A Categorical Approach to Ontology Alignment

Mihai Codescu, Till Mossakowski, and Oliver Kutz

Institute of Knowledge and Language Engineering
Otto-von-Guericke University of Magdeburg, Germany

Abstract. Ontology matching and alignment is a key mechanism for linking the diverse datasets and ontologies arising in the Semantic Web. We show that category theory provides the powerful abstractions needed for a uniform treatment at various levels: semantics, language design, reasoning and tools. The Distributed Ontology Language DOL is extended in a natural way with constructs for networks of ontologies. We in particular show how the three semantics of Zimmermann and Euzenat can be uniformly and faithfully represented using these DOL language constructs. Finally, we summarise how the DOL alignment features are currently being implemented in the Ontohub/Hets ecosystem, including support for the OWL and Alignment APIs.

1 Introduction

Ontology matching and alignment is a key mechanism for linking the diverse datasets and ontologies arising in the Semantic Web. Matching based on statistical methods is a relatively developed field, with yearly competitions since 2004 comparing the various strengths and weaknesses of existing algorithms [20].

Ontology alignments express semantic correspondences between the entities of different ontologies. The correspondences of an alignment can be various relations, like equivalence, subsumption, disjointness or instance between entities of the ontologies, which can be named entities, like classes, roles, individuals, function symbols etc. or even complex concepts or terms.

The problem of giving an interpretation to alignments in terms of the semantics of the ontologies is complicated by the fact that the domains of interpretation of the two ontologies may be incompatible. Different ways of dealing with this problem exist in the literature. The first solution, called simple semantics in [23], is to assume that the domain of interpretation of the ontologies is uniform [4, 5]. The second solution, called *integrated semantics* in [23], is to assume the existence of a universal domain together with functions relating the domains of individual ontologies to the universal domain. This approach has been introduced in [21], under the name of integrated distributed description logics (IDDL). Finally, the domains of the individual ontologies can be related among themselves directly instead via a unique universal domain. This approach gives rise to the third semantics, called *contextualised semantics* in [23]. It was introduced in [23] as an attempt to generalise a number of existing semantic formalisms (distributed first-order logics (DFOL) [10], distributed description logics (DDL) [2] and contextualised ontologies (C-OWL) [3]) and later corrected to a relational semantics in [22]. Package-based description logics (PDL) [1] also fall in this semantic category. Moreover, [23] discusses the implications of these possible interpretations of alignments with respect to reasoning and composition of alignments.

A major problem with these approaches is their diversity. There exist some attempts for unification, which however remain unsatisfactory: there is no common syntax, no common semantic framework, and no common tool support. In this work, we show how category theory can provide such a unifying framework at various levels, improving previous related work [24, 15, 22, 11] which did not spell out details, and did not make the step from abstract description and case studies to language design and implementation.

2 General approach

The general representation and reasoning framework that we propose includes: 1) a declarative language to specify networks of ontologies and alignments, with independent control over specifying local ontologies and complex alignment relations, 2) the possibility to align heterogeneous ontologies, and 3) in principle, the possibility to combine different alignment paradigms (simple/integrated/contextualised) within one network.

Through category theory, we obtain a unifying framework at various levels:

semantic level We give a uniform semantics for distributed networks of aligned ontologies, using the powerful notion of *colimit*, while reflecting properly the semantic variation points indicated above.

(meta) language level We provide a uniform notation (based on the distributed ontology language DOL) for distributed networks of aligned ontologies, spanning the different possible semantic choices.

reasoning level Using the notion of colimit, we can provide reasoning methods for distributed networks of aligned ontologies, again across all semantic choices.[1]

tool level The tool `ontohub.org` provides an implementation of analysis and reasoning for distributed networks of aligned ontologies, again using the powerful abstractions provided by category theory.

logic level Our semantics is given for the ontology language OWL, but due to the abstraction power of the framework, it easily carries over to other logics used in ontology engineering, like RDFS, first-order logic or F-logic.

This shows that category theory is not only a powerful abstraction at the semantic level, but can properly guide language design and tool implementations and thus provide useful abstraction barriers from a software engineering point of view.

The distributed ontology language DOL is a metalanguage in the sense that it enables the reuse of existing ontologies as building blocks for new ontologies using a variety of structuring techniques, as well as the specification of relationships between ontologies. One important feature of DOL is the ability to combine ontologies that are written in different languages without changing their semantics. A formal specification of the language can be found in [17]. However note that syntax and semantics of DOL alignments is introduced in this paper for the first time.

[1] We do not claim here that the reasoning methods we provide outperform more specialised alignment reasoning methods, say for DDL, or alignment debugging: our main contribution is the provision of a unifying framework that works simultaneously at the various levels.

The general picture is then as follows: existing ontologies can be integrated as-is into the DOL framework. With our new extended DOL syntax, we can specify different kinds of alignments. From such an alignment, we construct a graph of ontologies and morphisms between them—in a way depending on the chosen alignment framework. Sometimes, this step also involves transformations on the ontologies, such as relativisation of the (global) domain using predicates. A network of alignments can then be combined to an integrated alignment ontology via a so-called colimit. Reasoning in a network of aligned ontologies is then the same as reasoning in the combined ontology. Thus, in order to implement a reasoner, it is in principle sufficient to define the relativisation procedure for the local logics and the alignment transformation for each kinds of semantics.

3 Networks of ontologies and their semantics

In this section we recall networks of ontologies and their semantics introduced in [23, 8]. Networks of ontologies (here denoted NeO) [8], called distributed systems in [23], consist of a family $(O_i)_{i \in I}$ of ontologies over a set of indexes I interconnected by a set of alignments $(A_{ij})_{i,j \in I}$ between them. Alignments are sets of *correspondences* between the target ontology O_1 and source ontology O_2 of the alignment. Correspondences are triples (e_1, e_2, R) where e_1 and e_2 are entities built with the help of an entity language over O_1 and O_2, respectively, and R is a relation between entities from a set of relations $\mathfrak{R}$.

A semantics of networks of ontologies is given in terms of local interpretation of the ontologies and alignments it consists of. To be able to give such a semantics, one needs to give an interpretation of the relations between entities that are expressed in the correspondences. In the following three subsections let $S = \{(O_i)_{i \in I}, (A_{ij})_{i,j \in I}\}$ be a NeO over a set of indexes I.

Simple semantics In the simple semantics, the assumption is that all ontologies are interpreted over the same domain (or universe of interpretation) D. The relations in $\mathfrak{R}$ are interpreted as relations over D, and we denote the interpretation of $R \in \mathfrak{R}$ by R^D.

$$\begin{tikzcd}
O_1 \arrow[drr, "m_1"'] & O_2 \arrow[dr, "m_2"] & \dots & O_n \arrow[dl, "m_n"] \\
& & D &
\end{tikzcd}$$

If O_1, O_2 are two ontologies and $c = (e_1, e_2, R)$ is a correspondence between O_1 and O_2, we say that c is satisfied by interpretations m_1, m_2 of O_1, O_2 iff $m_1(e_1)\ R^D\ m_2(e_2)$. This is written $m_1, m_2 \models^S c$. A model of an alignment A between ontologies O_1 and O_2 is then a pair m_1, m_2 of interpretations of O_1, O_2 such that for all $c \in A$, $m_1, m_2 \models^S c$. We denote this by $m_1, m_2 \models^S A$. An interpretation of S is a family $(m_i)_{i \in I}$ of models m_i of O_i. A simple interpretation of S is an interpretation $(m_i)_{i \in I}$ of S over the same domain D.

Definition 1. *[23] A simple model of a S is a simple interpretation $(m_i)_{i \in I}$ of S such that for each $i, j \in I$, $m_i, m_j \models^S A_{ij}$. This is written $(m_i)_{i \in I} \models^S S$. We denote by $Mod^{sim}(S)$ the class of all simple models of S.*

Integrated Semantics Another possibility is to consider that the domain of interpretation of the ontologies of a NeO is not constrained, and a global domain of interpretation U exists, together with a family of *equalising functions* $\gamma_i : D_i \to U$, where D_i is the domain of O_i, for each $i \in I$. A relation R in $\mathfrak{R}$ is interpreted as a relation R^U on the global domain. Satisfaction of a correspondence $c = (e_1, e_2, R)$ by two models m_1 of O_1 and m_2 of O_2 means that $\gamma_i(m_i(e_1))R^U\gamma_j(m_j(e_2))$. We denote this by $m_1, m_2 \models^I_{\gamma_1,\gamma_2} c$ and by $m_1, m_2 \models^I_{\gamma_1,\gamma_2} A$ we denote that $m_1, m_2 \models^I_{\gamma_1,\gamma_2} c$ for each $c \in A$.

$$\begin{tikzcd}
O_1 \arrow[d, "m"'] & O_2 \arrow[d, "m_2"] & \cdots & O_n \arrow[d, "m_n"] \\
D_1 \arrow[drr, "\gamma_1"'] & D_2 \arrow[dr, "\gamma_2"] & \cdots & D_n \arrow[dl, "\gamma_n"] \\
 & & U &
\end{tikzcd}$$

An integrated interpretation of S is then $\{(m_i)_{i\in I}, (\gamma_i)_{i\in I}\}$ where $(m_i)_{i\in I}$ is an interpretation of S and $\gamma_i : D_i \to U$ is a function to a common global domain U for each $i \in I$. We here assume that the γ_i are inclusions.[2]

Definition 2. *[23] An integrated interpretation $\{(m_i), (\gamma_i)\}$ of S is an integrated model of S iff for each $i, j \in I$, $m_i, m_j \models^I_{\gamma_i,\gamma_j} A_{ij}$. We denote by $Mod^{int}(S)$ the class of all integrated models of a NeO S.*

Contextualised Semantics The functional notion of contextualised semantics in [23] is not very useful and has been replaced by a more flexible relational notion subsequently [8], closely related to the semantics of DDLs [2] and $\mathcal{E}$-connections [14].

$$\begin{tikzcd}
O_1 \arrow[d, "m"'] & O_2 \arrow[d, "m_2"] & \cdots & O_n \arrow[d, "m_n"] \\
D_1 \arrow[r, "r_{1,2}"] \arrow[rr, bend right, "r_{1,3}"] & D_2 \arrow[r, "r_{2,3}"] & \cdots & D_n \arrow[lll, bend left, "r_{n,1}"]
\end{tikzcd}$$

The idea is to relate the domains of the ontologies by a family of relations $r = (r_{ij})_{i,j\in I}$. The relations R in $\mathfrak{R}$ are interpreted in each domain of the ontologies in the NeO. Satisfaction of a correspondence $c = (e_1, e_2, R)$ by two models m_1 of O_1 and m_2 of O_2 means that $m_i(e_1)R^i r_{ji}(m_j(e_2))$, where R^i is the interpretation of R in D_i. We denote it by $m_1, m_2 \models^C_r c$, and extend this to alignments, denoted $m_1, m_2 \models^C_r A$ if all correspondences of the alignment are satisfied by m_1, m_2 w.r.t. r.

A contextualised interpretation of S is a pair $\{(m_i)_{i\in I}, (r_{ij})_{i,j\in I}\}$ where $(m_i)_{i\in I}$ is an interpretation of S and $(r_{ij})_{i,j\in I}$ is a family of domain relations such that r_{ij} relates the domain of m_i to the domain of m_j and r_{ii} is the identity (diagonal) relation. Further assumptions about domain relations can be added, thus restricting more the class of interpretations of a NeO.

Definition 3. *A contextualised model of the NeO S is a contextualised interpretation $((m_i)_{i\in I}, (r_{ij})_{i,j\in I})$ of S such that for each $i, j \in I$, $m_i, m_j \models^C_r A_{ij}$. We denote by $Mod^{con}(S)$ the class of all contextualised models of a NeO S.*

---
[2] The theory also works for injections without much change. Arbitrary, i.e. possibly non-injective maps, are conceptually not necessary: a local model can be quotiented by the kernel of a non-injective such map, and then be replaced by the quotient, leading to an injective map again.

4 DOL Alignments

In this section we start by introducing the DOL concepts necessary for giving semantics of alignments. We then introduce the syntax of alignments in DOL and illustrate with the help of an example involving OWL ontologies how the semantics of alignments can be given using diagrams and colimits. We then present the main result of the paper, showing how the categorical semantics of DOL alignments captures the three semantics of networks of ontologies.

4.1 DOL Diagrams and Combinations

The syntax for specifying diagrams in DOL is

`graph D =` $D_1, \ldots, D_m, O_1, \ldots, O_n, M_1, \ldots, M_p, A_1, \ldots, A_k$

where D_i are (sub-)diagrams, O_i are ontologies, M_i are morphisms and A_i are alignments. The user specifies a diagram D formed with the subgraphs given by diagrams D_i, extended with ontologies O_i and the morphisms M_i and the subdiagrams of the alignments A_i

DOL also provides means for combining a diagram of ontologies into a new ontology, such that the symbols related in the diagram are identified. The syntax of combinations is `ontology O = combine D`, where D is a diagram, named or specified as above. The semantics of a combination O is the class of models of the colimit ontology of the diagram specified in the combination. Under rather mild technical assumptions, this model class captures exactly the models of the diagram.

4.2 Syntax of DOL Alignments

DOL represents the general alignment format in a similar way to the Alignment API [7] as follows:

```
alignment A : O_1 to O_2 =
  s_1^1 REL^1 s_2^1, ..., s_1^n REL^n s_2^n
  assuming DOMAIN
end
```

where O_1 and O_2 are the ontologies to be aligned, s_1^i and s_2^i are O_1 and respectively O_2 symbols, for $i = 1, \ldots, n$, s_1^i REL^i s_2^i is a *correspondence* which identifies a relation between the ontology symbols, using one of the symbols $>$ (subsumes), $<$ (is subsumed), $=$ (equivalent), % (incompatible), $\in$ (instance) or $\ni$ (has instance) and `DOMAIN` records whether single, integrated or contextualised semantics is used, using the constant `SingleDomain, GlobalDomain` and `ContextualisedDomain` respectively.

Before starting to analyse the three semantics for NeOs in our setting, we can first define the diagram of a NeO in terms of the diagrams of its parts.

Definition 4. *The diagram of a* NeO $S = \{(O_i)_{i \in I}, (A_{ij})_{i,j \in I}\}$ *is obtained by putting together the diagrams of all alignments* A_{ij} *it consists of.*

The gap to be filled is the construction of the diagram associated with a single assignment, in all three possible assumptions about the semantics. Once this has been given, we can define the semantics of a NeO as the colimit ontology of its associated diagram.

Example 1. We illustrate the three approaches to semantics with the help of a simple example. Let us consider the following two ontologies:

ontology S = **Class:** *Person*
Individual: *alex* **Types:** *Person*
Class: *Child*

ontology T = **Class:** *HumanBeing*
Class: *Male* **SubClassOf:** *HumanBeing*
Class: *Employee*

together with the following correspondences: *S:Person* = *T:HumanBeing*, *S:alex* $\in$ *T:Male* and *S:Child* $\sqsubseteq \neg$ *T: Employee*.

Using the AlignmentAPI syntax, we can write this alignment as

alignment A : S **to** T = *Person* = *HumanBeing*,
alex $\in$ *Male*,
Child < $\neg$ *Employee*

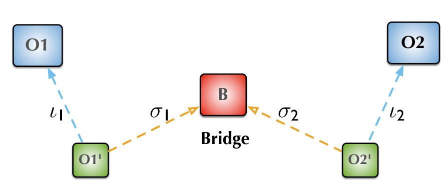
The assumption about the domains of S and T, which determines which of the three semantics is used, is left to be added in the specification of A.

In all three cases, the semantics of the alignment is the class of models of the colimit of the diagram of the alignment, which can be specified in DOL by writing `ontology C = combine A`.

4.3 Simple Semantics

In this simplest case, we simply turn the correspondences into OWL sentences to generate the bridge ontology. Moreover, for each entity occuring in an alignment we want to use both its axiomatisation in the original ontology as well as the bridge axioms introduced by the alignment. For this reason, we keep track of the dependency between the symbols of the bridge ontology and the ontology they have origin from by adding a common source in the diagram for these two occurences. This is a well-known construction, see [24].

Definition 5. *Let A be an alignment (using the notations of Sec. 4.2). The diagram of the alignment is of the following shape (a W-alignment in the sense of [24]):*



Its constituents are obtained as follows. The ontologies O'_1 and O'_2 collect, respectively, all the symbols s_1 and s_2 that appear in a correspondence $s_1 REL\ s_2$ in A, and have no sentences. The morphisms ι_i from O'_i to O_i, where $i = 1, 2$, are inclusions. The ontology B is constructed by turning the correspondences of the alignment into OWL *axioms. The morphisms σ_1 and σ_2 map the symbols occurring in correspondences to their counterpart in B. The alignment is ill-formed when it contains an equivalence between symbols of different kinds, or if B fails to be a well-formed ontology.*

Example 2. We start by adding the assumption that we have a shared domain for the ontologies in the alignment of Ex. 1:

alignment A : S **to** T = ...
    assuming SingleDomain

The diagram of A is then

$$
\begin{tikzcd}
S & & B & & T \\
& S' \arrow[ul, "\iota_1"] \arrow[ur, "\sigma_1"'] & & T' \arrow[ul, "\sigma_2"] \arrow[ur, "\iota_2"'] &
\end{tikzcd}
$$

where S' consists of the concepts *Person* and *Child* and the individual *alex* and T' consists of the concepts *HumanBeing*, *Employee* and *Male*, ι_1 and ι_2 are inclusions and σ_1 and σ_2 map, respectively, *Person* and *HumanBeing* to *Person_HumanBeing* and all other concepts and/or individuals identically.
The bridge ontology B is:

ontology B = **Class:** *Person_HumanBeing*
Class: *Employee*
Class: *Male*
Class: *Child* **SubClassOf:** ¬ *Employee*
Individual: *alex* **Types:** *Male*

The colimit ontology of the diagram of A is:

ontology C = **Class:** *Person_HumanBeing*
Class: *Employee*
Class: *Male* **SubClassOf:** *Person_HumanBeing*
Class: *Child* **SubClassOf:** ¬ *Employee*
Individual: *alex* **Types:** *Male*, *Person_HumanBeing*

4.4 Integrated Semantics

Capturing integrated semantics in DOL using families of models compatible with a diagram is more difficult, as compatibility with the diagram implies uniqueness of the domain. To remedy this, we use relativisation of an ontology where the universal concept becomes a new concept and thus can be interpreted as a subset of the relativised domain. Relativisations have previously been used in defining Common Logic modules [19] or in the re-encoding of DDL into OWL [6].

Definition 6. *Let O be an* OWL *ontology. We define the* relativisation *of O, denoted $\tilde{O}$, as follows. The concepts of $\tilde{O}$ are the concepts of O together with a new concept, denoted $\top_O$. The roles and individuals of $\tilde{O}$ are the same as in O. $\tilde{O}$ contains axioms stating that*

- *each concept C of O is subsumed by $\top_O$,*
- *each individual i of O is an instance of $\top_O$,*
- *each role r has its domain and range, if present, intersected with $\top_O$, otherwise they are $\top_O$.*

and the axioms of O where the following replacement of concepts is made:

- *each occurence of $\top$ is replaced by $\top_O$, and*
- *each concept $\neg C$ is replaced by $\top_O \sqcap \neg C$*
- *each concept $\forall R.C$ is replaced by $\top_O \sqcap \forall R.C$.*

Example 3. We add the assumption that we have a global domain where the domains of the ontologies in our alignment are included:

alignment A : S **to** T = ...
  assuming GlobalDomain

The diagram of A is then

$$\begin{tikzcd} \tilde{S} & & \tilde{B} & & \tilde{T} \\ & S' \arrow[ul, "\iota_1"] \arrow[ur, "\sigma_1"'] & & T' \arrow[ul, "\sigma_2"] \arrow[ur, "\iota_2"'] & \end{tikzcd}$$

where S' consists of the concepts $Thing_S$, *Person* and *Child* and the individual *alex* and T' consists of the concepts $Thing_T$, *HumanBeing*, *Employee* and *Male*, ι_1 and ι_2 are inclusions and σ_1 and σ_2 map *Person* and respectively *HumanBeing* to *Person_HumanBeing* and all other concepts and/or individuals identically.

The relativisations $\tilde{S}$ and $\tilde{T}$ of the ontologies S and T are

ontology $\tilde{S}$ = **Class:** $Thing_S$
Class: *Person* **SubClassOf:** $Thing_S$
Individual: *alex* **Types:** *Person*, $Thing_S$
Class: *Child* **SubClassOf:** $Thing_S$

ontology $\tilde{T}$ = **Class:** $Thing_T$
Class: *HumanBeing* **SubClassOf:** $Thing_T$
Class: *Male* **SubClassOf:** *HumanBeing*, $Thing_T$
Class: *Employee* **SubClassOf:** $Thing_T$

The relativised bridge ontology of an alignment is built by relativising the axioms that result from translating the correspondences of A to OWL sentences. Since we made the assumption that equalising functions are all inclusions, there is no need to introduce explicit symbols for them in the bridge ontology. In our case, the bridge ontology of A is

ontology $\tilde{B}$ = **Class:** $Thing_S$ **Class:** $Thing_T$
Class: *Person_HumanBeing* **SubClassOf:** $Thing_S$, $Thing_T$
Class: *Male* **Class:** *Employee*
Class: *Child* **SubClassOf:** $Thing_T$ **and** $\neg$ *Employee*
Individual: *alex* **Types:** *Male*

The colimit ontology of the relativised diagram of the alignment in Ex. 1 is:

ontology C = **Class:** *ThingS*
Class: *ThingT*
Class: *Person_ HumanBeing* **SubClassOf:** *ThingS*, *ThingC*
Class: *Male* **SubClassOf:** *Person_ HumanBeing*
Class: *Employee* **SubClassOf:** *ThingT*
Class: *Child* **SubClassOf:** *ThingS*
Class: *Child* **SubClassOf:** *ThingT* **and** ¬ *Employee*
Individual: *alex* **Types:** *Male*, *Person_ HumanBeing*

4.5 Contextualised Semantics

Here we need to introduce explicitly the relations between the domains in the language of the bridge ontology. The diagram of the alignment has thus the same shape as in Def. 5, but now the bridge ontology is computed differently and, as in the previous section, the ontologies are relativised. We denote the bridge ontology by $\overline{B}$ and define it to modify B as follows:

- r_{ji} is added to $\overline{B}$ as a role with domain $\top_T$ and range $\top_S$
- the correspondences are translated to axioms involving these roles:
  - $C_i = C_j$ becomes $C_i \equiv \exists r_{ji} \bullet C_j$
  - $a_i = a_j$ becomes $a_i \ r_{ji} \ a_j$
  - $a_i \in C_j$ becomes $a_i \in \exists r_{ji} \bullet C_j$
  - $C_i < C_j$ becomes $C_i \sqsubseteq \exists r_{ji} \bullet C_j$
  - $C_i \% C_j$ becomes $C_i \sqcap \exists r_{ji} \bullet C_j = \emptyset$
- the properties of the r_{ji} are added as axioms in $\overline{B}$.

Here we assume that the alignment A_{ij} contains no correspondence (r_i, r_j, R), where r_i and r_j are roles. Having such correspondences leads to sentences that cannot be expressed in OWL.

Example 4. We add the assumption that we have different domains for the ontologies, which are related by domain relations:

alignment A : S **to** T = ...
assuming ContextualisedDomain

The diagram of A is then

$$\begin{tikzcd}
\tilde{S} & & \overline{B} & & \tilde{T} \\
& S' \arrow[ul, "\iota_1"] \arrow[ur, "\sigma_1"'] & & T' \arrow[ul, "\sigma_2"] \arrow[ur, "\iota_2"'] &
\end{tikzcd}$$

where the constituents of the diagram, except $\overline{B}$, are as defined in Ex. 3. The bridge ontology of A now becomes:

ontology $\overline{B}$ = **Class:** *ThingS*
Class: *ThingT*
ObjectPropery: r_{TS} **Domain:** *ThingT* **Range:** *ThingS*
Class: *Person* **EquivalentTo:** r_{TS} **some** *HumanBeing*
Class: *Employee*
Class: *Male*
Class: *Child* **SubClassOf:** r_{TS} **some** ¬ *Employee*
Individual: *alex* **Types:** r_{TS} **some** *Male*

The colimit ontology of this diagram is:

ontology C = **Class:** *ThingS*
Class: *ThingT*
ObjectPropery: r_{TS} **Domain:** *ThingT* **Range:** *ThingS*
Class: *Person* **EquivalentTo:** r_{TS} **some** *HumanBeing*
Class: *Male* **SubClassOf:** *Person_HumanBeing*
Class: *Employee*
Class: *Child* **SubClassOf:** r_{TS} **some** ¬ *Employee*
Individual: *alex* **Types:** r_{TS} **some** *Male*, *Person*

4.6 The three semantics in DOL

In this section let $S = ((O_i)_{i\in I}, (A_{ij})_{i,j\in I})$ be a network of OWL ontologies. We denote $C(S)$ the colimit ontology of the diagram associated to S, regardless if the assumption about the alignments in S is that they use single, integrated or contextualised semantics. The model class of $C(S)$ is denoted $[\![C(S)]\!]$.

Theorem 1. [3]

1. *If the alignments of S use* `SingleDomain` *and the diagram of S is connected, then $[\![C(S)]\!]$ is in bijection with $Mod^{sim}(S)$.*
2. *If the alignments of S use* `GlobalDomain`*, then $[\![C(S)]\!]$ is in bijection with the class $Mod^{int}(S)$ of integrated models $((m_i), (\gamma_i))$ of S where γ_i are inclusions.*
3. *If the alignments of S use* `ContextualisedDomain`*, then $[\![C(S)]\!]$ is in bijection with $Mod^{con}(S)$.*

DOL is supported by Ontohub (`https://ontohub.org`), a Web-based repository engine for managing distributed heterogenous ontologies. The back-end of Ontohub is the Heterogeneous Tool Set HETS [18] which is used for parsing, static analysis and proof management of ontologies. HETS supports alignments and combinations: it generates the diagram of an alignment according to the assumption on the domain and can compute colimits of OWL ontologies automatically.

5 Conclusions and Future Work

Our theoretical contributions to the foundations of ontology alignment and combination have a potentially large impact on future alignment practices and reasoning. Regardless of the semantic paradigm employed, 'reasoning' with alignments involves at least three levels: (1) the finding/discovery of alignments (often based heavily on statistical methods), (2) the construction of the aligned ontology (the 'colimit'), and (3) reasoning over the aligned result, respectively debugging and repair, closing the loop to (1). Our contributions in this paper address levels (2) and (3).

Regarding (2), platforms such as Bioportal (with hundred thousands of mappings) illustrate that mappings between ontologies, ontology modules, and the concepts and definitions living in them, are of great importance to support re-use. The

[3] For a proof, see `http://iws.cs.uni-magdeburg.de/~mossakow/papers/onto-align.pdf`.

importance of alignment has also been well demonstrated for foundational ontologies in the repository ROMULUS [13]. In the case of Bioportal, the DOL language allows to declaratively manage sets of alignments, and to give precise semantics. In the case of ROMULUS, it allows to align ontologies such as Dolce or BFO expressed in first-order logic with OWL versions of the same ontology.

Regarding (3), alignment tools such as LogMap [12] and ALCOMO [16] employ reasoning over aligned ontologies and repair either parts of the input ontologies or revise the mappings (one technique to enable this is to re-encode the mappings into a global OWL ontology) to restore global consistency. Using DOL and the reasoning capabilities of the Hets/Ontohub ecosystem, such tools could be used to directly operate on a NeO, and to update the diagram structure accordingly.

The approach presented here provides an integration of the major paradigms of ontology alignment in one coherent framework. This includes standard alignment relations, DDLs, PD-L, IDDL, and $\mathcal{E}$-connections [14] which we currently study in more detail. Our construction assumes OWL as the local logic of the ontologies; however it can be generalised to an arbitrary logic by giving a (necessary logic-specific) relativisation procedure and alignment transformation. Moreover, DOL's support for heterogeneity allows us not only to handle heterogeneous alignment, but also to move to a more expressive logic when a bridge axiom cannot be expressed in the local logic of the ontologies. Thus we can remove the restriction on correspondences in the contextualised semantics.

Future work includes the combination of different alignment paradigms within one network (as principally enabled by our unifying framework) and an integration of techniques for the revision of NeOs [9] into DOL. In our setting, the propagation of detected repairs into a network could be done by updating the alignment mappings and re-computing the alignment diagrams. Further work is also needed for the problem of reasoning about the consequences of a NeO; here we expect module extraction to provide an increase in performance of proof search. At the tool level, the integration of the three semantics for alignments in Ontohub is currently in progress. Ontohub is already compatible with the OWL API, and its potential for interoperability is increased further by the integration of the Alignment API.

Acknowledgements We gratefully acknowledge the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open Grant number: 611553, project COINVENT.

References

1. J. Bao, G. Voutsadakis, G. Slutzki, and V. Honavar, 'Package-based description logics', in *Modular Ontologies*, 349–371, Springer, (2009).
2. A. Borgida and L. Serafini, 'Distributed Description Logics: Assimilating Information from Peer Sources', *Journal of Data Semantics*, **1**, 153–184, (2003).
3. P. Bouquet, F. Giunchiglia, F. van Harmelen, L. Serafini, and H. Stuckenschmidt, 'C-OWL: Contextualizing ontologies', in *ISWC*, pp. 164–179, (2003).
4. D. Calvanese, G. De Giacomo, and M. Lenzerini, 'Description logics for information integration', in *Computational Logic: Logic Programming and Beyond*, eds., A. C. Kakas and F. Sadri, volume LNCS 2408, pp. 41–60. Springer, (2002).

5. D. Calvanese, G. De Giacomo, M. Lenzerini, and R. Rosati, 'Logical foundations of peer-to-peer data integration', in *PODS*, eds., C. Beeri and A. Deutsch, pp. 241–251. ACM, (2004).
6. B. Cuenca-Grau and O. Kutz, 'Modular Ontology Languages Revisited', in *Proc. of the IJCAI'07 Workshop on Semantic Web for Collaborative Knowledge Acquisition (SWeCKa), Hyderabad, India, January 2007*, pp. 22–31, (2007).
7. J. David, J. Euzenat, F. Scharffe, and C. Trojahn dos Santos, 'The alignment API 4.0', *Semantic Web*, **2**(1), 3–10, (2011).
8. J. Euzenat and P. Shvaiko, *Ontology matching*, Springer, Heidelberg, 2nd edn., 2013.
9. Jérôme Euzenat, 'Foundations for revising networks of ontologies', in *3rd Int. Workshop on Debugging Ontologies and Ontology Mappings (WoDOOM-14)*, p. 1, (2014).
10. C. Ghidini and L. Serafini, 'Distributed first order logics', in *Frontiers Of Combining Systems 2, Studies in Logic and Computation*, pp. 121–140. Research Studies Press, (1998).
11. P. Hitzler, J. Euzenat, M. Krötzsch, L. Serafini, H. Stuckenschmidt, H. Wache, and A. Zimmermann, 'Integrated view and comparison of alignment semantics', Deliverable D2.2.5 (knowledgeweb), (2006).
12. E. Jiménez-Ruiz and B. Cuenca Grau, 'LogMap: Logic-based and Scalable Ontology Matching', in *Proc. of the 10th International Semantic Web Confernece (ISWC 2011)*, 273–288, Springer, (2011).
13. Z. Khan and M. Keet, 'The foundational ontology library ROMULUS', in *Model and Data Engineering*, 200–211, Springer, (2013).
14. O. Kutz, C. Lutz, F. Wolter, and M. Zakharyaschev, '$\mathcal{E}$-Connections of Abstract Description Systems', *Artificial Intelligence*, **156**(1), 1–73, (2004).
15. O. Kutz, T. Mossakowski, and M. Codescu, 'Shapes of Alignments - Construction, Combination, and Computation', in *Int. Workshop on Ontologies: Reasoning and Modularity (WORM-08)*, eds., U. Sattler and A. Tamilin, volume 348 of *CEUR-WS*, (2008).
16. C. Meilicke, H. Stuckenschmidt, and A. Tamilin, 'Repairing ontology mappings', in *AAAI*, volume 3, p. 6, (2007).
17. T. Mossakowski, O. Kutz, M. Codescu, and C. Lange, 'The Distributed Ontology, Modeling and Specification Language', in *Proc. of the 7th Int. Workshop on Modular Ontologies (WoMO-13)*, ed., C. Del Vescovo et al., volume 1081. CEUR-WS, (2013).
18. Till Mossakowski, Christian Maeder, and Klaus Lüttich, 'The Heterogeneous Tool Set', in *TACAS 2007*, eds., Orna Grumberg and Michael Huth, volume 4424 of *Lecture Notes in Computer Science*, pp. 519–522. Springer-Verlag Heidelberg, (2007).
19. F. Neuhaus and P. Hayes, 'Common Logic and the Horatio problem', *Applied Ontology*, **7**(2), 211–231, (2012).
20. P. Shvaiko and J. Euzenat, 'Ontology matching: State of the art and future challenges', *IEEE Trans. Knowl. Data Eng.*, **25**(1), 158–176, (2013).
21. A. Zimmermann, 'Integrated distributed description logics', in *Description Logics*, eds., D. Calvanese, E. Franconi, V. Haarslev, D. Lembo, B. Motik, A. Turhan, and S. Tessaris, volume 250 of *CEUR Workshop Proceedings*, (2007).
22. A. Zimmermann, 'Logical formalisms for agreement technologies', in *Agreement Technologies*, ed., S. Ossowski, volume 8 of *Law, Governance and Technology Series*, 69–82, Springer Netherlands, (2013).
23. A. Zimmermann and J. Euzenat, 'Three semantics for distributed systems and their relations with alignment composition', in *Proc. 5th International Semantic Web Conference (ISWC)*, LNCS 4273, pp. 16–29, Athens (GA US), (2006).
24. A. Zimmermann, M. Krötzsch, J. Euzenat, and P. Hitzler, 'Formalizing Ontology Alignment and its Operations with Category Theory', in *Proc. of FOIS-06*, pp. 277–288, (2006).