# Introduction to Description Logics and their usage in multimedia analysis

Stamatia Dasiopoulou

11 Nov 2009

# Outline

- Description Logics (DLs)
  - Basics (syntax, semantics, inference services)
  - OWL as a DL language
- DLs and multimedia analysis
  - Content representation (briefly)
  - Knowledge-based extraction / interpretation of content semantics

# What are Description Logics?

- A family of logic-based knowledge representation formalisms
  - Descendants of semantics networks, frame-based systems
- Distinguished by
  - Formal semantics
    - Decidable fragments of FOL
  - Inference
    - Sound & complete, highly optimized (implemented) algorithms

# What does it mean to "be" a formal language?

- Syntax
  - What expressions form valid sentences
- Semantics
  - What is the meaning of the expressed sentences
- Reasoning procedures
  - How is implicit knowledge derived from the explicitly stated one

(e.g. If Socrates is human and every human is mortal, I can derive the fact Socrates is mortal)

# DL Basics

- **Concept** names are equivalent to unary predicates
  - Interpreted as sets of objects e.g. Person, Student
- **Role** names are equivalent to binary predicates
  - Interpreted as binary relations on objects e.g. hasChild, likes
- **Individual** names equal constants e.g. Mary, John, India
- **Constructors**: concept and role forming operators

# DL family

- A given DL is defined by the set of allowed constructors
- Smallest propositionally closed DL is $\mathcal{ALC}$ , concept constructed using
  - $\sqcap, \sqcup, \neg$
  - restricted $\exists$ and $\forall$ quantifiers
  - atomic roles

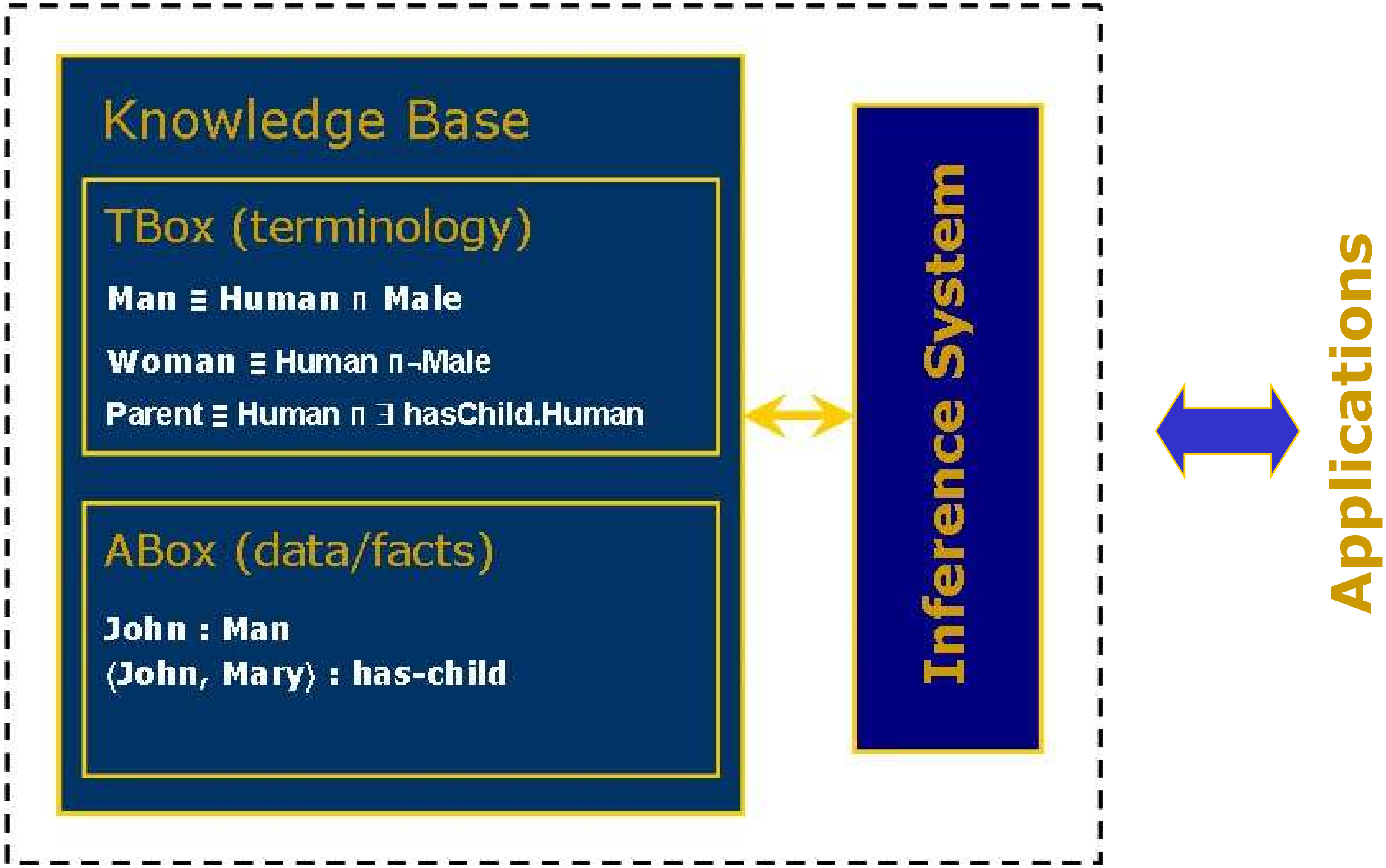
*E.g., Person all of whose children are either Doctors or have a child who is a Doctor:*

$$\text{Person} \sqcap \forall \text{hasChild}.(\text{Doctor} \sqcup \exists \text{hasChild}.\text{Doctor})$$

# DL family (cont.)

- Additional letters indicate other extension, e.g.:
  - H for role inclusion axioms (role hierarchy)
  - O for nominals (singleton classes)
  - I for inverse roles
  - N for number restrictions ($\geq nR$, $\leq nR$)
  - Q for qualified number restrictions ($\geq nR.C$, $\leq nR.C$)

# KR architecture based on DLs



# DL knowledge base (cont.)

- A DL Knowledge base $\mathcal{K}$ is a pair $\langle \mathcal{T}, \mathcal{A} \rangle$ where
  - $\mathcal{T}$ is a set of "terminological" axioms (the TBox)
  - $\mathcal{A}$ is a set of "assertional" axioms (the ABox)
- TBox axioms are of the form:
  $C \sqsubseteq D$, $C \equiv D$, $R \sqsubseteq S$, $R \equiv S$ and $R^+ \sqsubseteq R$

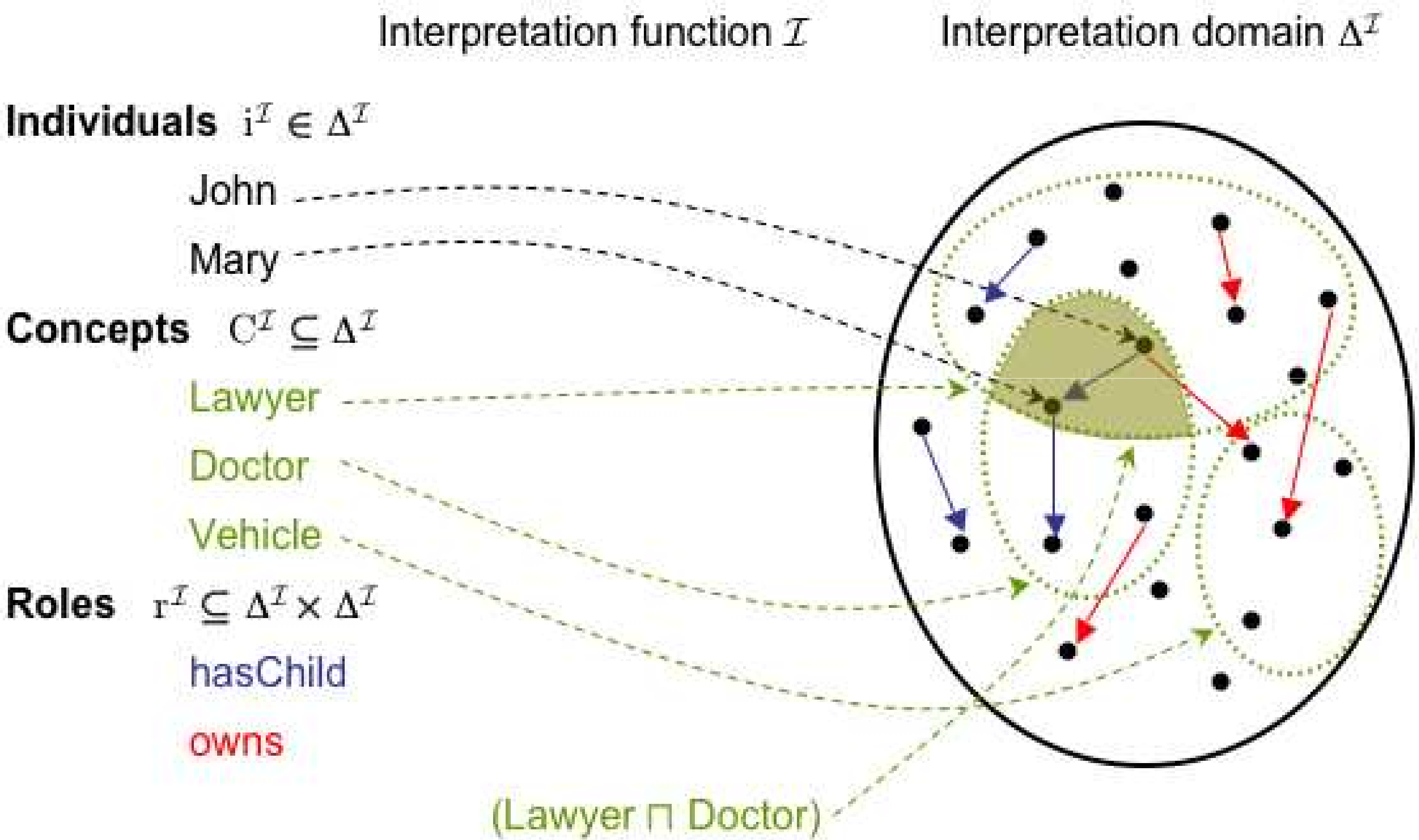
  where C, D concepts, R, S roles, and $R^+$ set of transitive roles
- ABox axioms are of the form:
  x:C, $\langle x,y \rangle$:R (concept/role assertion respectively)

  where x,y are individual names, C a concept and R a role

# DL semantics

- Semantics defined by interpretations
- An interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where
  - $\Delta^{\mathcal{I}}$ is the domain (a non-empty set)
  - $\cdot^{\mathcal{I}}$ is an interpretation function that maps:
    - Concept (class) name A $\rightarrow$ subset $A^{\mathcal{I}}$ of $\Delta^{\mathcal{I}}$
    - Role (property) name R $\rightarrow$ binary relation $R^{\mathcal{I}}$ over $\Delta^{\mathcal{I}}$
    - Individual name i $\rightarrow$ $i^{\mathcal{I}}$ element of $\Delta^{\mathcal{I}}$

# DL Semantics (cont.)




(FROM: I.Horrocks, OWL: A Description Logic Based Ontology Language, Seminar at the Centre for Intelligent Systems and their Applications, Uni. Of Edinburgh, Scotland, March, 2006)

# DL Semantics (cont.)

- Interpretation function ¢I extends to concept (and role) expressions

$$
\begin{array}{lcl}
(C \sqcap D)^{\mathcal{I}} & = & C^{\mathcal{I}} \cap D^{\mathcal{I}} \\
(C \sqcup D)^{\mathcal{I}} & = & C^{\mathcal{I}} \cup D^{\mathcal{I}} \\
(\neg C)^{\mathcal{I}} & = & \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}} \\
\{x\}^{\mathcal{I}} & = & \{x^{\mathcal{I}}\} \\
(\exists R.C)^{\mathcal{I}} & = & \{x \mid \exists y.\langle x, y\rangle \in R^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\} \\
(\forall R.C)^{\mathcal{I}} & = & \{x \mid \forall y.(x, y) \in R^{\mathcal{I}} \Rightarrow y \in C^{\mathcal{I}}\} \\
(\leqslant nR)^{\mathcal{I}} & = & \{x \mid \#\{y \mid \langle x, y\rangle \in R^{\mathcal{I}}\} \leqslant n\} \\
(\geqslant nR)^{\mathcal{I}} & = & \{x \mid \#\{y \mid \langle x, y\rangle \in R^{\mathcal{I}}\} \geqslant n\} \\
(R^{-})^{\mathcal{I}} & = & \{(x, y) \mid (y, x) \in R^{\mathcal{I}}\}
\end{array}
$$

# DL knowledge base Semantics

- An interpretation $\mathcal{I}$ satisfies (is a model) a TBox axiom A ($\mathcal{I} \models$ A):

  $\mathcal{I} \models C \sqsubseteq D$ iff $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$  $\mathcal{I} \models C \equiv D$ iff $C^{\mathcal{I}} = D^{\mathcal{I}}$

  $\mathcal{I} \models R \sqsubseteq S$ iff $R^{\mathcal{I}} \subseteq S^{\mathcal{I}}$  $\mathcal{I} \models R \equiv S$ iff $R^{\mathcal{I}} = S^{\mathcal{I}}$

  $\mathcal{I} \models R^{+} \sqsubseteq R$ iff $(R^{\mathcal{I}})^{+} \subseteq R^{\mathcal{I}}$
- $\mathcal{I}$ satisfies a TBox $\mathcal{T}$ ($\mathcal{I} \models \mathcal{T}$) iff $\mathcal{I}$ satisfies every axiom A in $\mathcal{T}$

- An interpretation $\mathcal{I}$ satisfies (models) an ABox axiom A ($\mathcal{I} \models$ A):

  $\mathcal{I} \models x{:}D$ iff $x^{\mathcal{I}} \in D^{\mathcal{I}}$  $\mathcal{I} \models \langle x,y \rangle{:}R$ iff $(x^{\mathcal{I}},y^{\mathcal{I}}) \in R^{\mathcal{I}}$
- $\mathcal{I}$ satisfies an ABox $\mathcal{A}$ ($\mathcal{I} \models \mathcal{A}$) iff $\mathcal{I}$ satisfies every axiom A in $\mathcal{A}$

- $\mathcal{I}$ satisfies an KB $\mathcal{K}$ ($\mathcal{I} \models \mathcal{K}$) iff $\mathcal{I}$ satisfies both $\mathcal{T}$ and $\mathcal{A}$

# CWA vs OWA

- ***Closed World*** *Assumption*: assumes that the available information is complete
  - If an assertion cannot be derived then its negation is deduced

- ***Open World*** *Assumption*: absence of information means lack of information
  - The assertion holds in some models, and doesn't hold in others

# Example

- $A_E$={hasChild(Iokaste,Edipus),
  hasChild(Iokaste,Polyneikes),
  hasChild(Edipus,Polyneikis),
  hasChild(Polyneikis,Thesandros),
  Patricide(Edipus), ¬Patricide(Thesandros)}.

$A_E$|={Iokaste:∃hasChild.(Patricide ⊓ ∃hasChild.¬Patricide)}
???

# Inference Services

- Every logical formalism provides its own reasoning services.

- Description Logics (DLs) provide reasoning services for TBoxes, ABoxes and Knowledge Bases (TBoxes and ABoxes together).

# Inference Services for TBoxes

- **Satisfiability:** A concept C is *satisfiable* w.r.t. a TBox T if there exists a model I of T such that $C^I \neq \emptyset$. Then I is called a *model* of C.
- **Subsumption:** A concept C is *subsumed* by a concept D w.r.t. T, written $C \sqsubseteq D$, if for every model I of T, $C^I \subseteq D^I$.
- **Equivalence:** Two concepts C and D are *equivalent* w.r.t. T, written $C \equiv D$, if for every model I of T, $C^I \equiv D^I$.
- **Disjointness:** Two concepts C and D are *disjoint*, w.r.t. T, written $C \neq D$, if for every model I of T, $C^I \neq D^I$.

# Some examples

- Is $Man \sqcap \neg Man$ satisfiable w.r.t. an empty TBox? **No**
- Is $Man \sqcap Woman$ satisfiable w.r.t. an empty TBox? **Yes**

- $Woman \sqsubseteq Person$ w.r.t. empty TBox?

  **No!** one can create an interpretation where

  $Woman^I \supseteq Person^I$
- $Man \subseteq Person$ w.r.t. $T=\{Person \equiv Man \sqcup Woman\}$?

  **Yes!** in all models I of T, $Person^I$ contains all objects of $Man^I$ plus all of $Woman^I$.

# Inference Services for ABoxes

- **Consistency:** An ABox A is *consistent* w.r.t. a TBox T if there exists a model of T which satisfies each assertion in A.

- **Entailment (Instance Checking):** An ABox A *entails* an assertion φ, written $A \models \varphi$ iff every interpretation that satisfies A also satisfies the assertion.

# Some examples

- Is A={Man(Jim), Woman(Jim)} consistent w.r.t. an empty TBox? **Yes**
- Is the above ABox A consistent w.r.t.

  T={Woman≡Person⊓Female,

  Man≡Person⊓¬Female}? **No**

# Automated reasoning

- State of the art DL systems typically use (highly optimised) tableaux algorithms

- Tableaux algorithms work by trying to construct a concrete example (model) consistent with KB axioms:
  - Start from ground facts (ABox axioms)
  - Explicate structure implied by complex concepts and TBox axioms
    - Syntactic decomposition using tableaux expansion rules
    - Infer constraints on (elements of) model

# Tableaux Expansions rules

| | |
|---|---|
| $\sqcap$-rule | if 1. $(C_1 \sqcap C_2) \in \mathcal{L}(v)$, $v$ is not indirectly blocked, and<br>2. $\{C_1, C_2\} \not\subseteq \mathcal{L}(v)$<br>then $\mathcal{L}(v) \rightarrow \mathcal{L}(v) \cup \{C_1, C_2\}$. |
| $\sqcup$-rule | if 1. $(C_1 \sqcup C_2) \in \mathcal{L}(v)$, $v$ is not indirectly blocked, and<br>2. $\{C_1, C_2\} \cap \mathcal{L}(v) = \emptyset$<br>then $\mathcal{L}(v) \rightarrow \mathcal{L}(v) \cup \{E\}$ for some $E \in \{C_1, C_2\}$ |
| $\exists$-rule | if 1. $\exists r.C \in \mathcal{L}(v_1)$, $v_1$ is not blocked, and<br>2. $v_1$ has no safe $r$-neighbour $v_2$ with $C \in \mathcal{L}(v_1)$,<br>then create a new node $v_2$ and an edge $\langle v_1, v_2 \rangle$<br>with $\mathcal{L}(v_2) = \{C\}$ and $\mathcal{L}(\langle v_1, v_2 \rangle) = \{r\}$. |
| $\forall$-rule | if 1. $\forall r.C \in \mathcal{L}(v_1)$, $v_1$ is not indirectly blocked, and<br>2. there is an $r$-neighbour $v_2$ of $v_1$ with $C \notin \mathcal{L}(v_2)$<br>then $\mathcal{L}(v_2) \rightarrow \mathcal{L}(v_2) \cup \{C\}$. |
| $\forall_+$-rule | if 1. $\forall r.C \in \mathcal{L}(v_1)$, $v_1$ is not indirectly blocked, and<br>2. there is some role $r'$ with $\mathsf{Trans}(r')$ and $r' \sqsubseteq^{*} r$<br>3. there is an $r'$-neighbour $v_2$ of $v_1$ with $\forall r'.C \notin \mathcal{L}(v_2)$<br>then $\mathcal{L}(v_2) \rightarrow \mathcal{L}(v_2) \cup \{\forall r'.C\}$. |
| *choose*-rule | if 1. $\leqslant n\, r.C \in \mathcal{L}(v_1)$, $v_1$ is not indirectly blocked, and<br>2. there is an $r$-neighbour $v_2$ of $v_1$ with $\{C, \dot{\neg} C\} \cap \mathcal{L}(v_2) = \emptyset$<br>then $\mathcal{L}(v_2) \rightarrow \mathcal{L}(v_2) \cup \{E\}$ for some $E \in \{C, \dot{\neg} C\}$. |
| $\geqslant$-rule | if 1. $\geqslant n\, r.C \in \mathcal{L}(v)$, $v$ is not blocked, and<br>2. there are not $n$ safe $r$-neighbours $v_1, \ldots, v_n$ of $v$<br>with $C \in \mathcal{L}(v_i)$ and $v_i \neq v_j$ for $1 \leqslant i < j \leqslant n$ |

# DLs Reasoners

- Mature, highly optimized implementations
  - Research implementations
    - FaCT++, Pellet, KAON2, CEL, HermiT…
  - Commercial implementations
    - Cerebra, RacerPro, SHER..

(http://www.cs.manchester.ac.uk/~sattler/reasoners.html)

# The Web Ontology Language (OWL) as a DL language

- OWL is W3C recommendation (i.e. a standard)
  - OWL DL is equivalent to the SHOIN
  - OWL Lite is equivalent to SHIF
  - Further connections issue from the recent OWL1.1 and OWL2 recommendations

- OWL exploits results of 15+ years of research in DLs
  - Well defined semantics
  - Complexity, decidability results
  - Reasoning algorithms
  - Implemented systems

# OWL class constructors

| Constructor | DL Syntax | Example | FOL Syntax |
|---|---|---|---|
| intersectionOf | $C_1 \sqcap \ldots \sqcap C_n$ | Human $\sqcap$ Male | $C_1(x) \wedge \ldots \wedge C_n(x)$ |
| unionOf | $C_1 \sqcup \ldots \sqcup C_n$ | Doctor $\sqcup$ Lawyer | $C_1(x) \vee \ldots \vee C_n(x)$ |
| complementOf | $\neg C$ | $\neg$Male | $\neg C(x)$ |
| oneOf | $\{x_1\} \sqcup \ldots \sqcup \{x_n\}$ | {john} $\sqcup$ {mary} | $x = x_1 \vee \ldots \vee x = x_n$ |
| allValuesFrom | $\forall P.C$ | $\forall$hasChild.Doctor | $\forall y.P(x,y) \rightarrow C(y)$ |
| someValuesFrom | $\exists P.C$ | $\exists$hasChild.Lawyer | $\exists y.P(x,y) \wedge C(y)$ |
| maxCardinality | $\leqslant nP$ | $\leqslant$1hasChild | $\exists^{\leqslant n} y.P(x,y)$ |
| minCardinality | $\geqslant nP$ | $\geqslant$2hasChild | $\exists^{\geqslant n} y.P(x,y)$ |

# OWL axioms

| OWL Syntax | DL Syntax | Example |
|---|---|---|
| subClassOf | $C_1 \sqsubseteq C_2$ | Human $\sqsubseteq$ Animal $\sqcap$ Biped |
| equivalentClass | $C_1 \equiv C_2$ | Man $\equiv$ Human $\sqcap$ Male |
| subPropertyOf | $P_1 \sqsubseteq P_2$ | hasDaughter $\sqsubseteq$ hasChild |
| equivalentProperty | $P_1 \equiv P_2$ | cost $\equiv$ price |
| transitiveProperty | $P^+ \sqsubseteq P$ | ancestor$^+$ $\sqsubseteq$ ancestor |

| OWL Syntax | DL Syntax | Example |
|---|---|---|
| type | $a : C$ | John : Happy-Father |
| property | $\langle a, b \rangle : R$ | $\langle$John, Mary$\rangle$ : has-child |

# Summing up DLs

- Represent world (and its semantics) in terms of concepts, roles, individuals

- Very expressive formal knowledge representation languages with well-defined inference services

- Allow to handle effectively incomplete knowledge and reason over it
  - e.g., $\text{Tall} \sqcup \text{Intelligent (Tom)}, \neg\text{Student(Jim)}$

- Reasoning amounts to constructing logical models, hence the added complexity
  - *Avoid using them to declare data structures and perform algorithmic computations*

# DLs in multimedia analysis related tasks

- DLs (ontologies) have been used for
  - Representation of domain specific knowledge for annotation purposes
  - Representation of media specific knowledge including low-level features and content structure/decomposition (so called multimedia ontologies)
  - Representation of domain-specific knowledge and linking with perceptual features in order to support the extraction of content semantics

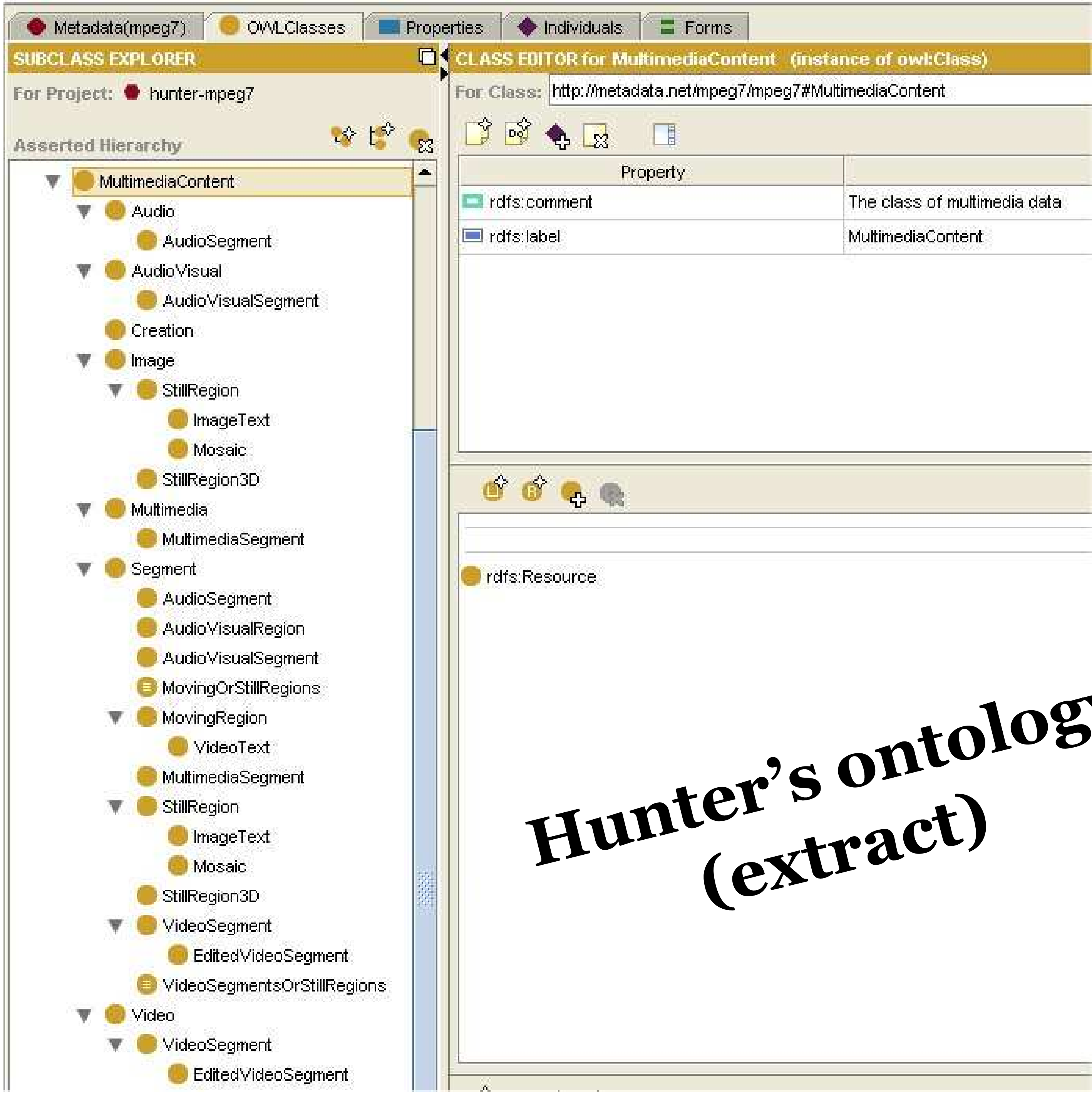
# Multimedia Ontologies (in a nutshell)

- Based in their majority on MPEG-7, multimedia ontologies aim to
  - Attach formal semantics to the XML-based Schema MPEG-7 definitions
  - Make explicit the normative specifications
  - Alleviate the ambiguities resulting from descriptions with multiple meanings

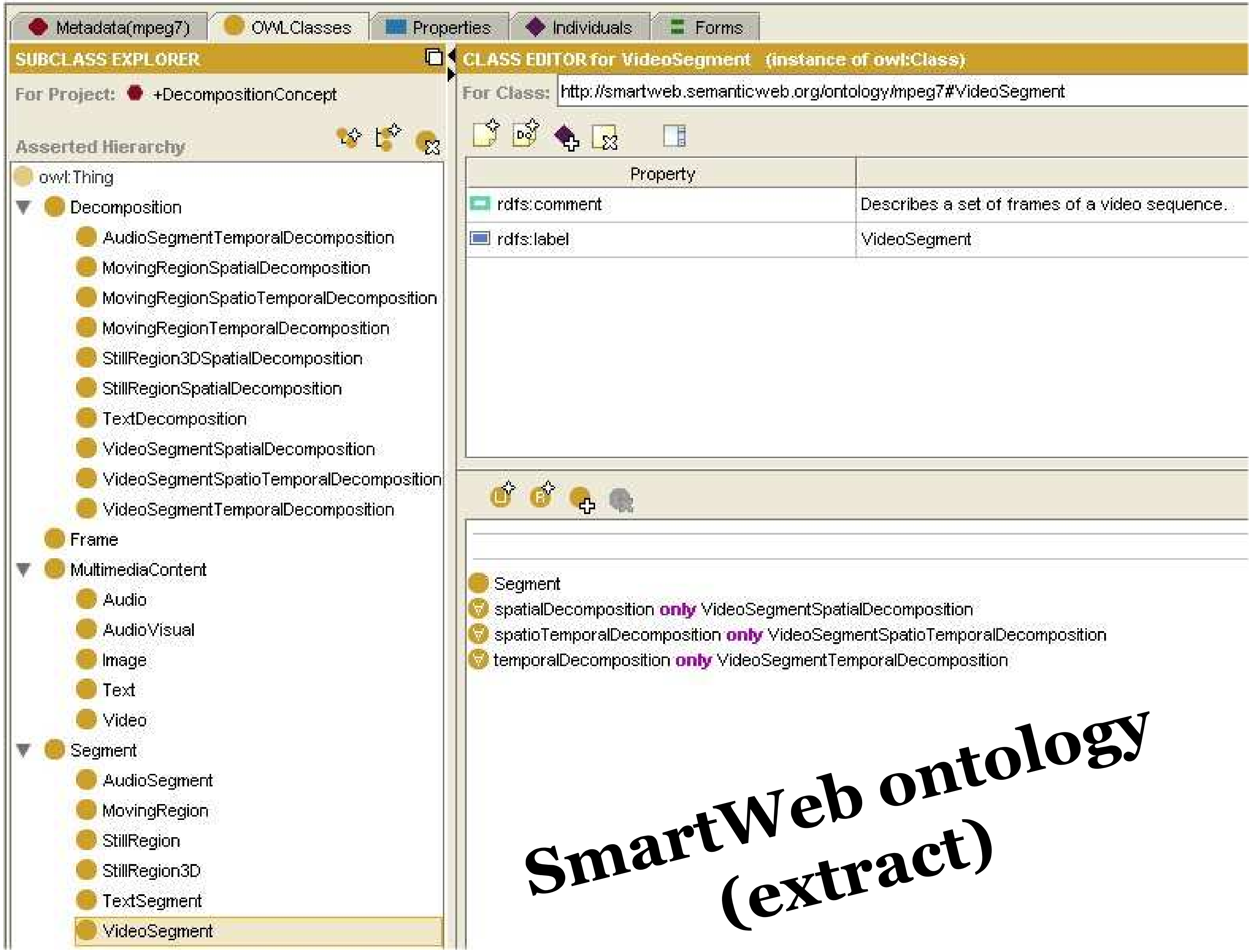
- Analysis, annotation, search, retrieval, presentation…

# Multimedia Ontologies (cont.)

- Hunter's ontology developed within the Harmony project (chronologically first initiative, 2001)
- The Multimedia Structure Ontology & Visual Descriptor Ontology, developed within aceMedia
- The Multimedia Content Ontology & Multimedia Descriptor Ontology, developed within BOEMIE
- SmartWeb, DS-MIRG, Rhizomik..
- The COMM (core) multimedia ontology (K-Space)

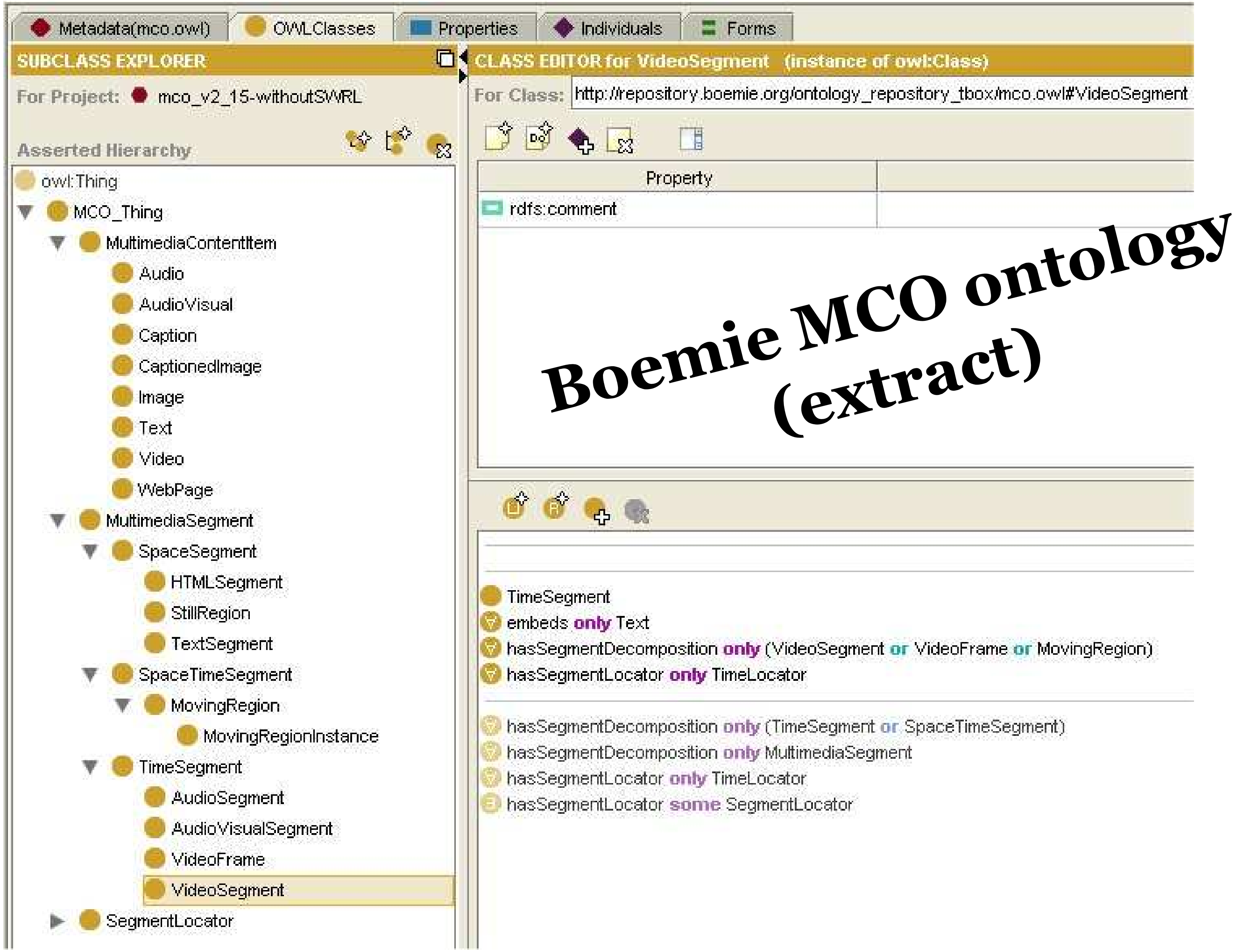



**Hunter's ontology (extract)**




**SmartWeb ontology (extract)**

# Boemie MCO ontology (extract)



# Multimedia Ontologies

- Differences in
  - Coverage
  - Modeling
  - Level of axiomatisation (semantics clarity)
  - Linking with domain ontologies

R.Troncy, O.Celma, S.Little, R.García, C.Tsinaraki, **MPEG-7 based Multimedia Ontologies: Interoperability Support of Interoperability Issue?**, in Workshop on MM Annotation and Retrieval enabled by Shared Ontologies (MAReSO'07), Genova, Italy, December 5, 2007.

S.Dasiopoulou. I.Kompatsiaris, M.G,Strintzis**, Enquiring MPEG-7 based Ontologies**, in Multimedia Tools and Appls., SI Data Semantics, 2009.

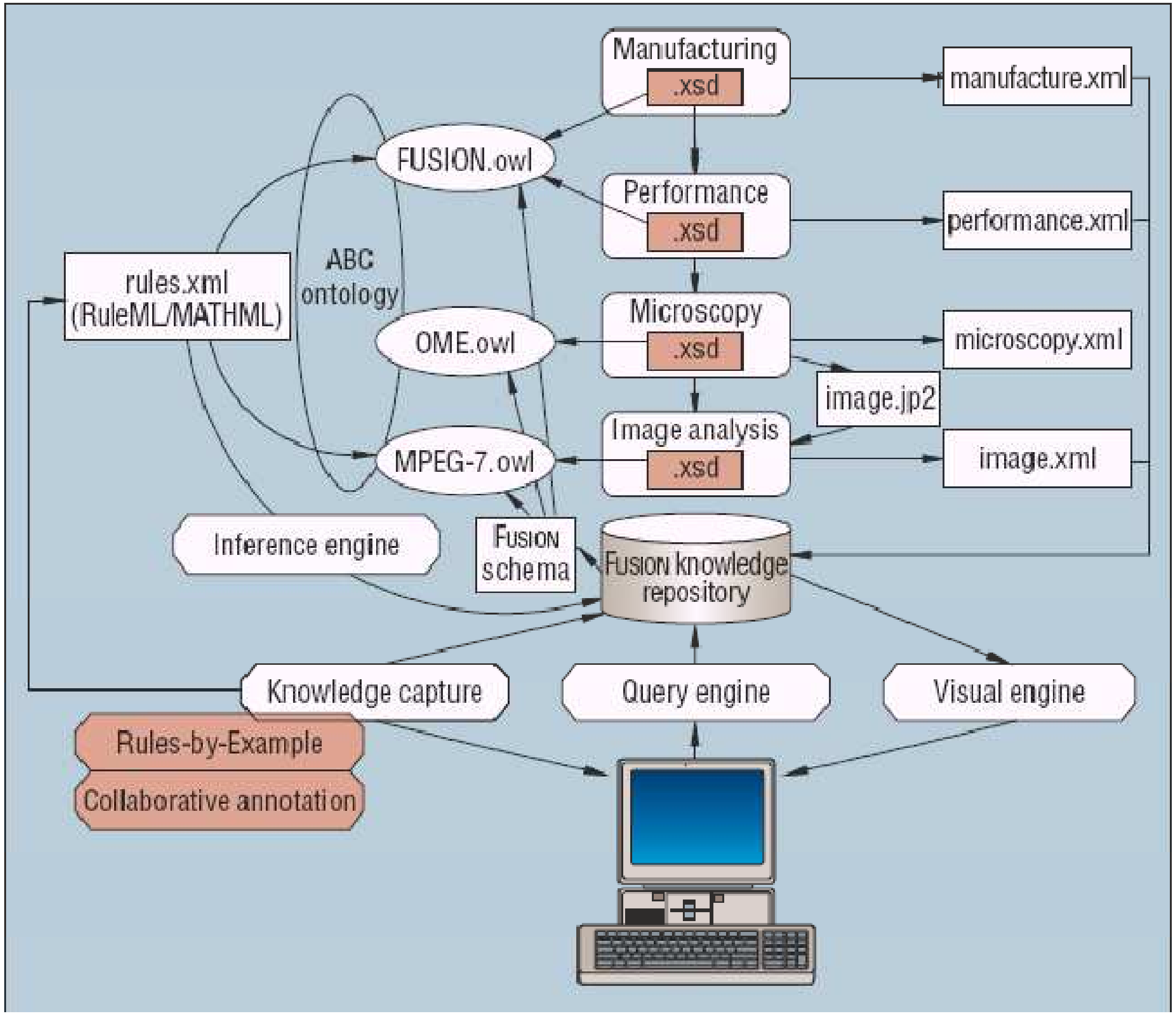
# Knowledge-Based Semantics Extraction

- Huge topic, vast literature (dates back to 1970s, AI..)
  - focus on recent DL related approaches only
  - present (some) representative examples

- TBox: background knowledge
  - defines valid (coherent) interpretation

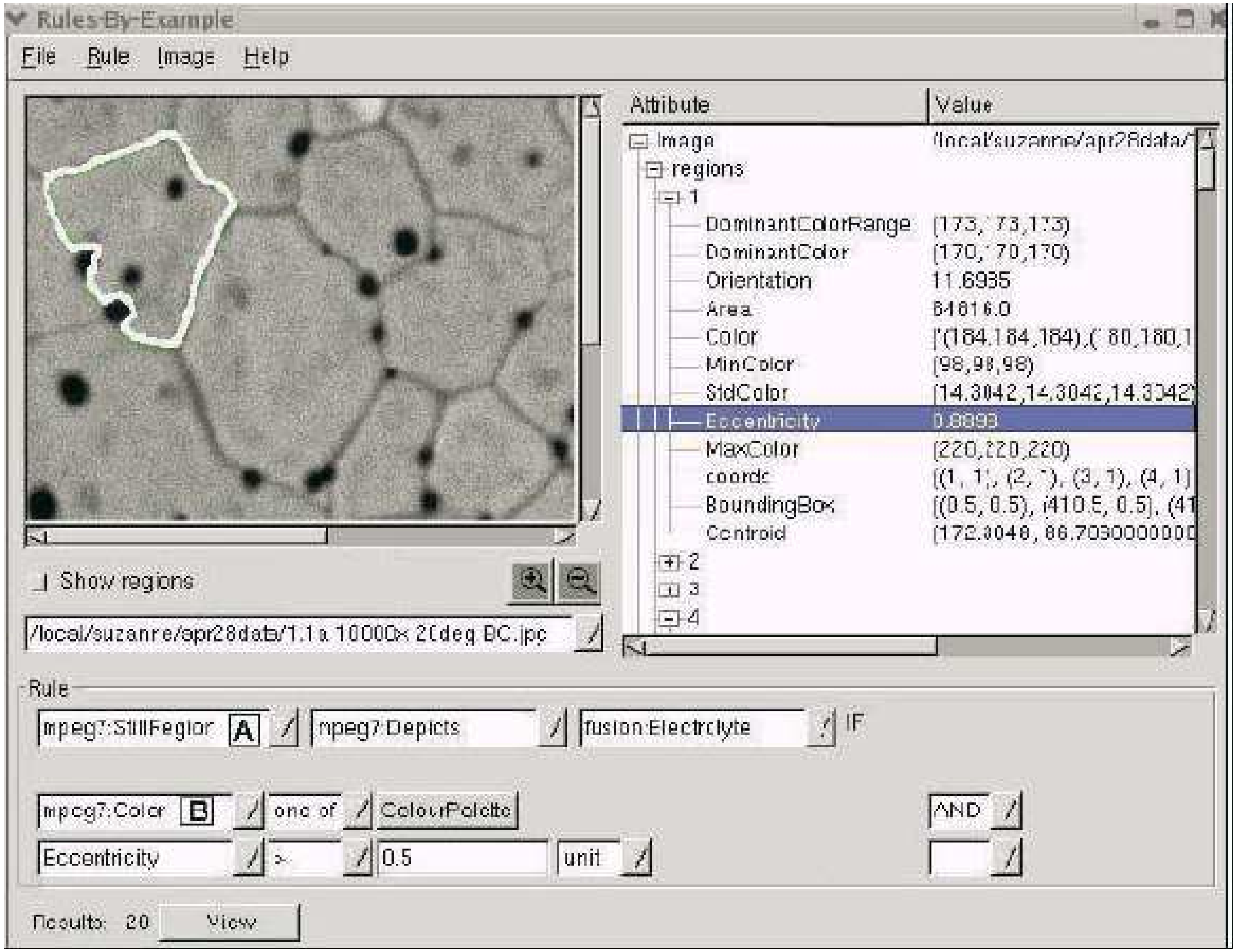
- ABox: extracted descriptions
  - analysis facts

# The FUSION project (*)

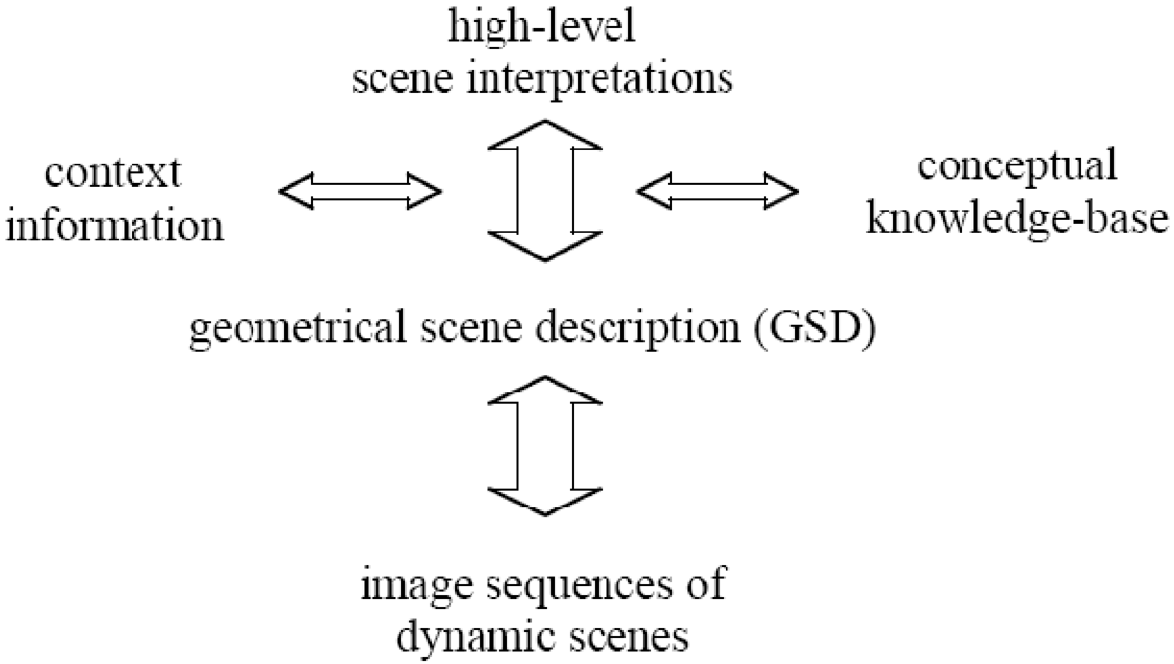



*J.Hunter,J.Drennan,S.Little, ***Realizing the Hydrogen Economy through Semantic Web Technologies***, in IEEE Intelligent Systemsm, S.I. on eScience, 2004.

# Domain knowledge definition and linking with low-level representations



# Scene Interpretation with DLs(*)

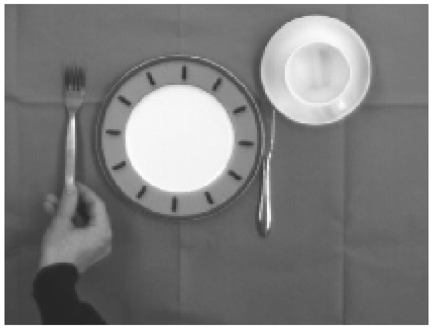



(*B.Neumann, R.Moller, On Scene Interpretation with Description Logics, in Cognitive Vision Systems, 2006.)

# Scene interpretation with DLs (cont.)

- Scene interpretation as model construction
- Logical aggregates capture complex objects/events

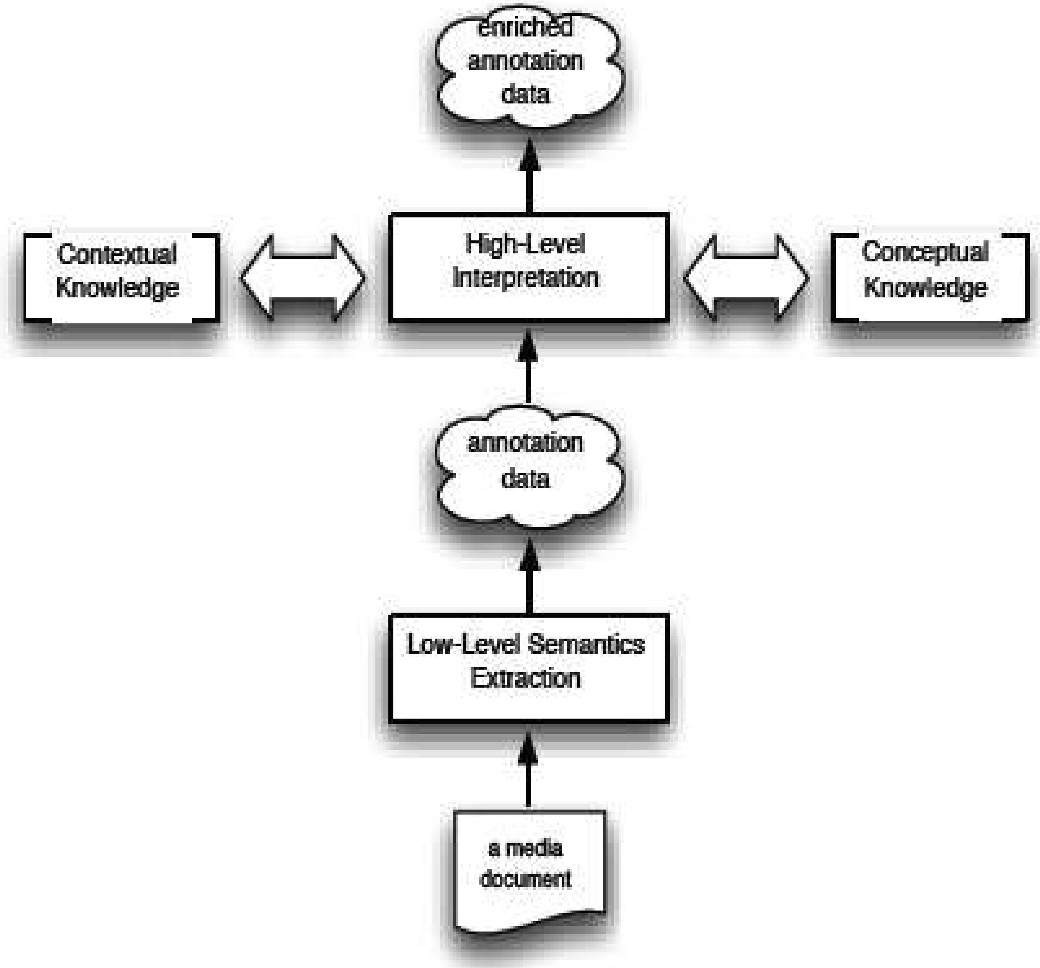
```
(equivalent cover
      (and configuration
            (exactly 1 cv-pl plate)
            (exactly 1 cv-sc (and saucer (some near plate)))
            (exactly 1 cv-cp (and cup (some on saucer)))
            (subset cv-pl (compose cv-sc near))
            (subset cv-sc (compose cv-cp on))))
```



# Scene interpretation with DLs (cont.)

- The available geometric descriptions are assumed to be
  - correct and non conflicting (which is not usually the case)
  - crisp (information loss)
  - complete (i.e. not missing)

# Abductive reasoning for multimedia interpretation(*)




(*S.Peraldi, A.Kaya, S.Meltzer, R.Moller & M.Wessel, **Towards a Media Interpretation Framework for the Semantic Web**,. In *Proc. of IEEE/WIC/ACM Conference on Web Intelligence (WI'07* ), 2007.)

# Abductive reasoning for multimedia interpretation (cont.)

- ABox abduction: given a background knowledge $\Sigma(T, A)$ and a set of observations $\Gamma$ derive a set of assertions $\Delta$ that **explain** $\Gamma$

  i.e. $\Sigma \sqcup \Delta \models \Gamma$

$\Gamma$ is divided into bona fiat (that need to be explained) and bona fide (that are taken to be true) assertions

# Abductive reasoning for multimedia interpretation (cont.)

$$
\begin{aligned}
Jumper &\sqsubseteq Human \\
Pole &\sqsubseteq Sports\_Equipment \\
Bar &\sqsubseteq Sports\_Equipment \\
Pole \sqcap Bar &\sqsubseteq \bot \\
Pole \sqcap Jumper &\sqsubseteq \bot \\
Jumper \sqcap Bar &\sqsubseteq \bot \\
Jumping\_Event &\sqsubseteq \exists_{\leq 1} hasParticipant.Jumper \\
Pole\_Vault &\sqsubseteq Jumping\_Event \sqcap \exists hasPart.Pole \sqcap \exists hasPart.Bar \\
High\_Jump &\sqsubseteq Jumping\_Event \sqcap \exists hasPart.Bar \\
near(Y,Z) &\leftarrow Pole\_Vault(X), hasPart(X,Y), Bar(Y), \\
&\quad\; hasPart(X,W), Pole(W), hasParticipant(X,Z), Jumper(Z) \\
near(Y,Z) &\leftarrow High\_Jump(X), hasPart(X,Y), Bar(Y), \\
&\quad\; hasParticipant(X,Z), Jumper(Z)
\end{aligned}
$$

$\Gamma \downarrow$

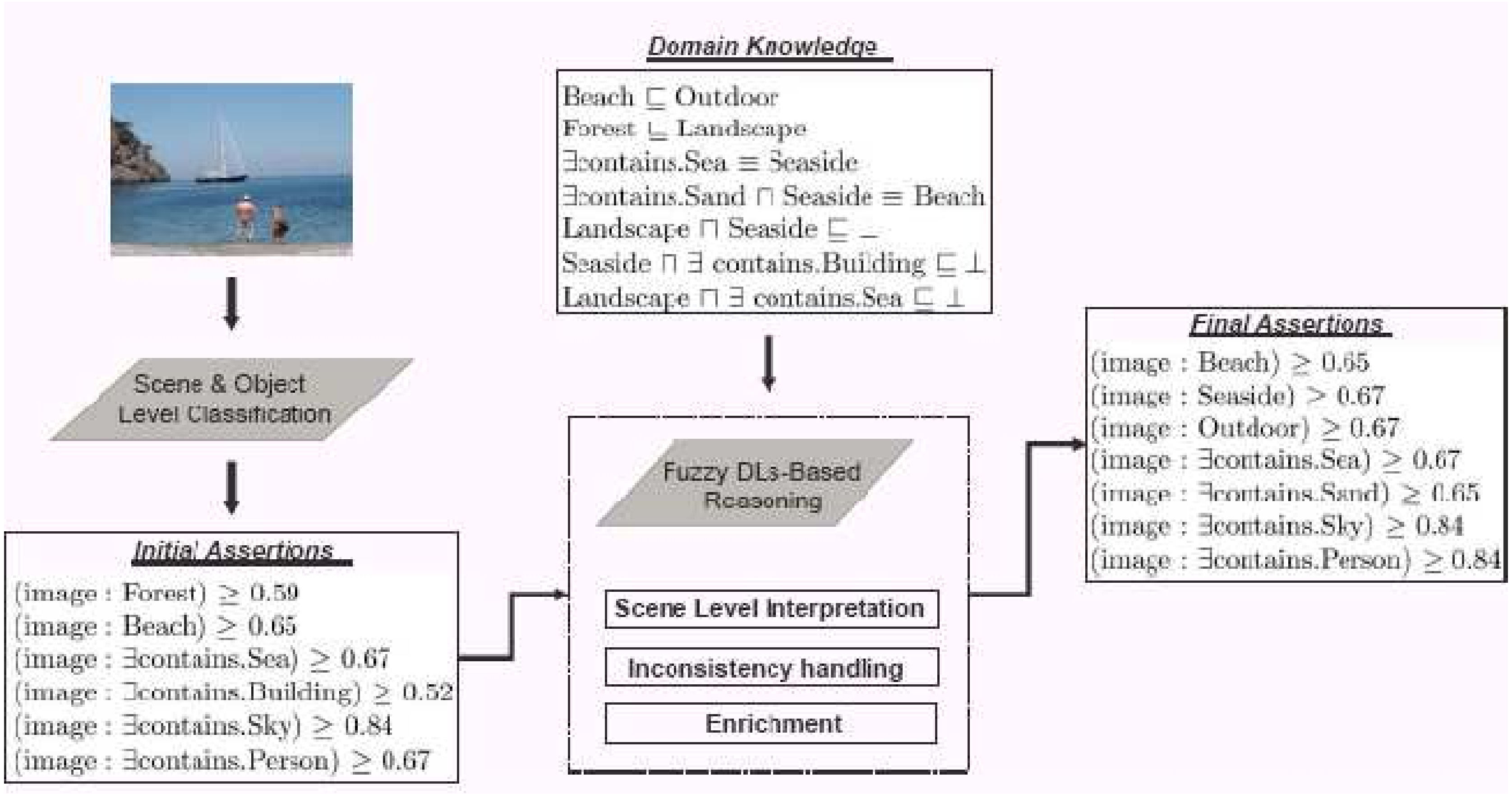
bona fiat

$$
\begin{aligned}
pole_1 &: Pole \\
human_1 &: Human \\
bar_1 &: Bar \\
(bar_1, human_1) &: near
\end{aligned}
$$

bona fide

- $\Delta_1 = \{new\_ind_1 : Pole\_Vault, (new\_ind_1, bar_1) : hasPart, (new\_ind_1, new\_ind_2) : hasPart, new\_ind_2 : Pole, (new\_ind_1, human_1) : hasParticipant, human_1 : Jumper\}$
- $\Delta_2 = \{new\_ind_1 : Pole\_Vault, (new\_ind_1, bar_1) : hasPart, (new\_ind_1, pole_1) : hasPart,$ $(new\_ind_1, human_1) : hasParticipant, human_1 : Jumper\}$
- $\Delta_3 = \{new\_ind_1 : High\_Jump, (new\_ind_1, bar_1) : hasPart, (new\_ind_1, human_1) : hasParticipant,$ $human_1 : Jumper\}$

# Enhancing Image Semantics Extraction using fuzzy DLs (*)




(*S.Dasiopoulou,I.Kompatisiaris, M.G.Strintzis, **Investigating fuzzy DLs-based Reasoning in Semantic Image Analysis**, in *Multimedia Tools and Apps.*, 2009.)

# Uncertainty Issues

- Machine learning provides now generic methodologies for supporting more than 100 concepts
  - captures conveniently complex associations between perceptual features and semantics
  - successful application examples, yet variable general performance

- **Semantics goes beyond perceptual manifestations**
  - *possibly contradictory (Mountain, Sand and Indoor)*
  - *possibly overlapping / complementary (Beach and Sea)*
  - *of restricted abstraction w.r.t. semantic expressiveness (face inside sea vs Swimmer)*

- Learning-based extracted annotations need to be ***semantically interpreted*** into a ***consistent*** description

# *Semantics goes beyond perceptual manifestations*

Discrepancy between semantic expressiveness

Discrepancy between intended and learned semantics

| | Best Possible | |
|---|---|---|
| **Search Topic** | *Best Detector* | *AP* |
| Two visible tennis players on the court | Athlete | 0.6501 |
| A goal being made in a soccer match | Stadium | 0.3429 |
| Basketball players on the court | Indoor Sports Venue | 0.2801 |
| A meeting with a large table and people | Furniture | 0.1045 |
| People with banners or signs | People Marching | 0.1013 |
| One or more military vehicles | Armored Vehicles | 0.0892 |
| Helicopter in flight | Helicopters | 0.0791 |
| A road with one or more cars | Car | 0.0728 |
| An airplane taking off | Classroom | 0.0526 |
| A tall building | Office Building | 0.0469 |
| A ship or boat | Cloud | 0.0427 |
| George Bush entering or leaving vehicle | Rocket Propelled Grenades | 0.0365 |
| Omar Karami | Chair | 0.0277 |
| Graphic map of Iraq, Baghdad marked | Graphical Map | 0.0269 |
| Condoleeza Rice | US National Flag | 0.0237 |
| One or more palm trees | Weapons | 0.0225 |

***Snoek et al., "Adding Semantics to Detectors for Video Retrieval", IEEE Multimedia, 2007***

# *Semantics goes beyond perceptual manifestations*

$(\text{image}:\exists\text{contains.Sand}) \geq 0.75$
$(\text{image}:\exists\text{contains.Sky}) \geq 0.87$
$(\text{image}:\exists\text{contains.Foliage}) \geq 0.76$
$(\text{image}:\exists\text{contains.Conifers}) \geq 0.88$
$(\text{image}:\text{Landscape}) \geq 0.92$

- Conifers detector semantics pertain to mountainous scenes
- Sand detector semantics pertains to beach scenes

$(\text{image}:\exists\text{contains.Sand}) \geq 0.75$
$(\text{image}:\exists\text{contains.Sea}) \geq 0.81$
$(\text{image}:\exists\text{contains.Person} \geq 0.67$
$(\text{image}:\exists\text{contains.Foliage}) \geq 0.76$
$(\text{image}:\exists\text{contains.Grass}) \geq 0.58$
$(\text{image}:\text{Beach}) \geq 0.85$
$(\text{image}:\text{Beach}) \geq 0.67$

- Sea and Sand detectors entail Beach scene
- Beach scenes entails both Natural and Outdoor scenes

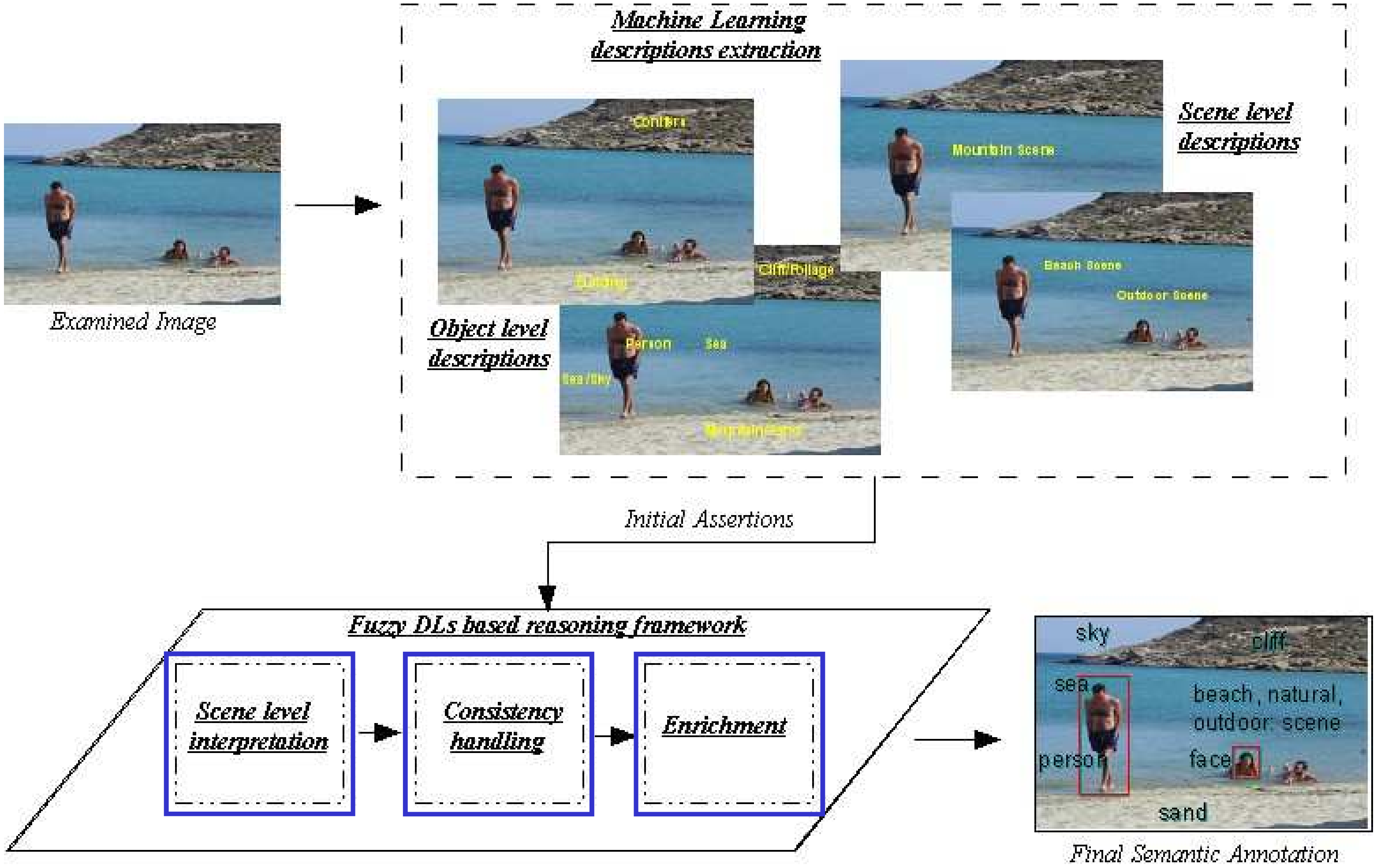
# Fuzzy DLs based approach

- **Goal:** enhance the robustness and completeness of learning-based extracted annotations

- **How:** semantics utilization
  - to interpret initial annotations
    - semantic integration
    - to detect and resolve inconsistencies
  - to enrich by means of entailment

- **Methodology:** fuzzy DL based reasoning
  - crisp TBox to conceptualize the domain semantics
  - fuzzy ABox to capture the uncertainty of initial annotations

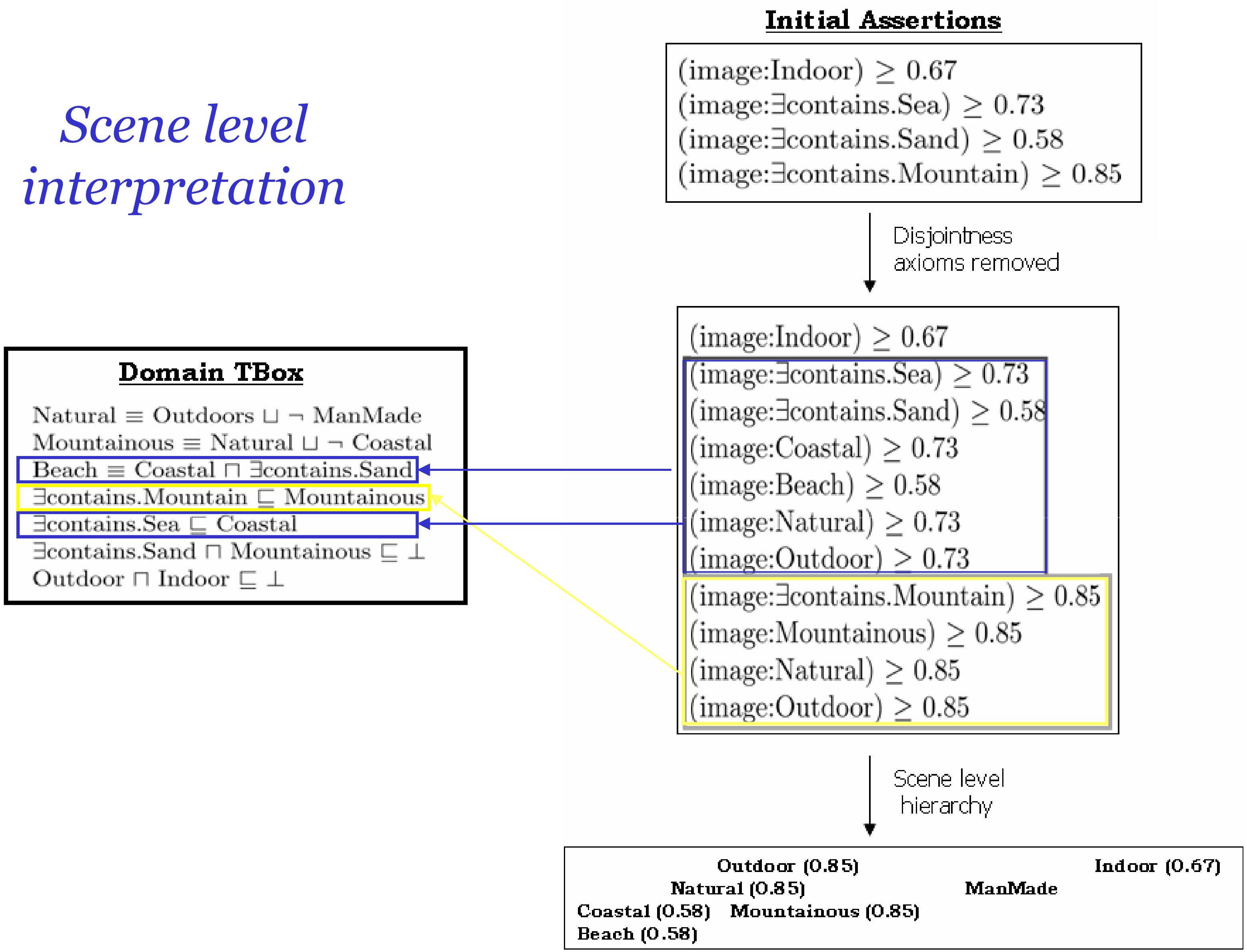
# General Framework



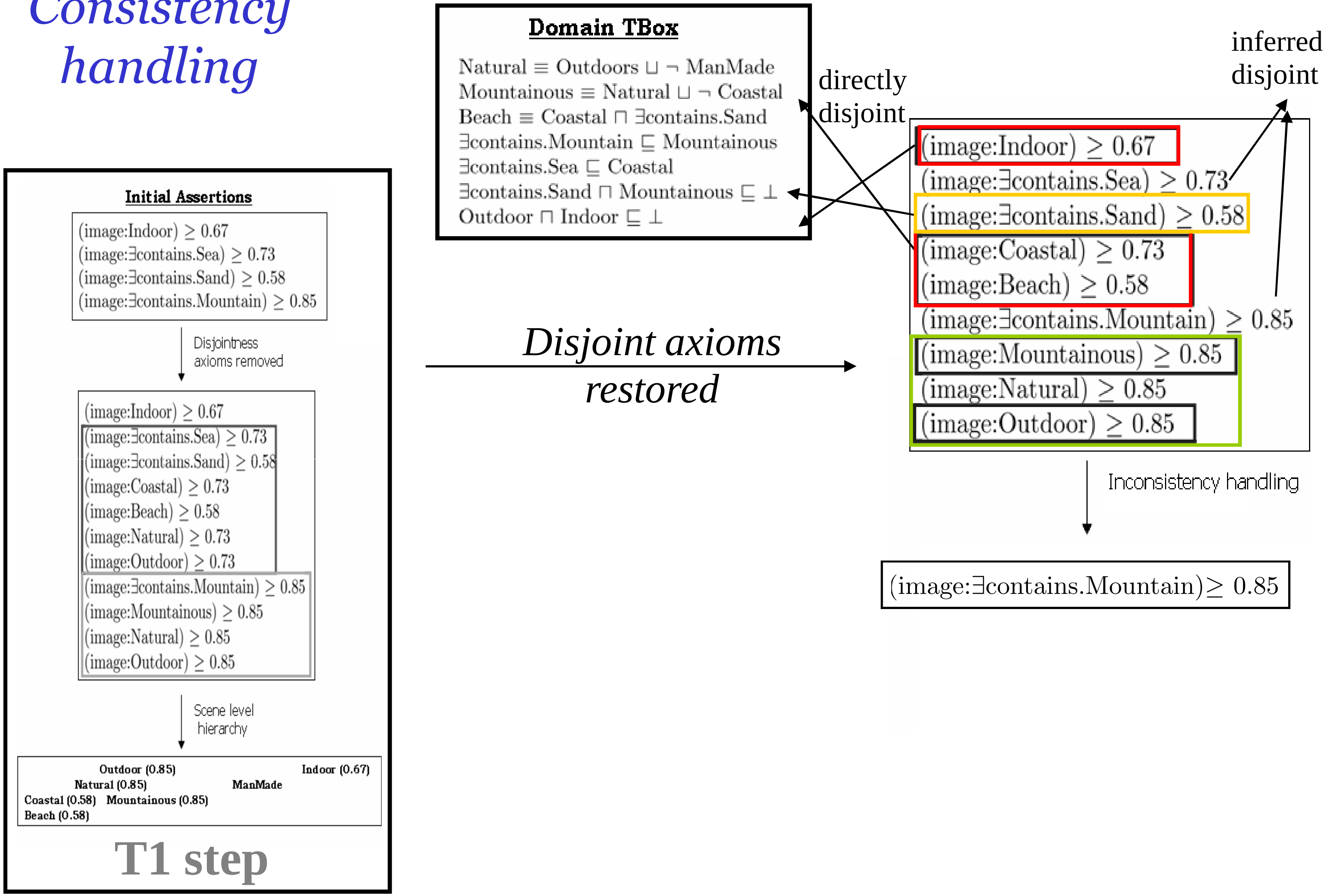
# Outdoor images TBox extract

$\text{Countryside\_buildings} \sqsubseteq \exists \text{contains.Buildings} \sqcap \exists \text{contains.Foliage}$
$\text{Countryside\_buildings} \sqsubseteq \text{Landscape}$
$\exists \text{contains.Forest} \sqcup \exists \text{contains.Grass} \sqcup \exists \text{contains.Tree} \sqsubseteq \exists \text{contains.Foliage}$
$\text{Rockyside} \sqsubseteq \exists \text{contains.Cliff}$
$\text{Rockyside} \sqsubseteq \exists \text{contains.Mountainous}$
$\text{Roadside} \sqsubseteq \exists \text{contains.Road}$
$\text{Roadside} \sqsubseteq \text{Landscape}$
$\exists \text{contains.Sea} \equiv \text{Coastal}$
$\text{Coastal} \sqsubseteq \text{Natural}$
$\exists \text{contains.Forest} \sqsubseteq \text{Landscape}$
$\text{Beach} \equiv \text{Coastal} \sqcap \exists \text{contains.Sand}$
$\text{Beach} \sqsubseteq \text{Natural}$
$\text{Cityscape} \sqsubseteq \text{ManMade}$
$\exists \text{contains.Sky} \sqsubseteq \text{Outdoor}$
$\exists \text{contains.Trunk} \sqsubseteq \exists \text{contains.Tree}$
$\text{Mountainous} \sqcap \text{Coastal} \sqsubseteq \bot$
$\text{Natural} \sqcap \text{ManMade} \sqsubseteq \bot$

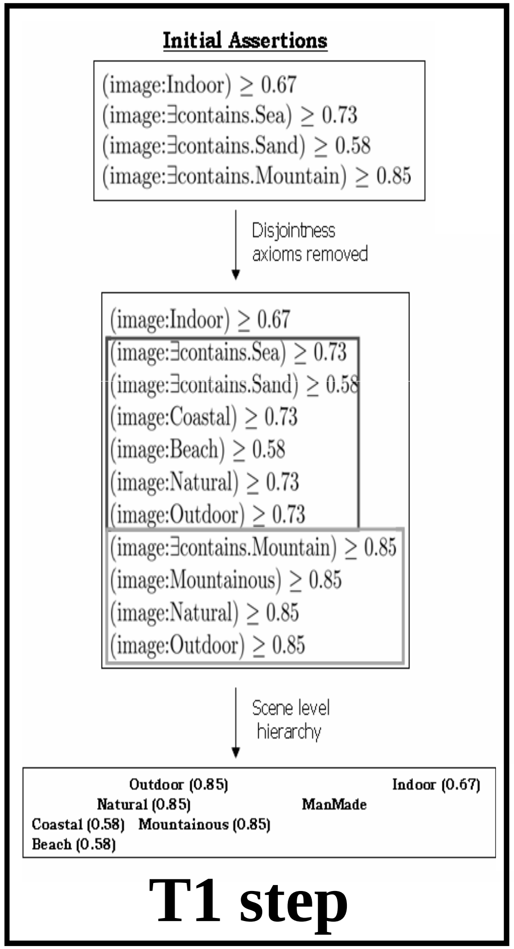
# *Scene level interpretation*

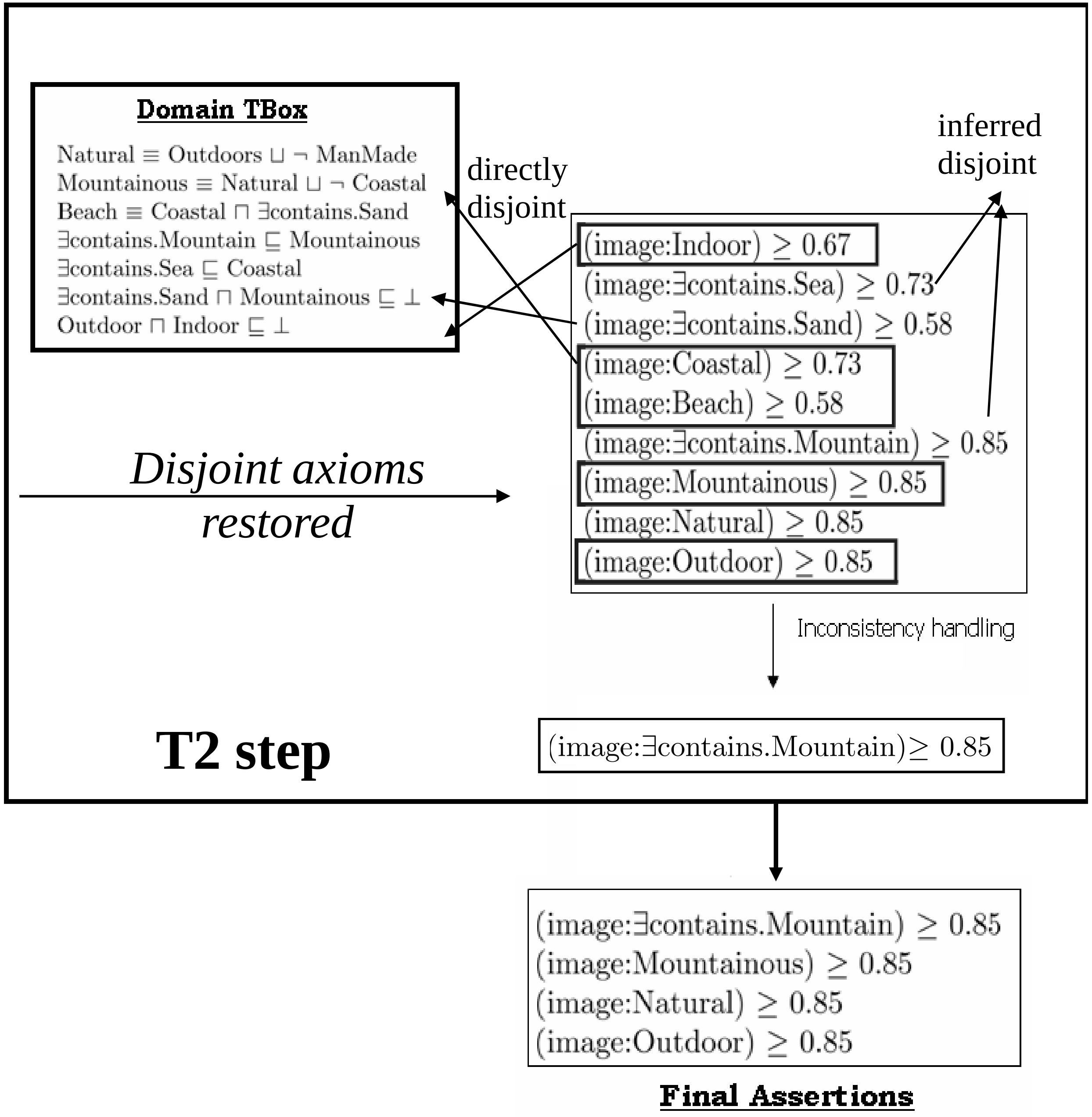


# *Consistency handling*



# *Enrichment*


Initial Assertions
(image:Indoor) ≥ 0.67
(image:∃contains.Sea) ≥ 0.73
(image:∃contains.Sand) ≥ 0.58
(image:∃contains.Mountain) ≥ 0.85
Disjointness axioms removed
(image:Indoor) ≥ 0.67
(image:∃contains.Sea) ≥ 0.73
(image:∃contains.Sand) ≥ 0.58
(image:Coastal) ≥ 0.73
(image:Beach) ≥ 0.58
(image:Natural) ≥ 0.73
(image:Outdoor) ≥ 0.73
(image:∃contains.Mountain) ≥ 0.85
(image:Mountainous) ≥ 0.85
(image:Natural) ≥ 0.85
(image:Outdoor) ≥ 0.85
Scene level hierarchy
Outdoor (0.85)
Indoor (0.67)
Natural (0.85)
ManMade
Coastal (0.58)
Mountainous (0.85)
Beach (0.58)
T1 step


Domain TBox
Natural ≡ Outdoors ⊔ ¬ ManMade
Mountainous ≡ Natural ⊔ ¬ Coastal
Beach ≡ Coastal ⊓ ∃contains.Sand
∃contains.Mountain ⊑ Mountainous
∃contains.Sea ⊑ Coastal
∃contains.Sand ⊓ Mountainous ⊑ ⊥
Outdoor ⊓ Indoor ⊑ ⊥
directly disjoint
inferred disjoint
(image:Indoor) ≥ 0.67
(image:∃contains.Sea) ≥ 0.73
(image:∃contains.Sand) ≥ 0.58
(image:Coastal) ≥ 0.73
(image:Beach) ≥ 0.58
(image:∃contains.Mountain) ≥ 0.85
(image:Mountainous) ≥ 0.85
(image:Natural) ≥ 0.85
(image:Outdoor) ≥ 0.85
Disjoint axioms restored
Inconsistency handling
(image:∃contains.Mountain)≥ 0.85
T2 step
(image:∃contains.Mountain) ≥ 0.85
(image:Mountainous) ≥ 0.85
(image:Natural) ≥ 0.85
(image:Outdoor) ≥ 0.85
Final Assertions

# Conclusions

- The use of explicit semantics is integral in multimedia semantics extractions; yet not the only necessary component
- Handling uncertainty is a critical factor
  - formal handling of annotations uncertainty semantics
  - utilization of domain semantics
  - consistent interpretations / descriptions
- Largely misestimated degrees/analysis descriptions can mislead the interpretation

# Future Directions

- Investigation of additional knowledge
  - probabilistic information in the form of co-occurrence patterns
  - spatial relations among object level concepts (aligning different segmentation masks)
- Investigation of intermediate representation level
  - link domain definitions with qualitative visual features
    - inconsistent at domain level interpretations are not simply rejected
- Experimentation with descriptions coming from other than image analysis sources
  - text, tags (expressed in ontological terms)
    - provenance-based weights

*Thank you for listening*

Any Questions ??

# Additional References


- F.Baader, D.Calvanese, D.L.McGuinness, D.Nardi,P.F.Patel-Schneider,The Description Logic Handbook: Theory, Implementation, and Applications, Cambridge University Press 2003.
- I.Horrocks, Hybrid Logics and Ontology Languages, Keynote at Inter. Workshop on Hybrid Logics, Seattle, USA, Aug. 2006.
- I.Horrocks, OWL: A Description Logic Based Ontology Language, Seminar at the Centre for Intelligent Systems and their Applications, Uni. Of Edinburgh, Scotland, March, 2006.
- J.Hunter, Adding Multimedia to the Semantic Web, Building an MPEG-7 Ontology, 1st Semantic Web Working Symposium, Stanford University, California, USA, 2001.
- C.Tsinaraki, P.Polydoros,S.Christodoulakis, Integration of OWL ontologies in MPEG-7 and TV Anytime Compliant Indexing, 16th Inter. Conf. on Advanced Information Systems Engineering (CAiSE), Riga, Latvia, 2004.
- R.Garcia, O.Celma, Semantic Integration and Retrieval of Multimedia Metadata, in Inter. Semantic Web Conference (ISWC), Galway, Ireland, 2005.
- R.Arndt, R.TRoncy, S.Staab, L.Hardman, M.Vacura, COMM: Designing a Well-Founded Multimedia Ontology for the Web, in Inter. Semantic Web Conference (ISWC), Busan,Korea, 2007.