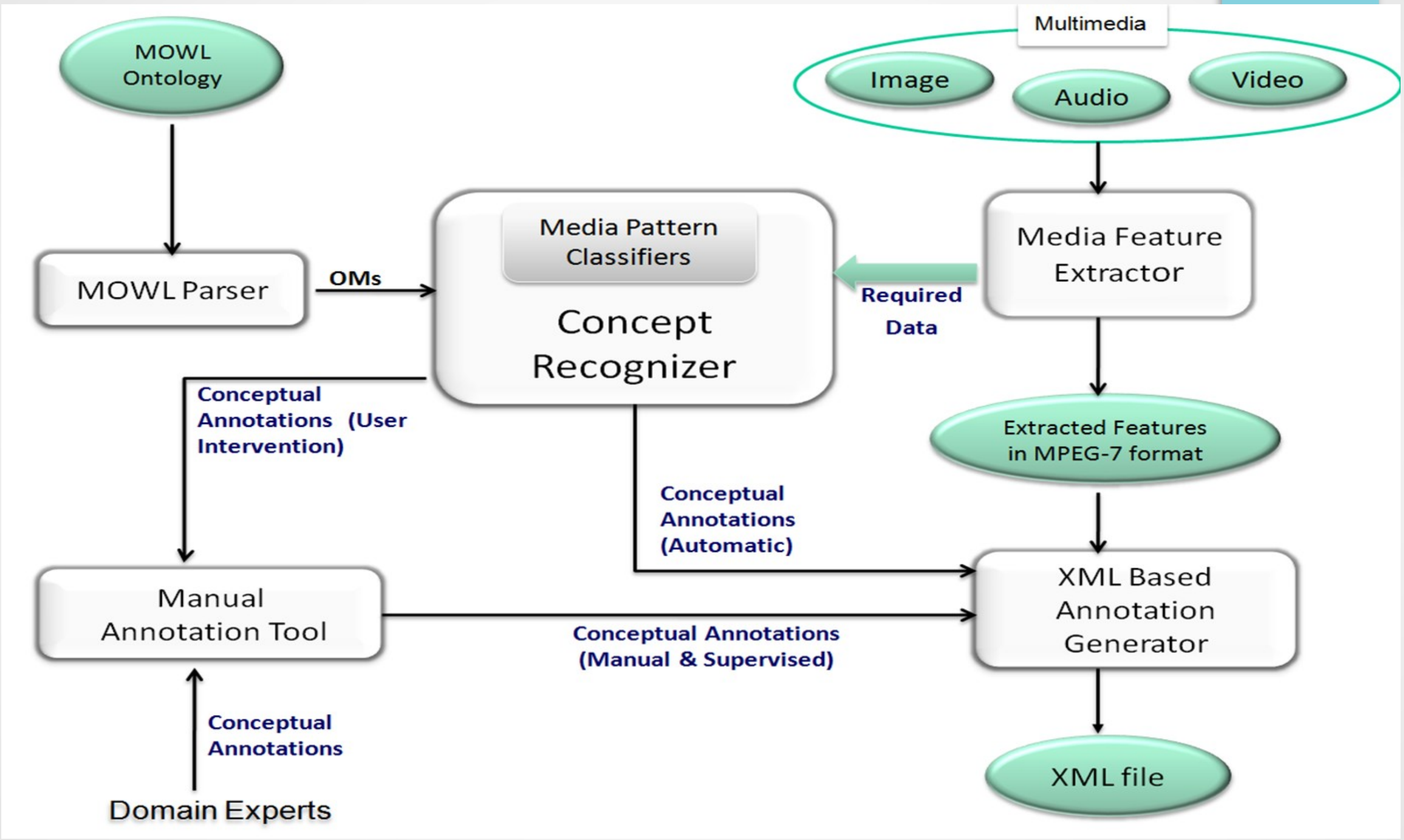
Mangalacharan: Opening action for Odissi dance



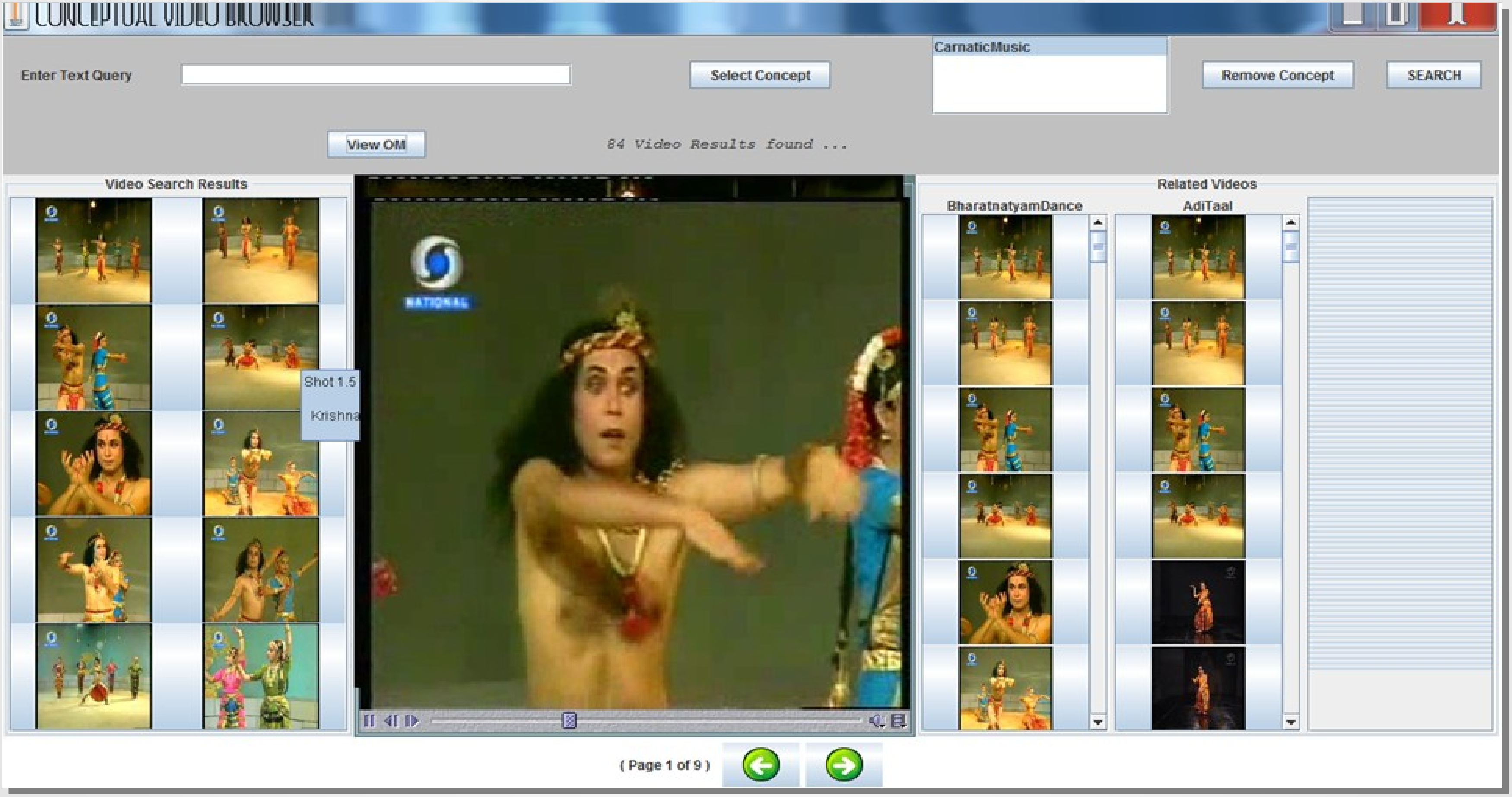
How does it help?

- Concept detection based on multiple cues
  - Robust against failure of one pattern recognizer
- Contextual information provides a powerful cue
- Low-level postures have more discriminating properties than high-level actions
  - Easier to build low-level posture detectors
  - Reasoning with temporal composition for concept recognition

Automatic Annotator

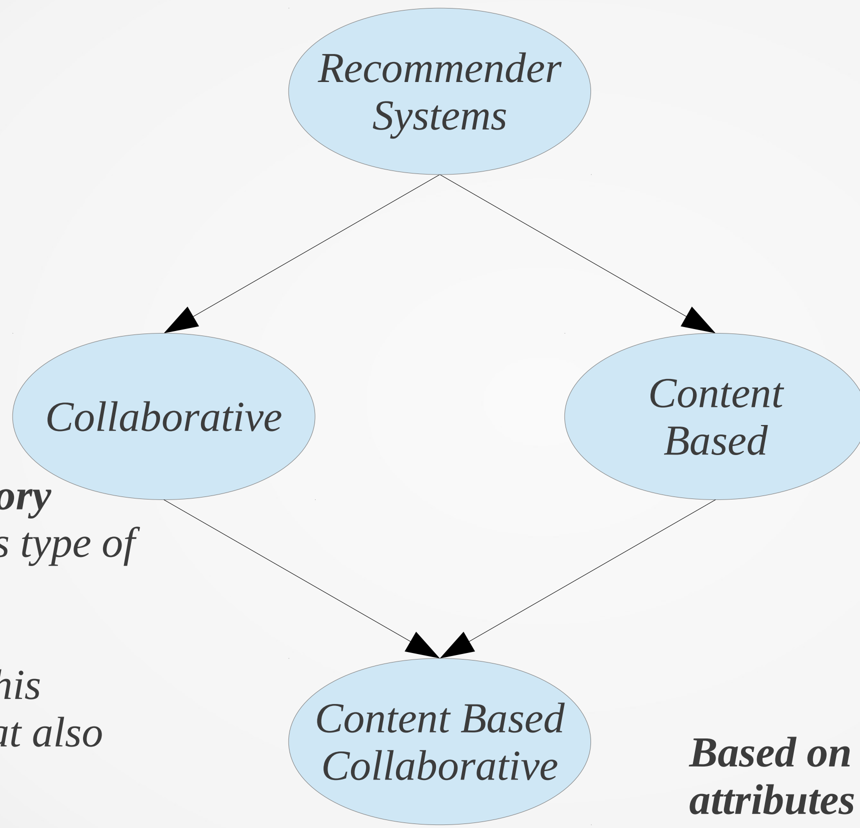


Conceptual browser



Apparel Recommender

Recommender systems



Based on product attributes and user profile

- *You are a EEL 806 Student; you need this book*
- *You spend heavily on credit card; gift this jewelry to your wife*

Based on purchase history

- *You generally buy this type of things*
- *This item is popular*
- *Your friends bought this*
- *Those who bought that also bought this*

Based on similarity of product attributes and purchase history

- *You bouught that movie; this one is by the same director*

Fashion Ontology

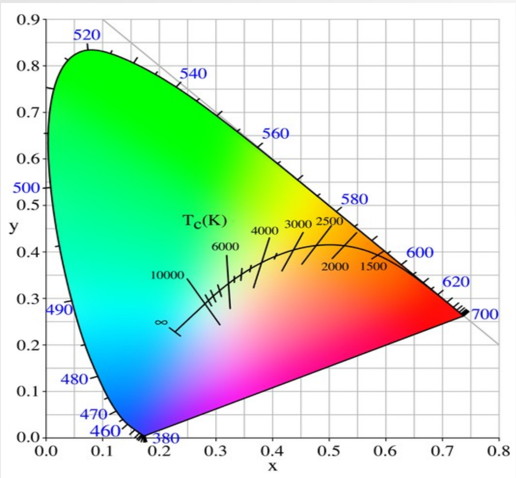
Typical garment recommendation rules

- Body colors influences choice of garment colors
- Body shape dictates material, styles
  - *Do not wear horizontal stripes, if you are short*
- Occasion dictates material, color, craft
  - *Wear silk for a party; cotton for casual*
  - *Wear black / white for mourning (culture-specific)*

The rules are soft:
Susceptable to personal choice, ethnicity, cultural background ...

Color season model

Kentner: 1979



Spring

Summer

Winter

Spring Color Chart

- Every season is characterized by some colors
  - *Related to color temperature, perceptual model?*
- A person's "Color Season" is characterized by her skin, hair and eye colors
- Depending on her color season, some colors suit her better

Causal model for color seasons

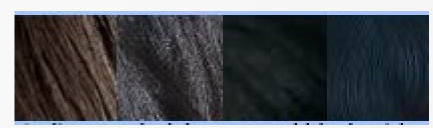
Color Season

{ Spring | Summer | ... }

Manifests in

Skin Colors

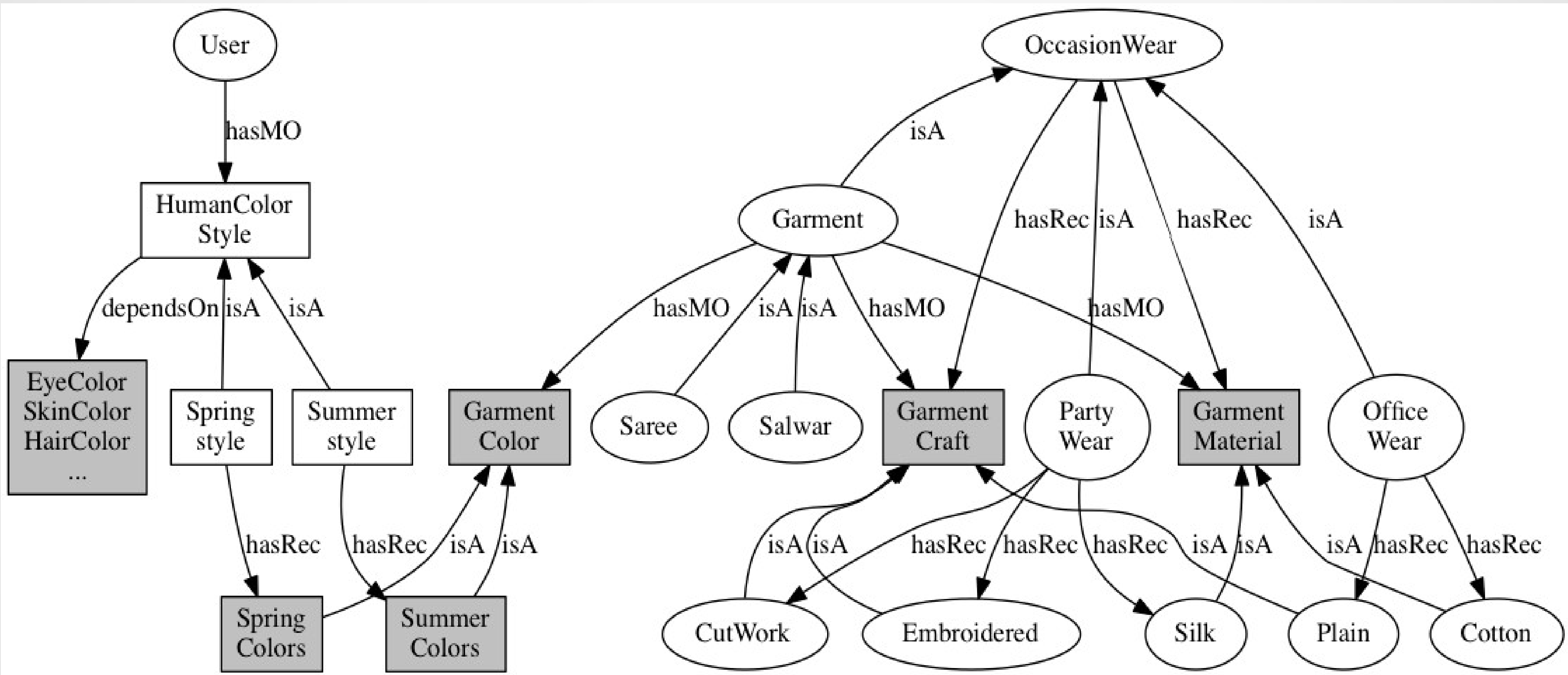
Eye Colors



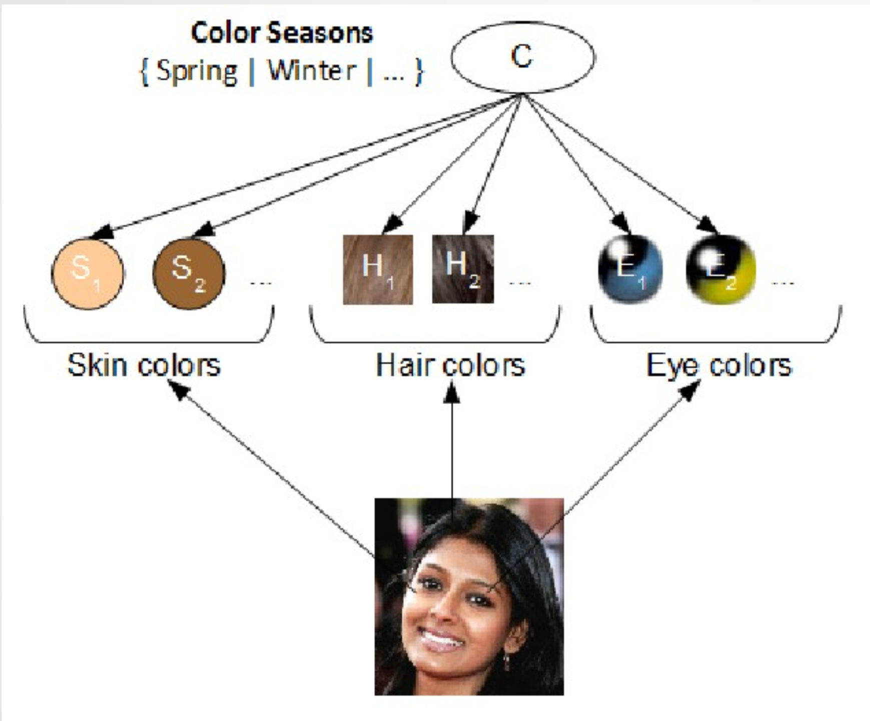
Hair Colors

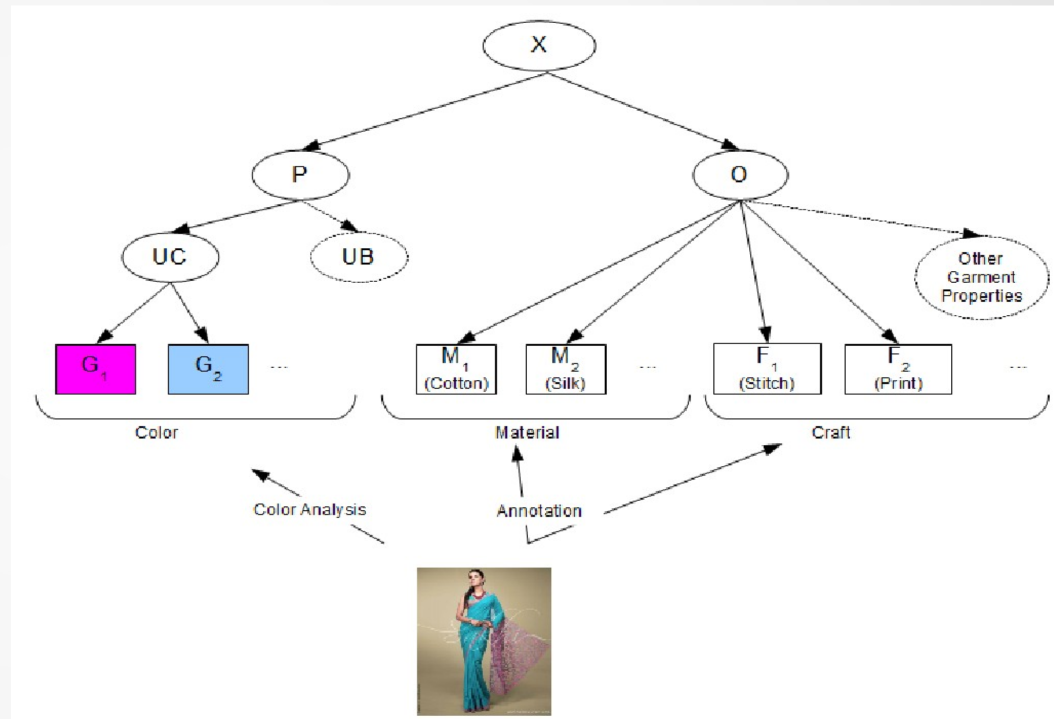
Other fashion concepts are modeled similarly

Modeling Fashion Ontology in MOWL



Reasoning for recommendation

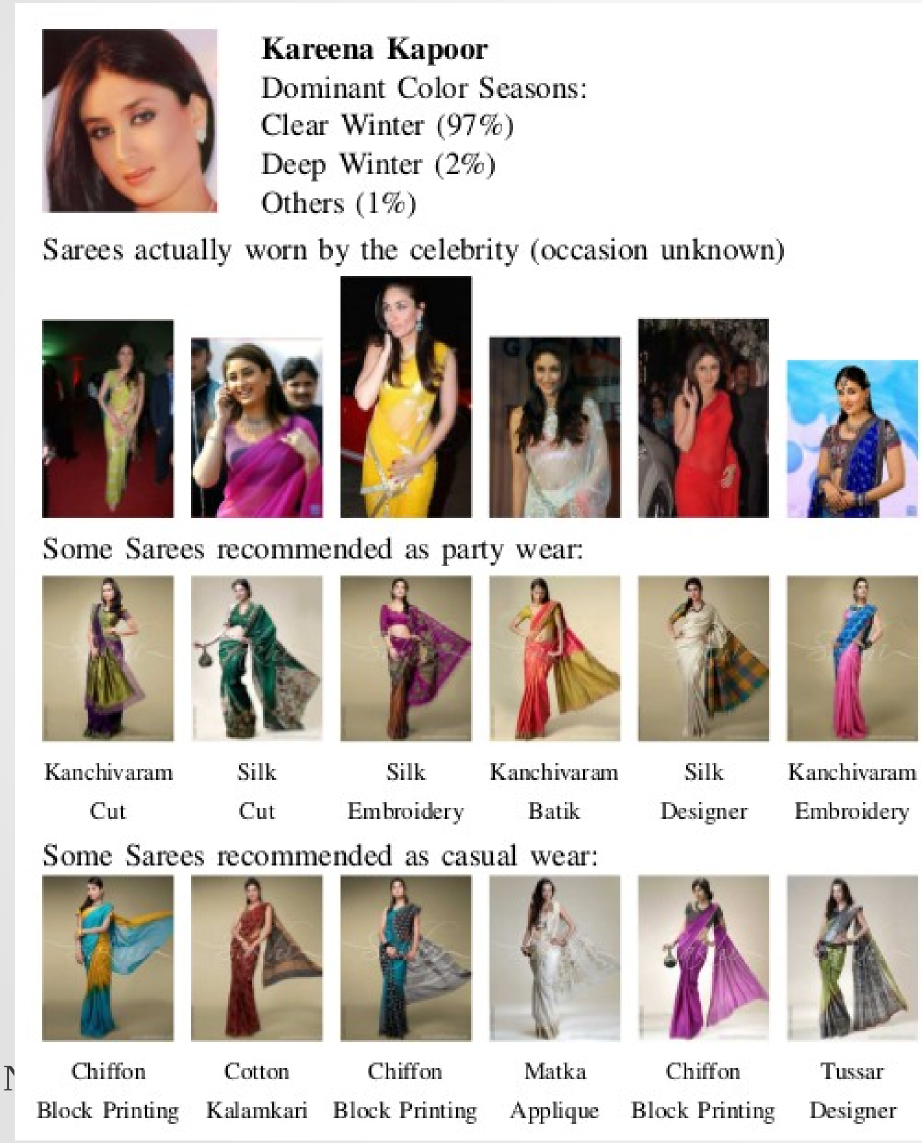




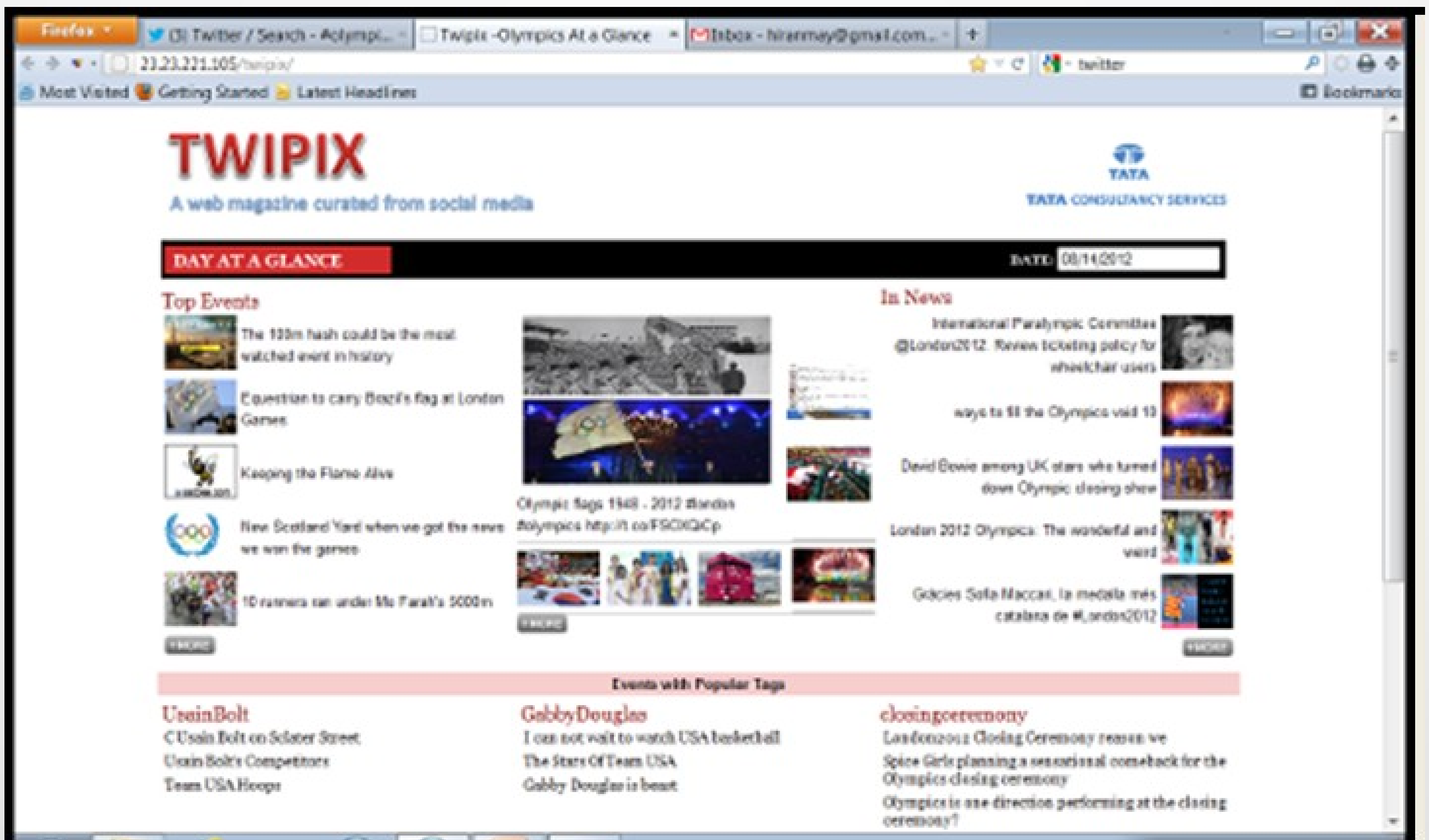
Infer personality attributes

Make recommendations

Recommendation results



Twipix:
Information Integration from Social media


TWIPIX
A web magazine curated from social media
TATA
TATA CONSULTANCY SERVICES
DAY AT A GLANCE
DATE 08/14/2012
Top Events
The 100m hash could be the most watched event in history
Equestrian to carry Brazil's flag at London Games
Keeping the Flame Alive
New Scotland Yard when we got the news we won the games
In News
International Paralympic Committee @London2012: Review ticketing policy for wheelchair users
David Bowie among UK stars who turned down Olympic closing show
London 2012 Olympics: The wonderful and weird
Events with Popular Tags
UsainBolt
Usain Bolt's Competitors
Team USA Hoops
GabbyDouglas
I can not wait to watch USA basketball
The Stars Of Team USA
Gabby Douglas is beast
closingceremony
Spice Girls planning a sensational comeback for the Olympics closing ceremony

Agenda

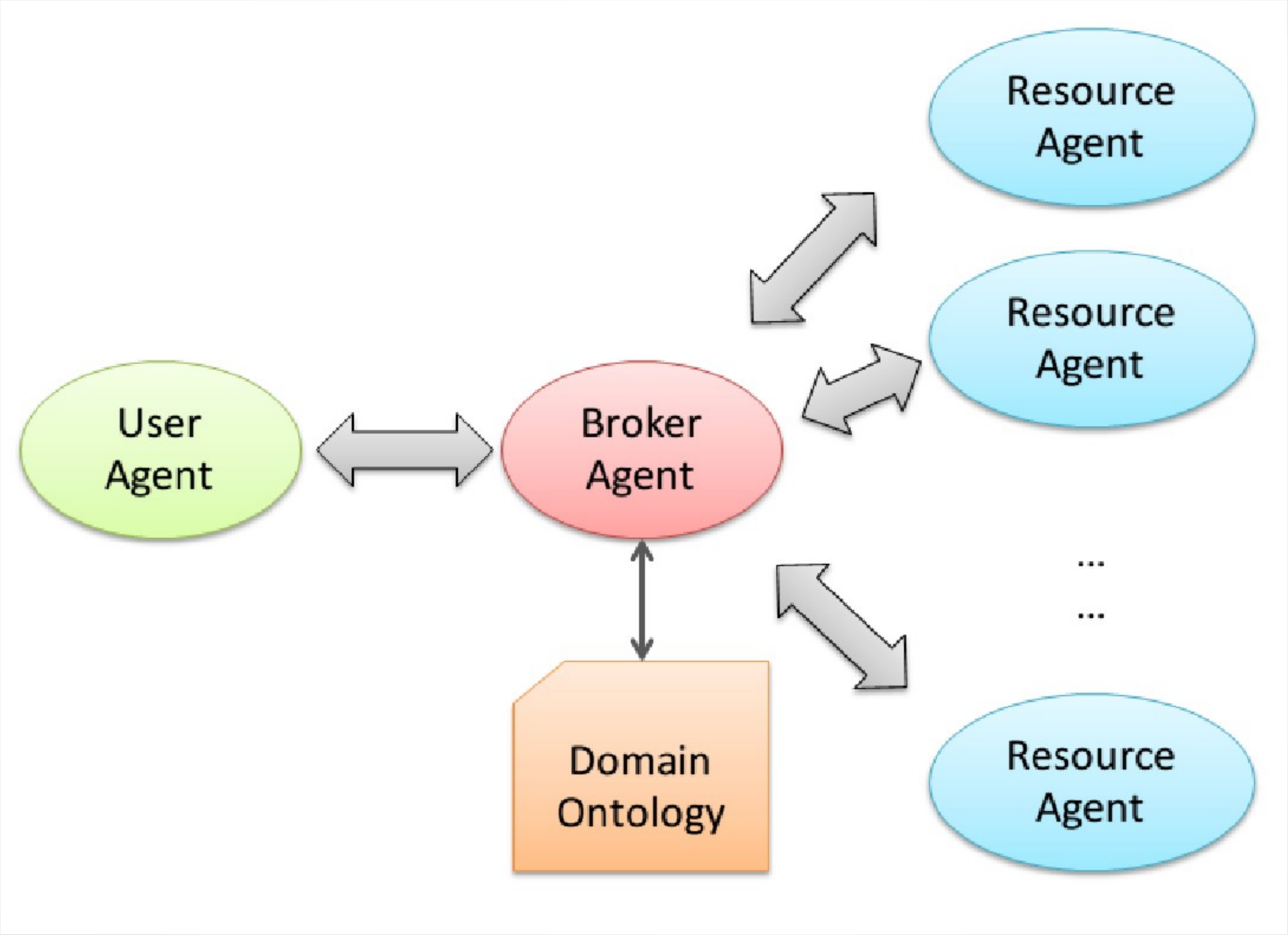
Part I

- **Introduction**
- Semantic Web and Ontology
- Semantic Multimedia Content Processing
- Ontology for Multimedia Data Interpretation

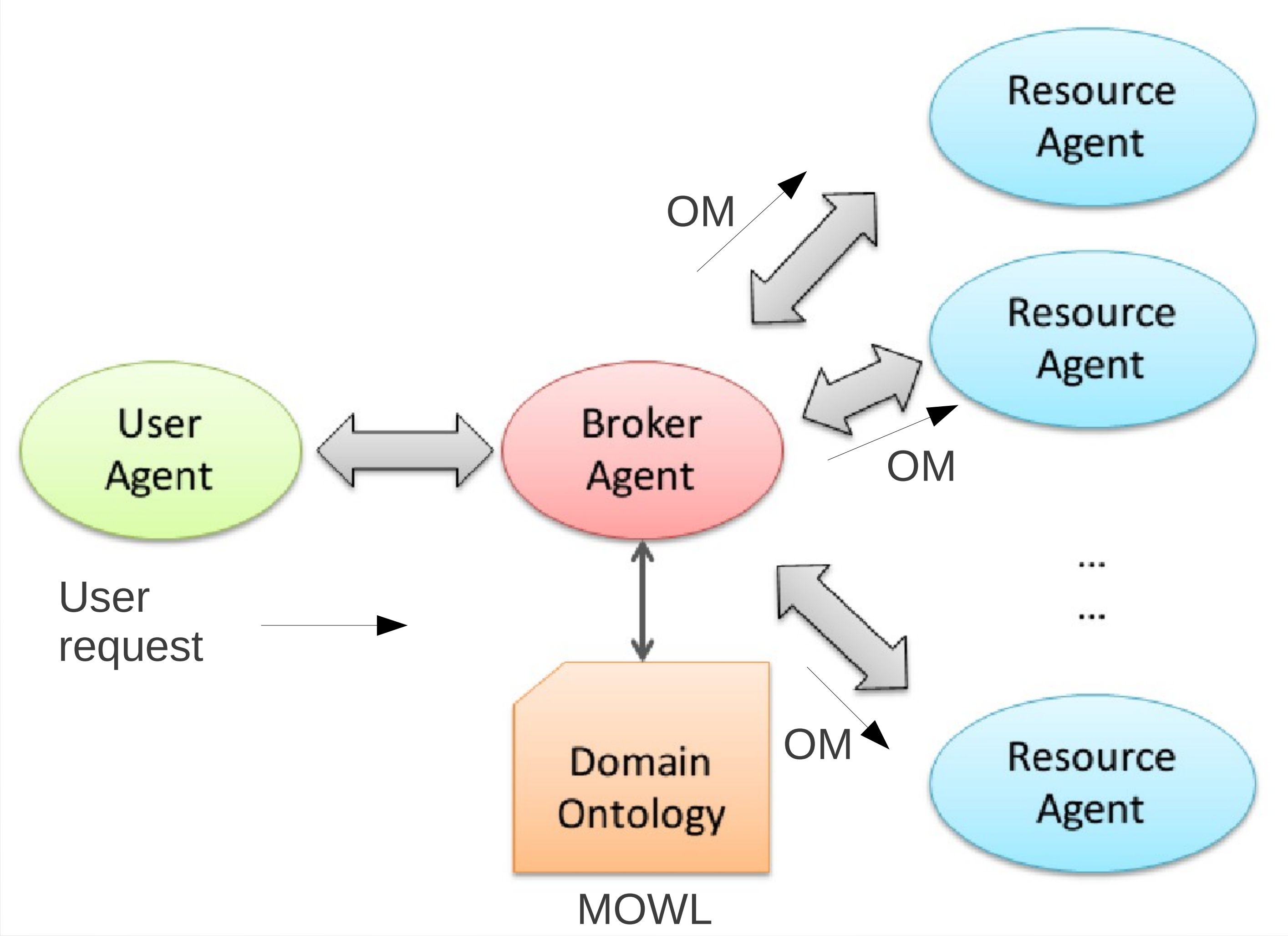
Part II

- Multimedia Web Ontology Language
- Application Examples
- **Distributed Multimedia Applications**
- Conclusion

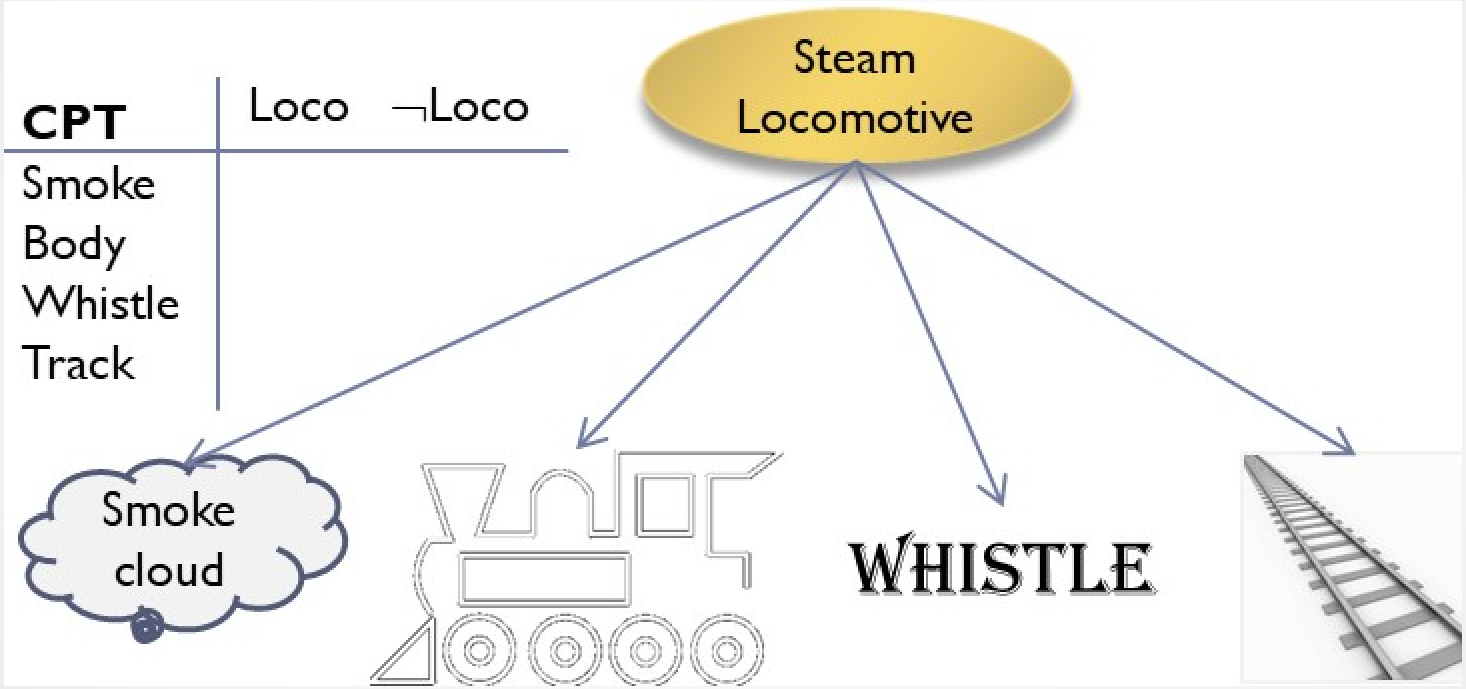
Multi-Agent architecture



Use in multimedia systems



Observation Model has redundancy



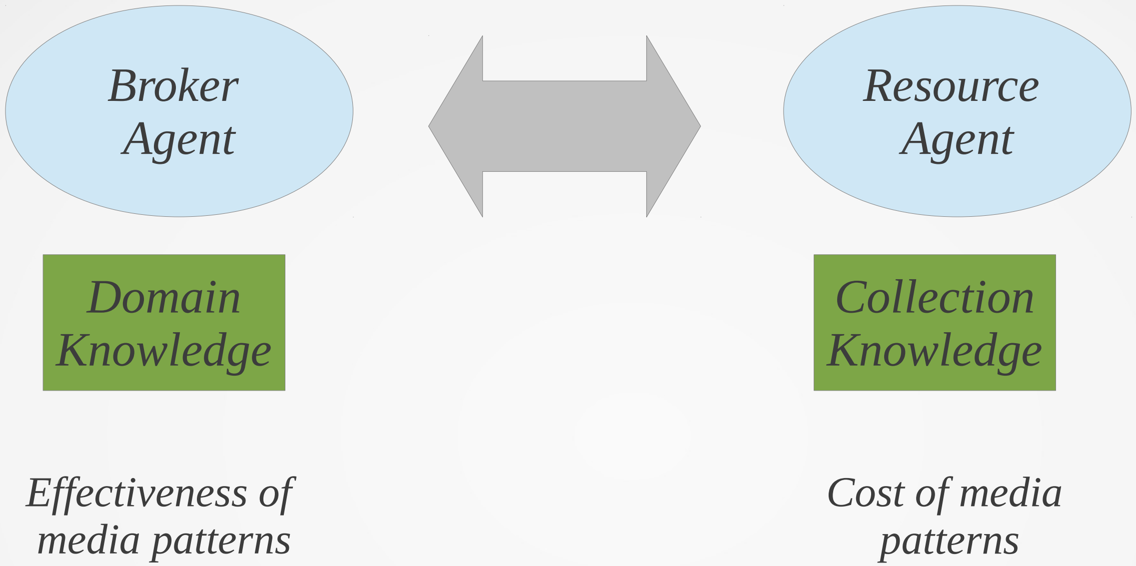
All media-patterns need not be checked for gaining sufficient confidence in the concept

How to choose?

- Different media patterns have different effectiveness in recognizing the concept
  - *Depend on their distinctiveness; reflect in CPT's*
  - Choose the most effective ones
- Different media patterns have diffent costs of computaion
  - Depends on the nature of the pattern, pre-processing and indexing structure of the repository, etc.
  - Can sometimes be infeasible
    - *e.g. Detecting an audio-pattern in an image library*
- Optimal observation plan necessary

Collaborative planning

Ghosh: 2004



<u>*Greedy algorithm*</u>

- *Broker: Recommend most effective pattern*
- *Resource: Compute total cost, effectiveness (confidence in results)*
  - *If satisfy both constraints: stop*
  - *If cost > permissible, reject highest cost pattern: recompute effectiveness*
  - *If effectiveness < permissible, iterate*

Advantages

- *Customized observation plan for each repository*
  - *All derived from the same ontology*
  - *Facilitates Information Integration*
- *Contextual interpretation of media patterns at multiple levels*
  - *Does not rely on pre-annotated concepts alone*
  - *Possible to use higher level patterns (media objects) for concept recognition, if such patterns are pre-computed*
- *Possible to integrate multimodal information*
  - *e.g. Audio, video and image resources*
- *Cross-modal queries possible*

Agenda

Part I

- **Introduction**
- Semantic Web and Ontology
- Semantic Multimedia Content Processing
- Ontology for Multimedia Data Interpretation

Part II

- Multimedia Web Ontology Language
- Application Examples
- Distributed Multimedia Applications
- **Conclusion**

Conclusion

- "*Semantic Gap*" still eludes researchers
  - Semantics of multimedia a subject of research
- Text is symbolic (conceptual); media is perceptual
  - Conventional ontology on top of media processing routines is not a solution
- MOWL is a step in the direction of "*perceptual modeling*" of domain
- Can be extended to different types of sensor data
  - Semantic Sensor Network

Further Reading

- F. Nack, *et al.* “That obscure object of desire: Multimedia metadata on the web, part 2,” IEEE MultiMedia, 12(1), pp. 54–63, Jan. 2005.
- J. R. Smith, “Largescale concept ontology for multimedia,” IEEE Multimedia Magazine, 13, (3), pp. 86–91, Jul–Sep 2006.
- S. Dasiopoulou, *et al.* “Enquiring mpeg-7 based multimedia ontologies,” Multimedia Tools Appl., 46(2-3), pp. 331–370, Jan. 2010.
- A. Mallik, *et al*, “MOWL: An ontology representation language for web based multimedia applications,” ACM Trans on Multimedia Comput., Commun. Appl, (In press).
- A. Mallik, et al, “Nrityakosha: Preserving the intangible heritage of indian classical dance,” J. Comput. Cult. Herit., 4 (3), pp. 11:1–11:25, Dec. 2011.
- S. Ajmani, *et al.* , “An ontology based personalized garment recommendation system,” in Workshop on Personalization, Recommender Systems and Social Media, November 2013.
- R. Bansal, *et al.*, “Twipix: a web magazine curated from social media,” in Proc 20th ACM int conf on Multimedia, 2012, pp. 1355–1356.
- H. Ghosh and S. Chaudhury, “Distributed and reactive query planning in R-MAGIC: An agent-based multimedia retrieval system,” IEEE Trans. on Knowl. and Data Eng., 16(9), pp. 1082–1095, Sep. 2004.
- H. Ghosh, *et al.* “Ontology for Multimedia Applications”. The IEEE Intelligent Informatics Bulletin, 14(1), Dec 2013

Thank you !