Non-standard inference services in Description Logics for Semantically Annotated Resource Retrieval

Tommaso **Di Noia**
Eugenio **Di Sciascio**
{t.dinoia | disciascio}@poliba.it

Politecnico di Bari

Italy

http://sisinflab.poliba.it

What is a Semantically Annotated Resource?

A semantically annotated resource is any kind of good, tangible or intangible (e.g. a document, a image, a product, a service) endowed of a description that refers to a shared ontology.

Roses all white

```
<owl:Class>
  <owl:intersectionOf rdf:parseType="Collection">
    <owl:Class rdf:about="rose"/>
    <owl:Restriction>
      <owl:onProperty rdf:resource="has_color"/>
      <owl:minCardinality
         rdf:datatype="#nonNegativeInteger">
        1
      </owl:minCardinality>
    </owl:Restriction>
    <owl:Restriction>
      <owl:onProperty rdf:resource="has_color"/>
      <owl:allValuesFrom>
        <owl:Class rdf:about="white"/>
      </owl:allValuesFrom>
    </owl:Restriction>
  </owl:intersectionOf>
</owl:Class>
```

What is a Semantically Annotated Resource?

A semantically annotated resource is any kind of good, tangible or intangible (e.g. a document, a image, a product, a service) endowed of a description that refers to a shared ontology.

We book roundtrip flights form BARI to MILTON KEYNES

```
<owl:Restriction>
  <owl:onProperty rdf:resource="flight_booking"/>
  <owl:allValuesFrom>
    <owl:Class>
      <owl:intersectionOf rdf:parseType="Collection">
        <owl:Class rdf:about="Round_Trip"/>
        <owl:Restriction>
          <owl:onProperty rdf:resource="departure"/>
          <owl:allValuesFrom>
            <owl:Class rdf:about="Bari"/>
          </owl:allValuesFrom>
        </owl:Restriction>
        <owl:Restriction>
          <owl:onProperty rdf:resource="destination"/>
          <owl:allValuesFrom>
            <owl:Class rdf:about="Milton_Keynes"/>
          </owl:allValuesFrom>
        </owl:Restriction>
      </owl:intersectionOf>
    </owl:Class>
  </owl:allValuesFrom>
</owl:Restriction>
```

Resource Retrieval

Given a set **R** of available resources and a request **q**, find and rank all the resources $\mathbf{r_i} \in \mathbf{R}$ such that they might satisfy the request **q**.

- Classical text based Information Retrieval can be considered a sub-problem of the Resource Retrieval one. In IR, resources are only documents.

ISSUES

- When a resource $\mathbf{r_i}$ might satisfy/match **q** ?
- Can we determine criteria useful for match ranking?

KR approach to Resource Retrieval

- **Text Based Information Retrieval**
  - The model for both the query and the document (resource) is extracted based on heuristic procedures for Natural Language.
  - Such procedures also evaluate metrics in order to establish the match degree between the two models
  - ISSUE: The extracted model is not related to the semantics of the document.

- **Semantic Based Resource Retrieval**
  - Semantic annotation rules out ambiguities of Natural Language. The model is explicit, well-defined, as it refers to a common ontology.
  - ISSUE: A problem arises about how well a single resource fits a particular request, or whether there is a pool of resources that - suitably composed – can fulfill a request.

Match: HOW TO evaluate?

In classical KR approaches to Resource Retrieval, typically, two inference services are exploited:

- **Classification**
  - "Is the resource **r** classified by the request **q** with respect to the common knowledge base?"
  - "Are all the request specifications within the resource description?"

  Given a knowledge base KB: $KB \models \mathbf{r} \Rightarrow \mathbf{q}$?

- **Consistency**
  - "Is the resource **r** consistent with the request **q** with respect to a common knowledge base?"
  - "Is in **r** any specification which contradicts the ones within **q**?"

  Given a knowledge base KB: $KB \models \mathbf{r} \wedge \mathbf{q} \equiv \bot$?

The need for explanation

- **Classification and Consistency limits**

Classification relation represents a full match between the request and the resource.

- A full match could not occur during the search process.
- A resource could be “almost” classified with respect to the request.
- Is it possible to **hypothesize an explanation** for the non-exact match?

In case of inconsistency, the requester might be interested in refining the request in order to retrieve an appealing resource.

- The user should know which part of the request is not compatible with the resource
- If a belief revision process is allowed, is it possible **to suggest explanation** to the user on which part of the request should be revised?

Explanation application

- **Query refinement**

  The user could refine the request based on incompatibility explanation or, in case of compatibility, on non-classification explanation hypothesis.

- **Negotiation in P2P scenarios**

  In a P2P scenario, the request description and the one related to available resources are not uniquely identified during a negotiation process. For instance, the supplier is able to refine his resource offer in order to make it more appealing with respect to the requester.

- **Resource ranking based on the explanation "length"**

  Having a metric based on explanation semantics, it is possible to compute a score, in case of non-exact match, representing how good is a resource with respect to a request. Such score can be used to rank different resources with respect to the same request.

Example

Request: graduated internal personnel, with experience in programming, networking

Resource1: Computer science engineer belonging to internal personnel

Resource2: Graduated programmer, working as consultant, expert in networking and SAP

Networking?

- **Resource1** does not completely match **Request**. The implication **Resource1** ⇒ **Request** does not hold. To reach an exact match, at least networking skills have to be hypothesized in **Resource1**.
- **Resource2** is not compatible with **Request** because of personnel type specification. In order to gain compatibility, whether internal personnel in **Request** or consultant in **Resource2** should be retracted.

Notice that once compatibility is reached, hypothesis can be formulated in order to make true the implication $\text{Resource2}_{\text{contracted}} \Rightarrow \text{Request}$ (or $\text{Resource2} \Rightarrow \text{Request}_{\text{contracted}}$)

Theoretical framework

Three Kinds of Belief Changes

- **Expansion**: add some information
  Why $KB \models \mathbf{r} \Rightarrow \mathbf{q}$ does not hold?
  Add (hypothesize) some information in **r**
- **Contraction**: retract on some information
  Why $KB \models \mathbf{r} \wedge \mathbf{q} \equiv \bot$?
  Contract whether **q** (revise your request) or **r**
- **Revision**: first retract and then add information

the INFORMATIONAL ECONOMY criterion

- Information is a precious good.
- We should change information as small as possible.
- Minimality criteria are necessary.

Expansion: Hypothesize explanation as small as possible.

Contraction: Contract your request as small as possible.

AS SMALL AS POSSIBLE: What does it mean?

- The explanation is dependent from the task.
- We could be interested in
  - explanation with the minimal **syntactical** length (e.g. containing the minimum number of terms)
  - explanation depending on some **semantic** criteria (e.g. contract only terms in the request representing concepts classified as NEGOTIABLE by the user)

Description Logics

A family of logic-based knowledge representation formalism.

Elements:

- **concept names** (unary predicates): sets of objects.
- **role names** (binary relations): relations between objects.
- **operators and quantifiers**: combine basic elements to form concept and role expressions.

Each DL allows a different set of operators and quantifiers:

$$\mathcal{AL}[\mathcal{C}][\mathcal{N}][\mathcal{H}][\mathcal{I}][\mathcal{Q}][\mathcal{O}]$$

Computational issues

OWL – DL : $\mathcal{SHOIN}$(D)   INTRACTABLE

Inference Services in DLs (1)

Standard Inferences

Given an ontology **T** and two concept expressions **q** and **r**

Subsumption (Classification)

Is **r** more specific than **q** ($\mathbf{r} \sqsubseteq \mathbf{q}$) with respect to the ontology?

Satisfiability (Consistency)

Is their conjunction ($\mathbf{q} \sqcap \mathbf{r}$) satisfiable with respect to the ontology?

Inference Services in DLs (2)

Non-Standard Inferences

Given an ontology T and two concept expressions **q** and **r**

Concept Abduction (Explanation Hypothesis)

If their conjunction ($\mathbf{r} \sqcap \mathbf{q}$) is satisfiable with respect to T, why the relation $\mathbf{r} \sqsubseteq \mathbf{q}$ does not hold?

May we hypothesize a concept expression **H** such that

$\mathbf{r} \sqcap \mathbf{H} \sqsubseteq \mathbf{q}$?

Concept Contraction (Inconsistency Explanation)

If their conjunction ($\mathbf{r} \sqcap \mathbf{q}$) is not satisfiable with respect to T, is it possible to formulate an inconsistency explanation for **r** (respectively for **q**) ?

What should be given up (**G**) and what should be kept (**K**) in **r** (respectively in **q**) in order to regain satisfiability ?

Why $\mathbf{r} \sqcap \mathbf{q}$ is not satisfiable?

Resource matchmaking via Contraction and Abduction: example

$\mathbf{T} = \{ \texttt{PC} \sqsubseteq \texttt{Computer} \sqcap \exists\texttt{hasOS}$

$\texttt{HomePC} \sqsubseteq \texttt{PC} \sqcap \forall\texttt{hasOS.MS} \sqcap \exists\texttt{pointer}$

$\texttt{HighLevel} \sqsubseteq \exists\texttt{cost} \sqcap \forall\texttt{cost.Expensive}$

$\texttt{Expensive} \sqsubseteq \neg\texttt{Cheap}$

$\texttt{MS} \sqsubseteq \neg\texttt{Unix}\}$

$q = \texttt{HomePC} \sqcap \exists\texttt{monitor} \sqcap \forall\texttt{pointer.}\forall\texttt{cost.Cheap}$

$r = \texttt{PC} \sqcap \forall\texttt{pointer.(Mouse} \sqcap \texttt{HighLevel)} \sqcap \forall\texttt{hasOS.Unix}$

r* vs. *q

Resource matchmaking via Contraction and Abduction: example

$T = \{\ PC \sqsubseteq Computer \sqcap \exists hasOS$

$HomePC \sqsubseteq PC \sqcap \forall hasOS.MS \sqcap \exists pointer$

$HighLevel \sqsubseteq \exists cost \sqcap \forall cost.Expensive$

$Expensive \sqsubseteq \neg Cheap$

$MS \sqsubseteq \neg Unix\ \}$

$q = HomePC \sqcap \exists monitor \sqcap \forall pointer.\forall cost.Cheap$

$r = PC \sqcap \forall pointer.(Mouse \sqcap HighLevel) \sqcap \forall hasOS.Unix$

$G_q = HomePC \sqcap \forall pointer.\forall cost.Cheap$

$K_q = PC \sqcap \exists pointer \sqcap \exists monitor$

$H_{K_q} = \exists monitor$

Proposed Minimality Criteria

- maximal under $\sqsubseteq_T$
- minimum length ($\leq$)
- minimal conjunctions ($\sqcap$)
- negotiable and strict constraints
- penalty functions (Calì et al. KES'04)

Concept Covering via Concept Abduction

Given an ontology T, a set of concept descriptions **R**={$\mathbf{r_1}$, $\mathbf{r_2}$, $\mathbf{r_3}$, …,$\mathbf{r_n}$} and a concept **q**.

A Concept Covering Problem(CCoP),is finding, if it exists, a subset of concept descriptions **Rc** $\subseteq$ **R**, such that

- the conjunction of all concepts $\mathbf{r_i} \in$ Rc is consistent w.r.t. the ontology T

$$T \nvDash \sqcap \mathbf{r_i} \equiv \bot$$

- a solution for the related Concept Abduction Problem

$$T \vDash (\sqcap \mathbf{r_i}) \sqcap \mathbf{H} \sqsubseteq \mathbf{q}$$

is such that $\mathbf{H} \sqsubseteq \mathbf{q}$ does not hold.

Concept Covering for composed resource retrieval

- Try to cover **q** as much as possible solving Concept Abduction Problems
- Choose the resource $\mathbf{r_i}$ with less explanation hypothesis
- Return the **composed resource** and what is yet to be hypotesized w.r.t. **q**

GREEDYsolveCCoP(R,q,T)
Input concepts q, ri ∈ *R*, i=1…k, q e ri satisfiable in T
Output ⟨*Rc*,H⟩
begin algorithm
Rc =∅;
q_uncovered = q;
Hmin = q;
do
    r_min = T
    for each ri ∈ *R*
        if *Rc* ∪ {ri} is a cover for q_uncovered **then**
            H = solveCAP(⟨L, ri, q_uncovered, T⟩);
            if H < H_min **then**
                r_min = ri;
                H_min = H;
            end if
        end if
    end for each
    if r_min ≠ T **then**
        R = *R* \ {ri};
        Rc = *Rc* ∪ {ri};
        q_uncovered = H_min;
    end if
while (r_min ≠ T);
return (*Rc*, q_uncovered);
end algorithm

Investigated scenarios

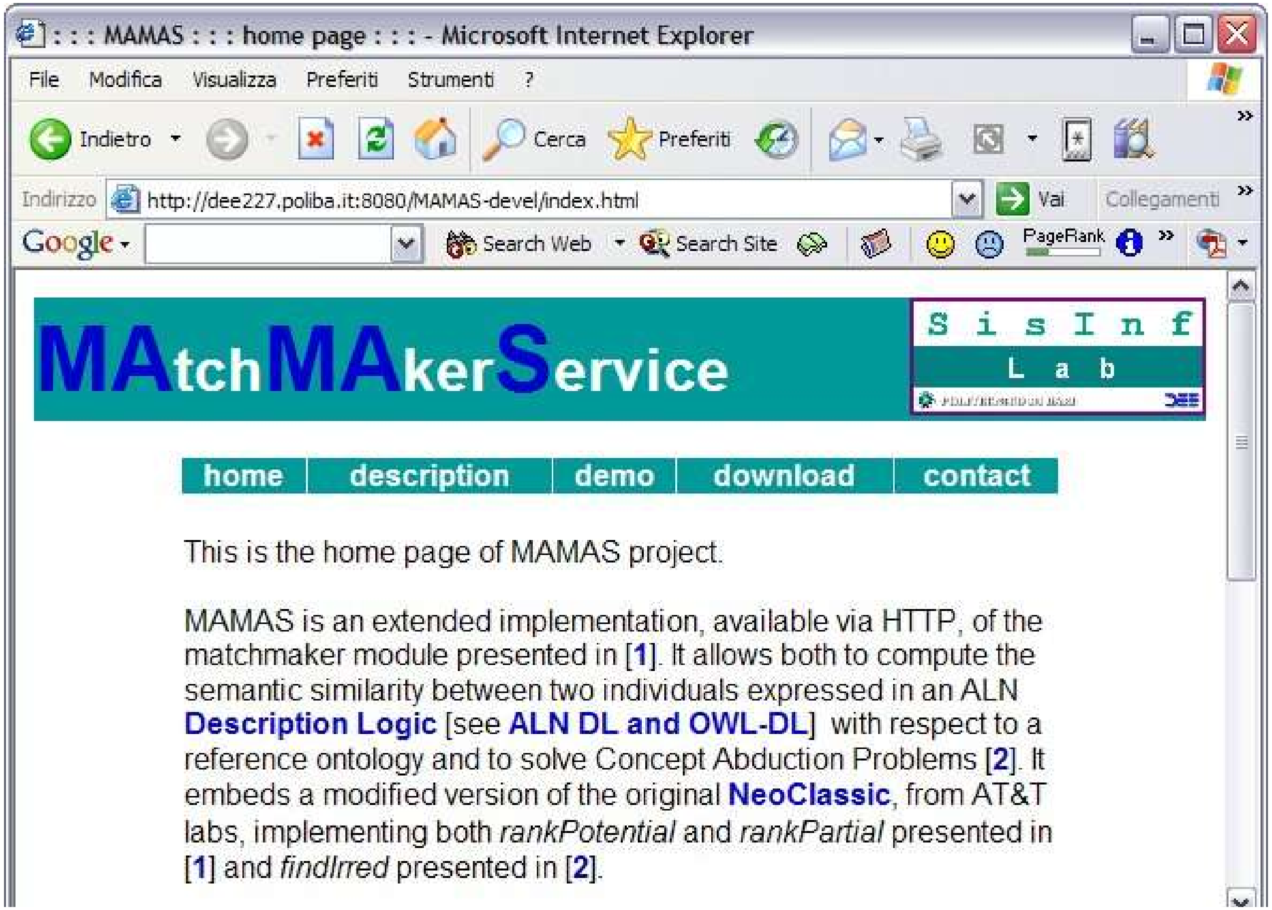
- **e-commerce**
  - negotiable and strict constraints in e-commerce for negotiation
  - demand vs. supply classification and ranking
- **skill matching**
  - automated job assignment and team composition
- **e-learning**
  - concept covering for automated courseware creation
- **automated Web Services orchestration**
  - create a ws execution flow with respect to a given request, using ws profile description

OWL – DL and $\mathcal{ALN}$

DL syntax	OWL syntax
C	<owl:Class rdf:ID = “C”/>
R	<owl:ObjectProperty rdf:ID = “R”/>
$\top$	<owl:Thing/>
$\bot$	<owl:Nothing/>
$A \sqsubseteq C$	<rdfs:subClassOf/>
$A \equiv C$	<owl:equivalentClass/>
$\leq n\ R$	<owl:maxCardinality/>
$\geq n\ R$	<owl:minCardinality/>
$= n\ R$	<owl:cardinality/>
$\neg A$	<owl:disjointWith/>
$\exists R$	<owl:someValuesFrom/>
$\forall R.C$	<owl:allValuesFrom/>
$C \sqcap D$	<owl:intersectionOf/>

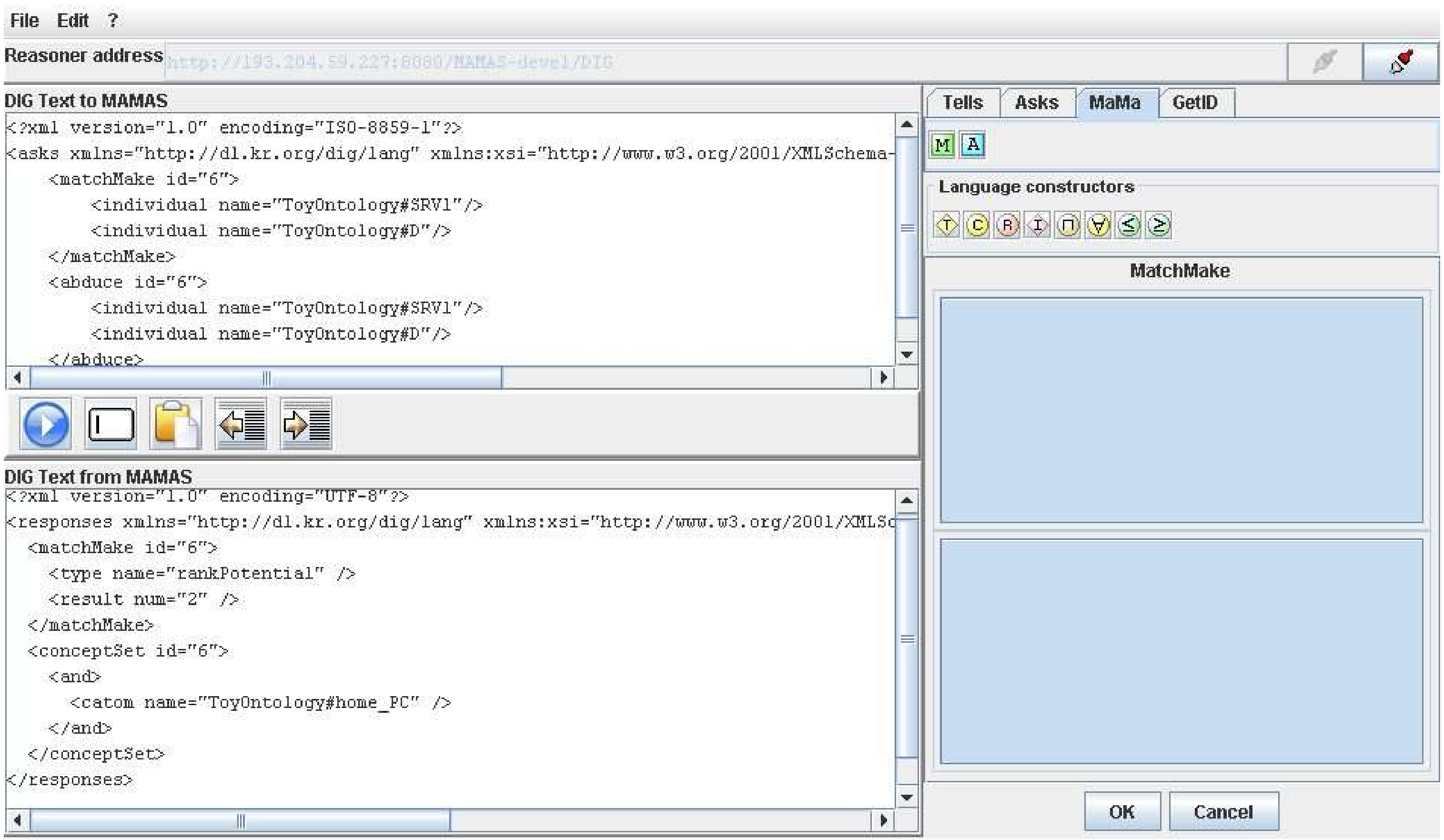
Polynomial in time

MaMaS: MatchMakerService



http://dee227.poliba.it:8080/MAMAS-devel/

MaMaS: MatchMakerService



Essential References

- S. Colucci, T. Di Noia, E. Di Sciascio, F.M. Donini, and M. Mongiello. Concept Abduction and Contraction for Semantic-based Discovery of Matches and Negotiation Spaces in an E-Marketplace. Electronic Commerce Research and Applications, 4(3), 2005.
- S. Colucci, T. Di Noia, E. Di Sciascio, F.M. Donini, and A. Ragone. Fully automated web services orchestration in a resource retrieval scenario. In Proc. of International Conference on web Services (ICWS '05). IEEE, 2005. to appear
- S. Colucci, T. Di Noia, E. Di Sciascio, F.M. Donini, and A. Ragone. Semantic-based automated composition of distributed learning objects for personalized e-learning. In 2nd European Semantic Web Conference (ESWC '05), 2005.
- S. Colucci, T. Di Noia, E. Di Sciascio, F.M. Donini, G. Piscitelli,and S. Coppi. Knowledge Based Approach to Semantic Composition of Teams in an Organization. In Proceedings ACM SAC'05, pages 1314-1319. March 2005.
- T. Di Noia, E. Di Sciascio and F.M. Donini. Extending Semantic-Based Matchmaking via Concept Abduction and Contraction. Proc. Of EKAW 2004, pp.307-320, 2004.
- T. Di Noia, E. Di Sciascio, F.M. Donini, and M. Mongiello. A system for principled Matchmaking in an electronic marketplace. International Journal of Electronic Commerce, vol. 8. No 4 pp. 9-37, 2004.
- S. Colucci, T. Di Noia, E. Di Sciascio, F.M. Donini, M. Mongiello, and M. Mottola. A formal approach to ontology-based semantic match of skills descriptions. Journal of Universal Computer Science (J.UCS), vol. 9. No.12, pp. 1437-1454, 2003.
- T. Di Noia, E. Di Sciascio, F.M. Donini, M. Mongiello. Abductive matchmaking using description logics. in Proceedings IJCAI '03, pp. 337-342, 2003.
- S. Colucci, T. Di Noia, E. Di Sciascio, F. M. Donini, M. Mongiello. Concept Abduction and Contraction in Description Logics. in proc. of 16th Intl. Workshop on Description Logics, Rome, Sept. 2003.