Multi-Agent and Semantic Web Systems: Ontology Matching

Fiona McNeill

School of Informatics

28 February 2013

Ontologies

A notion of relevant knowledge is highly subjective

- Which parts of the world it is important to talk about;
- How to segregate and organise the world;
- What terms to use.

Ontologies are designed by individuals: central control is impossible and undesirable.

Ontologies

- Therefore, ontologies are user- and domain-specific representations of knowledge.
- Ontologies represent only what is in their domain, otherwise they are too large
- Even within a domain, there is an enormous (unlimited?) number of ways of representing knowledge.

Ontologies

- Therefore, ontologies are user- and domain-specific representations of knowledge.
- Ontologies represent only what is in their domain, otherwise they are too large
- Even within a domain, there is an enormous (unlimited?) number of ways of representing knowledge.

Mismatch is common

Ontologies

- But what is mismatch?
  - Your ontology does not fully agree with someone with whom you wish to communicate;
  - Your ontology does not suitably reflect a physical world;
  - etc, …
- Occasionally, there will be a ‘correct’ version;
- Generally, this is completely subjective.

Ontologies

But ontological differences are desirable and essential:

- Freedom of expression;
- Ability to adapt to task;
- Changing environment.

Even direct contradictions can be desirable

- Is a tomato a fruit or a vegetable?

The crucial task is managing these differences.

Ontology Mismatch is Good!

We just need to know how to deal with it …

Ontological Mismatch

- Ontology mismatch occurs when your ontology does not accurately match your world:
  - Other agents with different ontologies
  - Physical world
- Often, this causes no problems – e.g., a robot can interact with the world even though it cannot fully represent the world.
- Sometimes it leads to serious problems.

Borges, "The Analytical Language of John Wilkins"

In a certain Chinese Encyclopedia it is written that animals are divided into:

1. those that belong to the Emperor,
2. embalmed ones
3. those that are trained
4. suckling pigs,
5. mermaids,
6. fabulous ones
7. stray dogs,
8. those included in the present classification,
9. those that tremble as if they were mad,
10. innumerable ones,
11. those drawn with a very fine camelhair brush,
12. others,
13. those that have just broken a flower vase,
14. those that from a long way off look like flies.......

The Problem of Meaning Text

- Meaning is subjective, context-dependent, vague …
- Good ontologies are hard to build; perfect ontologies are impossible to build.
- Ontologies are often built quickly by many different people and maintained and extended haphazardly.

The Problem of Meaning Text

- Meaning is subjective, context-dependent, vague …
- Good ontologies are hard to build; perfect ontologies are impossible to build.
- Ontologies are often built quickly by many different people and maintained and extended haphazardly.
- *Ontology matching will never work!*

Good enough matching

- Achieving perfect matching is impossible in all but tiny domains: not always clear what perfect matching is.

- We need to perform matching that is good enough to adequately meet our needs,
  - for example, to allow successful communication.
  - It may be necessary to only match parts of the ontologies.

- The more similar the ontologies are, the easier matching them will be.

Structural and Semantic matching

It has long been an implicit assumption that what needs to be matched is *words*.

Essentially, this is what you are concerned with when you match class hierarchies

This ignores the possibility that it may be the *representation* itself that may be wrong

This needs to be considered when you are matching structured (e.g., first-order) terms.

Matching of complex ontologies requires both.

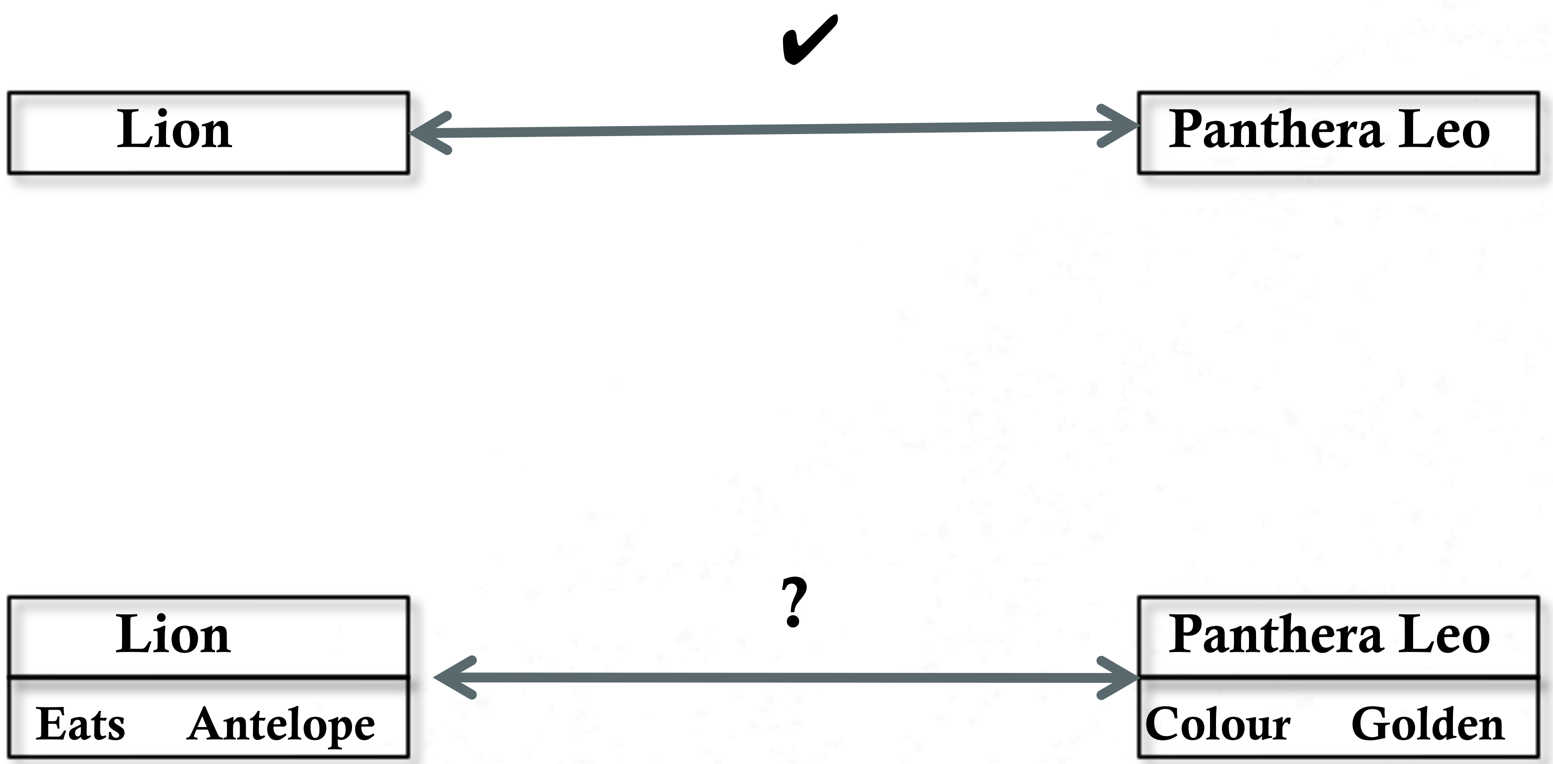
What do we want to match?

Anything in the ontology:

- From simple hierarchies
- To first-order relations, functions, axioms, planning rules, etc.

The more complex your ontology is, the harder the matching is!

What do we want to match?



What do we want to match?

The more complex ontological objects are
- classes, frames, relations, functions, axioms, rules .. the more complex the matching must be.

But ontologies are often very large
    If they are large enough, even very simple ontologies are hard match.

Even matching class hierarchies only is a hard – and unsolved – problem!

Different kinds of matching

Mapping – the most common task:

"the task of relating the vocabulary of two ontologies that share the same domain of discourse in such a way that the mathematical structure of ontological signatures and their intended interpretations, as specified by the ontological axioms, are respected"

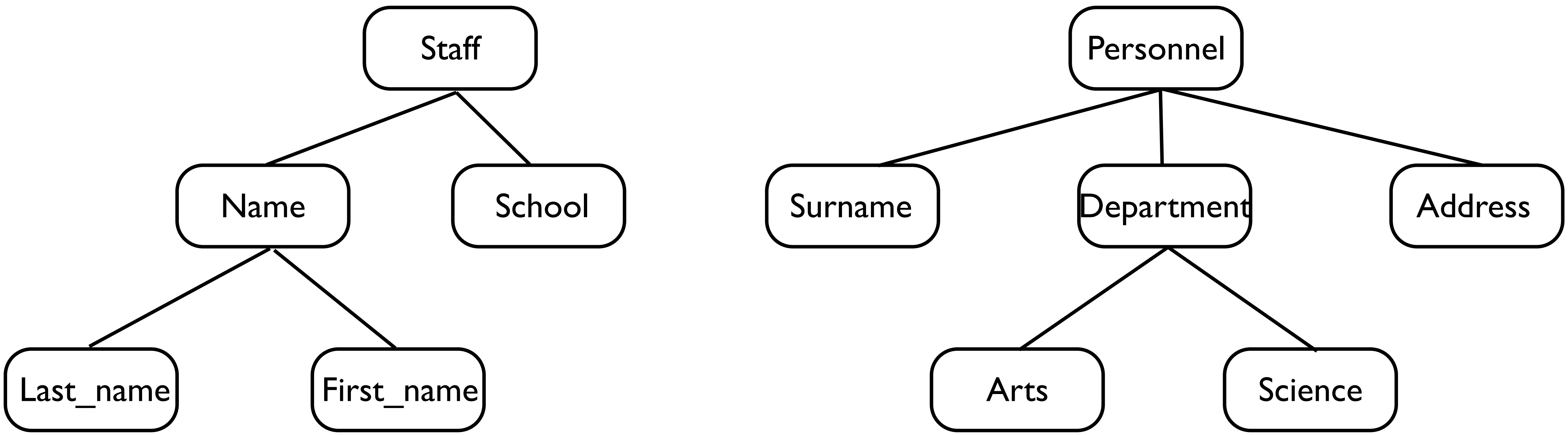
Kalfoglou and Schorlemmer (2003)

Mapping is usually pairwise between individual ontologies;

Many ontologies can be mapped to a single ontology (more scalable but less flexible);

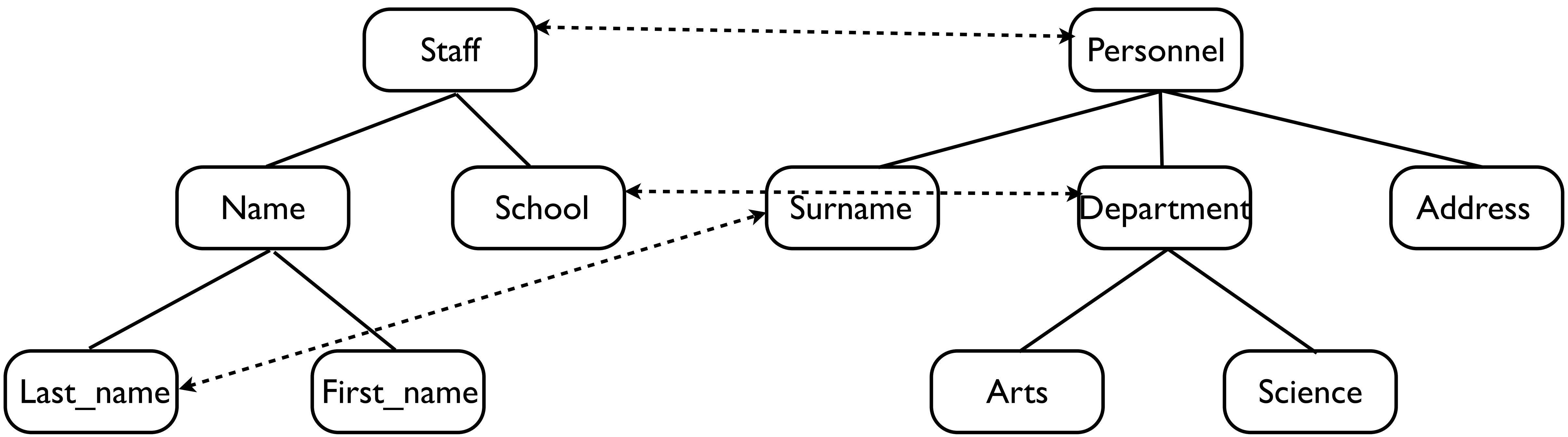
Can be done between 'ontology clusters'.

Different kinds of matching



Different but related ontologies

Different kinds of matching



Result of mapping the two ontologies

Different kinds of matching

Merging – creating a single new ontology.

"the act of building a new ontology by unifying several ontologies into a single one"

Hameed et al, 2003

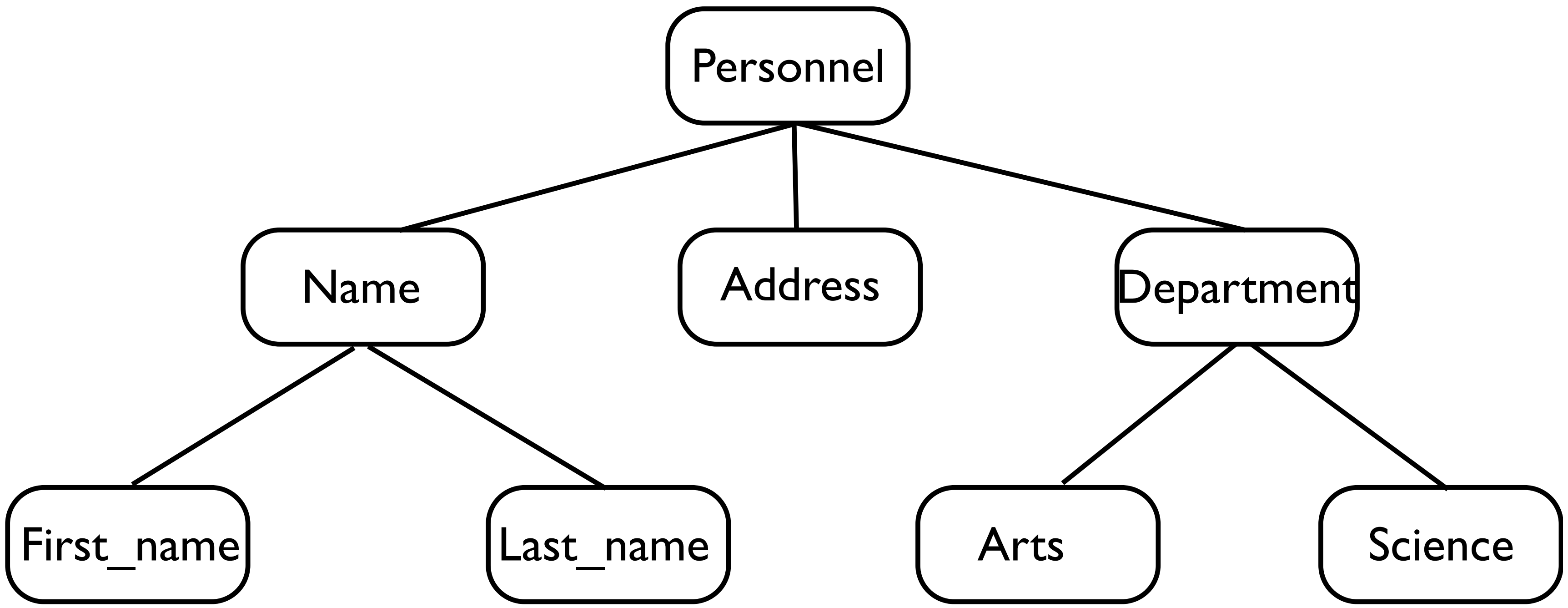
Much less common for agent interaction, but useful in other areas, e.g., database integration.

Generally thought not to be scalable:

"one monolithic information source is not feasible due to unresolvable inconsistencies between them"

Mitra et al, 2000

Different kinds of matching



One possible result of merging the two ontologies

Different kinds of matching

Alignment – a more complete form of mapping

> *"the sources must be made consistent and coherent with one another but kept separately"*
>
> Noy and Musen, 2000

It involves mapping concepts and relations to indicate equivalence.

It is applied to full ontologies.

A special case of mapping?

Different kinds of matching

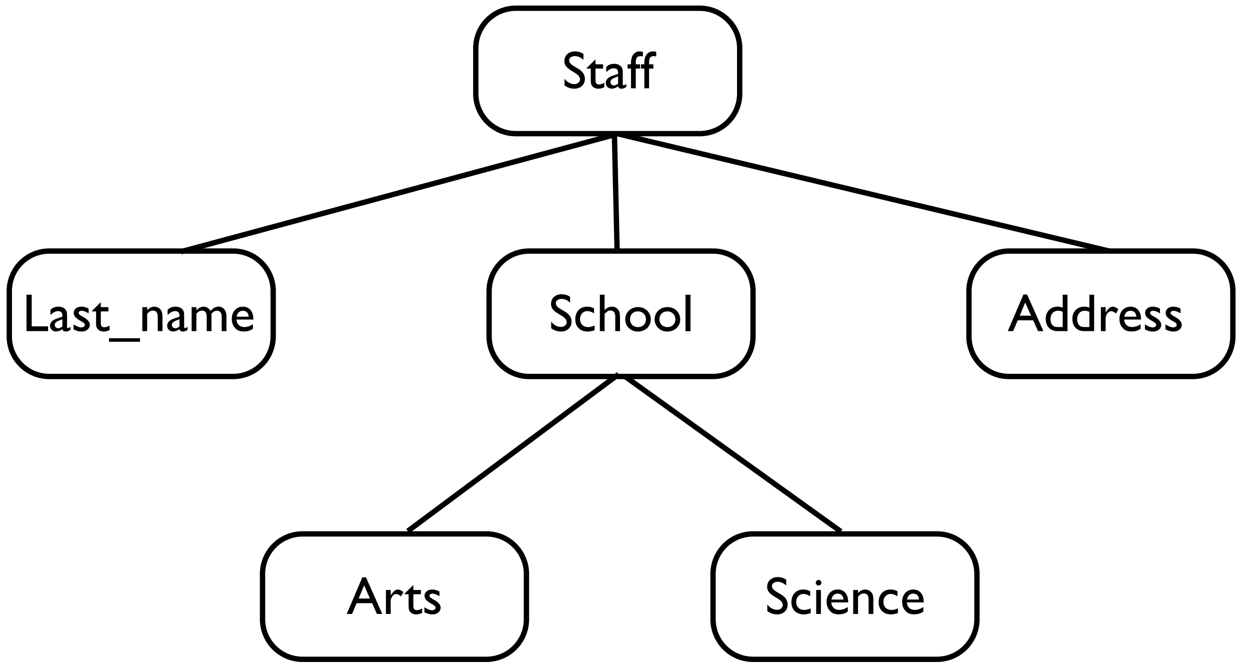
Translation – an implementation of the mapping process.

"whilst ontology mapping defines a collection of functions that specify which concepts and relations correspond to which other concepts and relations, ontology translation is the application of these mapping functions to the sentences based on one ontology into the other"

Kalfoglou and Schorlemmer, 2003

After mapping, your ontology remains unchanged; after translation, your ontology is different.

Different kinds of matching



One possible result of translating one of the ontologies

'Traditional' matching

- Assume full knowledge of both (all) ontologies;
- Assume ontologies are just class hierarchies;
- Match both (all) ontologies;
- Ignore time concerns: this is a slow process and is done offline.
- Much research on this: some successes but still huge difficulties

Dynamic matching

- Assume full knowledge of one ontology but very limited knowledge of other ontology (or world)
  - e.g., agent in MAS, robot in world

- Ontologies can be up to first-order or beyond;

- Match only when necessary, and only what is necessary

- Matching is done during interaction and is time critical

- Much less research: some encouraging early results, but must still to do!

When do we want to match?

Traditional approach: some need for these ontologies to be aligned is identified:

e.g., companies merging; need for frequent interactions between particular agents.

Dynamic approach: some kind of failure has occurred because the ontologies are mismatched:

e.g., agents not understanding one another; robots bumping into walls.

What is the aim of matching?

Traditional approach:

complete alignment between ontologies; no possibility of misunderstandings between agent using these ontologies.

Dynamic approach:

The mismatch that was leading to the observed failure is fixed so that that particular failure no longer occurs. Different failures are always possible.

Ontology consistency

- After mapping or alignment, ontologies have not changed, so no consistency problems are introduced (though they may exist already).

- After merging or translation, changes are introduced, so new inconsistencies may arise.

- If changes are purely semantic (e.g., exchanging one word for an equivalent one), there should be no problems. If the meaning is changed, this could be harder.

Traditional matching: how does it work?

- Find correspondences between nodes:R= $\{\equiv, \subseteq, \supseteq, \perp\}$;

- Relationships between nodes are expressed as: $<ID^{ij}, n_1{}^{i}, n_2{}^{j}, R^{ij}>$

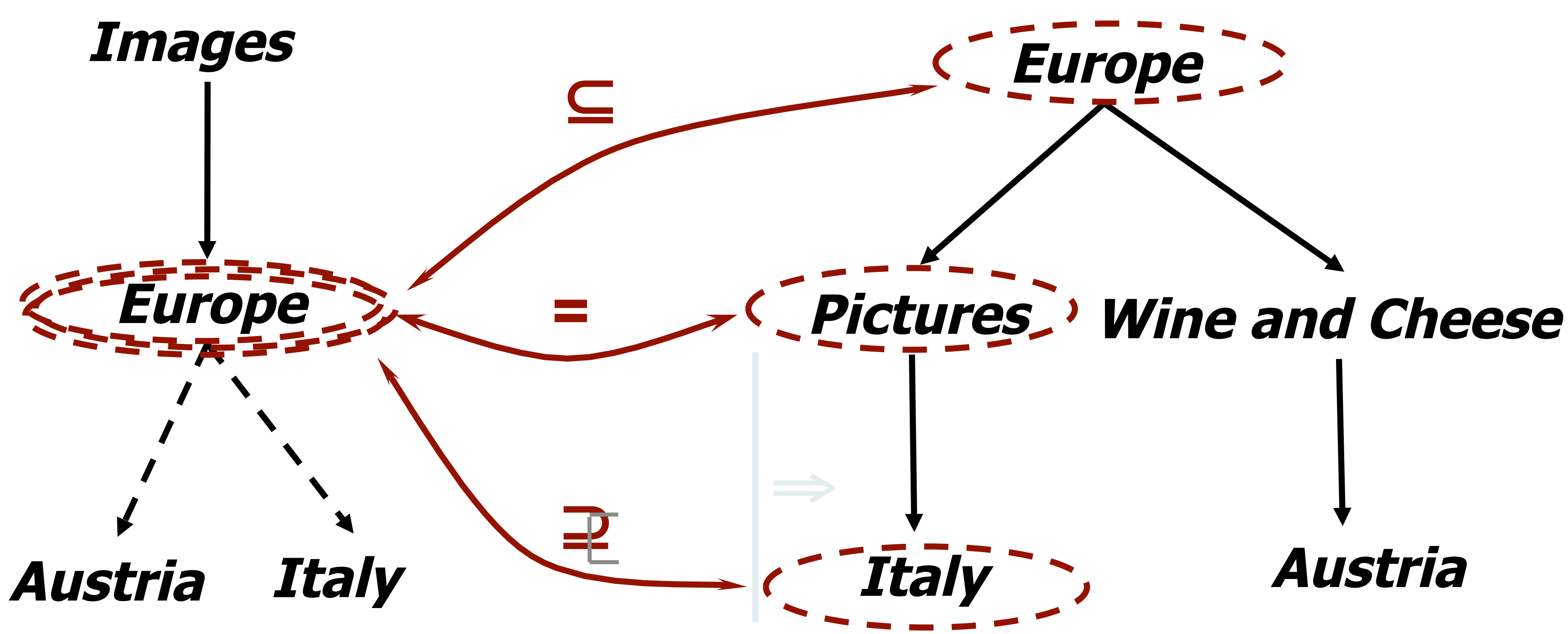
- For each node $n_1{}^{i} \in Ont_1$, find the strongest semantic relationship R' holding with $n_2{}^{j} \in Ont_2$.

- The overall match is determined by the maximising the strongest semantic relationship across all nodes.

Traditional matching: how does it work?

Sometimes the structure of the hierarchy is considered

Example from the S-Match system (Shvaiko, 2004)



Traditional matching: how does it work?

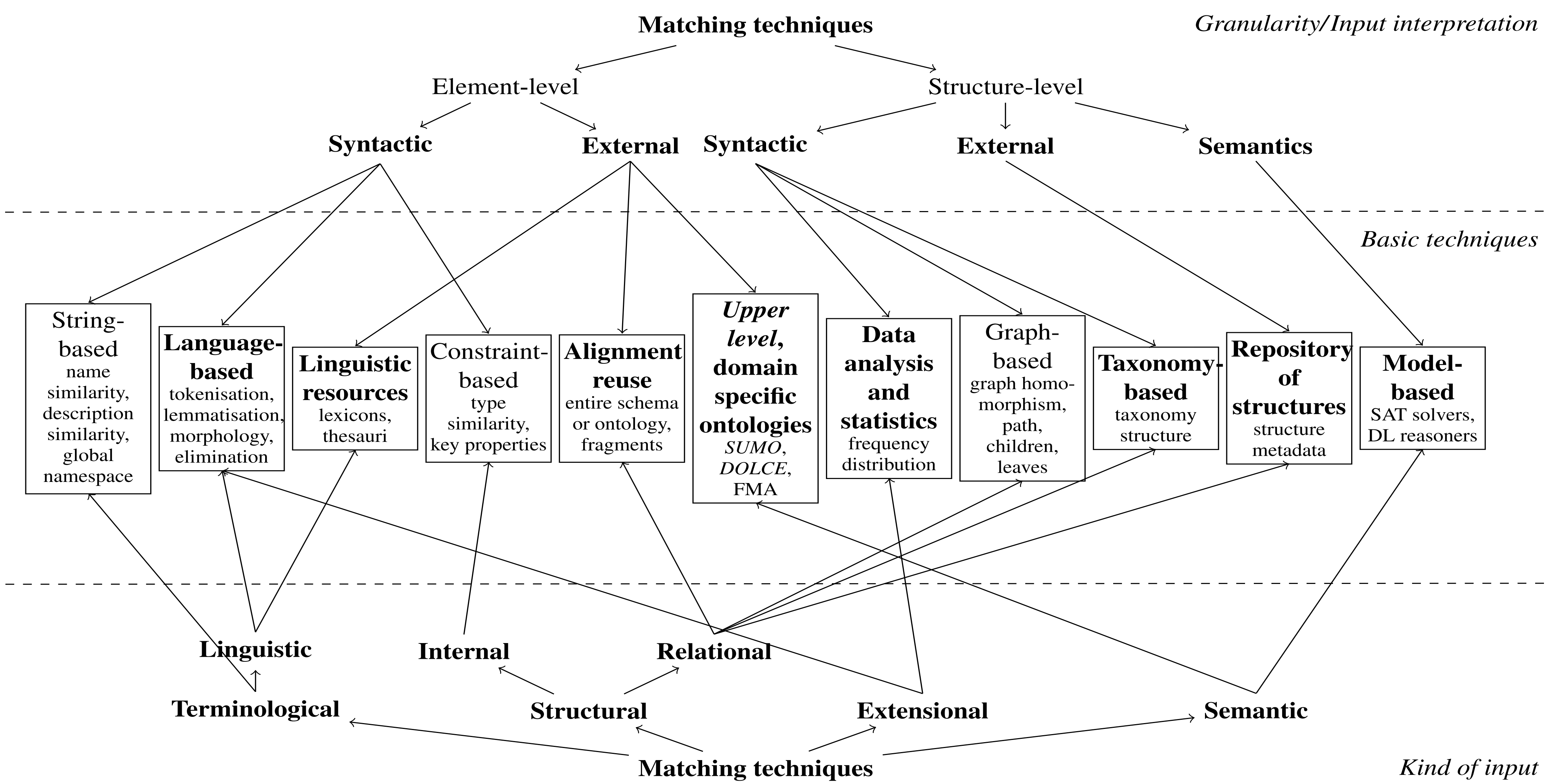


Fig. 3.1. The retained classifications of elementary matching approaches. The upper classification is based on granularity and input interpretation; the lower classification is based on the kind of input. The middle layer features classes of basic techniques. The novelty of this classification in comparison with our previous work in [Shvaiko and Euzenat, 2005] includes *extensional* category of techniques as well as *data analysis and statistics* class of methods.

Traditional matching: Ontology Alignment Evaluation Initiative

Tests run annually at the Ontology Matching workshop;

Systems must fully align sets of two large ontologies written in OWL-DL and RDF/XML;

Points are given for 'correct' matches and taken away for 'incorrect' matches;

Each system can choose which streams they wish to enter;

Only results are evaluated – timing is not considered.

Traditional matching: Ontology Alignment Evaluation Initiative

System	AgrMaker	AROMA	ASMOV	BLOOMS	CODI	Ef2Match	Falcon-AO	GeRMeSMB	LNR2	MapPSO	NBJLM	ObjectRef	RiMOM	SOBOM	TaxoMap	Total=15
Confidence	√	√	√	√		√	√	√	√	√	√	√	√	√	√	
benchmarks	√	√	√		√	√	√	√		√			√	√	√	11
anatomy	√		√	√	√	√		√			√			√	√	9
conference	√	√	√		√	√	√	√						√		8
directory			√					√		√					√	4
di												√	√			2
iimb+pr			√		√				√			√	√			5
Total	3	2	5	1	4	3	2	4	1	2	1	1	2	3	3	37

Dynamic matching

- Hope that our ontology will work;
- Only take action when it doesn't;
- Use the failure to help us identify what went wrong;
- Fix the ontology so as to avoid that particular problem;
- Keep going and hope for the best.

Dynamic matching: advantages

- Makes the problem more tractable
  - Only small parts of the ontology are matched
  - Failure-driven

- This means it can:
  - Be performed online
  - Match complex ontologies (up to first order – or even further)
  - Work in extremely complex environments (e.g. the real world) (theoretically)

- Focuses on an ontology's failure to accurately reflect some 'world' – so, flexible: can match to
  - Other ontologies (MAS, SW)
  - Real world (e.g., robotics)
  - Any other environment in which behaviour can be observed and feedback obtained.

Dynamic matching: what it requires

- Ability to observe failure

- Possibly, an oracle, which will provide some information about the world, e.g.,
    - In an MAS: the agent you are trying to communicate with – will provide answers to yes/no/don't know queries plus instantiations of predicates;
    - In robotics: the physical world, which provides testable feedback (but may be more limited in scope and harder to interpret).
    - Human user?

- Ability to reason about which parts of the ontology are likely to be at fault
- Ability to alter the ontology accordingly
- Ability to test changes through further interaction with the world

Summary: ontology mismatch

Ontology mismatch is inevitable

And brings advantages

We do not want everyone to have to use the same ontology

Ontology mismatch is often not a problem

And we want to ignore it if it isn't

Summary: repairing mismatches

The way this is tackled depends on the situation:

- How complete a solution do you need?
- What are you trying to match the ontology to?
- How large and complex is your ontology?
- How quickly do you need a solution?

Summary: traditional v dynamic matching

Traditional matching is where most of the research is currently focussed and is important

But is completely unusable in many contexts.

Dynamic matching is more flexible and solves problems traditional matching ignores

But can only fix small parts of ontologies

What can we hope to achieve?

Perfect matching is impossible

because of the complexities of language and representation

Sometimes it is not even clear what this means. Instead, we need *good enough* matching

- So agents understand one enough sufficiently for their current purposes;
- So robots can interact successfully with a dynamic and complex world

How close are we to achieving this?

Quite far!

- Traditional matching has been important for a few years (e.g., OM workshop 8 years old)
- Dynamic matching is still not widely studied.

The ability of the Semantic Web to take off is dependent on this problem being solved to a reasonable degree (in my opinion ...)

Robotics can (sometimes) fare better because it is easier to control the environment

But this is a big limitation: real world extremely complex

Reading and task

- Ontology matching: state of the art and future challenges

  Pavel Shvaiko and Jérôme Euzenat

- Create two simple hierarchies and try matching, merging and translating them.