World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:17, No:10, 2023

Merging and Comparing Ontologies Generically

Xiuzhan Guo, Arthur Berrill, Ajinkya Kulkarni, Kostya Belezko, Min Luo

Abstract—Ontology operations, e.g., aligning and merging, were studied and implemented extensively in different settings, such as, categorical operations, relation algebras, typed graph grammars, with different concerns. However, aligning and merging operations in the settings share some generic properties, e.g., idempotence, commutativity, associativity, and representativity, which are defined on an ontology merging system, given by a nonempty set of the ontologies concerned, a binary relation on the set of the ontologies modeling ontology aligning, and a partial binary operation on the set of the ontologies modeling ontology merging. Given an ontology repository, a finite subset of the set of the ontologies, its merging closure is the smallest subset of the set of the ontologies, which contains the repository and is closed with respect to merging. If idempotence, commutativity, associativity, and representativity properties are satisfied, then both the set of the ontologies and the merging closure of the ontology repository are partially ordered naturally by merging, the merging closure of the ontology repository is finite and can be computed, compared, and sorted efficiently, including sorting, selecting, and querying some specific elements, e.g., maximal ontologies and minimal ontologies. An ontology *V*-alignment pair is a pair of ontology homomorphisms with a common domain. We also show that the ontology merging system, given by ontology *V*-alignment pairs and pushouts, satisfies idempotence, commutativity, associativity, and representativity properties so that the merging system is partially ordered and the merging closure of a given repository with respect to pushouts can be computed efficiently.

Keywords—Ontology aligning, ontology merging, merging system, poset, merging closure, ontology *V*-alignment pair, ontology homomorphism, ontology *V*-alignment pair homomorphism, pushout.

I. Introduction

ONTOLOGY has been playing important roles in many areas, such as, database integration, peer-to-peer systems, e-commerce, semantic web services, social networks, etc. [1]. Individual ontologies, however, are not enough to provide the full functionalities. To reuse existing ontologies, one usually wants to integrate different ontologies together. This can be done by combining the ontologies in which they are merged, operated and transformed into one new ontology. In this case, the ontologies have to be aligned with mutual agreement from different ontologies with heterogeneous specifications [2].

Ontology operations, e.g., aligning and merging, are studied and implemented extensively in different settings, such as, categorical operations [2], [3], [4], [5], [6], [7], [8], [9], [10], relation algebras [11], typed graph grammars [12], [13]. In [2], the concept of ontology alignment was formalized by an ontology V-alignment pair $(r_1 : B \to O_1, r_2 : B \to O_2)$ with ontology homomorphisms $r_1 : B \to O_1$ and $r_2 : B \to O_2$ while merging two ontologies O_1 and O_2, for which there is an ontology V-alignment pair $(r_1 : B \to O_1, r_2 : B \to O_2)$, was obtained by the categorical colimit, the pushout of r_1 and r_2, in the category of the ontologies concerned there.

X. Guo, A. Berrill, K. Belezko, and M. Luo are with Royal Bank of Canada, Toronto, ON, Canada (Corresponding author: X. Guo, email: xiuzhan@gmail.com).
A. Kulkarni is with Dataiku.

Recall that a *groupoid* $(G, \circ)$ is a nonempty set G equipped with a binary operation $\circ$ that is a function from $G \times G$ to G and a *semigroup* $(G, \circ)$ is a groupoid $(G, \circ)$ with an *associative* binary operation $\circ$. In some cases, binary operations on a nonempty set G are *partial*, namely, the operations are defined only on a subset of $G \times G$. For example, given two matrices M and N, $M \cdot N$ is defined only when the number of columns in M equals the number of rows in N. Ontology merging can be viewed as a partial binary operation that is defined only on aligning ontology pairs and so the ontologies concerned, along with their merging, form a partial algebraic structure, e.g., a partial groupoid or a partial semigroup, depending on the associativity of the merging operation. Hence we can study the ontology aligning and merging operations together within the partial groupoid or semigroup using the properties the operations share.

A database can grow quickly from a finite set of data tables by generating new tables using the table operations, such as, join, select, combine, etc. Given an ontology *repository* or *instance* $\mathbb{O}$, a finite subset of the ontologies, it is natural to ask *how many new ontologies can be generated from the repository by merging and how to compare and rank these ontologies.* The main objective of this paper is to answer the natural questions by studying the properties ontology aligning and merging operations share and *ontology merging closures*. Even though ontology aligning and merging operations have been formalized in different settings, we shall not study the ontology aligning and merging operations with any specific setting and the internal details of the operations being required but we shall use the generic algebraic properties that ontology aligning and merging operations share, to provide a generic setting for studying ontology aligning and merging operations.

The rest of the paper is organized as follows: First, in Section II, we introduce ontology merging systems and investigate the properties defined on an ontology merging system, which ontology aligning and merging operations share, such as, idempotence, commutativity, associativity, and representativity, denoted by (I), (C), (A), and (R), respectively, described in Example 1.

In a given database, we may concern some specific tables, e.g., the tables containing maximal information, or smallest atomic tables. Given a set of ontologies, to query some specific ontologies from the set, we need to compare and rank the ontologies in the set by a *partial order* on the set. If the merging operation is associative, then the set of the ontologies concerned, along with merging, forms a partial semigroup.

Following partial order approaches on semigroups, in Section III we define the *natural partial order* given by the merging operation (Proposition 2). We show that a given partial order on the ontologies must be the natural partial order if some reasonable properties are satisfied (Theorem 1).

In Section IV, we view the repository $\mathbb{O}$ of ontologies as the generators in a partial algebraic environment and treat the merge closure of $\mathbb{O}$, the smallest set of ontologies, which contains $\mathbb{O}$ and is closed with respect to merging, as the subalgebraic structure generated by the repository $\mathbb{O}$. We prove that the merging closure of a given repository is partially ordered and is finite under some reasonable conditions (Theorem 3) so that we can compare, sort, and select these ontologies generated from the repository, using sorting and selection algorithms in partially ordered sets.

Section V shows that the ontology merging system, given by ontology V-alignment and pushouts, satisfies the algebraic properties introduced in Sections II and III (Theorem 4) so that there are efficient algorithms to compute merging closures and query specific ontologies generated by pushouts.

Finally, we provide our concluding remarks in Section VI.

II. Ontology Merging Systems and Properties

Ontology aligning and merging operations were studied and implemented in different settings extensively. However, the aligning and merging operations in the different settings share some algebraic properties. In this section, we shall model ontology aligning and merging operations together by *ontology merging systems* and introduce the algebraic properties that ontology aligning and merging operations share, defined on an ontology merging system. We shall not specify any ontology setting or any ontology internal details.

Let $\mathfrak{O}$ be the nonempty set of the ontologies concerned, e.g., the set of the objects of the category $\mathfrak{Ont}^+$ of the ontologies defined in [2], the ontology structures considered in [5], or the geospatial data ontologies in [14]. Let $\sim$ be a binary relation on $\mathfrak{O}$ that models a generic ontology alignment relationship and $\Cap$ a partial binary operation on $\mathfrak{O}$ that models a merging operation defined on alignment pairs: For all $O_1, O_2 \in \mathfrak{O}$, $O_1 \Cap O_2$ exists if $O_1 \sim O_2$ and $O_1 \Cap O_2$ is undefined, denoted by $O_1 \Cap O_2 = \uparrow$, otherwise. $(\mathfrak{O}, \sim, \Cap)$ forms an *ontology merging system*. $(\mathfrak{O}, \sim, \Cap)$ is *total* if $O_1 \Cap O_2 \neq \uparrow$ for all $O_1, O_2 \in \mathfrak{O}$. Throughout, $(\mathfrak{O}, \sim, \Cap)$ is an ontology merging system.

The *null extension* $(\mathfrak{O}^\uparrow, \sim^\uparrow, \Cap^\uparrow)$ of $(\mathfrak{O}, \sim, \Cap)$ is given by

- $\mathfrak{O}^\uparrow = \mathfrak{O} \dot{\cup} \{\uparrow\}$;
- for all $O_1, O_2, O \in \mathfrak{O}$, if $O_1 \sim O_2$ then $O_1 \sim^\uparrow O_2$, $O \not\sim^\uparrow \uparrow \not\sim^\uparrow O$, and $\uparrow \sim^\uparrow \uparrow$;
- for all $O_1, O_2 \in \mathfrak{O}$ such that $O_1 \sim O_2$,
$$O_1 \Cap^\uparrow O_2 = O_1 \Cap O_2.$$
for all $O_1, O_2, O \in \mathfrak{O}$ such that $O_1 \not\sim O_2$,
$$O_1 \Cap^\uparrow O_2 = O_2 \Cap^\uparrow O_1 = O \Cap^\uparrow \uparrow$$
$$= \uparrow \Cap^\uparrow O = \uparrow \Cap^\uparrow \uparrow = \uparrow .$$

We begin with the following algebraic properties on $(\mathfrak{O}, \sim, \Cap)$:

(I) *Idempotence*: For all $O \in \mathfrak{O}$, $O \sim O$ and $O \Cap O = O$. An ontology always aligns with itself, and merging it with itself yields the same ontology.

(C) *Commutativity*: For all $O_1, O_2 \in \mathfrak{O}, O_1 \sim O_2$ if and only if $O_2 \sim O_1$, and if $O_1 \sim O_2$ then $O_1 \Cap O_2 = O_2 \Cap O_1$. Aligning is symmetric and merging does not depend on the order of the ontologies being merged.

In general, the associativity of a binary operation means that rearranging the parentheses in an expression will produce the same result. There are many ways to specify the associative properties. In this paper, we are interested in the following three associative properties: (A), (CA), and (SA).

(A) *Associativity in [15]*: For all $O_1, O_2, O_3 \in \mathfrak{O}$ such that $O_1 \Cap (O_2 \Cap O_3)$ and $(O_1 \Cap O_2) \Cap O_3$ exist,
$$O_1 \Cap (O_2 \Cap O_3) = (O_1 \Cap O_2) \Cap O_3 \neq \uparrow .$$
Merging three ontologies O_1, O_2, O_3 yields the same result regardless of the orders in which O_1, O_2, O_3 are binarily grouped when both $O_1 \Cap (O_2 \Cap O_3)$ and $(O_1 \Cap O_2) \Cap O_3$ exist.

(CA) *Catenary associativity*: For all $O_1, O_2, O_3 \in \mathfrak{O}$ such that $O_1 \Cap O_2$ and $O_2 \Cap O_3$ exist,
$$O_1 \Cap (O_2 \Cap O_3) = (O_1 \Cap O_2) \Cap O_3 \neq \uparrow .$$
Merging three ontologies O_1, O_2, O_3 yields the same result regardless of the orders in which O_1, O_2, O_3 are binarily grouped when both $O_1 \Cap O_2$ and $O_2 \Cap O_3$ exist.

(SA) *Strong associativity:* For all $O_1, O_2, O_3 \in \mathfrak{O}$,
$$O_1 \Cap (O_2 \Cap O_3) = (O_1 \Cap O_2) \Cap O_3$$
in the sense that both sides of the equation yield the same value or both are simultaneously undefined. It is equivalent to either of $(O_1 \Cap O_2) \Cap O_3$ and $O_1 \Cap (O_2 \Cap O_3)$ is defined so is the other and they are equal.

We consider the following representativity properties (R_l), (R_r), and (R).

(R_l) *Left representativity*: Let $O_1, O_2, O_3 \in \mathfrak{O}$ such that $O_3 = O_1 \Cap O_2$. For all $O \in \mathfrak{O}$ such that $O \sim O_1$, $O \sim O_3$.

(R_r) *Right representativity*: Let $O_1, O_2, O_3 \in \mathfrak{O}$ such that $O_3 = O_1 \Cap O_2$. For all $O \in \mathfrak{O}$ such that $O_1 \sim O$, $O_3 \sim O$.

(R) *Representativity*: (R_l) and (R_r). The ontology O_3 obtained from merging ontologies O_1 and O_2, represents the original O_1 and O_2: the ontology O that has matched O_1 will also match O_3 from both left and right sides.

Some relationships of the above ontology properties are summarized in Proposition 1. To improve the readability, the nontrivial proofs for the proposition and subsequent propositions and theorems will be provided in Appendix B while the lemmas needed will be shown in Appendix A, without the noise of these lengthy proofs in the main text. The main results will be restated in Appendix B.

Proposition 1. *Let $(\mathfrak{O}, \sim, \mathbin{\wedge\!\!\!\wedge})$ be an ontology merging system.*

(i) If $(\mathfrak{O}, \sim, \mathbin{\wedge\!\!\!\wedge})$ satisfies (SA), then $(\mathfrak{O}, \sim, \mathbin{\wedge\!\!\!\wedge})$ satisfies (A);
(ii) If $(\mathfrak{O}, \sim, \mathbin{\wedge\!\!\!\wedge})$ satisfies (CA), then $(\mathfrak{O}, \sim, \mathbin{\wedge\!\!\!\wedge})$ satisfies (A);
(iii) If $(\mathfrak{O}, \sim, \mathbin{\wedge\!\!\!\wedge})$ satisfies (C), then (CA) is equivalent to (A) and (R);
(iv) $(\mathfrak{O}, \sim, \mathbin{\wedge\!\!\!\wedge})$ satisfies (SA) if and only if the null extension $(\mathfrak{O}^{\uparrow}, \sim^{\uparrow}, \mathbin{\wedge\!\!\!\wedge}^{\uparrow})$ is a total ontology merging system such that $(O_1 \mathbin{\wedge\!\!\!\wedge}^{\uparrow} O_2) \mathbin{\wedge\!\!\!\wedge}^{\uparrow} O_3 = O_1 \mathbin{\wedge\!\!\!\wedge}^{\uparrow} (O_2 \mathbin{\wedge\!\!\!\wedge}^{\uparrow} O_3)$ for all $O_1, O_2, O_3 \in \mathfrak{O}^{\uparrow}$, namely, $(\mathfrak{O}^{\uparrow}, \mathbin{\wedge\!\!\!\wedge}^{\uparrow})$ is a semigroup.

Example 1. 1. Define $(\mathfrak{O}, \asymp, \uplus)$ by
- for all $O_1, O_2 \in \mathfrak{O}$, $O_1 \asymp O_2$,
- $O_1 \uplus O_2$ is the disjoint union of O_1 and O_2.

It is straightforward to verify that $(\mathfrak{O}, \asymp, \uplus)$ is an ontology merging system satisfying (C), (A), (CA), (SA), and (R) but not (I).

2. Let $\mathfrak{G}$ be the set of the graphs that represent the ontologies concerned as in [12]. Define $(\mathfrak{G}, \sim, \between)$ by
   - for all $G_1, G_2 \in \mathfrak{G}$, $G_1 \sim G_2$ if and only if there exists an overlapping subgraph S (up to graph isomorphism) of G_1 and G_2,
   - $G_1 \between G_2$ is the union of G_1 and G_2 by identifying their overlapping subgraph S when $G_1 \sim G_2$.

   Then $(\mathfrak{G}, \sim, \between)$ is an ontology merging system satisfying (I), (C), (A), (CA), (SA), and (R).
3. In [14], Sun et al. defined

$$\mathbf{GeoDataOnt} =$$
$$\{(E, R_{(E_i, E_j)}) \mid E_i, E_j \in E, 0 \le i, j \le |E|\},$$

   where E is the set of the geographic entities concerned and R the set of relations between the entities from E. **GeoDataOnt** can be represented as a table of geographic entities and their relations, labeled by E and R, respectively. Let $\mathfrak{Geo}$ be set of such tables. For $\mathbb{G}_1, \mathbb{G}_2 \in \mathfrak{Geo}$, define

$$\mathbb{G}_1 \smile \mathbb{G}_2 \text{ for all } \mathbb{G}_1, \mathbb{G}_2 \in \mathfrak{Geo}$$

   and

$$\mathbb{G}_1 \bowtie \mathbb{G}_2 = \mathbb{G}_1 \text{ full join } \mathbb{G}_2.$$

   Then $(\mathfrak{Geo}, \smile, \bowtie)$ is an ontology merging system satisfying all above properties: (I), (C), (A), (CA), (SA), and (R).
4. Let $(\mathfrak{O}, \approx, \sqcup)$ be the ontology merging system, which we shall define formally in Definition 3 below, given by
   - for all $O_1, O_2 \in \mathfrak{O}$, $O_1 \approx_B O_2$ if and only if there is a pair of ontology homomorphisms $(r_1 : B \to O_1, r_2 : B \to O_2)$,
   - if $O_1 \approx_B O_2$, then $O_1 \sqcup_B O_2$ is given by the pushout of r_1 and r_2.

   Then $(\mathfrak{O}, \approx, \sqcup)$ satisfies (I), (C), (A), (CA), (SA), and (R), which will be proved in Theorem 4 below.
5. In [5], Cafezeiro and Haeusler defined an ontology homomorphism between ontology structures introduced in [16], as a pair of functions (f, g), where f is a function between the concepts and g a function between relations, which preserve the ontology structures. They aligned ontologies by V-alignment pairs and merged aligned ontologies by pushouts in the category of ontology structures. Similarly, by Theorem 4 below, the ontology merging system satisfies (I), (C), (A), (CA), (SA), and (R).

III. Natural Partial Order Given by Merging

A partial order on a semigroup $(S, \circ)$ is called *natural* if it is defined by means of the operation $\circ$. Recall that the set E_S of all idempotents of S is partially ordered by

$$e \le_E f \text{ if and only if } e = e \circ f = f \circ e.$$

In [17], Mitsch defined a natural partial order $\le_M$ that extends the ordering $\le_E$ from E_S to S by:

$$\text{for all } a, b \in S, a \le_M b \text{ if and only if there exist}$$
$$s, t \in S^1 \text{ such that } a = s \circ b = b \circ t, s \circ a = a,$$

where S^1 is the semigroup obtained by adjoining the identity 1 to S. Mitsch proved that $\le_M$ is equal to $\le_E$ when restricted to E_S [17].

Ontology merging $\mathbin{\wedge\!\!\!\wedge}$ aims to obtain more information by combining the aligned ontologies together. The natural ontology partial order $O_1 \le_{\wedge\!\!\wedge} O_2$ is defined if merging O_1 to O_2 does not yield the more information than O_2. Clearly $\le_{\wedge\!\!\wedge}$ is the dual order of $\le_E$ when restricted to the idempotents.

Definition 1. *For all $O_1, O_2 \in \mathfrak{O}$, $O_1 \le_{\wedge\!\!\wedge} O_2$ if and only if $O_1 \sim O_2$, $O_2 \sim O_1$, and $O_1 \mathbin{\wedge\!\!\!\wedge} O_2 = O_2 \mathbin{\wedge\!\!\!\wedge} O_1 = O_2$.*

Proposition 2 blow shows that $(\mathfrak{O}, \le_{\wedge\!\!\wedge})$ is a *partially ordered set* (*poset*), namely, $\le_{\wedge\!\!\wedge}$ is a *partial order*: a *reflexive, antisymmetric, and transitive binary relation*, on $\mathfrak{O}$, when (I) and (CA) are satisfied.

Proposition 2. *If $(\mathfrak{O}, \sim, \mathbin{\wedge\!\!\!\wedge})$ satisfies (I) and (CA), then $\le_{\wedge\!\!\wedge}$ is a partial order on $\mathfrak{O}$ and so $(\mathfrak{O}, \le_{\wedge\!\!\wedge})$ is a poset.*

Since $(\mathfrak{O}, \le_{\wedge\!\!\wedge})$ is a poset if (I) and (CA) are satisfied, we can compare ontologies in $\mathfrak{O}$ and apply the sorting and selection algorithms in posets and for transitive relations [18] to query the specific ontologies in $\mathfrak{O}$, e.g., maximal ontologies, minimal ontologies.

Let $\preceq$ be a given partial order on $\mathfrak{O}$. We consider the following properties with respect to $\preceq$:

(LU) *Merge gives the least upper bound*: For all ontologies O_1, O_2 such that $O_1 \sim O_2$, $O_1 \mathbin{\wedge\!\!\!\wedge} O_2$ is the least upper bound of O_1 and O_2 with respect to $\preceq$.

(CP_l) *$\preceq$ is left compatible with $\sim$ and $\mathbin{\wedge\!\!\!\wedge}$*: For all ontologies O_1, O_2, O such that $O_1 \preceq O_2$ and $O \sim O_1$, $O \sim O_2$ and $O \mathbin{\wedge\!\!\!\wedge} O_1 \preceq O \mathbin{\wedge\!\!\!\wedge} O_2$.

(CP_r) *$\preceq$ is right compatible with $\sim$ and $\mathbin{\wedge\!\!\!\wedge}$*: For all ontologies O_1, O_2, O such that $O_1 \preceq O_2$ and $O_1 \sim O$, $O_2 \sim O$ and $O_1 \mathbin{\wedge\!\!\!\wedge} O \preceq O_2 \mathbin{\wedge\!\!\!\wedge} O$.

(CP) *$\preceq$ is compatible with $\sim$ and $\mathbin{\wedge\!\!\!\wedge}$*: If $\preceq$ is both left and right compatible with $\sim$ and $\mathbin{\wedge\!\!\!\wedge}$.

A partial order on $(\mathfrak{O}, \sim, \mathbin{\wedge\!\!\!\wedge})$, where $\sim$ is reflexive and commutative, must be the natural partial order $\le_{\wedge\!\!\wedge}$ if (LU) and (CP) are satisfied, shown in Theorem 1.

Theorem 1. *Let $\sim$ be reflexive and commutative and $\preceq$ a partial order on $(\mathfrak{O}, \sim, \wedge\!\!\wedge)$. Then $(\mathfrak{O}, \sim, \wedge\!\!\wedge)$ satisfies (LU) and (CP) with respect to $\preceq$ if and only if $(\mathfrak{O}, \sim, \wedge\!\!\wedge)$ satisfies (I), (C), (A), and (R) and $\preceq = \leq_{\wedge\!\!\wedge}$.*

IV. Ontology Merging Closures and Finiteness

An ontology *repository or instance* in $(\mathfrak{O}, \sim, \wedge\!\!\wedge)$ is a finite subset $\mathbb{O} \subseteq \mathfrak{O}$. Given an ontology repository $\mathbb{O}$, it is natural to ask: *How many new ontologies can be generated from $\mathbb{O}$ by merging operation $\wedge\!\!\wedge$ and how to compare and rank the ontologies?* In this section, we shall provide our answers to the questions by introducing the merging closure of $\mathbb{O}$.

Definition 2. *Given a repository $\mathbb{O} \subseteq \mathfrak{O}$, the* merging closure *of $\mathbb{O}$, denoted by $\widehat{\mathbb{O}}$, is the smallest set $\mathbb{P} \subseteq \mathfrak{O}$ such that*

1. $\mathbb{O} \subseteq \mathbb{P}$,
2. *$\mathbb{P}$ is closed with respect to merging: for all $O_1, O_2 \in \mathbb{P}$ such that $O_1 \sim O_2$, $O_1 \wedge\!\!\wedge O_2 \in \mathbb{P}$.*

The existence and uniqueness of the merging closure of a given ontology repository are given by Theorem 2.

Theorem 2. *Given a repository $\mathbb{O} \subseteq \mathfrak{O}$, the merging closure $\widehat{\mathbb{O}}$ exists and it is unique.*

Constructively, $(\mathfrak{O}, \wedge\!\!\wedge)$ being considered as a partial groupoid, $\widehat{\mathbb{O}}$ is the partial subgroupoid generated by $\mathbb{O}$ with the partial binary operation $\wedge\!\!\wedge$. Hence

$$\widehat{\mathbb{O}} = \bigcup_{n=1}^{+\infty} \mathbb{O}^n,$$

where $\mathbb{O}^n$ is defined inductively by:

- if $n = 1$, $\mathbb{O}^n = \mathbb{O}$,
- if $n > 1$, $O \in \mathbb{O}^n$ if and only if

$$\left\{ \begin{array}{l} \text{there is a positive integer } k : 1 \leq k < n \\ \text{there is } O_1 \in \mathbb{O}^k \\ \text{there is } O_2 \in \mathbb{O}^{n-k} \end{array} \right\}$$

  such that $O = O_1 \wedge\!\!\wedge O_2$.

Obviously, the product $\mathbb{O}^n$ is the set of all possible values

$$O_1 \wedge\!\!\wedge \cdots (O_i \wedge\!\!\wedge O_{i+1}) \cdots \wedge\!\!\wedge O_n,$$

where all $O_1 \wedge\!\!\wedge \cdots O_i \cdots \wedge\!\!\wedge O_n$ are binarily grouped with the parentheses "(" and ")" being inserted and $O_i \in \mathbb{O}, i = 1, \cdots, n$.

To compute $\widehat{\mathbb{O}}$ efficiently, we hope that it is finite, which is implied by the properties (I), (C), (A), and (R) due to Corollary A.1 [15]. Also, by Proposition 2, $(\widehat{\mathbb{O}}, \leq_{\wedge\!\!\wedge})$ is a poset. Hence we have:

Theorem 3. *For each ontology repository $\mathbb{O}$, if $\widehat{\mathbb{O}}$ satisfies (I), (C), (A), and (R), then $\widehat{\mathbb{O}}$ is a finite poset with the natural order $\leq_{\wedge\!\!\wedge}$ given by $\wedge\!\!\wedge$.*

Remark 1.
1. The study of finiteness conditions of semigroups consists in giving some conditions to assure the finiteness of such semigroups. In this study, one of the conditions which is generally required is that of being *finitely generated*. A semigroup is *periodic* if only finitely many elements can be generated from each element e, namely, $\{e, e^2, \cdots, e^n, \cdots\}$ is finite. *Burnside problem for semigroups* asks if a periodic finitely generated semigroup is finite. This problem has a negative answer in general [19], [20]. However, the answer is positive if the property of commutativity or idempotence is required: If the finitely generated semigroup is commutative, then its finiteness is trivial. For the finiteness of a finitely generated idempotent semigroup, see [21].
2. Given an ontology merging system $(\mathfrak{O}, \sim, \wedge\!\!\wedge)$ and an ontology repository $\mathbb{O} \subseteq \mathfrak{O}$, $(\mathfrak{O}, \wedge\!\!\wedge)$ is a partial groupoid or semigroup, depending on the associativity of $\wedge\!\!\wedge$. Hence the finiteness of $\widehat{\mathbb{O}}$ requires the *Burnside problem for partial groupoids or semigroups*.

V. Ontology Merging Given by Pushouts

Recall that given two sets X and Y, the diagram

$$\begin{array}{ccc} X \cap Y & \hookrightarrow & Y \\ \downarrow & & \downarrow \\ X & \hookrightarrow & X \cup Y \end{array} \qquad (1)$$

with all inclusion functions, is a pushout in the category **Set** of sets and functions between sets. If $X \cap Y = \emptyset$, then the pushout square (1) turns out to be the following pushout:

$$\begin{array}{ccc} \emptyset & \hookrightarrow & Y \\ \downarrow & & \downarrow \\ X & \hookrightarrow & X \coprod Y \end{array}$$

where $\emptyset$ is the initial object in **Set** and the coproduct $X \coprod Y = X \cup Y$.

For a bit more general case, let $f : B \to X$ and $g : B \to Y$ be two functions. When f and g are injective, $f(B)$ and $g(B)$ can be viewed as the different labels of B in X and Y, respectively. One can form the disjoint union $X \uplus Y$ and then identify $f(b)$ with $g(b)$ for $b \in B$ as they are the different labels of B in X and Y, respectively. Hence one has the following pushout:

$$\begin{array}{ccc} B & \xrightarrow{g} & Y \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle \iota_Y} \\ X & \xrightarrow{\iota_X} & X \sqcup_B Y \end{array} \qquad (2)$$

where $X \sqcup_B Y = X \uplus Y / \doteq$ and $\doteq$ is the smallest equivalent relation generated by

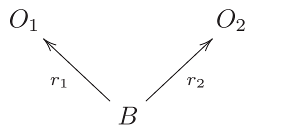
$$\{(f(b), g(b)) \mid b \in B\}.$$

Now let's return to ontology merging operation. Generally, if two ontologies are entirely unrelated, clearly they can be combined by their disjoint union as in (1). On the other hand, merging two ontologies which *overlap*, up to relabeling, should lead to a new ontology that identifies the overlapping elements and keeps related elements apart as far as it is possible without violating the requirements on the ontology structures [2]. This is similar to the case shown in (2) above.

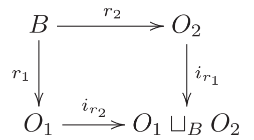
Hence the merge of ontologies O_1 and O_2 can be obtained by a categorical colimit, pushout, shown in [2], [3], [5], [13], when there exists an ontology V-alignment pair between O_1 and O_2.

Recall the definition of V-alignment pairs, in which ontologies O_1 and O_2 overlap after relabeling B in O_1 and O_2, respectively:

Definition 3. 1. *An ontology* V-alignment pair $(r_1 : B \to O_1, r_2 : B \to O_2)$ *is given by a pair of ontology homomorphisms with a common domain:* $r_1 : B \to O_1$ *and* $r_2 : B \to O_2$*, displayed as*



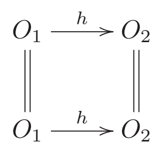
2. *Let the merging system* $(\mathfrak{O}, \approx, \sqcup)$ *be given by*
   - $O_1 \approx_B O_2$ *if and only if there is an ontology* V-*alignment pair* $(r_1 : B \to O_1, r_2 : B \to O_2)$,
   - *if* $O_1 \approx_B O_2$ *then* $O_1 \sqcup_B O_2$ *is given by the pushout:*



Given two ontologies O_1 and O_2, notice that both definitions of $\approx_B$ and $\sqcup_B$ in Definition 3 above depend on a base ontology B. Hence the merge of ontologies O_1 and O_2 given by pushouts depends on the base ontology B. We assume that the overlaps of O_1 and O_2 are identified *as many as possible* when merging O_1 and O_2 using pushouts.

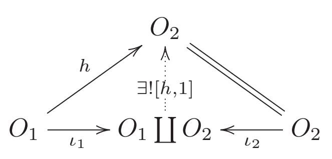
If there exists an initial object I in the category of the ontologies concerned, then there is an ontology V-alignment pair $(!_{O_1} : I \to O_1, !_{O_2} : I \to O_2)$ and $O_1 \sqcup_I O_2 = O_1 \coprod O_2$, the coproduct of O_1 and O_2.

If there is an ontology homomorphism $h : O_1 \to O_2$, then there is an ontology V-alignment pair $(1_{O_1} : O_1 \to O_1, h : O_1 \to O_2)$ and $O_1 \sqcup_{O_1} O_2 = O_2$ since



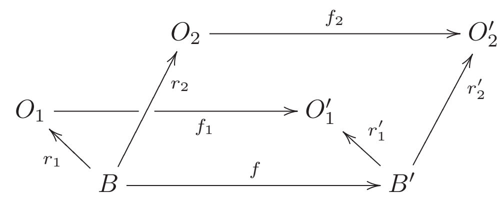
is a pushout. Clearly, there is a unique ontology homomorphism $[h, 1] : O_1 \sqcup_I O_2 \to O_1 \sqcup_{O_1} O_2 = O_2$ as

$$O_1 \sqcup_I O_2 = O_1 \coprod O_2 \to O_2 = O_1 \sqcup_{O_1} O_2 :$$



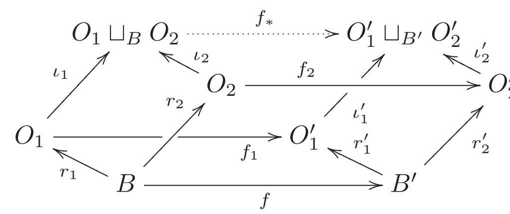
To understand how an ontology homomorphism $O_1 \sqcup_B O_2 \to O_1 \sqcup_{B'} O_2$ is induced after the base being changed from B to B', let's define a *homomorphism* between ontology V-alignment pairs.

Definition 4. *A* homomorphism *from an ontology* V*-alignment pair* $(r_1 : B \to O_1, r_2 : B \to O_2)$ *to an ontology* V*-alignment pair* $(r'_1 : B' \to O'_1, r'_2 : B' \to O'_2)$ *is given by a triple of ontology homomorphisms* $(f : B \to B', f_1 : O_1 \to O'_1, f_2 : O_2 \to O'_2)$ *such that*



commutes.

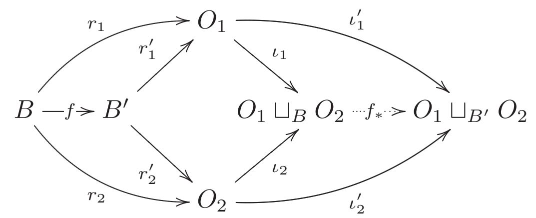
By the universal property of a pushout, a homomorphism $(f : B \to B', f_1 : O_1 \to O'_1, f_2 : O_2 \to O'_2)$ between ontology V-alignment pairs gives rise to a unique ontology homomorphism $f_* : O_1 \sqcup_B O_2 \to O'_1 \sqcup_{B'} O'_2$ making


(3)

commute. Hence an ontology homomorphism $f_* : O_1 \sqcup_B O_2 \to O'_1 \sqcup_{B'} O'_2$ is induced when there is a homomorphism from $(r_1 : B \to O_1, r_2 : B \to O_2)$ to $(r'_1 : B' \to O'_1, r'_2 : B' \to O'_2)$. Therefore, we have:

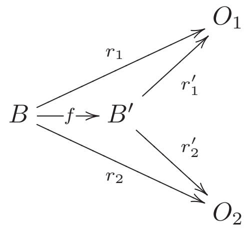
Proposition 3. *Each homomorphism* $(f : B \to B', f_1 : O_1 \to O'_1, f_2 : O_2 \to O'_2)$ *between ontology* V*-alignment pairs gives rise to a unique ontology homomorphism* $f_* : O_1 \sqcup_B O_2 \to O'_1 \sqcup_{B'} O'_2$ *making the diagram* (3) *commute.*

If both f_1 and f_2 are identity homomorphisms in the commutative diagram (3), then we have the following commutative diagram:


(4)

If f is epic, namely, $xf = yf$ implies $x = y$, then the induced homomorphism f_* is an isomorphism:

Proposition 4. *If* f *is an epic ontology homomorphism such* *that*



commutes in the category of the ontologies concerned, then $O_1 \sqcup_B O_2 \cong O_1 \sqcup_{B'} O_2$.

Proposition 5 shows that $\leq_\sqcup$ can be characterized by the existence of an ontology homomorphism.

Proposition 5. *In the ontology merging system* $(\mathfrak{O}, \approx, \sqcup)$*, there is an ontology* B *such that* $O_1 \leq_{\sqcup_B} O_2$ *if and only if there is an ontology homomorphism* $h : O_1 \to O_2$.

By Lemmas 3, 4, 5, 6, and 7 in Appendix A, and Theorem 1, clearly $(\mathfrak{O}, \approx, \sqcup)$ is actually an ontology merging system satisfying all properties introduced Sections II and III and $(\mathfrak{O}, \sqcup)$ is a partial idempotent commutative semigroup. Therefore, we have:

Theorem 4. *The merging system* $(\mathfrak{O}, \approx, \sqcup)$ *satisfies* (LU)*,* (CP)*,* (I)*,* (C)*,* (A)*,* (CA)*,* (SA)*, and* (R) *and* $(\mathfrak{O}, \sqcup)$ *is a partial idempotent commutative semigroup.*

Remark 2. Given an ontology merging system $(\mathfrak{O}, \approx, \sqcup)$ and an ontology repository $\mathbb{O} \subseteq \mathfrak{O}$, the finiteness of $\widehat{\mathbb{O}}$ can also be obtained by (SA).

By Lemma 5 in Appendix A, $(\mathfrak{O}, \approx, \sqcup)$ satisfies (SA) and so, by Proposition 1(iv), the null extension $(\mathfrak{O}^\uparrow, \sqcup^\uparrow)$ is a total semigroup and

$$\widehat{\mathbb{O}} \cup \{\uparrow\} = \bigcup_{i=1}^{+\infty} \left(\mathbb{O} \cup \{\uparrow\}\right)^i,$$

that is, $\widehat{\mathbb{O}} \cup \{\uparrow\}$ is the idempotent finitely generated semigroup given by $\mathbb{O} \cup \{\uparrow\}$. Hence the finiteness of $\widehat{\mathbb{O}}$ in $(\mathfrak{O}, \approx, \sqcup)$ can be replied by the finiteness of an idempotent finitely generated semigroup $\langle \mathbb{O} \cup \{\uparrow\} \rangle$, which is finite as shown in [21].

VI. Conclusions

We introduced ontology merging systems and studied the ontology aligning and merging operations by the properties the operations shared in an ontology merging system, such as, idempotence (I), commutativity (C), associativity (A), and representativity (R) without any specific setting or internal ontology details being considered.

Following the natural partial order approaches on semigroups, we introduced the natural partial order given by the merging operation to the ontologies concerned so that the ontologies, in the ontology merging system satisfying (I) and (CA), form a poset. Hence the ontologies can be compared, sorted, and selected by sorting and selection algorithms in posets and for the transitive relations.

An ontology repository $\mathbb{O}$ is a finite subset of the ontologies concerned. We viewed the repository $\mathbb{O}$ as generators in a partial algebraic environment and treated the merging closure of $\mathbb{O}$, the smallest set of ontologies, which is closed with respect to merging, as the subalgebraic structure generated by $\mathbb{O}$. We proved that the merging closure of a given repository is partially ordered and finite under some reasonable conditions: (I), (C), (A), and (R), so that we can compute, compare, and query these ontologies generated from the repository by merging.

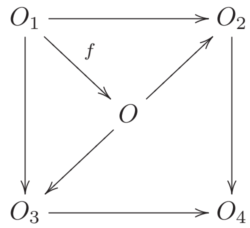
We also showed that the ontology merging system, given by ontology V-alignment pairs and pushouts, satisfies all properties introduced in Sections II and III and so there are efficient algorithms to compute merging closures and query the specific ontologies generated by pushouts.

Appendix A
Lemmas

The minimum requirements of the categorical notions and ontology concepts for the paper include: category, homomorphism, isomorphism, coproduct, pullback, pushout, monic, epic, injection, initial object, ontology, ontology homomorphism. For the notions and a systematic introduction to category theory, the reader may consult, for instance, [2], [3], [22], [23].

Lemmas 1 and 2 will be shown in a general category.

Lemma 1. *Assume that the diagram*

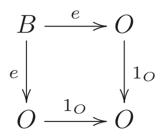


commutes.

1. *If outer square* O_1, O_2, O_3, O_4 *is a pushout, then so is* O, O_2, O_3, O_4.
2. *If* O, O_2, O_3, O_4 *is a pushout and* f *is an epic, then the outer square* O_1, O_2, O_3, O_4 *is a pushout.*

Proof: Both are clear. ■

Lemma 2. $e : B \to O$ *is an epic if and only if*



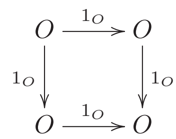
is a pushout.

Proof: It is routine to verify both directions. ■

Lemmas 3, 4, and 5 below are proved on the ontology merging system $(\mathfrak{O}, \approx, \sqcup)$.

Lemma 3. $(\mathfrak{O}, \approx, \sqcup)$ *satisfies* (I).

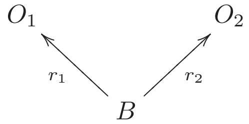
Proof: Since $1_O : O \to O$ is an epic, by Lemma 2



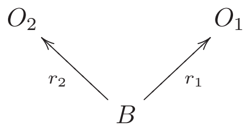
is a pushout. Then $O \sqcup_O O = O$ and so (I) holds. ■

Lemma 4. $(\mathfrak{O}, \approx, \sqcup)$ *satisfies* (C).

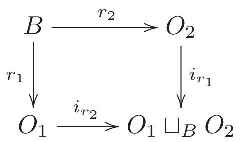
Proof: Given an ontology V-alignment pair:



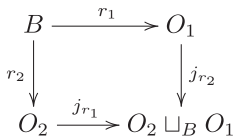
we have the following ontology V-alignment pair:



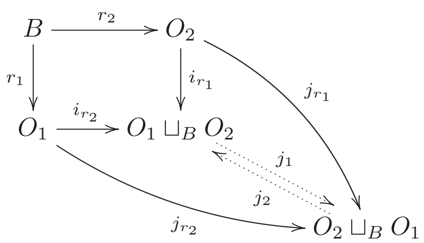
and the merges $O_1 \sqcup_B O_2$ and $O_2 \sqcup_B O_1$ are given by the following pushouts:



and



respectively. Hence there exist unique homomorphisms $j_1 : O_1 \sqcup_B O_2 \to O_2 \sqcup_B O_1$ and $j_2 : O_2 \sqcup_B O_1 \to O_1 \sqcup_B O_2$ making the following diagram



commutes serially. Hence

$$(j_2 j_1) i_{r_1} = j_2 j_{r_1} = i_{r_1}$$

and

$$(j_2 j_1) i_{r_2} = j_2 j_{r_2} = i_{r_2}$$

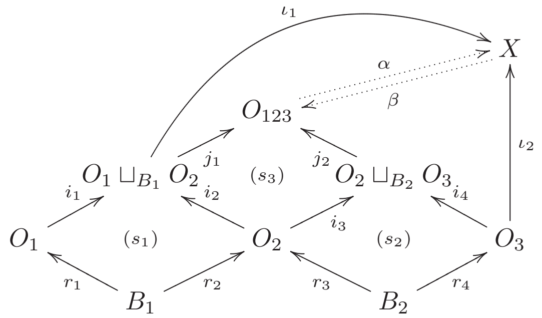
and therefore $j_2 j_1 = 1_{O_1 \sqcup_B O_2}$.

Similarly, we have $j_1 j_2 = 1_{O_2 \sqcup_B O_1}$. Thus, $O_1 \sqcup_B O_2 \cong O_2 \sqcup_B O_1$. That is, $O_1 \sqcup_B O_2 = O_2 \sqcup_B O_1$, up to ontology isomorphism, as desired. ■

Lemma 5. $(\mathfrak{O}, \approx, \sqcup)$ *satisfies* (SA).

Proof: For all ontologies $O_1, O_2, O_3 \in \mathfrak{O}$, clearly either of $O_1 \sqcup (O_2 \sqcup O_3)$ and $(O_1 \sqcup O_2) \sqcup O_3$ is defined so is the other. Given ontology V-alignment pairs $(r_1 : B_1 \to O_1, r_2 : B_1 \to O_2)$ and $(r_3 : B_2 \to O_2, r_4 : B_2 \to O_3)$ that form a W-alignment diagram, it suffices to show that $O_1 \sqcup_{B_1} (O_2 \sqcup_{B_2} O_3) = (O_1 \sqcup_{B_1} O_2) \sqcup_{B_2} O_3$, up to ontology isomorphism.

Form pushouts (s_1), (s_2), and (s_3) and then the pushout $B_2, O_3, O_1 \sqcup_{B_1} O_2, X$:



where

$$O_{123} = (O_1 \sqcup_{B_1} O_2) \sqcup_{O_2} (O_2 \sqcup_{B_2} O_3)$$

and

$$X = (O_1 \sqcup_{B_1} O_2) \sqcup_{B_2} O_3.$$

Since $(s_2) + (s_3)$ is a pushout, there is a unique ontology homomorphism $\alpha : O_{123} \to X$ such that

$$\iota_1 = \alpha j_1 \text{ and } \iota_2 = \alpha j_2 i_4.$$

Since $B_2, O_3, O_1 \sqcup_{B_1} O_2, X$ is a pushout, there is a unique ontology homomorphism $\beta : X \to O_{123}$ such that

$$j_1 = \beta \iota_1 \text{ and } j_2 i_4 = \beta \iota_2.$$

Hence

$$\alpha\beta\iota_1 = \alpha j_1 = \iota_1 \text{ and } \alpha\beta\iota_2 = \alpha j_2 i_4 = \iota_2$$

and

$$\beta\alpha j_1 = \beta\iota_1 = j_1 \text{ and } \beta\alpha j_2 i_4 = \beta\iota_2 = j_2 i_4$$

and therefore

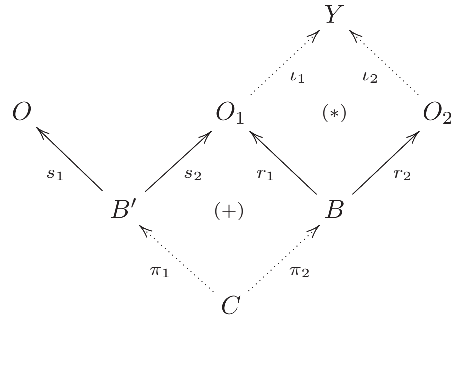
$$\alpha\beta = 1_X \text{ and } \beta\alpha = 1_{O_{123}}.$$

So $(O_1 \sqcup_{B_1} O_2) \sqcup_{B_2} = X \cong O_{123}$.

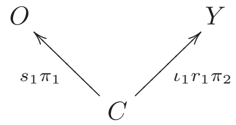
Similarly, we have $O_1 \sqcup_{B_1} (O_2 \sqcup_{B_2} O_3) \cong O_{123}$. Thus, $O_1 \sqcup_{B_1} (O_2 \sqcup_{B_2} O_3) = (O_1 \sqcup_{B_1} O_2) \sqcup_{B_2} O_3$, up to ontology isomorphism, as desired. ■

Lemma 6. $(\mathfrak{O}, \approx, \sqcup)$ *satisfies* (R)*: Given ontologies* O_1*,* O_2*, and* O *such that* $O_1 \sqcup_B O_2$ *exists,* $O \approx O_1$ *implies* $O \approx O_1 \sqcup_B O_2$ *and* $O_1 \approx O$ *implies* $O_1 \sqcup_B O_2 \approx O$*.*

Proof: Since $O_1 \sqcup_B O_2$ exists and $O \approx O_1$, $O_1 \sqcup_B O_2 = Y$ is given by the pushout $(*)$ and there is an ontology V-alignment pair $(s_1 : B' \to O, s_2 : B' \to O_1)$:



here $(+)$ is a pullback. Hence there is an ontology V-alignment pair:



and therefore $O \approx O_1 \sqcup_B O_2$. ■

Lemma 7. $(\mathfrak{O}, \sqcup)$ *is a partial idempotent commutative semigroup.*

Proof: By Lemmas 3, 4, and 5. ■

Appendix B
Proofs

Proposition 1. *Let $(\mathfrak{O}, \sim, \curlywedge)$ be an ontology merging system.*

(i) If $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (SA), then $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (A);

(ii) If $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (CA), then $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (A);

(iii) If $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (C), then (CA) is equivalent to (A) and (R);

(iv) $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (SA) if and only if the null extension $(\mathfrak{O}^{\uparrow}, \sim^{\uparrow}, \curlywedge^{\uparrow})$ is a total ontology merging system such that $(O_1 \curlywedge^{\uparrow} O_2) \curlywedge^{\uparrow} O_3 = O_1 \curlywedge^{\uparrow} (O_2 \curlywedge^{\uparrow} O_3)$ for all $O_1, O_2, O_3 \in \mathfrak{O}^{\uparrow}$, namely, $(\mathfrak{O}^{\uparrow}, \curlywedge^{\uparrow})$ is a semigroup.

Proof: (i) and (ii) are obvious.

(iii): Suppose that (CA) holds true. By (ii), (A) follows. If $O_3 = O_1 \curlywedge O_2$ and $O \sim O_1$, then, by (CA),

$$O \curlywedge (O_1 \curlywedge O_2) = (O \curlywedge O_1) \curlywedge O_2 \neq \uparrow .$$

Hence $O \sim O_1 \curlywedge O_2$ and therefore (R_r) is satisfied. Similarly, (R_l) is also satisfied. Thus, (R) holds true.

Conversely, suppose that both $O_1 \curlywedge O_2$ and $O_2 \curlywedge O_3$ exist. By (R) and (C), $O_1 \curlywedge O_2 \sim O_3$ and $O_1 \sim O_2 \curlywedge O_3$. Hence $(O_1 \curlywedge O_2) \curlywedge O_3$ and $O_1 \curlywedge (O_2 \curlywedge O_3)$ exist and therefore, by (A),

$$(O_1 \curlywedge O_2) \curlywedge O_3 = O_1 \curlywedge (O_2 \curlywedge O_3) \neq \uparrow .$$

So (CA) holds true.

(iv): If $(\mathfrak{O}^{\uparrow}, \curlywedge^{\uparrow})$ is a semigroup, namely, $\curlywedge^{\uparrow}$ is an associative operation on $\mathfrak{O}^{\uparrow}$, then, for all $O_1, O_2, O_3 \in \mathfrak{O} \subseteq \mathfrak{O}^{\uparrow}$,

$$(O_1 \curlywedge^{\uparrow} O_2) \curlywedge^{\uparrow} O_3 = O_1 \curlywedge^{\uparrow} (O_2 \curlywedge^{\uparrow} O_3) = O \in \mathfrak{O}^{\uparrow}.$$

If $O = \uparrow$, then

$$(O_1 \curlywedge O_2) \curlywedge O_3 = O_1 \curlywedge (O_2 \curlywedge O_3) = \uparrow .$$

If $O \neq \uparrow$, then

$$(O_1 \curlywedge O_2) \curlywedge O_3 = O_1 \curlywedge (O_2 \curlywedge O_3) \neq \uparrow .$$

Hence $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (SA).

Conversely, suppose now that $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (SA). For all $O_1, O_2, O_3 \in \mathfrak{O}^{\uparrow}$, if $\uparrow \in \{O_1, O_2, O_3\}$, then

$$(O_1 \curlywedge^{\uparrow} O_2) \curlywedge^{\uparrow} O_3 = O_1 \curlywedge^{\uparrow} (O_2 \curlywedge^{\uparrow} O_3) = \uparrow \in \mathfrak{O}^{\uparrow}.$$

If $\uparrow \notin \{O_1, O_2, O_3\}$, then, by (SA),

$$(O_1 \curlywedge O_2) \curlywedge O_3 = O_1 \curlywedge (O_2 \curlywedge O_3) = \uparrow \quad \text{or} \quad \neq \uparrow .$$

Hence

$$(O_1 \curlywedge^{\uparrow} O_2) \curlywedge^{\uparrow} O_3 = O_1 \curlywedge^{\uparrow} (O_2 \curlywedge^{\uparrow} O_3) \in \mathfrak{O}^{\uparrow}$$

and therefore $\curlywedge^{\uparrow}$ is associative. Thus, $(\mathfrak{O}^{\uparrow}, \curlywedge^{\uparrow})$ is a semigroup, as desired. ■

Proposition 2. *If $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (I) and (CA), then $\leq_{\curlywedge}$ is a partial order on $\mathfrak{O}$ and so $(\mathfrak{O}, \leq_{\curlywedge})$ is a poset.*

Proof: It suffices to prove that $\leq_{\curlywedge}$ is reflexive, transitive, and antisymmetric.

By (I), $O \curlywedge O = O$ and so $O \leq_{\curlywedge} O$ for all $O \in \mathfrak{O}$. Hence $\leq_{\curlywedge}$ is reflexive.

Assume that $O_1 \leq_{\curlywedge} O_2$ and $O_2 \leq_{\curlywedge} O_3$. Then $O_1 \sim O_2$, $O_2 \sim O_1$, $O_2 \sim O_3$, $O_3 \sim O_2$, $O_1 \curlywedge O_2 = O_2 \curlywedge O_1 = O_2$, and $O_2 \curlywedge O_3 = O_3 \curlywedge O_2 = O_3$. By (CA), $O_1 \curlywedge O_3 = O_1 \curlywedge (O_2 \curlywedge O_3) = (O_1 \curlywedge O_2) \curlywedge O_3 = O_2 \curlywedge O_3 = O_3$ and $O_3 \curlywedge O_1 = (O_3 \curlywedge O_2) \curlywedge O_1 = O_3 \curlywedge (O_2 \curlywedge O_1) = O_3 \curlywedge O_2 = O_3$. Hence $O_1 \sim O_3$, $O_3 \sim O_1$, and $O_1 \leq_{\curlywedge} O_3$ and therefore $\leq_{\curlywedge}$ is transitive.

If $O_1 \leq_{\curlywedge} O_2$ and $O_2 \leq_{\curlywedge} O_1$, then $O_1 \curlywedge O_2 = O_2 \curlywedge O_1 = O_2$ and $O_1 \curlywedge O_2 = O_2 \curlywedge O_1 = O_1$ and so $O_1 = O_1 \curlywedge O_2 = O_2$. Thus, $\leq_{\curlywedge}$ is antisymmetric. ■

Theorem 1. *Let $\sim$ be reflexive and commutative and $\preceq$ a partial order on $(\mathfrak{O}, \sim, \curlywedge)$. Then $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (LU) and (CP) with respect to $\preceq$ if and only if $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (I), (C), (A), and (R) and $\preceq = \leq_{\curlywedge}$.*

Proof: "only if": For each ontology $O \in \mathfrak{O}$, since $\sim$ is reflexive, $O \sim O$. By (LU), $O \preceq O \curlywedge O$. Since $O \preceq O$ and $O \curlywedge O$ is the least upper bound of O and O, $O \curlywedge O \preceq O$. Since $\preceq$ is a partial order, $O \curlywedge O = O$. Hence (I) holds.

If $O_1 \sim O_2$, then, clearly, $O_2 \sim O_1$ as $\sim$ is commutative. Since $O_2 \preceq O_1 \curlywedge O_2$ and $O_1 \preceq O_1 \curlywedge O_2$, by (LU) we have $O_2 \curlywedge O_1 \preceq O_1 \curlywedge O_2$. Similarly, $O_1 \curlywedge O_2 \preceq O_2 \curlywedge O_1$. Then $O_1 \curlywedge O_2 = O_2 \curlywedge O_1$ and so (C) is satisfied.

Suppose that $O_1 \curlywedge O_2$ exists and $O_1 \sim O$. Since $O_1 \preceq O_1 \curlywedge O_2$, $O_1 \curlywedge O_2 \sim O$ by (CP_r). Hence (R_r) holds. Similarly, (R_l) holds true. Thus, (R) is satisfied.

Assume that both $(O_1 \curlywedge O_2) \curlywedge O_3$ and $O_1 \curlywedge (O_2 \curlywedge O_3)$ exist. Since $O_3 \preceq O_2 \curlywedge O_3$, by (CP_l),

$$(O_1 \curlywedge O_2) \curlywedge O_3 \preceq (O_1 \curlywedge O_2) \curlywedge (O_2 \curlywedge O_3).$$

Clearly, $O_2 \curlywedge O_3 \preceq (O_1 \curlywedge O_2) \curlywedge O_3$ and $O_1 \curlywedge O_2 \preceq (O_1 \curlywedge O_2) \curlywedge O_3$. So $(O_1 \curlywedge O_2) \curlywedge (O_2 \curlywedge O_3) \preceq (O_1 \curlywedge O_2) \curlywedge O_3$ by (LU). Hence $(O_1 \curlywedge O_2) \curlywedge O_3 = (O_1 \curlywedge O_2) \curlywedge (O_2 \curlywedge O_3)$. Symmetrically, $O_1 \curlywedge (O_2 \curlywedge O_3) = (O_1 \curlywedge O_2) \curlywedge (O_2 \curlywedge O_3)$. Then $O_1 \curlywedge (O_2 \curlywedge O_3) = (O_1 \curlywedge O_2) \curlywedge O_3$ and so (A) is satisfied.

Assume now that $O_1 \leq_{\curlywedge} O_2$. Then $O_2 = O_1 \curlywedge O_2 = O_2 \curlywedge O_1$. Hence $O_1 \preceq O_1 \curlywedge O_2 = O_2$ by (LU). On the other hand, suppose that $O_1 \preceq O_2$. Then, by (LU), $O_2 \preceq O_1 \curlywedge O_2$. Hence $O_2 = O_1 \curlywedge O_2$ as $O_1 \curlywedge O_2 \preceq O_2 \curlywedge O_2 = O_2$ by (CP_r). Similarly, $O_2 = O_2 \curlywedge O_1$. Thus, $O_1 \leq_{\curlywedge} O_2$, as desired. So $\preceq = \leq_{\curlywedge}$.

"if": Suppose now that $\preceq = \leq_{\curlywedge}$ and $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (I), (C), (A), and (R). By Proposition 1(iii), $(\mathfrak{O}, \sim, \curlywedge)$ satisfies (CA) and so $\leq_{\curlywedge}$ is a partial order by Proposition 2.

Let $O_1, O_2 \in \mathfrak{O}$ be such that $O_1 \sim O_2$. Then $O_1 \curlywedge O_2$ and $O_2 \curlywedge O_1$ exist and equal by (C). Since

$$O_1 \curlywedge (O_1 \curlywedge O_2) = (O_1 \curlywedge O_1) \curlywedge O_2 = O_1 \curlywedge O_2$$

and $(O_1 \curlywedge O_2) \curlywedge O_1 = O_1 \curlywedge (O_2 \curlywedge O_1) = O_1 \curlywedge (O_1 \curlywedge O_2) = (O_1 \curlywedge O_1) \curlywedge O_2 = O_1 \curlywedge O_2$, we have $O_1 \leq_{\curlywedge} O_1 \curlywedge O_2$. Similarly, $O_2 \leq_{\curlywedge} O_1 \curlywedge O_2$. Hence $O_1 \curlywedge O_2$ is an upper bound of O_1 and O_2.

Assume that $O_1 \leq_{\curlywedge} O$ and $O_2 \leq_{\curlywedge} O$. Then $O_1 \curlywedge O = O \curlywedge O_1 = O$ and $O_2 \curlywedge O = O \curlywedge O_2 = O$. Hence

$$O = O_1 \curlywedge O = O_1 \curlywedge (O_2 \curlywedge O) = (O_1 \curlywedge O_2) \curlywedge O$$

and $O = O \curlywedge O_1 = (O \curlywedge O_2) \curlywedge O_1 = O \curlywedge (O_2 \curlywedge O_1) = O \curlywedge (O_1 \curlywedge O_2)$ and therefore $O_1 \curlywedge O_2 \leq_{\curlywedge} O$. Then $O_1 \curlywedge O_2$ is the least upper bound of O_1 and O_2 and so (LU) holds.

For all ontologies O_1, O_2, O such that $O_1 \leq_{\curlywedge} O_2$ and $O \sim O_1$, obviously $O \sim O_1 \curlywedge O_2 = O_2$ by (R). Hence (CP_r) is satisfied. Similarly, (CP_l) is also satisfied. Thus, (CP) holds true. ■

Theorem 2. *Given a repository $\mathbb{O} \subseteq \mathfrak{O}$. the merging closure $\widehat{\mathbb{O}}$ exists and it is unique.*

Proof: Let

$$\mathcal{C} \stackrel{\text{def}}{=} \{\mathbb{P} \mid \mathbb{O} \subseteq \mathbb{P} \text{ and } \mathbb{P} \text{ is closed with respect to merging}\}.$$

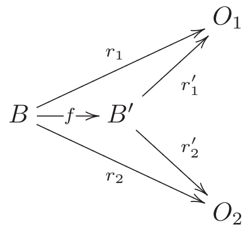
Clearly, $\mathfrak{O} \in \mathcal{C} \neq \emptyset$ and $\mathcal{C}$ is partially ordered by $\subseteq$. For each chain $\cdots \subseteq \mathbb{P}_\alpha \subseteq \cdots \subseteq \mathbb{P}_\beta \cdots$, $\cap \mathbb{P}_\alpha \in \mathcal{C}$ and it is a lower bound of the totally ordered subset $\{\cdots, \mathbb{P}_\alpha, \cdots, \mathbb{P}_\beta, \cdots\}$. Thus, $\mathcal{C}$ has a minimal element, $\cap_{\mathbb{P} \in \mathcal{C}} \mathbb{P}$.

Let $\mathbb{P}_1$ and $\mathbb{P}_2$ be two minimal elements for $(\mathcal{C}, \subseteq)$. Then $\mathbb{P}_1 \cap \mathbb{P}_2$ is closed with respect to ontology merging $\curlywedge$ and smaller than $\mathbb{P}_1$ and $\mathbb{P}_2$ and so $\mathbb{P}_1 = \mathbb{P}_1 \cap \mathbb{P}_2 = \mathbb{P}_2$. The uniqueness follows. The unique minimal element of $(\mathcal{C}, \subseteq)$ is the merging closure $\widehat{\mathbb{O}}$:

$$\begin{aligned}\widehat{\mathbb{O}} &= \bigcap_{\mathbb{P} \in \mathcal{C}} \mathbb{P} \\ &= \bigcap \{\mathbb{P} \mid \mathbb{O} \subseteq \mathbb{P} \text{ and } \mathbb{P} \text{ is closed with respect to merging}\} \\ &\neq \emptyset.\end{aligned}$$

■

Proposition 4. *If f is an epic ontology homomorphism such that*



(5)

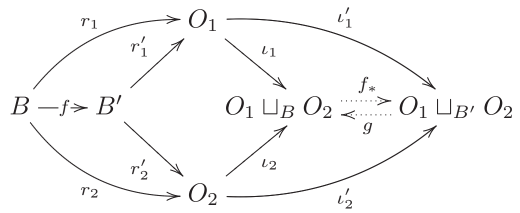
commutes in the category of the ontologies concerned, then $O_1 \sqcup_B O_2 \cong O_1 \sqcup_{B'} O_2$.

Proof: Since the diagram (5) commutes, $(1_{O_1}, 1_{O_2}, f) : (r_1 : B \to O_1, r_2 : B \to O_2) \to (r'_1 : B' \to O_1, r'_2 : B' \to O_2)$ is a homomorphism between ontology V-alignment pairs. Hence there is a unique ontology homomorphism

$$f_* : O_1 \sqcup_B O_2 \to O_1 \sqcup_{B'} O_2$$

such that the diagram (4) commutes. If f is an epic, then, by Lemma 1, $B', O_1, O_1 \sqcup_B O_2, O_2$ is commutative and a pushout and so there is a unique ontology homomorphism $g : O_1 \sqcup_{B'} O_2 \to O_1 \sqcup_B O_2$ such that

$$g\iota'_1 = \iota_1 \text{ and } g\iota'_2 = \iota_2.$$



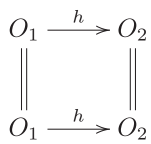
Hence

$$f_* g \iota'_1 = f_* \iota_1 = \iota'_1 \text{ and } f_* g \iota'_2 = f_* \iota_2 = \iota'_2$$

and therefore $f_* g = 1_{O_1 \sqcup_{B'} O_2}$. Similarly, we have $g f_* = 1_{O_1 \sqcup_B O_2}$ from $f_* \iota_1 = \iota'_1$ and $f_* \iota_2 = \iota'_2$. Then f_* is an isomorphism and so $O_1 \sqcup_B O_2 \cong O_1 \sqcup_{B'} O_2$. ■

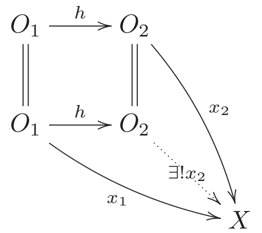
Proposition 5. *In the ontology merging system $(\mathfrak{O}, \approx, \sqcup)$, there is an ontology B such that $O_1 \leq_{\sqcup_B} O_2$ if and only if there is an ontology homomorphism $h : O_1 \to O_2$.*

Proof: "*if*": Suppose that there exists an ontology homomorphism $h : O_1 \to O_2$. Then we have a commutative square:



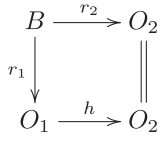
(6)

For any ontology homomorphisms $x_1 : O_1 \to X$ and $x_2 : O_2 \to X$ such that $x_1 = x_2 h$, clearly there is a unique ontology homomorphism $x_2 : O_2 \to X$ making



commute. Hence (6) is a pushout and therefore $O_1 \sqcup_{O_1} O_2 = O_2 \sqcup_{O_1} O_1 = O_2$. Thus, there exists an ontology O_1 such that $O_1 \leq_{\sqcup_{O_1}} O_2$, as desired.

"*only if*": Assume now that $O_1 \leq_{\sqcup_B} O_2$. Then $O_1 \approx_B O_2$ and $O_1 \sqcup_B O_2 = O_2 \sqcup_B O_1 = O_2$ and so there is a pushout in the category of the ontologies concerned:



Hence there exists an ontology homomorphism $h : O_1 \to O_2$. ■

REFERENCES

[1] J. Euzenat and P. Shvaiko, *Ontology Matching*, 2nd ed. Springer, 2013.
[2] A. Zimmermann, M. Krotzsch, J. Euzenat, and P. Hitzler, “Formalizing ontology alignment and its operations with category theory,” in *Proc. 4th International Conference on Formal Ontology in Information Systems (FOIS)*. Baltimore, USA: IOS Press, 2006, pp. 277–288.
[3] P. Hitzler, M. Krötzsch, M. Ehrig, and Y. Sure, “What is ontology merging? - A category theoretical perspective using pushouts,” in *Proceedings of the First International Workshop on Contexts and Ontologies: Theory, Practice and Applications (C&0). Workshop at the 20th National Conference on Artificial Intelligence AAAI-05*. Pittsburgh, Pennsylvania, Jul. 2005.
[4] C. Antunes and M. Abel, “Ontologies in category theory: A search for meaningful morphisms,” in *SEMINAR ON ONTOLOGY RESEARCH IN BRAZIL, PROCEEDINGS*. São Paulo, 2018.
[5] I. Cafezeiro and E. H. Haeusler, “Semantic interoperability via category theory,” in *Proceedings of the 26th International Conference on Conceptual Modeling*. Auckland, New Zealand, 2007, pp. 197–202.
[6] M. Codescu, T. Mossakowski, and O. Kutz, “A categorical approach to ontology alignment,” in *Proceedings of the 9th International Workshop on Ontology Matching collocated with the 13th International Semantic Web Conference (ISWC 2014)*. Riva del Garda, Trentino, Italy, 2014.
[7] M. Codescu, T. Mossakowski, and O. Kutz, “A categorical approach to networks of aligned ontologies,” *Journal on Data Semantics*, vol. 6, no. 4, pp. 155–197, 2017.
[8] L. Hu and J. Wang, “Geo-ontology integration based on category theory,” in *International Conference On Computer Design and Applications*. Qinhuangdao, 2010, V1-5–V1-8.
[9] N. Kibret, W. Edmonson, and S. Gebreyohannes, “Category theoretic based formalization of the verifiable design process,” in *IEEE International Systems Conference (SysCon)*. Orlando, FL, USA, 2019, pp. 1–8.
[10] M. Mendonca, J. Aguilar, and N. Perozo, “Application of category theory,” *Ingénierie des systèmes d’information*, vol. 23, no. 2, 2018, pp. 11–38.
[11] J. Euzenat, “Algebras of ontology alignment relations,” In *Sheth A. et al. ed. International Semantic Web Conference - ISWC 2008, Lecture Notes in Computer Science, Vol. 5318*. Berlin, Heidelberg: Springer, 2008, pp. 387–402.
[12] M. Mahfoudh, L. Thiry, G. Forestier, and M. Hassenforder, “Algebraic graph transformations for merging ontologies,” in *Model & Data Engineering, 4th International Conference, MEDI 2014*. Larnaca, Cyprus, Sep. 2014, pp. 154–168.
[13] M. Mahfoudh, G. Forestier, and M. Hassenforder, “A benchmark for ontologies merging assessment,” in *Proc. of the International Conference on Knowledge Science, Engineering and Management (KSEM)*. 2016, pp. 555–566.
[14] K. Sun, Y. Zhu, P. Pan, Z. Hou, D. Wang, W. Li, and J. Song, “Geospatial data ontology: the semantic foundation of geospatial data integration and sharing,” *Big Earth Data*, vol. 3, no. 3, pp. 269–296, 2019.
[15] O. Benjelloun, D. Garcia-Molina, D. Menestrina, Q. Su, S. E. Wang, and J. Widom, “Swoosh: A generic approach to entity resolution,” *The VLDB Journal*, vol. 18, no. 1, pp. 255–276, 2009.
[16] A. Maedche and S. Staab, “Ontology learning for the semantic web,” *IEEE Intelligent systems*, vol. 16, no. 2, pp. 72–79, 2001.
[17] H. Mitsch, “A natural partial order for semigroups,” *Proceedings of the American Mathematical Society*, vol. 97, no. 3, pp. 384–388, 1986.
[18] C. Daskalakis, R. M. Karp, E. Mossel, S. J. Riesenfeld, and E. Verbin, “Sorting and selection in posets,” *SIAM Journal on Computing*, vol. 40, no. 3, pp. 392–401, 2011.
[19] A. de Luca and S. Varricchio, *Finiteness and regularity in semigroups and formal languages*. Springer Science & Business Media, 2012.
[20] M. Morse and G. A. Hedlund, “Unending chess, symbolic dynamics and a problem in semigroups,” *Duke Math. J.*, vol. 11, no. 1, pp. 1–7, 1944.
[21] T. C. Brown and E. Lazerson, “On finitely generated idempotent semigroups,” *Semigroup Forum*, vol. 78, no. 1, pp. 183–186, 2009.
[22] J. Adámek, H. Herrlich, and G. E. Strecker, *Abstract and Concrete Categories: The Joy of Cats*. Dover Publications, 2009.
[23] S. Mac Lane, *Categories for the Working Mathematician*, 2nd ed. Berlin: Springer Verlag, 1997.