Knowledge web Heterogeneity

Jérôme Euzenat

Jerome.Euzenat@inrialpes.fr

Outline

1. Knowledge web heterogeneity work package;
2. Ontology alignment framework;
3. (Standard) alignment format.

The problem with heterogeneity

Resources being expressed in different ways must be reconciled before being used.

The mismatch between formalized knowledge can occur when:

— different languages are used;

— different terminologies are used;

— different modelling is used.

Reconciliation can be achieved online or offline with different constraints.

General objectives of KW2.2

- Provide the tools for dealing with these three cases of heterogeneity;

- Gather most of the European research people on providing solutions; be a worldwide forum for ontology alignment research.

=> The more people involved, the better.

Precise goals of KW 2.2

Contribute to improving ontology alignment techniques.

One way: evaluating existing techniques

- Common framework;
- Common formats;
- Benchmarking technology;
- Tools for using alignment results.

Reconciling heterogeneous resources

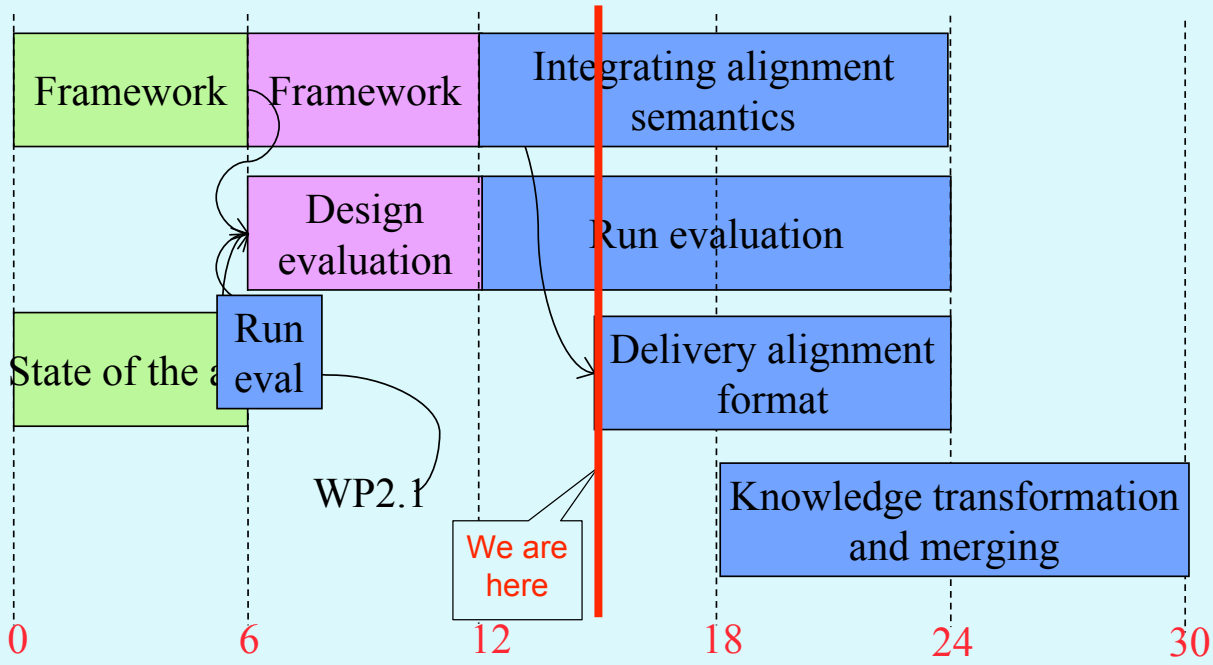
1) Finding the correspondence (between languages, terms or models);
2) Apply it to knowledge (message, ontology…):

— translating from one language to another;
— adding “bridge axioms” between two ontologies;
— creating database-like views;
— …

Benefits of this separation

- many algorithms can be used for the matching;
- many application can use one algorithm;
- we can publish alignments;
- we can compare alignment results;

JPA Update



Timeline (cont'd)

After 36 months

Specification of delivery alignment format

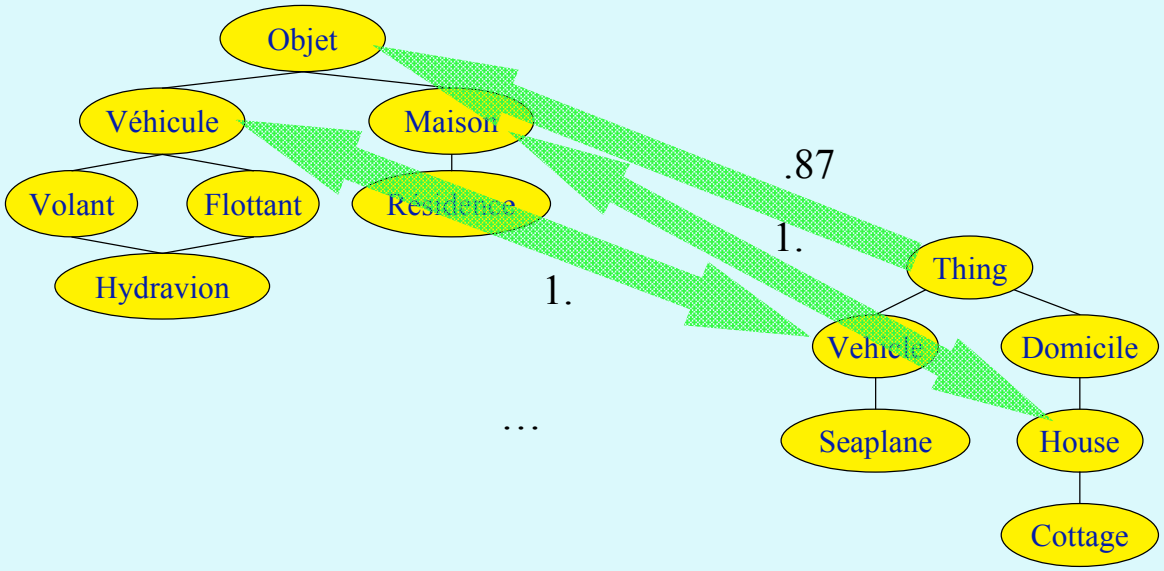
=> deliver to VISWE

After 48 months

Implementation of knowledge transformation and merging tools taking input from alignment

Framework

Ontologies and alignments



Diversity

- Variety of relations;
- Variety of strength;
- Variety of semantics (transitive or not).

Correspondences

A correspondence is an object of the following form:

$$\langle e, e', n, R \rangle$$

where:

e and *e'* are the entities between which a relation is asserted by the correspondence (e.g., formulas, terms, classes, individuals);

n is a degree of confidence in the correspondence

R is the relation associated with the correspondence, holding between *e* and *e'*.

Correspondences as mappings

Crisp mappings: mappings whose degree of confidence is 1.

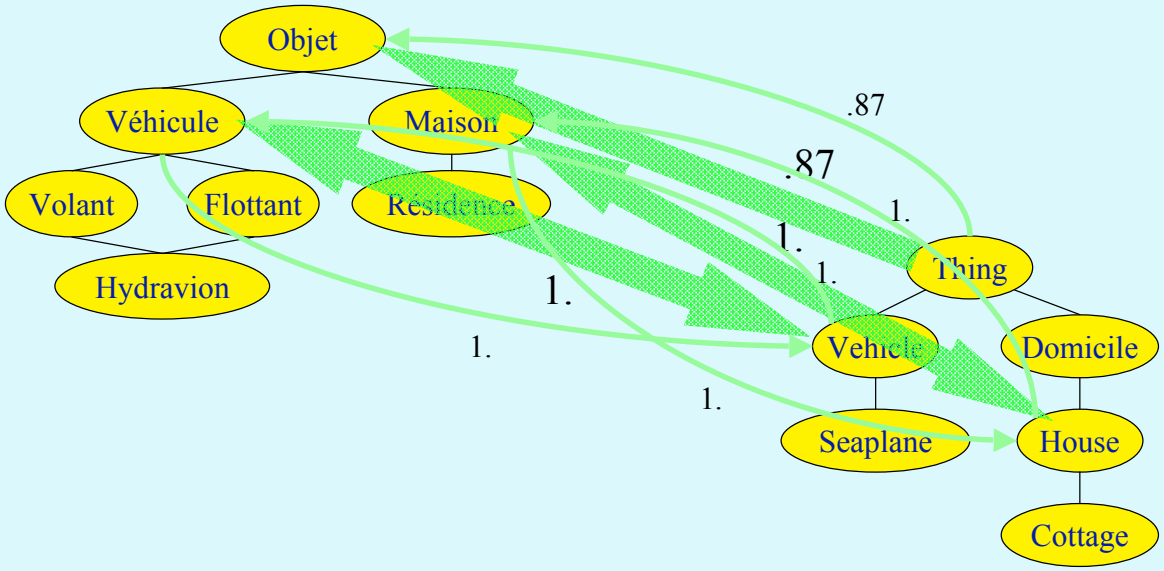
Two subcases:

1. *Classical Mapping*: mappings are like axioms that can be reasoned about
2. *Rule-Based Mapping*: mappings are cross-ontology rules that can only be applied on available knowledge

Fuzzy mappings: mappings whose degree of confidence is a (real) number between 0 and 1

[See deliverable for details on the formal semantics of the three types of mapping]

Correspondences as mappings



Rule-based expression

$$i{:}\phi(x) \rightarrow_{\alpha} j{:}\psi(x)$$

With global or peer (rule based) semantics?

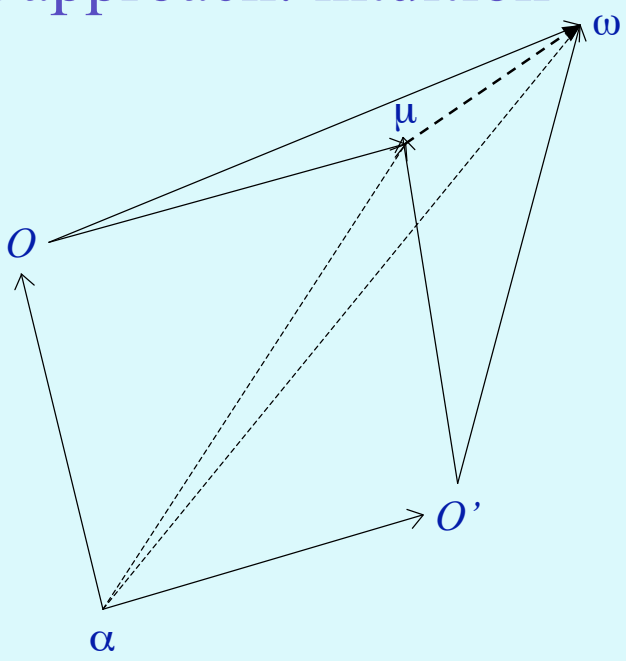
$i{:}\phi(x) \rightarrow j{:}\psi(x)$

$j{:}\psi(x) \rightarrow k{:}\gamma(x)$

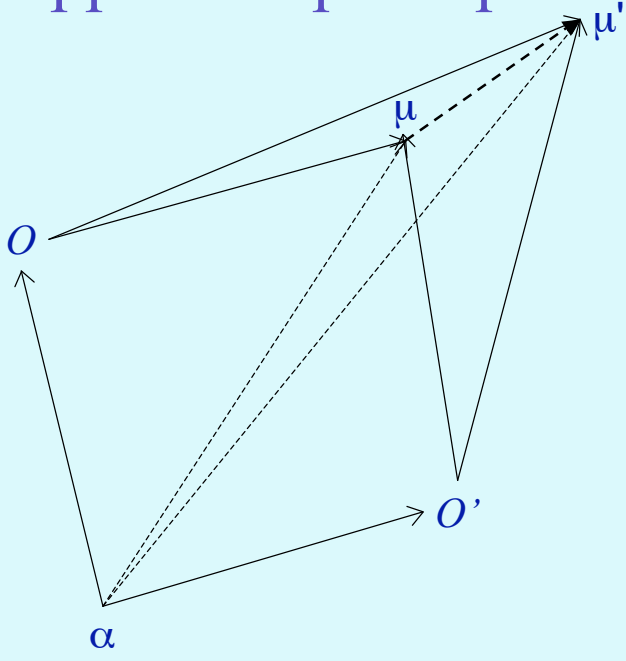
$i{:}\phi(a)$

$k{:}\gamma(a)$?

Categorical approach: intuition

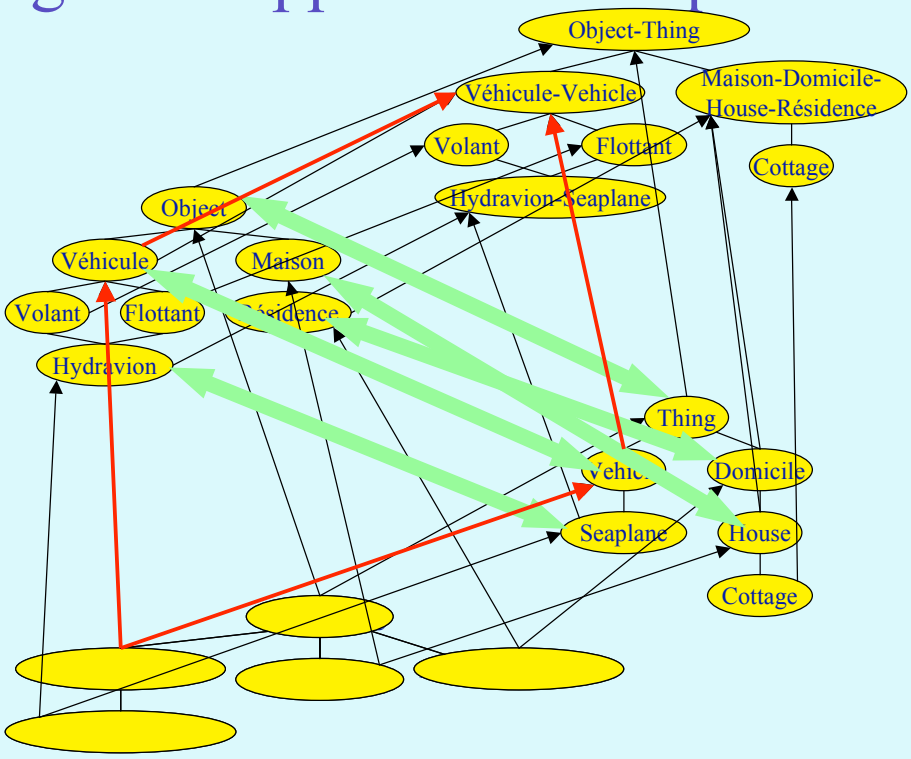


Categorical approach: principles

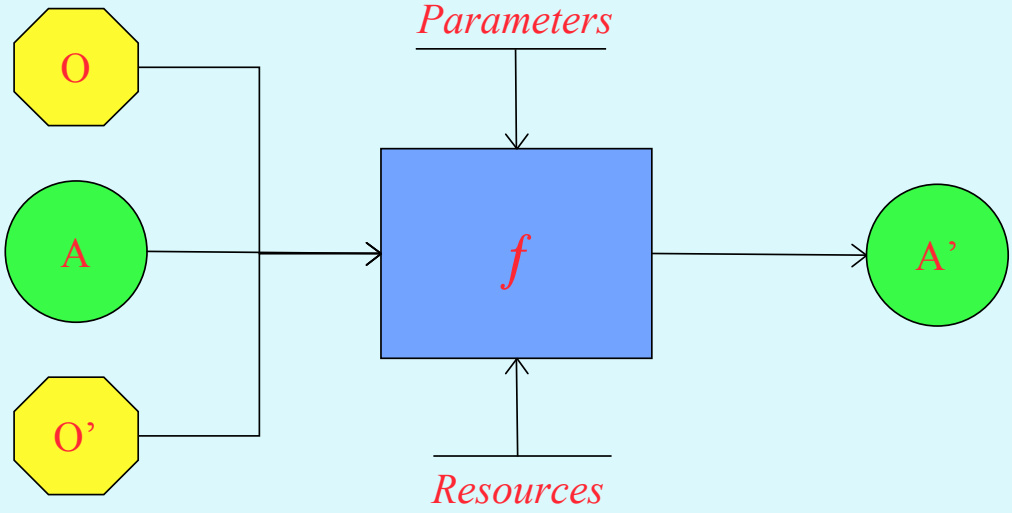


Push-out

Categorical approach: example



Alignment process



From framework to format

What's in an alignment (Level 0)

An alignment is a set of pairs, containing:

- Two aligned entities (class, object, relation…);
- The alignment relation (equality, specificity, aggregation, fuzzy, whatever);
- A confidence measure (qualifying the correspondence, not the relation).

$$\{ uri: < e, e', \alpha, R >^* \}$$

What's in an alignment (container)

- Level indicator;
- Description of the alignment: 1-1, 1-*, complete, etc.;
- Pair of ontology URIs;
- Set of pairs;

What else?

Level 1

Instead of two aligned entities, are two aligned sets of entities.

Level 2 correspondence (OWL…/SQL/…/F-Logic)

(blank nodes universaly quantified)

(unbound blank nodes existentially quantified)

other example, employee.name = person.first+“ ”+person.last

The “SEKT Mapping language” falls into this.

Level 2 (OWL…/SQL/…/F-Logic)

OWL expression
(blank nodes universaly quantified)
=R>
OWL expression
(blank nodes existentially quantified)

[[e.g., employee.name = person.first+“ ”+person.last]]

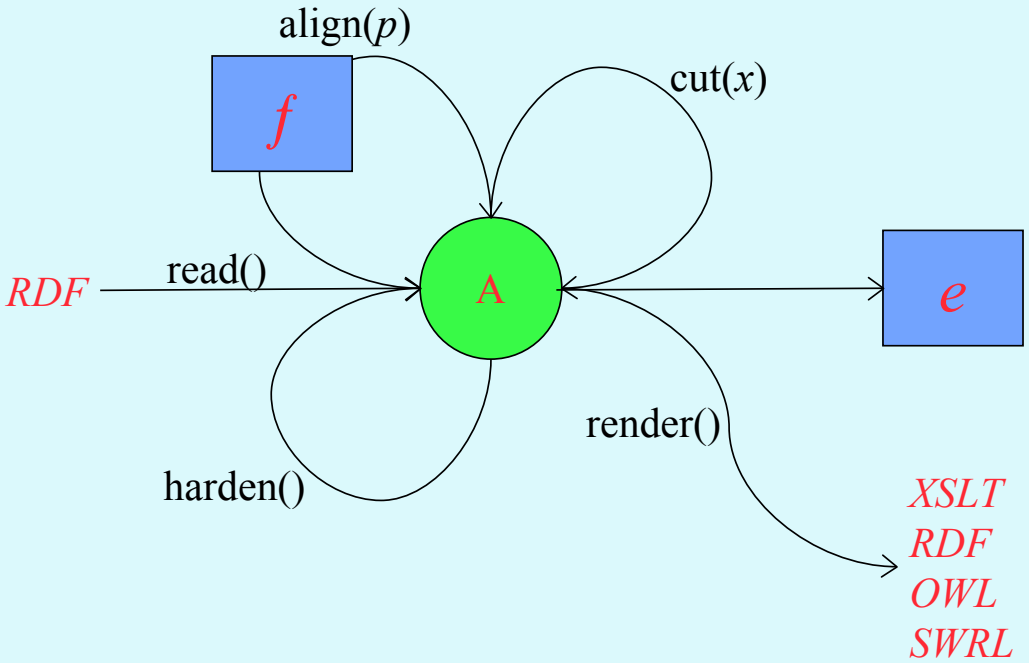
The “SEKT Mapping language” falls into this.

What is an alignment for?

- Storing, finding, and floating around;
- Piping alignments (Improving an existing alignment);
- Manipulating (thresholding and hardening);
- Generating processing output (transformations, axioms, rules);
- Comparing alignments.

  => already developed an API which does that [Euzenat2004f].

Alignment API

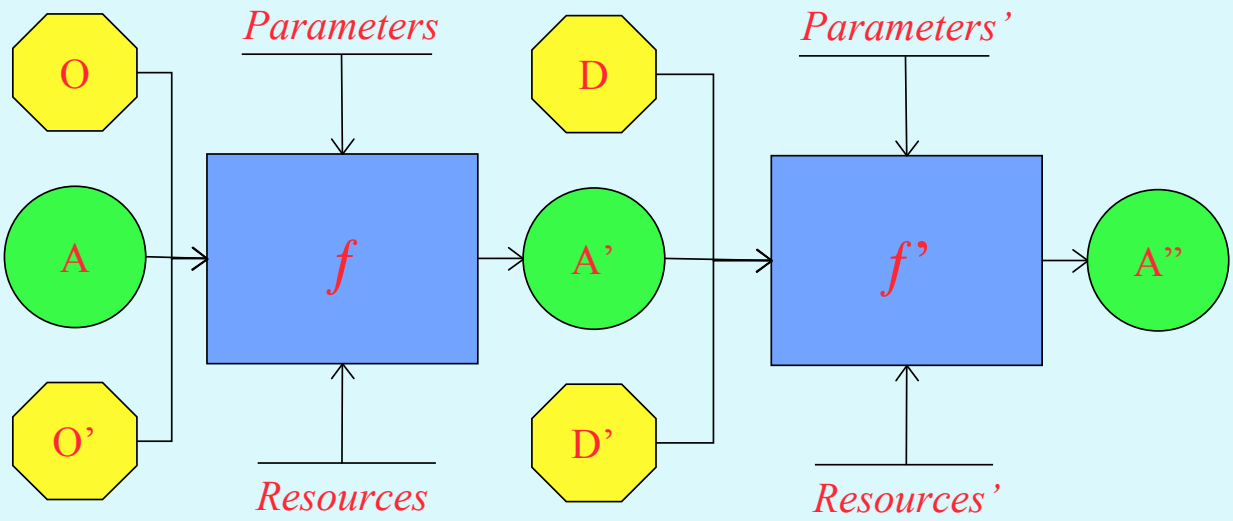


Current state of the alignment API

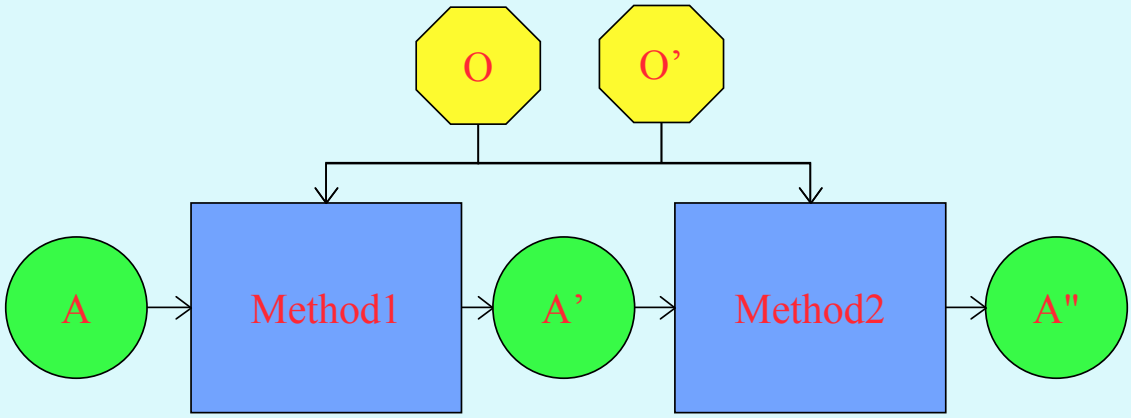
- Described in [Euzenat2004f];
- API as a set of Java interfaces (Alignment, Cell, Relation, Evaluator…);
- Unique rendering format in RDF/XML described by DTD, RDF(S) and OWL ontology;
- Implementation with the OWL API (available with many examples);
- Used in the EON contest.

A (standard) alignment format

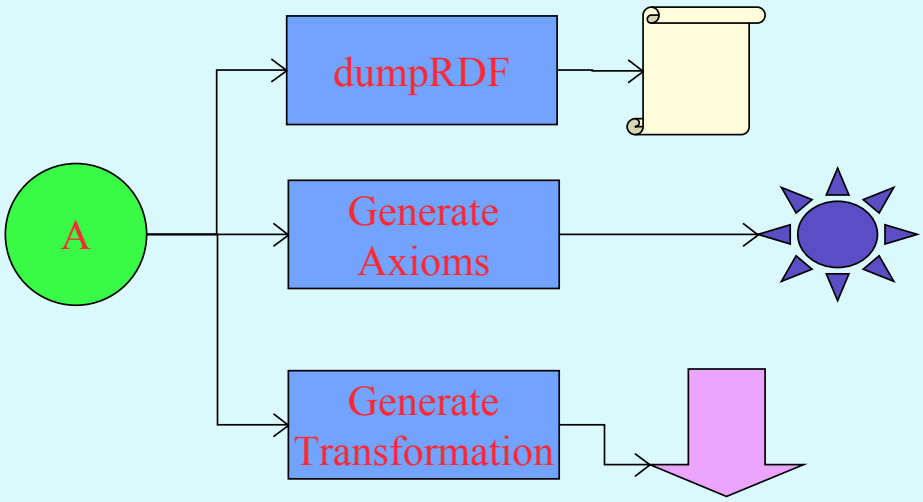
Data integration process



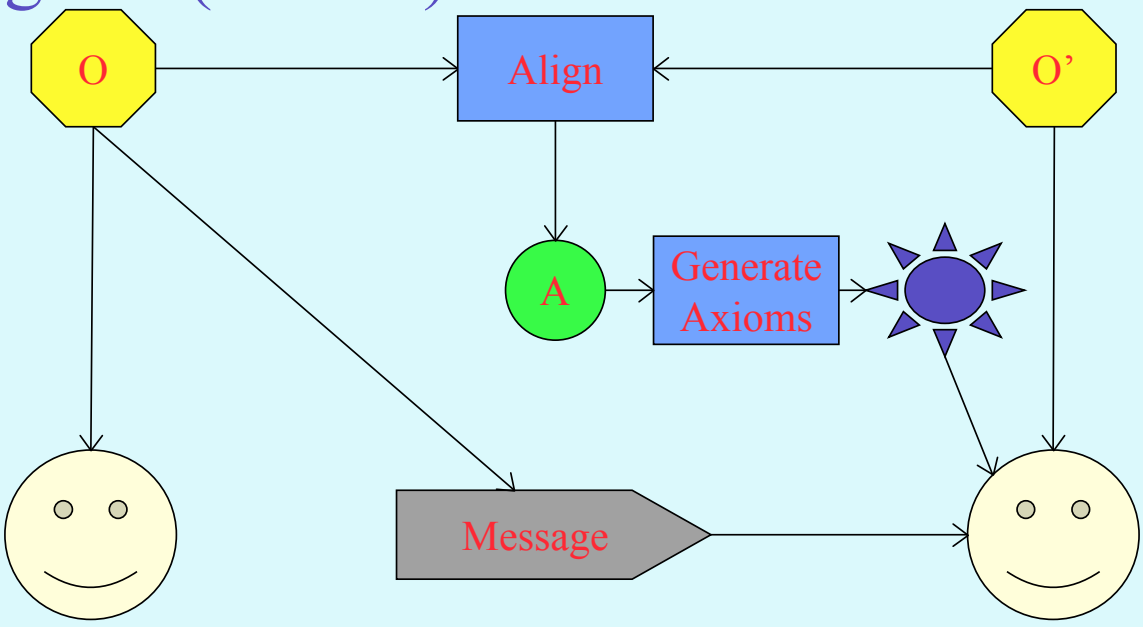
Standardizing from a practical standpoint: composition



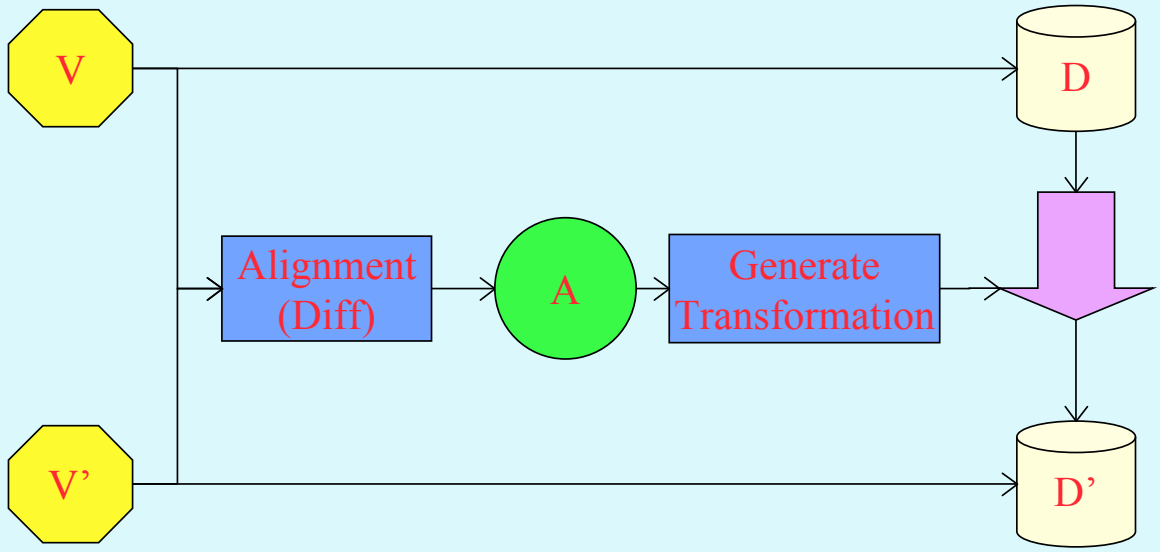
Standardizing from a practical standpoint: adapter generation



Applications: meaning negotiation in agents (WP2.3)

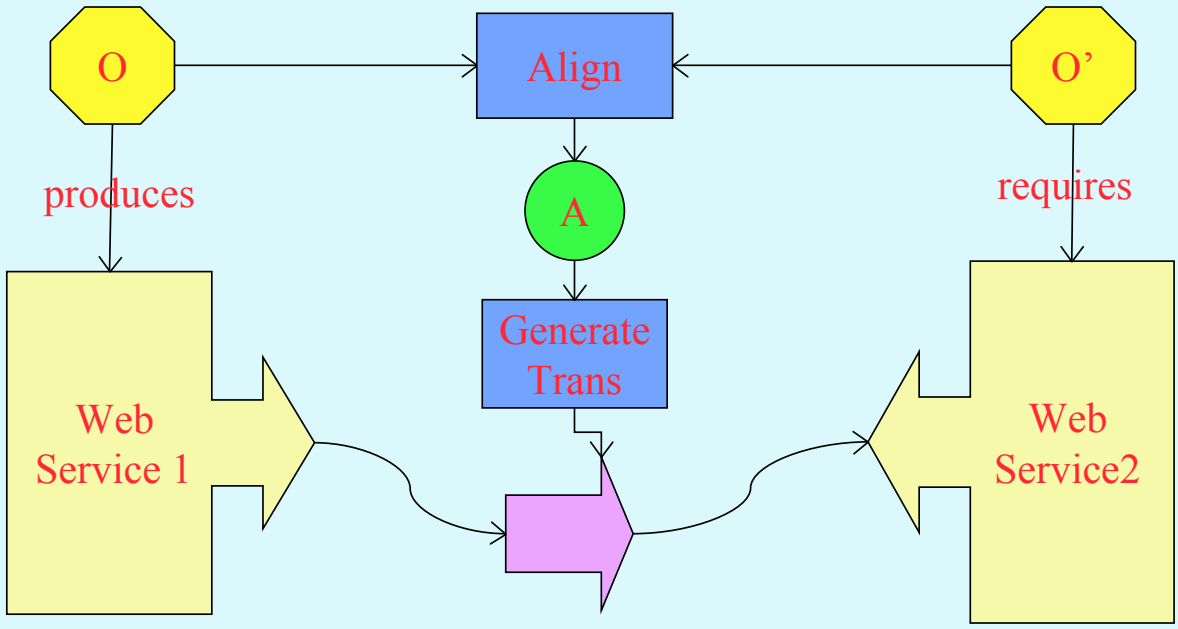


Application: version alignment (WP2.3) [+ module interface (WP2.1)]

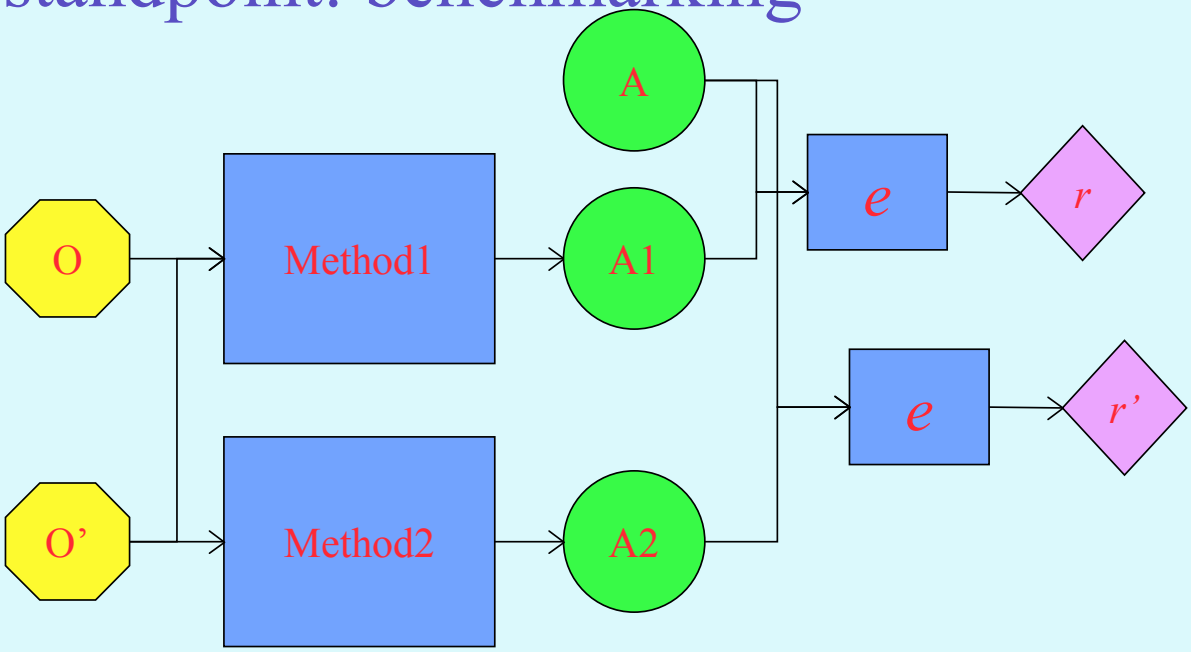


Natural candidate for upgrading/downgrading versions

Application: semantic web services (WP2.4)



Standardizing from an engineering standpoint: benchmarking



Setting up evaluation of ontology alignment

Evaluation: why?

- Comparing the possible solutions;
- Detecting the best methods;
- Finding out where we are weak.

Various goals:

For the developer: improving the solutions;

For the user: choosing the best tools;

For both: testing compliance with a norm.

What has been done?

- Information Interpretation and Integration Conference (I^3CON), held at the NIST Performance Metrics for Intelligent Systems (PerMIS) Workshop: focused on "real-life" test cases and comparing algorithm global performance.
  - led by Todd Hughes (Lockheed Martin Advanced Technology Laboratories)
- The Ontology Alignment Contest at the 3rd Evaluation of Ontology-based Tools (EON) Workshop, to be held the International Semantic Web Conference (ISWC): aims at defining a proper set of benchmark tests for assessing feature-related behavior.
  - led by myself (INRIA)

Why two challenges?

The idea of evaluating tools and methods has been out for long…

…but there has just been two occasions.

As a result:

— the setup has been quick&dirty;

— two different evaluations;

— but we can learn lessons from that!

Evaluation: how?

Several alternatives:

- Take real life cases and set the deadline;
- Take several cases, normalize them;
- Take simple cases identifying what they highlight (Benchmark suite);
- Build a challenge (MUC, TREC).

Ontologies

- 1 ontology and 20 variations ;
- 33 concepts, 60 properties, 57 instances;
- Narrow topic: bibliographic references;
- Linked with other resources (FOAF, iCalendar…);
- Expressed in OWL-DL (supposedly);
- Target alignment (made on purpose) published.

Concept tests (1xx)

Stupid tests such as aligning the reference ontology with:

- itself;
- the wine ontology of the OWL guide;
- itself restricted to it OWL-Lite part or its OWL-Lite generalization.

Systematic tests (2xx)

Ontology generated from the reference one by discarding one feature at a time:

- instances, properties, restrictions, etc.
- adding super-classes, collapsing the hierarchy, etc.

Initially planned to cover the full space of restrictions…

Real world tests (3xx)

4 ontologies on the same topic found on the web (UMBC, MIT, Karlsruhe, INRIA) without change!

Alignments

- 1:1 or rather ?:1;
- made by hand by myself (but most of them were made by construction);
- all weights at 1;
- only relations between classes and properties;
- all relations, but very few, at =.

Tools: Alignment API implementation

- Allows to manipulate the alignment output.
- Evaluators for comparing with reference alignments (provide precision/recall and other measures).
- Sample implementations of simple aligners.

More about that tomorrow

Other matters

Open workshop (EON);

Pre-publishing of the tests;

Ask for a paper, with comments on the tests and on the achieved results (as well as the results in normalized format).

Results and lessons

Few participants…

…especially with regard to the number of published papers on the topic!

We expected 5 participants but had only 4.

Several teams told me that they started playing with the contest.

Format nightmare (expected)

- OWL vs. RDFS;
- OWL species;
- XML/RDF vs. XML vs. N3;
- Jena vs. OWL-API;
- …and don't forget the bloody namespace!

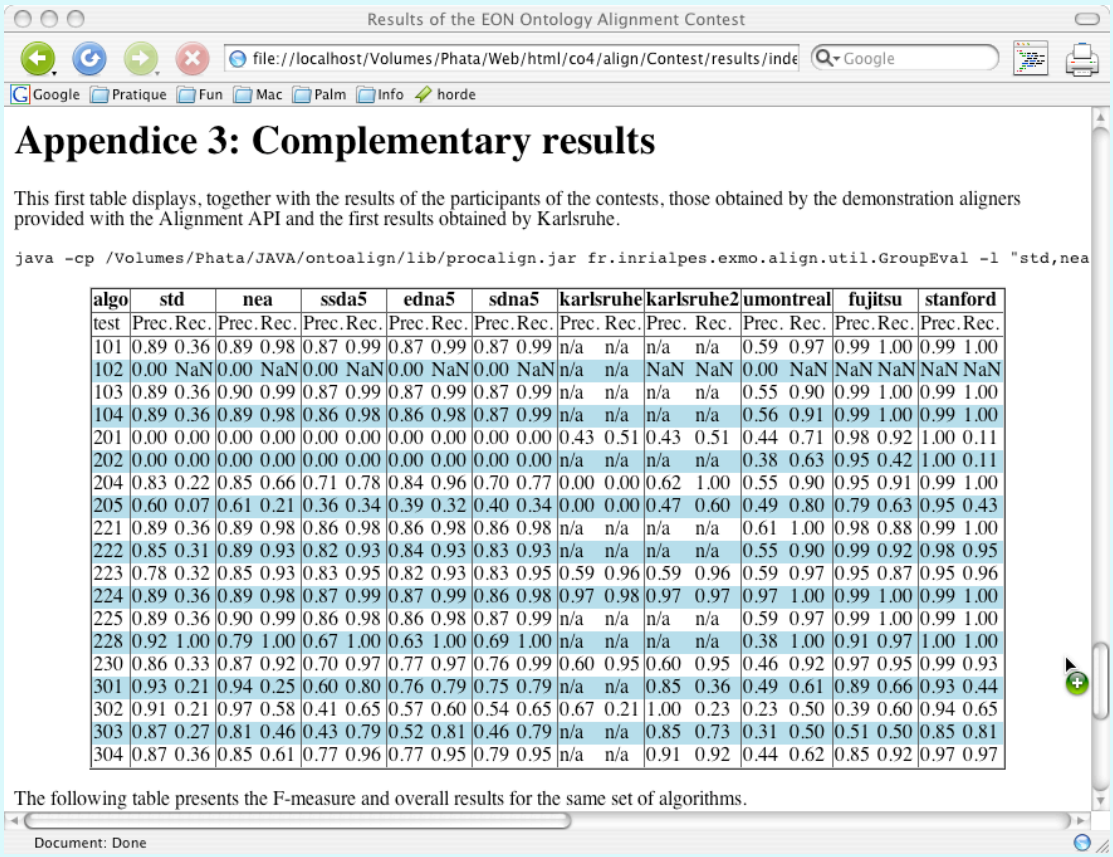
But we can overcome this.

Criteria

Precision/recall of pairs found with regard to the reference alignment;

A few irrelevant pairs have been discarded;

I forgot to count for 1/2 the ≤ which have been found =.



algo	std		nea		ssda5		edna5		sdna5		karlsruhe		karlsruhe2		umontreal		fujitsu		stanford	
test	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.
101	0.89	0.36	0.89	0.98	0.87	0.99	0.87	0.99	0.87	0.99	n/a	n/a	n/a	n/a	0.59	0.97	0.99	1.00	0.99	1.00
102	0.00	NaN	0.00	NaN	0.00	NaN	0.00	NaN	0.00	NaN	n/a	n/a	NaN	NaN	0.00	NaN	NaN	NaN	NaN	NaN
103	0.89	0.36	0.90	0.99	0.87	0.99	0.87	0.99	0.87	0.99	n/a	n/a	n/a	n/a	0.55	0.90	0.99	1.00	0.99	1.00
104	0.89	0.36	0.89	0.98	0.86	0.98	0.86	0.98	0.87	0.99	n/a	n/a	n/a	n/a	0.56	0.91	0.99	1.00	0.99	1.00
201	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.43	0.51	0.43	0.51	0.44	0.71	0.98	0.92	1.00	0.11
202	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	n/a	n/a	n/a	n/a	0.38	0.63	0.95	0.42	1.00	0.11
204	0.83	0.22	0.85	0.66	0.71	0.78	0.84	0.96	0.70	0.77	0.00	0.00	0.62	1.00	0.55	0.90	0.95	0.91	0.99	1.00
205	0.60	0.07	0.61	0.21	0.36	0.34	0.39	0.32	0.40	0.34	0.00	0.00	0.47	0.60	0.49	0.80	0.79	0.63	0.95	0.43
221	0.89	0.36	0.89	0.98	0.86	0.98	0.86	0.98	0.86	0.98	n/a	n/a	n/a	n/a	0.61	1.00	0.98	0.88	0.99	1.00
222	0.85	0.31	0.89	0.93	0.82	0.93	0.84	0.93	0.83	0.93	n/a	n/a	n/a	n/a	0.55	0.90	0.99	0.92	0.98	0.95
223	0.78	0.32	0.85	0.93	0.83	0.95	0.82	0.93	0.83	0.95	0.59	0.96	0.59	0.96	0.59	0.97	0.95	0.87	0.95	0.96
224	0.89	0.36	0.89	0.98	0.87	0.99	0.87	0.99	0.86	0.98	0.97	0.98	0.97	0.97	0.97	1.00	0.99	1.00	0.99	1.00
225	0.89	0.36	0.90	0.99	0.86	0.98	0.86	0.98	0.87	0.99	n/a	n/a	n/a	n/a	0.59	0.97	0.99	1.00	0.99	1.00
228	0.92	1.00	0.79	1.00	0.67	1.00	0.63	1.00	0.69	1.00	n/a	n/a	n/a	n/a	0.38	1.00	0.91	0.97	1.00	1.00
230	0.86	0.33	0.87	0.92	0.70	0.97	0.77	0.97	0.76	0.99	0.60	0.95	0.60	0.95	0.46	0.92	0.97	0.95	0.99	0.93
301	0.93	0.21	0.94	0.25	0.60	0.80	0.76	0.79	0.75	0.79	n/a	n/a	0.85	0.36	0.49	0.61	0.89	0.66	0.93	0.44
302	0.91	0.21	0.97	0.58	0.41	0.65	0.57	0.60	0.54	0.65	0.67	0.21	1.00	0.23	0.23	0.50	0.39	0.60	0.94	0.65
303	0.87	0.27	0.81	0.46	0.43	0.79	0.52	0.81	0.46	0.79	n/a	n/a	0.85	0.73	0.31	0.50	0.51	0.50	0.85	0.81
304	0.87	0.36	0.85	0.61	0.77	0.96	0.77	0.95	0.79	0.95	n/a	n/a	0.91	0.92	0.44	0.62	0.85	0.92	0.97	0.97

Raw results

algo	Karlsruhe		Montréal +INRIA		Fujitsu +Tokyo		Stanford	
test	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.
1xx	.	.	0,57	0,93	0,99	1,00	0,99	1,00
2xx	0,61	0,83	0,55	0,89	0,95	0,86	0,98	0,77
3xx	0,90	0,56	0,37	0,56	0,66	0,67	0,92	0,72
total	0,73	0,72	0,51	0,82	0,89	0,84	0,97	0,80

There were indeed three set of tests…

1xx were pretty easy;

2xx more difficult

3xx even more difficult.

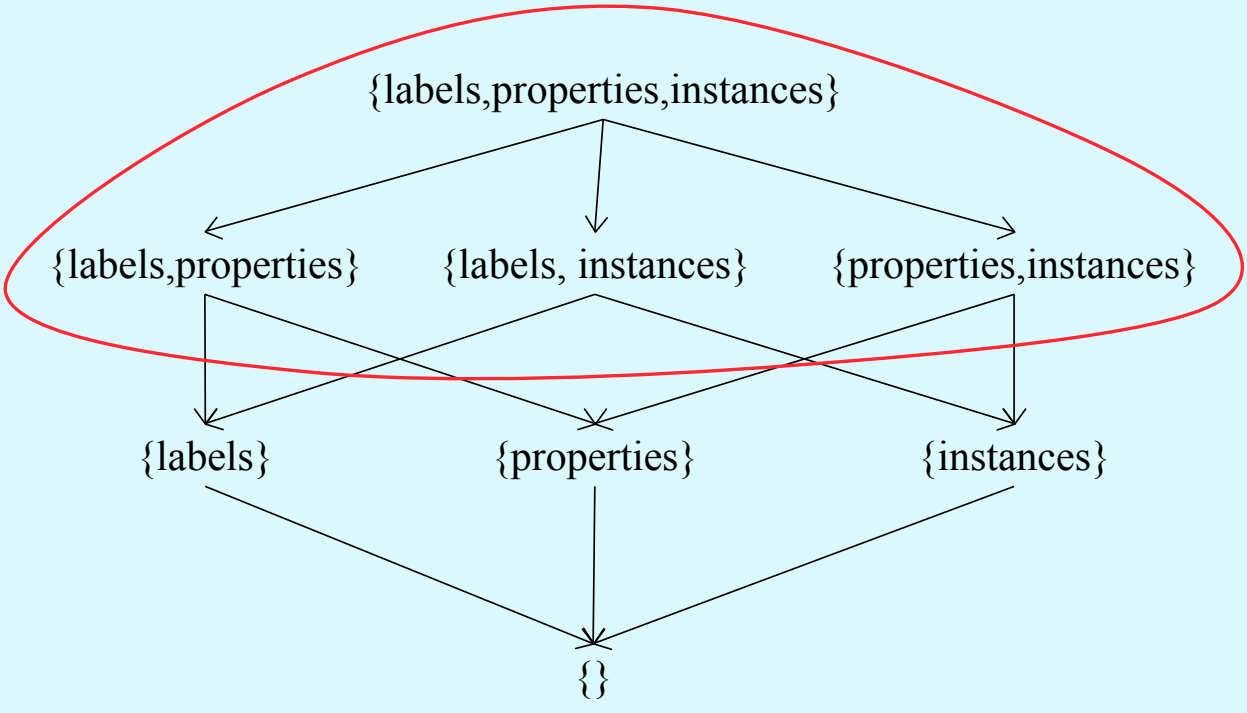
And within each single system these differences seem visible.

…and two classes of participants

Two where heavily based on threatment of labels and string. They took advantage of the fact that these where very often preserved.

Two tried to balance the influence of many different factors: so, they missed this advantage…

Range of tests



Future plans

What's next?

- Combine both initiatives in a yearly stable contest:

  benchmark suite + real world challenge?

  (known results) + (unknown result);

- Provide more automated procedures;

- Need to build some "consensus committee".

Future: alignment competition

- Use in addition real-world ontologies.

Currently under organization.

Automation

- *Standard alignment formats;*
- *Alignment manipulating tools;*
- *Measuring tools;*
- Iterating tools;
- Packaging;
- Test generating tools (random and variability).

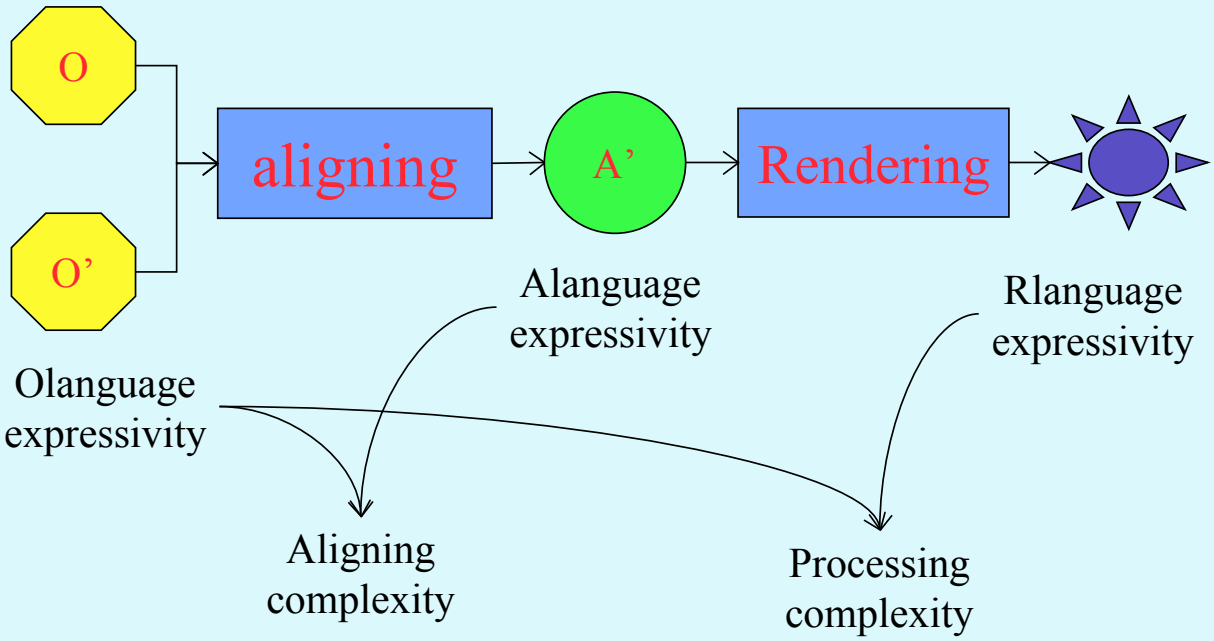
Variability

- Input ontologies: heterogeneity, language, number;
- Input alignment: completion, multiplicity;
- Parameters: Oracle, training, parameters;
- Output alignment: multiplicity, strictness;
- Alignment process: resource constraints, language restrictions, properties (between O and A);

One ultimate goal

- Have stable sets of benchmarks that people having algorithm can compare to;
- Publish results, which are certified once three independent experts have run the soft and found the results;
- Do not want to hear/read, we have got the greatest algorithm ever (and you do not dare citing us) from people not entering evaluations…

Trade-off



http://knowledgeweb.semanticweb.org

http://co4.inrialpes.fr/align

?

http://www.inrialpes.fr/exmo

Jerome.Euzenat@inrialpes.fr