Interprétation d’images
Apports des ontologies et des logiques de description

Jamal Atif - Isabelle Bloch - Céline Hudelot

January 30, 2017
Semantic image interpretation and annotation

Questions

What is the semantic content of these images? What do they represent?
Semantic image interpretation and annotation

Increasing structural complexity

Single label

Dog

Multiple labels

Dog, tree, leaf

Localization

An happy shaggy airdale poses in the autumn forest

Description

Source: T Berg
Semantic image interpretation and annotation

A hard problem for machines in spite of the increasing performance of sensors and the computing capacities.

Issues  [Smeulders 00, Snoek 10]

- Sensory gap.
- Semantic gap.
- Scaling gap: balance between expressivity/complexity and scaling of models.
Image and semantics

Semantic image interpretation and annotation

Sensory gap

Image = projection of a reality, often in 3D and continuous, into a discrete and 2D representation.

Numerous advances [Lowe 04, Dalal 05]
**Semantic image interpretation and annotation**

Scale gap

Semantic image interpretation and annotation

Semantic gap

Definition

Lack of coincidence between the information that one can extract from the visual data and the interpretation of these data by a user in a given situation [Smeulders 00]. Known as symbol grounding [Harnad 99] in AI and robotics.
Image and semantics

What is the semantics of this image?

- A white object on a green background.
- An insect.
- A white fly on a rose leaf.

- Image semantics is not inside the image.
- Image interpretation depends on a priori knowledge.
- Image interpretation depends on the user objectives.
- Importance of contextual and structural information.
Image and semantics
A multi-level paradigm

Since the early years of CV

D. Marr hierarchy [Marr 82]

Semantic pyramid [Jaimes 00]

Niveau de la scène
Générique : Paysage de montagne, rallye
Spécifique : Chypre
Abstrait : Sport, Divertissement

Niveau de l'objet
Générique : voiture, voiture de rallye
Spécifique : citroen de Sebastien Loeb
Image and semantics

- a) Type/Technique
  - Complex color graphic
  - Color Photograph

- b) Global Distribution
  - Similar texture, color

- c) Local Structure
  - Dark spots in x-ray
  - Lines in microscopic image

- d) Global Composition
  - Centered object, diagonal leading line

- e) Generic Objects
  - Car
  - Woman

- f) Generic Scene
  - Outdoors countryside
  - Outdoors, beach

- g) Specific Objects
  - Alex, Player No. 9
  - Chrysler building

- h) Specific Scenes
  - Washington D.C.
  - Paris

- i) Abstract Objects
  - Law
  - Arts

- j) Abstract Scene
  - Agreement, Business
  - Industry, Pollution
Image and semantics

A multi-level paradigm

Even in the recent representation learning with deep learning approaches.
Several semantics acceptations: from object semantics to structural description semantics.

Car: present
Cow: present
Bike: not present
Horse: not present
...

This is a photograph of one person and one brown sofa and one dog. The person is against the brown sofa. And the dog is near the person, and beside the brown sofa.

[Yao 10, Kulkarni 11, Farhadi 10, Farhadi 13, Karpathy 14]
Image and semantics

Importance of contextual and spatial information

Source: [Parikh 12]

Source: [Galleguillos 10]
Importance of spatial relations in image interpretation

- Spatial reasoning
- Carry an important structural information
- More stable and reliable than object features
Image and semantics
Importance of prior knowledge

Semantics = a property that *emerges* from the interaction between data and knowledge [Hanson 78, Santini 01, Hudelot 03].

⇒ Interest of ontologies
Outline

1. Image and semantics
2. What is an ontology?
3. Ontologies for image understanding: overview
4. Description Logics
5. Description Logics for image understanding
6. Conclusion
What is an ontology?

Example from F. Gandon, WIMMICS Team, INRIA

What is the last document that you have read?

Documents

Your answer is based on a shared ontology

I can understand

You can reason

Document

Book

Novel

Short story
Ontologies: Definition

Ontology

ethymology: ontos (being, that which is) + logos (science, study, theory)

- Philosophy
  - Study of the nature of being, becoming and reality.
  - Study of the basic categories of being and their relations.

- Computer Science
  - Formal representation of a domain of discourse.
  - Explicit specification of a conceptualization [Gruber 95].

Ref: [Guarino 09]
What is an ontology?

Ontologies: Definition

ontology

Formal, explicit (and shared) specification of a conceptualization [Gruber 95, Studer 98]

- Formal, explicit specification:
  - a formal language is used to refer to the elements of the conceptualization, e.g. description logics
- Conceptualization:
  - Objects, concepts and other entities and their relationships

Concept

Denoted by:
- a name
- a meaning (intensional definition)
- a set of denoted objects (extensional definition)

Relation

Denoted by:
- a name
- an intension
- an extension
The different types of ontologies

According to their expressivity

Source: [Uschold 04]
The different types of ontologies

According to their abstraction level

- **Top (or Upper)-level ontology**: very general concepts that are the same across all knowledge domains [Wikipedia] (e.g. DOLCE).
- **Core ontology**: minimal set of concepts and relations used to structure and describe a given domain (e.g. Dublin Core).
- **Domain ontology**: concepts and relations of a specific domain (e.g. FMA).
Outline

1. Image and semantics
2. What is an ontology?
3. Ontologies for image understanding: overview
4. Description Logics
5. Description Logics for image understanding
6. Conclusion
Ontologies for image interpretation

A growing interest since 2001

Various objectives:

- Providing an **unified vocabulary** for the description and annotation of image content.
  - e.g. MPEG-7 ontologies.

- **Structuring** the vocabulary and the database for **large-scale** image problems.
  - e.g. visual ontologies (LabelMe, ImageNet, Visipedia).

- Representing the application domain knowledge for **reasoning** and for **guiding** the interpretation process.
  - e.g. formal ontologies based on description logics.
Ontologies for an unified and standardized description of image content

- MPEG-7 ontologies: Boemie, AceMedia, Rhizomik... (see [Dasiopoulou 10b] for a recent review).
  Main motivation: interoperability between applications.
- LSCOM (Large Scale concept ontology for multimedia) [Naphade 06], MediaMill [Habibian 13].
  Main motivation: common vocabulary for video shot description.

Mainly focused on the descriptive part of ontologies.
Ontologies for image understanding: overview

Ontologies for structuring the vocabulary and the learning database (1/3)

Main motivation: image classification, annotation and retrieval at large scale [Liu 07, Deng 10].

- Ontologies based on lexical resources (e.g. Wordnet) populated with images:
  - ImageNet [Russakovsky 15], LabelMe [Russell 08], Visipedia [Belongie 16], Visual Genome [Krishna 16]...

Which concepts are closer?

Adequacy of the lexical resources for image interpretation problems?
Mainly lightweight ontologies (non-formal, without reasoning capabilities).
Ontologies for structuring the vocabulary and the learning database (2/3)

Main motivation: hierarchical image classification.

- Visual concept hierarchies inferred from image datasets: [Fei-Fei 05, Marszalek 08, Griffin 08, Sivic 08, Bart 08, Gao 11].

- Mainly hierarchies (no other semantic relations than *is-a*).
- Concepts without semantics (except the leaves).
- Mainly lightweight ontologies (non-formal, without reasoning capabilities).
Main motivation: image classification and annotation.

- Ontologies combining text and visual knowledge: [Li 10, Wu 12, Bannour 14, Krishna 16].

Image hierarchy [Li 10]  

VCNet [Wu 12]

- Dedicated knowledge models.
- Mainly lightweight ontologies (non-formal, without reasoning capabilities).
Main motivation: image captionning.
More and more approaches, under the dynamics of image captioning to represent objects, attributes of objects and relationships between objects: Scene graphs [?], Visual Genome [Krishna 16], Visipedia [Belongie 16]
Ontologies for image understanding: overview

Main motivation: image captioning.
More and more approach, under the dynamics of image captioning to represent objects, attributes of objects and relationships between objects: Scene graphs, Visual Genome [Krishna 16], Visipedia [Belongie 16]

Visual Genome

<table>
<thead>
<tr>
<th>Regions</th>
<th>Attributes</th>
<th>Relationships</th>
</tr>
</thead>
<tbody>
<tr>
<td>The sky is blue</td>
<td>sky is blue</td>
<td>chair ON sand</td>
</tr>
<tr>
<td>the ocean is blue</td>
<td>ocean is blue</td>
<td>person under an umbrella</td>
</tr>
<tr>
<td>7 umbrellas are pictured</td>
<td>umbrella is yellow</td>
<td>umbrella ON sand</td>
</tr>
<tr>
<td>the umbrellas are yellow</td>
<td>sand is brown</td>
<td>person standing on sand</td>
</tr>
<tr>
<td>the sand is brown</td>
<td>structure is open</td>
<td>person sitting on sand</td>
</tr>
<tr>
<td>the shade structure is open</td>
<td>chair is white</td>
<td>person sitting on a chair</td>
</tr>
<tr>
<td>white chairs are on the beach</td>
<td>tree is green</td>
<td></td>
</tr>
<tr>
<td>people are sitting under the umbrellas</td>
<td>structure is blue</td>
<td></td>
</tr>
</tbody>
</table>

Question Answers
- When was this picture taken?
  - During the day.
- Where are the umbrellas?
  - On the beach.
- Why are there blue tents on the beach?
  - To help protect people from the sun.
- How is the weather in the scene?
  - Sunny and warm.
- Why do people come to the beach?
  - To enjoy the sand, sun and ocean.
Image interpretation as an ontological driven inference approach

Main motivation: explicit and formal representation of domain and contextual knowledge used to reason and infer the interpretation.

- Annotation and interpretation refinement using basic DLs inference services: [Simou 08, Dasiopoulou 09, Dasiopoulou 10a, Bannour 14].

- Ontologies to narrow the semantic gap: [Town 06, Bagdanov 07, Hudelot 08]

- Image interpretation as a non-monotonic reasoning process:
  - Image interpretation as a default reasoning service [Möller 99a, Neumann 08].
  - Abductive reasoning for image interpretation [Peraldi 07, Möller 08, Atif 14, Donadello 14].

Often based on Description Logics (DLs).
Outline

1. Image and semantics
2. What is an ontology?
3. Ontologies for image understanding: overview
4. Description Logics
5. Description Logics for image understanding
6. Conclusion
**Descriptions logics**

- Family of logics for representing structured knowledge.
- Well understood semantics.
- Defined by a set of concepts and role forming operators.
- Compact and expressive and basis of OWL language to represent ontologies.

**A description logic system**

- **Tbox**
  - Defines the terminology of the application domain
- **Abox**
  - States facts about a specific world

**Knowledge Base**
Description logics: the description language
Syntax of $\mathcal{ALC}$: attributive language with complement

Basic language $\mathcal{ALC}$ + constructors ($C$ for the complement $\neg$ operator)

- Signature $\Sigma = (N_C, N_R)$, disjoint sets of concept names and role names respectively.
- Concept descriptions in $\mathcal{ALC}$ are formed according to the following syntax rule:

$$ C, D \rightarrow A \mid \text{(atomic concepts)} $$
$$ \top \mid \text{(universal concept)} $$
$$ \bot \mid \text{(bottom concept)} $$
$$ \neg C \mid \text{(negation)} $$
$$ C \sqcap D \mid \text{(conjunction)} $$
$$ C \sqcup D \mid \text{(disjunction)} $$
$$ \forall r.C \mid \text{(value restriction)} $$
$$ \exists r.C \mid \text{(existential restriction)} $$

$A \in N_C$ and $r \in N_R$
Description logics: the description language

Examples of $\mathcal{ALC}$-concept descriptions

- Atomic concepts: $Person$, $Female$, $Tutorial$, $Boring$
- Atomic role: $attends$
- $\mathcal{ALC}$-descriptions:

\[
\begin{align*}
\text{Person} &\sqcap \text{Female} \\
\text{Person} &\sqcap \neg \text{Female} \\
\text{Person} &\sqcap \exists \text{attends}.\text{Tutorial} \\
\text{Person} &\sqcap \forall \text{attends}.(\text{Tutorial} \sqcap \neg \text{Boring})
\end{align*}
\]
Description logics : the description language
Semantics of $\mathcal{ALC}$: attributive language with complement

An interpretation $\mathcal{I} = \langle \Delta^\mathcal{I}, \cdot^\mathcal{I} \rangle$

- $\Delta^\mathcal{I}$: a non-empty set, the domain of interpretation
- $\cdot^\mathcal{I}$: an interpretation function, which assigns to:
  - every atomic concept $A \in N_C$, a set $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$,
  - every atomic role $r \in N_R$, a binary relation $r^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$.

Extension to concept descriptions

$$\top^\mathcal{I} = \Delta^\mathcal{I}$$
$$\bot^\mathcal{I} = \emptyset$$
$$(\neg C)^\mathcal{I} = \Delta^\mathcal{I} \setminus C^\mathcal{I}$$
$$(C \cap D)^\mathcal{I} = C^\mathcal{I} \cap D^\mathcal{I}$$
$$(C \cup D)^\mathcal{I} = C^\mathcal{I} \cup D^\mathcal{I}$$
$$(\forall r.C)^\mathcal{I} = \{ a \in \Delta^\mathcal{I} | \forall b.(a, b) \in r^\mathcal{I} \rightarrow b \in C^\mathcal{I} \}$$
$$(\exists r.C)^\mathcal{I} = \{ a \in \Delta^\mathcal{I} | \exists b.(a, b) \in r^\mathcal{I} \land b \in C^\mathcal{I} \}$$
The basic description language $\mathcal{AL}$

Semantics

Equivalence:

\[ C \equiv D \text{ if } C^\mathcal{I} = D^\mathcal{I} \text{ for all interpretations } \mathcal{I} \]

Example

\[ \forall \text{hasChild}. \text{Female} \sqcap \forall \text{hasChild}. \text{Student} \text{ and } \forall \text{hasChild}. (\text{Female} \sqcap \text{Student}) \text{ are equivalent.} \]
The family of $AL$ languages

$AL[U][E][C][N][Q], \cdots$

Many additional constructors have been introduced.
The family of $\mathcal{AL}$ languages

$\mathcal{ALEN}$ example

$\text{Person} \sqcap (\leq 1 \text{hasChild} \sqcup (\geq 3 \text{hasChild} \sqcap \exists \text{hasChild}.\text{Female}))$
Description logics: terminological knowledge

Terminological axioms

- General Concept Inclusion (GCI)
  \[ C \sqsubseteq D \]
  where \( C, D \) are concept descriptions

- Concept definition\(^a\)
  \[ A \equiv C \]
  where \( A \) is a concept name, and \( C \) is a concept description

\(^a\) abbreviation for \( A \sqsubseteq C \) and \( C \sqsubseteq A \)

TBox

A TBox is a finite set of GCIs
Description logics: terminological knowledge

- An interpretation $\mathcal{I}$ satisfies a GCI $C \sqsubseteq D$ iff $C^\mathcal{I} \subseteq D^\mathcal{I}$
  \[ \mathcal{I} \models (C \sqsubseteq D) \iff C^\mathcal{I} \subseteq D^\mathcal{I} \]

- An interpretation $\mathcal{I}$ satisfies an equality $C \equiv D$ if $C^\mathcal{I} \equiv D^\mathcal{I}$
  \[ \mathcal{I} \models (C \equiv D) \iff C^\mathcal{I} \equiv D^\mathcal{I} \]

- The interpretation $\mathcal{I}$ is a model of a TBox $\mathcal{T}$ iff it satisfies all the GCIs in $\mathcal{T}$

- Two TBoxes are equivalent if they have the same model.
Description logics: terminological knowledge

TBox example

\[
\text{Woman} \equiv \text{Person} \sqcap \text{Female} \\
\text{Man} \equiv \text{Person} \sqcap \neg \text{Woman} \\
\text{Mother} \equiv \text{Woman} \sqcap \exists \text{hasChild}. \text{Person} \\
\text{Father} \equiv \text{Man} \sqcap \exists \text{hasChild}. \text{Person} \\
\text{Parent} \equiv \text{Father} \sqcup \text{Mother} \\
\text{Grandmother} \equiv \text{Mother} \sqcap \exists \text{hasChild}. \text{Parent} \\
\text{MotherWithManyChildren} \equiv \text{Mother} \sqcap \geq 3 \text{hasChild}
\]
Description logics: assertional knowledge

**Assertional axioms**

- Concept assertion: $C(a)$
- Role assertion: $R(a, b)$
  
  $C$ a concept description, $a, b$ are individuals names from a set $N_I$

**ABox**

An ABox is a finite set of assertions

**Interpretation**

- Given $\mathcal{I}$, each individual $a$ is mapped to an element $a^\mathcal{I} \in \Delta^\mathcal{I}$
- Unique name assumption: $a^\mathcal{I} \neq b^\mathcal{I}$
- $\mathcal{I}$ is a model of the ABox $\mathcal{A}$ if it satisfies all its assertions:
  - $a^\mathcal{I} \in C^\mathcal{I}$ for all $C(a) \in \mathcal{A}$
  - $(a^\mathcal{I}, b^\mathcal{I}) \in R$ if for all $R(a, b) \in \mathcal{A}$
Description logics : knowledge base

A knowledge base $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ consists of a TBox $\mathcal{T}$ and an ABox $\mathcal{A}$.

The interpretation $\mathcal{I}$ is a model of the knowledge base $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ iff it is a model of $\mathcal{T}$ and a model of $\mathcal{A}$. 
Description logics for knowledge representation

Example in the medical domain

Knowledge in brain imaging

- caudate nucleus: a deep gray nucleus of the telencephalon involved with control of voluntary movement
- the left caudate nucleus is inside the left hemisphere
- it is close to the lateral ventricle
- it is outside (left of) the left lateral ventricle

Excerpt of a corresponding TBox

- AnatomicalStructure ⊑ SpatialObject
- LV ⊑ AnatomicalStructure
- GN ⊑ AnatomicalStructure
- CN ⊑ GN
- LV ≡ RLV ⊔ LLV
- CN ≡ RCN ⊔ LCN
- LCN ≡ GN ⊓∃ closeTo.(LLV) ⊓∃ leftOf.(LLV)
- etc.
Description logics: exercice 1

On the blackboard
Students are Persons. Students are following some UEs. Curious students are following the UE Image Interpretation. Axel is a curious student. Persons are descendants of others persons.

1. What are the concepts, the relations, the instances?
2. Write the corresponding TBox with $ALCN$.
3. Write the corresponding ABox with $ALCN$. 
Description logics: exercice 2

On the blackboard

We consider the following interpretation

- $\Delta^\mathcal{I} = \{t_1, t_2, f_1, f_2, c_1, c_2, j, k, l, m, n\}$
- $\text{Person}^\mathcal{I} = \{j, k, l, m, n\}$
- $\text{Car}^\mathcal{I} = \{t_1, t_2, f_1, f_2, c_1, c_2\}$
- $\text{Ferrari}^\mathcal{I} = \{f_1, f_2\}$
- $\text{Toyota}^\mathcal{I} = \{t_1, t_2\}$
- $\text{likes}^\mathcal{I} = \{(j, f_1), (k, f_1), (k, t_2), (l, c_1), (l, c_2), (m, c_1), (m, t_2), (n, f_2), (n, c_2)\}$

Give the interpretations of:

- $\exists \text{likes. Ferrari} \sqcap \exists \text{likes. Toyota}$
- $\exists \text{likes. Ferrari} \sqcap \forall \text{likes. Toyota}$
- $\exists \text{likes. Ferrari} \sqcap \exists \text{likes.} \neg \text{Ferrari}$
- $\exists \text{likes. Cars} \sqcap \forall \text{likes.} \neg (\text{Toyota} \sqcup \text{Ferrari})$
Description logics: concrete domains

- A way to integrate *concrete and quantitative qualities* (integers, strings,...) of real world objects with conceptual knowledge [Baader,91].
- A pair \((\Delta_D, \Phi_D)\) where \(\Delta_D\) is a set and \(\Phi_D\) a set of predicates names on \(\Delta_D\). Each predicate name \(P\) is associated with an arity \(n\) and an \(n\)-ary predicate \(P^D \subseteq \Delta_D^n\).

Examples

- **Concrete domain \(N\):**
  - Domain: non negative integers.
  - Predicates: \(\leq\) (binary predicate) \(\leq n\) unary predicate.
  - \(\text{Person} \sqcap \exists \text{age.} \leq 20\) denotes a person whose age is less than 20.

- **Concrete domain \(AL\), Allen’s interval calculus:**
  - Domain: intervals.
  - Predicates: built from Allen’s basic interval relations.
Description logics: reasoning services

⇒ Infer implicit knowledge from explicitly one.
  - Terminological reasoning.
  - Assertional reasoning.
Description logics: reasoning services

Terminological reasoning

**Satisfiability**

C is satisfiable w.r.t. a TBox $\mathcal{T}$ iff $C^\mathcal{I} \neq \emptyset$ for some model $\mathcal{I}$ of $\mathcal{T}$.

**Subsumption**

C is subsumed by $D$ w.r.t. a TBox $\mathcal{T}$ ($C \subseteq_\mathcal{T} D$) iff $C^\mathcal{I} \subseteq D^\mathcal{I}$ for all models $\mathcal{I}$ of $\mathcal{T}$.

**Equivalence**

C is equivalent to $D$ w.r.t. a TBox $\mathcal{T}$ ($C \equiv_\mathcal{T} D$) iff $C^\mathcal{I} = D^\mathcal{I}$ for all models $\mathcal{I}$ of $\mathcal{T}$.

**Disjointness**

Two concepts $C$ and $D$ are disjoint with respect to $\mathcal{T}$ if $C^\mathcal{I} \cap D^\mathcal{I} = \emptyset$ for every model $\mathcal{I}$ of $\mathcal{T}$.
Reduction to subsumption

For concepts $C, D$ we have

- $C$ is unsatisfiable $\iff C$ is subsumed by $\bot$;
- $C$ and $D$ are equivalent $\iff C$ is subsumed by $D$ and $D$ is subsumed by $C$;
- $C$ and $D$ are disjoint $\iff C \sqcup D$ is subsumed by $\bot$.

The statements also hold with respect to a TBox.
Reduction to Unsatisfiability

For concepts $C$, $D$ we have

- $C$ is subsumed by $D$ $\iff$ $C \cap \neg D$ is unsatisfiable;
- $C$ and $D$ are equivalent $\iff$ both $C \cap \neg D$ and $\neg C \cap D$ are satisfiable;
- $C$ and $D$ are disjoint $\iff$ $C \cap D$ is unsatisfiable.

The statements also hold with respect to a TBox.
Reducing Unsatisfiability

Let $C$ be a concept. Then the following are equivalent:

- $C$ is unsatisfiable;
- $C$ is subsumed by $\bot$;
- $C$ and $\bot$ are equivalent;
- $C$ and $\bot$ are disjoint.

The statements also hold with respect to a TBox.
**Description logics: reasoning services**

**Assertional reasoning**

Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be an ontology.

**Consistency**

$\mathcal{A}$ is consistent with respect to a TBox $\mathcal{T}$, if there is an interpretation that is a model of both $\mathcal{A}$ and $\mathcal{T}$.

**Instance checking**

$a$ is an instance of $C$ w.r.t. $\mathcal{T}$ iff $a^\mathcal{I} \in C^\mathcal{I}$ for all models $\mathcal{I}$ of $\mathcal{T}$. We also write $\mathcal{A} \models C(a)$. The same holds for roles.

**Retrieval problem**

Given an ABox $\mathcal{A}$ and a concept $C$, find all individuals $a$ such that $\mathcal{A} \models C(a)$.

**Realization problem (dual to the retrieval problem)**

Given an individual $a$ and a set of concepts, find the most specific concepts (msc) $C$ from the set such that $\mathcal{A} \models C(a)$. The mscs are the concepts that are minimal with respect to the subsumption ordering $\sqsubseteq$. 
Reduction

\( A \models C(a) \) iff \( A \cup \{ \neg C(a) \} \) is inconsistent; 
\( C \) is satisfiable iff \( \{ C(a) \} \) is consistent.
Subsumption checking

- Structural subsumption
- Semantic tableaux
- etc.
Open world, Closed world

Closed World Assumption
Limitations to what is expressed

- example: ABox: hasChild(anne, paul)
- anne has only one child: paul

Open World Assumption: description logics
Open world: no limitations to what is expressed

- example: ABox: hasChild(anne, paul)
- Anne can have other child than paul
- ($\leq 1\text{hasChild})(anne)$
Tableau based reasoning

Principle

To prove $F$: build a tree with:

- The root is labeled with $\neg F$.
- The nodes are labeled by the concepts.
- Node successors are built by some expansion rules.
- A clash at the end of a path if:
  - $C(x) \in A$ and $\neg C(x) \in A$
  - $C(x) \in A$ and $\neg C(y) \in A$ and ($x = y$ or $y = x$)
  - $\bot(x) \in A$
Tableau based reasoning

□ rule

Conditions

\( A \) contains \((C_1 \cap C_2)(x)\) and does not contain \(C_1(x)\) and \(C_2(x)\)

Action

Prolongation: \( \mathcal{A}' = A \cup \{C_1(x), C_2(x)\} \)
Tableau based reasoning

□ rule

Conditions

A contains \((C_1 \sqcup C_2)(x)\) and does not contain \(C_1(x)\) and \(C_2(x)\)

Action

Branching: \(A' = A \cup \{C_1(x)\}\) and \(A'' = A \cup \{C_2(x)\}\)
Tableau based reasoning

∃ rule

Conditions

A contains \((\exists R.C)(x)\) and there is no individual \(z\) such as \(R(x, z)\) and \(C(z)\) are also in \(A\)

Action

\(A' = A \cup \{R(x, y), C(y)\}\) where \(y\) is an individual name which is not in \(A\)
∀ rule

Conditions

$A$ contains $(\forall R.C)(x)$ and $R(x, y)$ but does not contain $C(y)$

Action

$A' = A \cup \{C(y)\}$
Outline

1. Image and semantics
2. What is an ontology?
3. Ontologies for image understanding: overview
4. Description Logics
5. Description Logics for image understanding
   - Ontologies for interpretation refinement
     - Narrowing the semantic gap
     - Non-monotonic reasoning for image interpretation
       - Default reasoning
       - Abductive reasoning
6. Conclusion
Interpretation refinement using basic DLs inference services

Main principles

- Application domain knowledge is encoded into a TBox.
- A first interpretation of the targeted image is built using computer vision algorithms and translated into ABox assertions.
- Basic reasoning services of DLs such as consistency handling are used to revise the interpretation.
- Fuzzy DLs are used to take into account the imprecision of computer vision algorithms results.

Investigating fuzzy DLs-based reasoning in semantic image analysis [Dasiopoulou 10a].
Building and using fuzzy multimedia ontologies for semantic image annotation [Bannour 14].
Interpretation refinement using basic DLs inference services

**Domain Knowledge**
- Beach ⊆ Outdoor
- Forest ⊆ Landscape
- ∃contains.Sea ⊆ Seaside
- ∃contains.Sand ∩ Seaside ⊆ Beach
- Landscape ∩ Seaside ⊆ ⊥
- Seaside ∩ ∃contains.Building ⊆ ⊥
- Landscape ∩ ∃contains.Sea ⊆ ⊥

**Final Assertions**
- (image : Beach) ≥ 0.65
- (image : Seaside) ≥ 0.67
- (image : Outdoor) ≥ 0.67
- (image : ∃contains.Sea) ≥ 0.67
- (image : ∃contains.Sand) ≥ 0.65
- (image : ∃contains.Sky) ≥ 0.84
- (image : ∃contains.Person) ≥ 0.84

**Initial Assertions**
- (image : Forest) ≥ 0.59
- (image : Beach) ≥ 0.65
- (image : ∃contains.Sea) ≥ 0.67
- (image : ∃contains.Building) ≥ 0.52
- (image : ∃contains.Sky) ≥ 0.84
- (image : ∃contains.Person) ≥ 0.67

Dasiopoulou et al. [Dasiopoulou 10a]
Outline

1. Image and semantics
2. What is an ontology?
3. Ontologies for image understanding: overview
4. Description Logics
5. Description Logics for image understanding
   - Ontologies for interpretation refinement
   - Narrowing the semantic gap
   - Non-monotonic reasoning for image interpretation
     - Default reasoning
     - Abductive reasoning
6. Conclusion
Narrowing the semantic gap

Main approaches

- Building a dedicated visual concept ontology as an intermediate level between image features and application domain concepts:
  [Town 06, Bagdanov 07, Maillot 08, Porello 13, Mezaris 04].

- Using concrete domains to link high level concepts to their specific representations into the image domain:
  [Hudelot 08, Hudelot 14].

⇒ operational ontologies for image interpretation.
A spatial relation ontology for semantic image interpretation
Hudelot et al. [Hudelot 08, Hudelot 14]
Ontologies, concrete domains and semantic gap

Idea

Each application domain concept is linked to its representation in the image domain: use of concrete domains.

Hudelot et al. [Hudelot 08]
A spatial relation ontology

Hudelot et al. [Hudelot 08]
Formal representation of spatial relations

Abox:

- \( y: \text{SpatialObject}; x: \text{SpatialObject} \)
- \( \text{Right	extunderscore Of}_y \equiv \text{Right	extunderscore Of} \sqcap \exists \text{hasReferentObject} \cdot \{y\} \)
- \( x: \text{SpatialObject} \sqcap \exists \text{hasSpatialRelation}. \text{Right	extunderscore Of}_y \) and \( x: \text{SpatiallyRelatedObject} \)
- \( C_0 \equiv \text{SpatialRelation} \sqcap \exists \text{hasReferentObject} \cdot \{y\} \sqcap \exists \text{hasTargetObject} \cdot \{x\} \)

x is to the right of y: true
A dedicated logic for spatial reasoning: $\mathcal{ALC}(\mathbf{F})$

Instantiation of the description logic $\mathcal{ALC}\mathcal{RP}(D)$ with the concrete domain $\mathbf{F} = (\Delta_F, \Phi_F)$.

$\Delta_F = (\mathcal{F}, \leq_F, \wedge, \vee, \emptyset_F, 1_F, t, I)$

A residuated lattice of fuzzy sets defined over the image space $\mathcal{S}$, $\mathcal{S}$ being typically $\mathbb{Z}^2$ or $\mathbb{Z}^3$ for 2D or 3D images, with $t$ a t-norm (fuzzy intersection) and $I$ its residuated implication.

Main predicates of $\Phi_F$:

- $\mu_X$: degree of belonging to the spatial representation of the object $X$ in the spatial domain.
- $\nu_R$: fuzzy structuring element representing the fuzzy relation $R$ in the spatial domain.
- $\delta_{\nu_R}^{\mu_X}$: fuzzy dilation.
- $\varepsilon_{\nu_R}^{\mu_X}$: fuzzy erosion.
Application to brain imaging

Objective:
Progressive recognition of anatomical structures using spatial information.
Description of anatomical knowledge

Tbox:

- AnatomicalStructure ⊑ SpatialObject
- GN ⊑ AnatomicalStructure
- RLV ≡ AnatomicalStructure ⊓ ∃ hasFR.μRLV
- LLV ≡ AnatomicalStructure ⊓ ∃ hasFR.μLLV
- LV ≡ RLV △ LLV
- LV ≡ RLV △ LLV
- Right_of ≡ DirectionalRelation ⊓ ∃ hasFR.νIN_DIRECTION_0
- Close_to ≡ DistanceRelation ⊓ ∃ hasFR.νCLOSE_TO
- Right_of_RLV ≡ DirectionalRelation ⊓ ∃ hasReferentObject.RLV ⊓ ∃ hasFR.δμRLV νIN_DIRECTION_0
- Close_To_RLV ≡ DistanceRelation ⊓ ∃ hasReferentObject.RLV ⊓ ∃ hasFR.δμRLV νCLOSE_TO
- RCN ≡ GN ⊓ ∃ hasSR.(Right_of_RLV △ Close_To_RLV)
- CN ≡ GN ⊓ ∃ hasSR.(Close_To_LV)
- CN ≡ RCN △ LCN
Example

Abox:

- $c_1$: RLV, $(c_1, \mu_{S_1})$: hasFR
- $r_1$: Right_of, $(r_1, \nu_{IN\_DIRECTION\_0})$: hasFR
- $r_2$: Close_to, $(r_2, \nu_{CLOSE\_TO})$: hasFR
Example

Objective:

- Find some spatial constraints in the image domain on an instance $c_2$ of the Left Caudate Nucleus.
- $⇒$ Find constraints on concrete domains to ensure the satisfiability of the assertions $c_2$: RCN, $(c_2, \mu_{S_2})$: hasFR

Results using inference and properties

\[(\mu_{S_2})^F \leq \mathcal{F} (\delta_{\nu_{IN\_DIRECTION\_0}}^{\mu_{S_1}})^F \land (\delta_{\nu\_CLOSE\_TO}^{\mu_{S_1}})^F\]
Inference details:

\[ A \cup \{ c_2 : GN \sqcap \exists \text{hasSR.}(\text{Right_of_RLV} \sqcap \text{Close_to_RLV}), (c_2, \mu_{S_2}) : \text{hasFR} \} \]

\[ \sqcap \text{−rule} \]

\[ c_2 : GN, c_2 : \exists \text{hasSR.}(\text{Right_of_RLV} \sqcap \text{Close_to_RLV}) \]

\[ \sqcap \exists \text{−rule} \]

\[ c_3 : \text{Right_of_RLV} \sqcap \text{Close_to_RLV}, (c_2, c_3) : \text{hasSR}, (c_3, \mu_{S_3}) : \text{hasFR} \]

\[ \sqcap \text{Spatial Object Conjunction Rule } \mathcal{R} \sqcap \]

\[ (\mu_{\text{Right_of_RLV}} \sqcap d (\mu_{\text{Close_to_RLV}}))^F \]

\[ \sqcap \text{Spatial Object Conjunction Rule } \mathcal{R} \sqcap \]

\[ c_3 : \text{Right_of_RLV}, c_3 : \text{Close_to_RLV} \]

\[ \sqcap \text{Spatial Relation Rule } \mathcal{R}_{2RX} \]

\[ \mu_{S_3} = \delta_{\nu_{\text{IN DIRECTION}_0}} \sqcap d \delta_{\nu_{\text{CLOSE TO}}} \]

\[ \sqcap \text{spatial constraints} \]

\[ \text{fit}(\mu^{F}_{S_2}, \mu^{F}_{S_3}) = \text{fit}(\mu^{F}_{S_2}, (\delta_{\nu_{\text{IN DIRECTION}_0}} \sqcap d \delta_{\nu_{\text{CLOSE TO}}})^F) = 1 \]

\[ \sqcap \text{spatial constraints} \]

\[ (\mu^{F}_{S_2}) \leq \mathcal{F} (\delta_{\nu_{\text{IN DIRECTION}_0}})^F \wedge (\delta_{\nu_{\text{CLOSE TO}}})^F \]
Outline

1. Image and semantics
2. What is an ontology?
3. Ontologies for image understanding: overview
4. Description Logics
5. Description Logics for image understanding
   - Ontologies for interpretation refinement
   - Narrowing the semantic gap
   - Non-monotonic reasoning for image interpretation
     - Default reasoning
     - Abductive reasoning
6. Conclusion
Non-monotonic reasoning for image interpretation

Main principles:

Image interpretation is modeled as a non-monotonic reasoning process.

- **Default reasoning:**
  Non-monotonic logic to formalize reasoning with default assumptions [Reiter 80].

- **Abductive reasoning:**
  Backward reasoning: from observations to explanations, Charles Sanders Peirce in the late 19th century.
Image interpretation as a default reasoning service

Default rule

\[ \alpha : \beta_1, \ldots, \beta_n \]

\[ \gamma \]

- \( \alpha \): precondition of the rule.
- \( \beta_i \): justifications.
- \( \gamma \): consequent.

Intuitive explanation

Starting with a world description \( \alpha \) of what is known to be true, i.e. deducible and it is consistent to assume \( \beta_i \) then conclude \( \gamma \).

Example

\[ \forall x, \text{plays\_instruments}(x) : \text{improvises}(x) / \text{jazz\_musician}(x)^a \]

\[ \text{For all } x, \text{ if } x \text{ plays an instrument and if the fact that } x \text{ can improvise is consistent with all other knowledge then we can conclude that } x \text{ is a jazz musician.} \]
Default reasoning in DL

Reiter’s default theory [Reiter 80]

A pair \((\mathcal{W}, \mathcal{D})\) where \(\mathcal{W}\) is a set of closed first-order formulae (the world description) and \(\mathcal{W}\) a set of default rules.

Terminological default theory [Baader 92]

A pair \((\mathcal{A}, \mathcal{D})\) where:

- \(\mathcal{A}\): an ABox.
- \(\mathcal{D}\): a finite set of default rules whose preconditions, justifications and consequents are concept terms.

Maintaining decidability

- Default rules have to be closed over the ABox (instanciation with explicitly mentioned ABox individuals).
- Closed default rules: \(\alpha, \beta_i, \gamma\) are ABox concept axioms (no use of free variables, i.e. TBox concept axioms).
Spatioterminological default reasoning
Moller et al. approach [Möller 99b, Neumann 08]

Use case: topological reasoning for aerial image interpretation

Main idea
- Defaults are used for hypothesis generation regarding the classification of areas in an image.
- Default reasoning generates ABox extensions (hypothesized classifications) consistent with the rest of the knowledge base.

Preliminaries
The description logic $\mathcal{ALCRP}(S_2)$ for spatial information modeling and reasoning: $\mathcal{ALC}$ with:
- predicate existence restriction: $\exists u_1, \ldots, u_n.P$ with $P$ a predicate name from $S_2$ with arty $n$ and $u_1, \ldots, u_n$ feature chains.
- a concrete domain $S_2$ defined w.r.t. the topological space $\langle \mathbb{R}^2, 2^{\mathbb{R}^2} \rangle$. 
Spatioterminological default reasoning
Moller et al. approach [Möller 99b, Neumann 08]

The concrete domain $S_2$ over the topological space $\langle \mathbb{R}^2, 2^{\mathbb{R}^2} \rangle$

- $\Delta S_2$: set of non-empty, regular closed subsets of $\mathbb{R}^2$: regions
- Set of predicate names:
  - Predicate `is_region` with `is_region^{S_2} = \Delta S_2` and its negation `is_no_region` with `is_no_region^{S_2} = 0^{S_2}`
  - 8 basic predicates `dc`, `ec`, `po`, `tpp`, `ntpp`, `tppi`, `eq` (RCC-8 relations)
  - Predicates to name disjunctions of base relations: $p_1 - \ldots - p_n$
  - The predicate `dc-ec-po-tpp-ntpp-tppi-ntppi-eq` is called `spatially_related`
  - A binary predicate `inconsistent_relation` with `inconsistent_relation^{S_2} = \emptyset` (negation of `spatially_related`).
Spatioterminological default reasoning
Moller et al. approach [Möller 99b, Neumann 08]

Example

Interpretation problem: generate hypotheses for object b.

$S_2$ predicates formalization

\[
\begin{align*}
\text{inside} & \equiv \exists (\text{has\_area})(\text{has\_area}) . \text{tpp} - \text{ntpp} \\
\text{contains} & \equiv \exists (\text{has\_area})(\text{has\_area}) . \text{tppi} - \text{ntppi} \\
\text{overlaps} & \equiv \exists (\text{has\_area})(\text{has\_area}) . \text{po} \\
\text{touches} & \equiv \exists (\text{has\_area})(\text{has\_area}) . \text{ec} \\
\text{disjoint} & \equiv \exists (\text{has\_area})(\text{has\_area}) . \text{dc}
\end{align*}
\]
Spatioterminological default reasoning
Moller et al. approach [Moller 99b, Neumann 08]

<table>
<thead>
<tr>
<th>TBox</th>
</tr>
</thead>
<tbody>
<tr>
<td>area \equiv \exists (has_area).is_region</td>
</tr>
<tr>
<td>natural_region \equiv \neg\text{administrative_region}</td>
</tr>
<tr>
<td>country_region \sqsubseteq \text{administrative_region} \sqcap</td>
</tr>
<tr>
<td>\text{large_scale} \sqcap area</td>
</tr>
<tr>
<td>city_region \sqsubseteq \text{administrative_region} \sqcap</td>
</tr>
<tr>
<td>\neg\text{large_scale} \sqcap area</td>
</tr>
<tr>
<td>lake_region \sqsubseteq \text{natural_region} \sqcap area</td>
</tr>
<tr>
<td>river_region \sqsubseteq \text{natural_region} \sqcap area</td>
</tr>
</tbody>
</table>

Example

| country \equiv country\_region \sqcap                                |
| \forall\text{contains}.\neg country\_region \sqcap                  |
| \forall\text{overlaps}.\neg country\_region \sqcap                  |
| \forall\text{inside}.\neg country\_region                           |
| city \equiv city\_region \sqcap                                     |
| \exists\text{inside\_country\_region}                              |
| lake \equiv lake\_region                                             |
| river \equiv river\_region \sqcap                                   |
| \forall\text{overlaps}.\neg lake\_region \sqcap                     |
| \forall\text{contains}.\bot \sqcap                                   |
| \forall\text{inside}.\neg lake\_region                             |
Spatioterminological default reasoning
Moller et al. approach [Möller 99b, Neumann 08]

Example

Abox
\{a : country, b : area, (a, b) : contains, (b, a) : inside\}

Spatioterminological default rules

\[ d_1 = \frac{area : city}{city} \quad d_2 = \frac{area : lake}{lake} \quad d_3 = \frac{area : city}{city} \]

Closed spatioterminological default rules, \( d_i(\text{ind}) \)
e.g.

\[ d_1(a) = \frac{\{a : area\} : \{a : city\}}{\{a : city\}} \]

6 different closed defaults can be obtained \((d_1(a), d_1(b), d_2(a), d_2(b), d_3(a), d_3(b))\)
Spatiotermotinological default reasoning

Moller et al. approach [Möller 99b, Neumann 08]

Default rules reasoning

\[ d_1 = \frac{\text{area} : \text{city}}{\text{city}} \]

- \( d_1(a) \): cannot be applied. Contradiction between \( a : \text{city} \) and \( a : \text{country} \) in the Abox. \textit{country_region} and \textit{city_region} are disjoint in the TBox (due to \textit{large_scale} and \textit{¬large_scale}).

- \( d_1(b) \): can be applied. Abox extension:

\[ \{a : \text{country}, b : \text{area}, b : \text{city}, (a, b) : \text{contains}, (b, a) : \text{inside}\} \]
Spatioterminological default reasoning
Moller et al. approach [Möller 99b, Neumann 08]

Default rules reasoning

\[ d_2 = \frac{\text{area} : \text{lake}}{\text{lake}} \]

- \( d_2(a) \): cannot be applied. Contradiction between \( a : \text{lake} \) and \( a : \text{country} \) in the Abox. \textit{administrative\_region} and \textit{natural\_region} are disjoint.
- \( d_2(b) \): can be applied. Abox extension:

\[
\{ a : \text{country}, b : \text{area}, b : \text{lake}, (a,b) : \text{contains}, (b,a) : \text{inside} \}
\]

But if Abox contains \( d_1(a), d_2(b) \) cannot be applied \( \implies \) two possible extensions.
Spatioterminological default reasoning
Moller et al. approach [Möller 99b, ?]

Default rules reasoning, cont’d

\[ d_3 = \frac{\text{area : country}}{\text{country}} \]

- \( d_3(a) \) cannot be applied. Its conclusion is already entailed by the ABox.
- \( d_3(b) \) cannot be applied. The consequent \( b : \text{country} \) makes the Abox inconsistent because \( a \) is already known as a country.

\[ \mathcal{A} \models (a : \forall \text{contains} \neg \text{country_region}) \]

\( (a, b) : \text{contains}, b : \text{country} \implies b : \text{country_region} \)
Spatioterminological default reasoning.
Moller et al. approach [Möller 99b, ?]

Abox

\{a : country, b : area, (a, b) : overlaps, (b, a) : overlaps\}

⇒ the default rule \(d_1(b)\) cannot be applied to conclude that object \(b\) is a city.

Example 2

Subtle inferences due to topological constraints
Spatioterminological default reasoning.

Moller et al. approach [Möller 99b, ?]

Example 2

\[ A = \{ a : \text{country}, b : \text{area}, (a, b) : \text{overlaps}, (b, a) : \text{overlaps} \} \]

\[(b, a) : \text{overlaps}, b : \text{city} \implies b : \text{city}_\text{region} \sqcap \exists \text{inside}.\text{country}_\text{region} \implies \neg (a : \text{country}_\text{region}) \]  
(since \((b, a) : \text{overlaps})

Remark

- Due to \(\exists\), there exists an implicit individual \(c\) which is a \text{country}_\text{region} such that \((b, c) : \text{inside}\) holds.
- Impossible due to topological constraints (\(b\) inside \(c\) and \(c\) not overlap with \(a\) or does not contain \(a\)).
- No way to conclude that \(b\) could possibly be a city.
Spatioterminological default reasoning.
Moller et al. approach [Möller 99b, ?]

Example 3

We can conclude that the spatial relationship between the river and the lake is either $ec$ or $dc$. 
Spatioterminological default reasoning.
Moller et al. approach [Möller 99b, ?]

Example 3

Incomplete spatial information

Restricted default theories with ABox patterns

\[ d_4 = \frac{\{x : \text{lake}, y : \text{river}, (x, y) : \text{spatially-related} : \text{country}\} : \{(x, y) : \text{disjoint}\}}{\{(x, y) : \text{disjoint}\}} \]

\[ d_4 = \frac{\{x : \text{lake}, y : \text{river}, (x, y) : \text{spatially-related} : \text{country}\} : \{(x, y) : \text{touches}\}}{\{(x, y) : \text{touches}\}} \]

Closing the patterns yields 8 different closed defaults.
Abductive reasoning

- Abduction using safe rules (Peraldi et al. [Peraldi 09]).
- Concept abduction (Atif et al. [Atif 14]).
Abductive reasoning

Sort of *backward reasoning* from a set of observations to a cause.

**Definition**

Given a knowledge base $\mathcal{K}$ and a formula $\mathcal{O}$ representing an *observation* with $\mathcal{K} \not\models \mathcal{O}$, we look for an *explanation* formula $\mathcal{H}$ such that $\mathcal{H}$ is satisfiable w.r.t. $\mathcal{K}$ and

$$\mathcal{K} \cup \mathcal{H} \models \mathcal{O}$$

holds.

**Case of image interpretation**

- Scene = observation.
- Interpretation = look for the *best* explanation considering a terminological knowledge part about the scene context.
Abductive reasoning and description logics

Distinct abductive problems [Elsenbroich 06]

Let $\mathcal{L}$ be a DL, $\mathcal{K}$ a knowledge base in $\mathcal{L}$

- Concept abduction
- ABox abduction
- TBox abduction
- Knowledge base abduction
Abduction using safe rules

Multimedia interpretation as abduction
Peraldi et al. [Peraldí 09]

Ontology-based reasoning techniques for multimedia interpretation and retrieval
Möller et al. [Möller 08]
Multimedia interpretation as an abduction problem.

Main idea: Abduction as a non-standard retrieval inference service

Observations are used to constitute queries that have to be answered by acquiring what should be added to the knowledge base in order to positively answer to a query.

Use of conjunctive queries

Structure of the form \{head | body\}:

\{(X_1, \cdots, X_n) | atom_1, \cdots, atom_m\}, \text{ with } \\
atom = C(X), R(X, Y), (X = Y)

- head: list of variables for which we like to compute bindings
- body: query atoms

Example: \{x | \exists y \exists z (ChildOf(x, y) \land ChildOf(x, z) \land Married(y, z))\}

Query answer: set of bindings for variables in the head
Abduction inference

Given a set of ABox assertions $\Gamma$ in form of a query and a KB, $\Sigma = (\mathcal{T}, \mathcal{A})$, derive all sets of ABox assertions $\Delta$ (explanations) such that $\Delta$ is consistent w.r.t the ontology $\Sigma$ ($\Sigma \cup \Delta$ is satisfiable) and:

- $\Sigma \cup \Delta \models \Gamma$.
- $\Delta$ is a minimal explanation for $\Gamma$, i.e. there exists no other explanation $\Delta'$ in the solution set that is not equivalent to $\Delta$ and it holds that $\Sigma \cup \Delta' \models \Delta$. 

Peraldi et al. [Peraldi 09]
Formalisation

Multimedia abduction:

- $\Sigma = (\mathcal{T}, \mathcal{A})$, a knowledge base on the application domain with $\mathcal{A}$ assumed empty.
- $\Gamma = \Gamma_1 \cup \Gamma_2$, set of Abox assertions, encoding low level extracted information from images (objects and their spatial relationships):
  - $\Gamma_1$: bona fide assertions, assumed to be true by default.
  - $\Gamma_2$: assertions requiring fiats (aimed to be explained).
- Abduction process: compute $\Delta$, a set of ABox explanations, such that

$$\Sigma \cup \Gamma_1 \cup \Delta \models \Gamma_2$$

The process is implemented as (boolean) query answering.
Description Logics for image understanding

Non-monotonic reasoning for image interpretation

Illustration on an example

Peraldi et al. [Peraldi 09]

ABox $\Gamma$ : low-level image analysis results

\[
\begin{align*}
\text{pole}_1 & : \text{Pole} \\
\text{human}_1 & : \text{Human} \\
\text{bar}_1 & : \text{Bar} \\
\{\text{bar}_1, \text{human}_1\} & : \text{near}
\end{align*}
\]

$\Sigma$, a Tbox and DL-safe rules on the athletics domain

\[
\begin{align*}
\text{Jumper} & \sqsubseteq \text{Human} \\
\text{Pole} & \sqsubseteq \text{Sports_Equipment} \\
\text{Bar} & \sqsubseteq \text{Sports_Equipment} \\
\text{Pole} \sqcap \text{Bar} & \sqsubseteq \text{Pole} \sqcap \text{Jumper} \\
\text{Pole} \sqcap \text{Jumper} & \sqsubseteq \text{Jumper} \sqcap \text{Bar} \\
\text{Jumping_Event} & \sqsubseteq \exists \leq 1 \text{hasParticipant.Jumper} \\
\text{Pole_Vault} & \sqsubseteq \text{Jumping_Event} \sqcap \exists \text{hasPart.Pole} \sqcap \exists \text{hasPart.Bar} \\
\text{High_Jump} & \sqsubseteq \text{Jumping_Event} \sqcap \exists \text{hasPart.Bar} \\
\text{near}(Y, Z) & \rightarrow \exists \text{Pole_Vault}(X), \exists \text{hasPart}(X, Y), \exists \text{Bar}(Y), \exists \text{hasPart}(X, W), \exists \text{Pole}(W), \exists \text{hasParticipant}(X, Z), \exists \text{Jumper}(Z) \\
\text{near}(Y, Z) & \rightarrow \exists \text{High_Jump}(X), \exists \text{hasPart}(X, Y), \exists \text{Bar}(Y), \exists \text{hasParticipant}(X, Z), \exists \text{Jumper}(Z)
\end{align*}
\]
Illustration on an example

ABox $\Gamma$ : low-level image analysis results

\begin{align*}
\text{pole}_1 & : \text{Pole} \\
\text{human}_1 & : \text{Human} \\
\text{bar}_1 & : \text{Bar} \\
\{\text{bar}_1, \text{human}_1\} & : \text{near}
\end{align*}

\begin{itemize}
\item $\Gamma_1 = \{\text{pole}_1 : \text{Pole}, \text{human}_1 : \text{Human}, \text{bar}_1 : \text{Bar}\}$
\item $\Gamma_2 = \{(\text{bar}_1, \text{human}_1) : \text{near}\}$
\item Boolean query $Q_1 := \{() \mid \text{near}(\text{bar}_1, \text{human}_1)\}$
\end{itemize}
Possible explanations:

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

Preference score:

\[
S_p(\Delta) := S_i(\Delta) - S_h(\Delta), \quad \text{with}
\]
\[
S_i(\Delta) := |\{ i | i \in inds(\Delta) \text{ and } i \in inds(\Sigma \cup \Gamma_1) \}| 
\]
\[
S_h(\Delta) := |\{ i | i \in inds(\Delta) \text{ and } i \in new\_inds \}| 
\]
\( \Delta_1 \) incorporates \( \text{human}_1 \) and \( \text{bar}_1 \) from \( \Gamma_1 \), then \( S_i(\Delta_1) = 2 \).

\( \Delta_1 \) hypothesizes two new individuals: \( \text{new\_ind}_1, \text{new\_ind}_2 \), then
\[
S_h(\Delta_1) = 2.
\]

\[
S_p(\Delta_1) = 0
\]

\[
S_p(\Delta_2) = 3 - 1 = 2.
\]

\[
S_p(\Delta_3) = 2 - 1 = 1.
\]

\[ \Rightarrow \] \( \Delta_2 \) represents the ‘preferred’ explanation:

\[
\Delta_2 = \{ \text{new\_ind}_1 : \text{Pole\_Vault}, (\text{new\_ind}_1, \text{bar}_1) : \text{hasPart}, (\text{new\_ind}_1, \text{pole}_1) : \text{hasPart}, (\text{new\_ind}_1, \text{human}_1) : \text{hasParticipant}, \text{human}_1 : \text{Jumper} \}
\]

The image should better be interpreted as showing a pole vault and not a high jump.
Multimedia interpretation as concept abduction

Explanatory reasoning for image understanding using formal concept analysis and description logics.

Atif et al. [Atif 14]
Brain image understanding

Image interpretation

Pathological brain with small deforming peripheral tumor

Interpretation as an abduction process

\[ \mathcal{K} \models (\gamma \rightarrow \varphi) \]

*Computing of the best explanation* from observations \( \varphi \) given some a priori expert knowledge \( \mathcal{K} \) encoded in description logics.
Knowledge representation

\[
\begin{align*}
\text{CerebralHemisphere} & \sqsubseteq \text{BrainAnatomicalStructure} \\
\text{PeripheralCerebralHemisphere} & \sqsubseteq \text{CerebralHemisphereArea} \\
\text{SubCorticalCerebralHemisphere} & \sqsubseteq \text{CerebralHemisphereArea} \\
\text{GreyNuclei} & \sqsubseteq \text{BrainAnatomicalStructure} \\
\text{LateralVentricle} & \sqsubseteq \text{BrainAnatomicalStructure} \\
\text{BrainTumor} & \sqsubseteq \text{Disease} \\
\text{SmallDeformingTumor} & \equiv \text{BrainTumor} \\
\text{LargeDefTumor} & \equiv \text{BrainTumor} \sqcap \\
\text{PeripheralSmallDeformingTumor} & \equiv \text{BrainTumor} \sqcap \\
\text{SubCorticalSmallDeformingTumor} & \equiv \text{SmallDeformingTumor} \sqcap \\
\text{Initial ABox } A_1 \{ t_1 : \text{BrainTumor}; e_1 : \text{NonEnhanced}; l_1 : \text{LateralVentricle}; p_1 : \text{PeripheralCerebralHemisphere}; (t_1, e_1) : \text{hasEnhancement}; (t_1, l_1) : \text{farFrom}; (t_1, p_1) : \text{hasLocation}; \ldots \}.
\end{align*}
\]

\[
\begin{align*}
\text{BrainTumor} & \sqcap \\
\text{NonEnhanced} & \sqcap \\
\text{LateralVentricle} & \sqcap \\
\text{PeripheralCerebralHemisphere} & \sqcap \\
\text{hasEnhancement} & \sqcap \\
\text{farFrom} & \sqcap \\
\text{hasLocation} & \sqcap \\
\text{Brain} & \sqcap \\
\text{CerebralHemisphere} & \sqcap \\
\text{Edema} & \sqcap \\
\text{Necrosis} & \sqcap \\
\text{Enhanced} & \sqcap \\
\end{align*}
\]

\[
\begin{align*}
\exists \text{hasLocation}. \text{CerebralHemisphere} \\
\exists \text{hasComponent}. \text{Edema} \\
\exists \text{hasComponent}. \text{Necrosis} \\
\exists \text{hasEnhancement}. \text{Enhanced}
\end{align*}
\]
Interpretation as a concept abduction process

\[ \mathcal{K} \models \gamma \subseteq O, \text{ with } O, \text{ main specific concept of } t_1, \text{ defined as} \]

\[
\begin{align*}
&\text{BrainTumor} \sqcap \exists \text{hasEnhancement}.\text{NonEnhanced} \sqcap \\
&\exists \text{farFrom}.\text{LateralVentricle} \sqcap \\
&\exists \text{hasLocation}.\text{PeripheralCerebralHemisphere}
\end{align*}
\]

A set of possible explanations is:

\[
\{ \text{DiseasedBrain, SmallDeformingTumoralBrain, PeripheralSmallDeformingTumoralBrain} \}
\]

The preferred solution according to minimality constraints is:

\[
\gamma \equiv \text{PeripheralSmallDeformingTumoralBrain}
\]
Abduction and logics

Description logics

Where are we?
- Only a few works
- Rewriting approach (Modal logics - Description Logics)

Propositional logics (morpho-logics, Bloch et al. [Bloch 02])

\[
\begin{align*}
\lbrack \varepsilon(\varphi) \rbrack &:= \varepsilon(\lbrack \varphi \rbrack), \\
\lbrack \delta(\varphi) \rbrack &:= \delta(\lbrack \varphi \rbrack)
\end{align*}
\]

Successive erosions of the set of models

- Erosion of the conjunction of the theory with the formula to be explained

- Erosion of the theory while maintaining the coherence with the formula to be explained
Proposed approach

Enrichment of description logics with abductive reasoning services

⇒ Association between three theories:

- Concept lattices
- Models
- Description logics

\[(P(X), C)\]
Global scheme

**Application domain**
(e.g. brain imaging)

**DL**
- Generic knowledge
  (e.g. from anatomical textbooks)
- Concepts
  TBox

**FCA**
- Exploration algorithm
  (includes complex concepts)

**Interpretation of a particular case**
- Observations
  (resulting from image processing algorithms applied on a specific image)
- ABox
  most specific concept in the TBox
  operators in C

- TBox
- M
- M₀
- M, G
- C

**J. Atif, I. Bloch, C. Hudelot**

Image Understanding
Concept lattice induced from $K_{\text{brain}}$.
Erosion path leading to compute a preferred explanation
Outline

1. Image and semantics
2. What is an ontology?
3. Ontologies for image understanding: overview
4. Description Logics
5. Description Logics for image understanding
6. Conclusion
Conclusion

Ontologies and logic-based approaches for image interpretation

- A growing interest in the literature.
- Main advantages: explicit knowledge encoding for reuse and reasoning processes.
- Need for more convergence between computer vision, machine learning and logics community.
Thanks for your attention
Lourdes Agapito, Michael M. Bronstein & Carsten Rother, editeurs.

Explanatory reasoning for image understanding using formal concept analysis and description logics.

F. Baader & B. Hollunder.
Embedding Defaults into Terminological Knowledge Representation Formalisms.

Andrew D. Bagdanov, Marco Bertini, Alberto Del Bimbo, Giuseppe Serra & Carlo Torniai.
Semantic annotation and retrieval of video events using multimedia ontologies.


Conclusion

Gustavo Carneiro, Antoni B. Chan, Pedro J. Moreno & Nuno Vasconcelos.

*Supervised Learning of Semantic Classes for Image Annotation and Retrieval.*

Navneet Dalal & Bill Triggs.

*Histograms of Oriented Gradients for Human Detection.*

Stamatia Dasiopoulou, Ioannis Kompatsiaris & Michael Strintzis.

*Applying Fuzzy DLs in the Extraction of Image Semantics.*

Stamatia Dasiopoulou & Ioannis Kompatsiaris.

*Trends and Issues in Description Logics Frameworks for Image Interpretation.*

Stamatia Dasiopoulou, Vassilis Tzouvaras, Ioannis Kompatsiaris & Michael G. Strintzis. 
*Enquiring MPEG-7 based multimedia ontologies.* 

Jia Deng, Alexander C. Berg, Kai Li & Li Fei-Fei. 
*What does classifying more than 10,000 image categories tell us?* 

Chabane Djeraba. 
*Association and Content-Based Retrieval.* 
IEEE Transaction on Knowledge and Data Engineering (TDDE’03), vol. 15, pages 118–135, January 2003.

Ivan Donadello & Luciano Serafini. 
*Mixing Low-Level and Semantic Features for Image Interpretation - A Framework and a Simple Case Study.* 
In Agapito et al. [Agapito 15], pages 283–298.

Corinna Elsenbroich, Oliver Kutz & Ulrike Sattler.  
*A Case for Abductive Reasoning over Ontologies.*  

Dumitru Erhan, Christian Szegedy, Alexander Toshev & Dragomir Anguelov.  
*Scalable Object Detection Using Deep Neural Networks.*  

Ali Farhadi, Mohsen Hejrati, Mohammad Amin Sadeghi, Peter Young, Cyrus Rashtchian, Julia Hockenmaier & David Forsyth.
Every Picture Tells a Story: Generating Sentences from Images.

Ali Farhadi & Mohammad Amin Sadeghi.  
Phrasal Recognition.  

L. Fei-Fei & P. Perona.  
A Bayesian hierarchical model for learning natural scene categories.  

Carolina Galleguillos & Serge J. Belongie.  
Context based object categorization: A critical survey.  

Tianshi Gao & Daphne Koller.  
Discriminative learning of relaxed hierarchy for large-scale visual recognition.  
Conclusion

Gregory Griffin & Pietro Perona.  
*Learning and Using Taxonomies for Fast Visual Categorization.*  

Thomas R. Gruber.  
*Toward principles for the design of ontologies used for knowledge sharing.*  

Nicola Guarino, Daniel Oberle & Steffen Staab.  
*What Is an Ontology?*  

*Recommendations for Video Event Recognition Using Concept Vocabularies.*  
Allen R Hanson & Edward M Riseman.
*VISIONS: A computer system for interpreting scenes.*

Stevan Harnad.
*The Symbol Grounding Problem.*

C. Hudelot & M. Thonnat.
*A cognitive vision platform for automatic recognition of natural complex objects.*

*Fuzzy Spatial Relation Ontology for Image Interpretation.*

Céline Hudelot, Jamal Atif & Isabelle Bloch.
*ALC(Ł): A New Description Logic for Spatial Reasoning in Images.*
In Agapito et al. [Agapito 15], pages 370–384.

Alejandro Jaimes & Shih fu Chang.


*Baby talk: Understanding and generating simple image descriptions.*

V. Lavrenko, R. Manmatha & J. Jeon.

*A model for learning the semantics of pictures.*

Li-Jia Li, Chong Wang, Yongwhan Lim, David M. Blei & Fei-Fei Li.

*Building and using a semantivisual image hierarchy.*

Ying Liu, Dengsheng Zhang, Guojun Lu & Wei-Ying Ma.

*A survey of content-based image retrieval with high-level semantics.*

David G. Lowe.

*Distinctive Image Features from Scale-Invariant Keypoints.*
Nicolas Maillot & Monique Thonnat.
*Ontology based complex object recognition.*

David Marr.
Vision: A computational investigation into the human representation and processing of visual information.

Marcin Marszalek & Cordelia Schmid.
*Constructing Category Hierarchies for Visual Recognition.*

Vasileios Mezaris, Ioannis Kompatsiaris & Michael G. Strintzis.
*Region-Based Image Retrieval Using an Object Ontology and Relevance Feedback.*

Ralf Möller & Bernd Neumann.
Ontology-based reasoning techniques for multimedia interpretation and retrieval, pages 55–98.

Ralf Möller, Bernd Neumann & Michael Wessel.
Towards Computer Vision with Description Logics: Some Recent Progress.
Integration of Speech and Image Understanding, ICCV Workshop on, vol. 0, page 101, 1999.

Ralf Möller, Bernd Neumann & Michael Wessel.
Towards Computer Vision with Description Logics: Some Recent Progress.
Integration of Speech and Image Understanding, ICCV Workshop on, vol. 0, page 101, 1999.

Ralf Möller & Bernd Neumann.
Ontology-Based Reasoning Techniques for Multimedia Interpretation and Retrieval.

Milind Naphade, John R. Smith, Jelena Tesic, Shih-Fu Chang, Winston Hsu, Lyndon Kennedy, Alexander Hauptmann & Jon Curtis.
Conclusion

Large-Scale Concept Ontology for Multimedia.

Bernd Neumann & Ralf Möller.
On scene interpretation with description logics.

Devi Parikh, C. Lawrence Zitnick & Tsuhan Chen.

Multimedia Interpretation as Abduction.
In International Workshop on Description Logics (DL’07), 2007.

Irma Sofía Espinosa Peraldí, Atila Kaya & Ralf Möller.
Formalizing Multimedia Interpretation based on Abduction over Description Logic Aboxes.
In Bernardo Cuenca Grau, Ian Horrocks, Boris Motik & Ulrike Sattler, editeurs, Proceedings of the 22nd International Workshop on Description
Daniele Porello, Marco Cristani & Roberta Ferrario.  
*Integrating Ontologies and Computer Vision for Classification of Objects in Images.*  

Raymond Reiter.  
*A Logic for Default Reasoning.*  

*ImageNet Large Scale Visual Recognition Challenge.*  

*LabelMe: A Database and Web-Based Tool for Image Annotation.*
Simone Santini, Amarnath Gupta & Ramesh Jain.  
*Emergent Semantics through Interaction in Image Databases.*  

N. Simou, Th Athanasiadis, G. Stoilos & S. Kollias.  
*Image indexing and retrieval using expressive fuzzy description logics.*  

*Unsupervised Discovery of Visual Object Class Hierarchies.*  

Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta & Ramesh Jain.  
*Content-based image retrieval at the end of the early years.*  
Conclusion

Cees G. M. Snoek & Arnold W. M. Smeulders.

*Visual-Concept Search Solved?*


Rudi Studer, V. Richard Benjamins & Dieter Fensel.

*Knowledge Engineering: Principles and Methods.*


Christopher Town.

*Ontological inference for image and video analysis.*


Michael Uschold & Michael Gruninger.

*Ontologies and Semantics for Seamless Connectivity.*


Lei Wu, Xian-Sheng Hua, Nenghai Yu, Wei-Ying Ma & Shipeng Li.

*Flickr Distance: A Relationship Measure for Visual Concepts.*


B.Z. Yao, Xiong Yang, Liang Lin, Mun Wai Lee & Song-Chun Zhu.
I2T: Image Parsing to Text Description.