A Novel Approach to Ontological View-Based Semantic Mapping in Decentralized Environments

Fateh Mohamed Ali Adhnouss[a], Husam M. Ali El-Asfour and Kenneth McIsaac
Dept. of Electrical & Computer Engineering, University of Western, London, Canada

Keywords: Ontological View, Extensional Semantics, Intensional Semantics, Semantic Integration.

Abstract: Semantic integration and interoperability are vital for effective communication and data exchange in information systems. This paper explores the significance of shared referential semantics, ontological views (OVs), and intensional semantics in achieving semantic integration. It addresses the challenges arising from divergent OVs and emphasizes the role of accessibility relations in bridging gaps between different systems. The paper presents theorems establishing relationships between accessibility relations, the overlap of non-logical symbols, and the consistency and accessibility of OVs. It also examines mapping possibilities in decentralized environments by considering scenarios with shared intended models, overlapping intended models, or no intersection of intended models. The paper uses the healthcare domain as an illustrative example of applying intensional semantics and semantic interoperability. To ensure efficient semantic integration and interoperability, systems must consider the shared meaning of the vocabulary, particularly non-logical symbols, along with the underlying conceptualizations and intensional semantics.

1 INTRODUCTION

Ontologies represent a fundamental aspect of solutions tailored towards achieving semantic integration(de Mello et al., 2022). They enable diverse systems to comprehend and share information effectively. However, ontologies are conventionally formed within limited environments. When multiple system designers approach a single domain, they each manifest a unique perspective corresponding to their specific interests, which inevitably results in the development of multiple models (Adhnouss et al., 2022; Xue et al., 2012). Given that each perspective offers a different view of the domain, there exists no singular universally agreed-upon ontology. Instead, we see the formation of unique, formally defined OVs. Crucially, explicit ontologies are often missing from information systems, with inherent semantics typically incorporated within the information model (Wang et al., 2009). This model reflects a particular conceptualization view, thereby defining an implicit Ov.

Current semantic integration strategies rooted in extensional models are ill-suited to a decentralized environment (Majkić and Prasad, 2018; Ali and McIsaac, 2020; Adhnouss et al., 2022) because they fail to adequately address its fluid nature, characterized by a set of entities and their evolving relationships. Thus, the requirement arises for an encompassing semantic integration model that accounts for the fluidity of a decentralized environment by capturing the dynamic entity relations and shifts therein.

In Artificial Intelligence, OV pertains to an engineering artifact, characterized by a distinctive vocabulary depicting a particular reality, supplemented by explicit assumptions about the intended interpretation of the vocabulary. Each OV for any logical language constitutes a set of carefully crafted axioms that faithfully mirror the language's intended models at the intensional semantic level. Therefore, OVs only indirectly 'specify' a conceptualization, approximating a conceptualization when applied to a logical language, assuming an intensional level exists where the language's intended models are encapsulated within the OV's models(Adhnouss et al., 2023).

Consequently, an OV is seen as a logical theory expressing the intended meaning of a formal vocabulary, with the intended models of a language using such vocabulary in a way that is confined by its intensional semantic level. The OV indirectly denotes this commitment (and the underlying conceptualization) by closely mirroring these intended models.

Anchored in the fundamental principles of in-

[a] https://orcid.org/0000-0003-4191-6894

Adhnouss, F., El-Asfour, H. and McIsaac, K.
A Novel Approach to Ontological View-Based Semantic Mapping in Decentralized Environments.
DOI: 10.5220/0012157600003598
In *Proceedings of the 15th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2023) - Volume 2: KEOD*, pages 155-163
ISBN: 978-989-758-671-2; ISSN: 2184-3228

tensional and extensional semantic levels delineated in(Adhnouss et al., 2022), this paper narrows its focus on semantic integration to contexts devoid of a comprehensive global perspective. This ongoing research ambitiously endeavors to reveal the pivotal properties that have their roots in the methodological structure proposed in the aforementioned work. Such discoveries could potentially serve as structural pillars, enabling the establishment of integrative links amidst varying OVs. While the focus of the previous contributions has been on the notion of representation, this paper takes a sharp turn towards integration - an equally significant facet- thereby interpreting semantics through an innovative lens.

The remainder of this paper is structured as follows: In Section 2, we provide a comprehensive review of relevant literature on semantic integration, exploring existing approaches, techniques, and models proposed in the field. Section 3 delves into the importance of shared referential semantics in semantic integration and discusses the concept of OVs. We then explore the relationship between OV equivalence and shared semantics in Section 4. In Section 5, we examine accessibility relations and their impact on the consistency and overlap of non-logical symbols in OVs. Section 6 and 7 focuses on the formal definitions and mapping possibilities in decentralized environments, discussing different scenarios and proposing an approach for achieving completeness and usability of mapped OVs. In Section 8, we apply the concepts of intensional semantics and semantic interoperability in the healthcare domain. Finally, Section 9 concludes the paper by summarizing the key findings and contributions, highlighting the importance of shared understanding for semantic interoperability, and providing guidance for achieving effective integration and integration itself in information systems.

2 LITERATURE REVIEW

The traditional database systems have faced ongoing challenges due to semantic heterogeneity, which is essentially a result of differences in the representation of real-world entities across various models developed independently (Xue et al., 2012). These variations in structure, terminology, and interpretation have proved to be hurdles in the semantically coherent interoperation of such models. Consequently, researchers and developers have proposed numerous solutions over the years (Adhnouss et al., 2022; Wang et al., 2009). These solutions, however, largely fall into two categories- structure-based and semantics-based- with their effectiveness heavily dependent on the presence of a pre-defined schema or ontology.

In the realm of information integration, solutions often resort to establishing semantic correspondences or mappings between vocabularies originating from distinct data sources. Various methods such as linguistic term analysis (Melluso et al., 2022), mapping to a common reference ontology (Cao et al., 2022), and the utilization of heuristics to discern concept definition patterns (Zhang et al., 2022), have been employed. These strategies, however, invariably call for a comprehensive global knowledge representation or some form of human intervention.

Concurrently, numerous studies have delved into the different aspects of ontology, with early research efforts primarily focusing on its definition and mathematical representation (Gruber, 1995; Guarino et al., 2009; Wang et al., 2009). Ontology was hence conceptualized as a meticulously constructed artifact, employing a specific vocabulary to depict a selected reality, bolstered by explicit assumptions regarding the meanings of the vocabulary symbols.

Within this framework, the concept of 'intensional semantics'—introduced by (Adhnouss et al., 2022; Adhnouss et al., 2023)—emerges as a cornerstone. This concept shapes the intrinsic meanings and the internal content of the vocabulary symbols, thereby turning ontology into a logical theory that unfolds these intended meanings, expressing an intensional semantic to a specific world view.

Despite this, it is crucial to highlight that fundamentally, ontology resides in 'extensional semantics.' This aspect is embodied by the ontology's 'intended models,' representing specific instances or scenarios that the ontology aims to encapsulate (Adhnouss et al., 2022).

Ontology thus evolves from being a mere reflection of reality to a sophisticated system of extensional semantics, encapsulating a unique world interpretation. Further research has been carried out to investigate ontology classification (Dhakal et al., 2022), development of ontology languages (Chandrashekar et al., 2023), construction (Du et al., 2022), and reuse (Polenghi et al., 2022). Efforts have also been made to explore methodologies for ontology integration (Cao et al., 2022).

In more contained environments, ontologies are typically agreed upon by consensus. However, this consensus often breaks down in decentralized environments due to their inherent characteristics. The lack of a global ontology and the impracticality of frequent human intervention in decentralized information systems underline these systems' unique attributes. Existing solutions largely depend on a global ontology, human intervention, or both, highlighting the need to broaden inquiry in this area to cater to the specialized requirements of decentralized systems.

Our investigation pivots around defining and mathematically portraying partial knowledge related to a particular domain. After an extensive exploration, we found that the OV definition proposed by (Adhnouss et al., 2022) aligns seamlessly with our objectives. This work offers a top-down formal model as a theoretical framework where each possible world presents unique views of the overarching conceptualization, reflecting information system modeling.

This methodology underscores the importance of semantics within decentralized environments where a global view might be purely theoretical. It emphasizes the theoretical aspects of these notions, utilizing epistemology as a conduit to convey intensional semantics, which can be actualized into multiple OVs (extensions), leading to extensional semantics.

Our research stands out by focusing on semantic integration in the absence of a global view. We aim to pinpoint shared axioms or pivotal properties of OVs identified in the top-down approach that could serve as foundational elements to bridge these possible worlds. Our work capitalizes on a detailed understanding and analysis of diverse OVs within the system, referred to as ”Epistemology” (Adhnouss et al., 2023). As a result, our focus shifts from representation to integration as we study semantics across different views without a global viewpoint—a perspective that provides a distinct and equally vital insight.

3 REFERENTIAL SEMANTICS IN SEMANTIC INTEGRATION

Proceeding from our literature review, it becomes apparent that shared referential semantics hold vital importance in the realm of semantic integration. An OV, as discussed earlier, doesn't necessarily represent a single, objective reality. Instead, it portrays a unique interpretation or perspective of a particular domain. Therefore, different designers or organizations might conceive distinct OVs for the same domain, influenced by differing assumptions, interpretations, or objectives.

For instance, consider two information systems designed for the same domain, but adopting unique OVs. Each OV encapsulates a specific viewpoint on the underlying concepts and relationships, reflecting the distinct goals, assumptions, or interpretations of the designers. Thus, within a particular business domain, each system's semantics presents a unique ontological viewpoint, or an OV.

Semantics, in many domains, can be articulated at various explicit and implicit levels within the domain knowledge, design, and the structure of the information model. This inherent complexity can make seamless integration a daunting task, especially when semantics are not immediately apparent or accessible.

Nevertheless, every information system harbors a conceptualization of the observed (modeled) domain, which can serve as a bedrock for semantic-based integration. When different systems have overlapping intended models of the same domain, this shared understanding can be harnessed to facilitate integration, ensuring consistency in data meaning and interpretation. This process is intrinsically tied to our exploration of intensional and extensional semantics, ensuring continuity and coherence in our research.

4 OV EQUIVALENCE AND SEMIOTICS

Delving into the complex intersections of information semantics and OVs, we must comprehend the significance of semantic equivalence between OVs. While a shared vocabulary forms the bedrock of communication in an ontological scenario, understanding the deeper interconnections between terms unlocks the shared conceptualization that an OV encapsulates.

Semantic equivalence between OVs arises when the ontologies delineate identical concepts and relationships within the domain. This implies that despite differences in their symbolic or linguistic representations, semantic equivalence holds if the ontologies revolve around the same subject or entity.

Nevertheless, it is crucial to distinguish between logical and semantic equivalence. Logically equivalent ontologies retain the same logical structure and consistent truth values. However, differences in the specific concepts and relationships they symbolize can create scenarios where ontologies are logically but not semantically equivalent.

To discern whether ontologies are semantically equivalent, it is necessary to examine both the logical constructs and the unique concepts and relationships they represent. If ontologies convey the same meanings and relate to the same subject or entity, they are semantically equivalent. Conversely, ontologies that express dissimilar meanings or pertain to different subjects or entities are not semantically equivalent, irrespective of their logical equivalence.

To illustrate, consider two sets of formulas - $OV1$ and $OV2$, representing a patient treatment domain and expressed using the formalism of first-order logic (FOL).

Given these axiom sets:

$OV1$:

- $\forall x(\text{patient}(x) \rightarrow \text{human}(x))$
- $\forall x(\text{has-disease}(x,y) \rightarrow \text{patient}(x))$
- $\forall x(\text{has-symptoms}(x,y) \wedge \text{has-disease}(y,z) \rightarrow \text{prescribed-treatment}(x,z))$
- $\forall x(\text{has-treatment}(x,y) \rightarrow \text{patient}(y))$
- $\forall x(\text{patient}(x) \wedge \text{has-disease}(x,y) \wedge \text{has-treatment}(y,z) \rightarrow \text{cured}(x))$
- $\text{patient}(\text{John})$
- $\text{has-disease}(\text{John}, \text{Cancer})$
- $\text{has-symptoms}(\text{John}, \text{Headache})$

And another set:

$OV2$:

- $\forall x(\text{P1}(x) \rightarrow \text{P3}(x))$
- $\forall x(\text{P4}(x,y) \rightarrow \text{P1}(x))$
- $\forall x(\text{P5}(x,y) \wedge \text{P4}(y,z) \rightarrow \text{P2}(x,z))$
- $\forall x(\text{P6}(x,y) \rightarrow \text{P1}(y))$
- $\forall x(\text{P1}(x) \wedge \text{P4}(x,y) \wedge \text{P6}(y,z) \rightarrow \text{P7}(x))$
- $\text{P1}(\text{A})$
- $\text{P4}(\text{B}, \text{A})$
- $\text{P5}(\text{A}, \text{B})$

In the context of both $OV1$ and $OV2$, it's evident that the non-logical symbols they contain, though represented differently, refer to the same entities in the patient treatment domain. This situation gives rise to a deeper level of equivalence—semantic equivalence—when non-logical symbols are interpreted within the ontology.

A set of formulas is semantically equivalent if, apart from demonstrating logical equivalence, they refer to the same domain and convey identical meanings for the non-logical symbols. These symbols embody specific concepts and relationships in the domain, hence their interpretation directly influences the semantic correspondence.

In our case, despite differences in predicates and individuals, the non-logical symbols in both $OV1$ and $OV2$ signify the same concepts and relationships. Thus, even if the syntactic presentation differs, the semantic content they encapsulate remains identical. Therefore, $OV1$ and $OV2$ are semantically equivalent.

Understanding semantic equivalence is paramount for assessing the alignment of different OVs, as it captures the essence of meanings embodied by non-logical symbols, transcending mere syntactic and logical equivalence. This understanding is vital for ensuring effective semantic integration, especially in environments where multiple OVs represent the same domain in a decentralized information system.

5 ACCESSIBILITY RELATIONS AND NON-LOGICAL SYMBOLS IN OVs

In the previous discussion, we recognized the critical role of intensional and extensional semantics in achieving semantic equivalence and successful integration. Now, we extend this exploration to discern how accessibility relations in the context of different OV can influence the consistency and overlap of non-logical symbols. Understanding this relationship is key for effective semantic integration in decentralized systems.

Reflecting upon the insights presented in (Adhnouss et al., 2023; Adhnouss et al., 2022), we acknowledge that each possible world represents an OV of the domain. The accessible relation denotes that all possible worlds share some overlapping conceptualizations, a feature that can be utilized for fostering semantic integration and interoperability between systems. Our approach involves establishing a shared understanding of the domain by first defining high-level concepts, relationships, and terminology. This shared understanding can then be tailored and specialized for different possible worlds, which may have specific requirements or nuances.

Consequently, a multitude of ontological perspectives are created, with the most general concepts positioned at the top, becoming increasingly specific as we delve deeper.

As we traverse the intricate landscape of information semantics in decentralized systems, it becomes paramount to understand how accessibility relations impact the underlying structures within these environments. This understanding paves the way for our first theorem, which explores the conditions under which an accessible relation between two distinct worlds suggests the consistency of non-logical symbols in these worlds.

Theorem 1 (Consistency via Accessibility Relations)**.** *Let $W_1 = (E_1, \Sigma_1)$ and $W_2 = (E_2, \Sigma_2)$ be two distinct worlds, where E_i is the set of entities and Σ_i is the set of non-logical symbols in W_i, for $i = 1, 2$. Let $R : E_1 \times E_2 \rightarrow \{0, 1\}$ be an accessible relation between W_1 and W_2, where $R(e_1, e_2) = 1$ if and only if the entity e_1 in E_1 can be accessed or understood as e_2 in E_2. Furthermore, let $f : E_1 \rightarrow E_2$ be a bijective mapping function from entities in E_1 to entities in E_2, such that for every entity e_1 in E_1, $f(e_1)$ is a corresponding entity e_2 in E_2 that preserves certain properties represented by the non-logical symbols. Formally, we require that for every symbol $s \in \Sigma_1$ that represents a property of e_1, there exists a symbol $s' \in \Sigma_2$ that represents the same property of $e_2 = f(e_1)$, and* *vice versa.*

Then the theorem states: If for all $e_1 \in E_1$ there exists an $e_2 \in E_2$ such that $R(e_1, e_2) = 1$ and f preserves the properties of e_1 and e_2 as described, then Σ_1 and Σ_2 are derived from the same set of non-logical symbols in Θ, ensuring consistency between W_1 and W_2. Formally,

$$\forall e_1 \in E_1, \exists e_2 \in E_2 : \\ R(e_1, e_2) = 1 \wedge (\forall s \in \Sigma_1, \exists s' \in \Sigma_2, s' = f(s)) \\ \rightarrow (\Sigma_1 \cap \Sigma_2 \neq \emptyset).$$

Theorem 2 (Overlap via Accessibility Relations). *Let $W_1 = (E_1, \Sigma_1)$ and $W_2 = (E_2, \Sigma_2)$ be two distinct worlds, where E_i is the set of entities and Σ_i is the set of non-logical symbols in W_i, for $i = 1, 2$. Define $R(W_1, W_2)$ to be an accessible relation between W_1 and W_2. We say that: If $R(W_1, W_2)$ holds, then there is a non-empty intersection between Σ_1 and Σ_2, signifying an overlap of non-logical symbols between the two worlds. Formally,*

$$R(W_1, W_2) \rightarrow (\Sigma_1 \cap \Sigma_2 \neq \emptyset).$$

Theorem 3 (Accessibility Relations via Overlap). *Let $W_1 = (E_1, \Sigma_1)$ and $W_2 = (E_2, \Sigma_2)$ be two distinct worlds, where E_i is the set of entities and Σ_i is the set of non-logical symbols in W_i, for $i = 1, 2$. Define $R(W_1, W_2)$ to be an accessible relation between W_1 and W_2. We then assert: Given the same sets of non-logical symbols Σ_1 and Σ_2 for worlds W_1 and W_2, if there is a non-empty intersection between Σ_1 and Σ_2, it implies an accessibility relation $R(W_1, W_2)$ between the two worlds, meaning the entities and properties in W_1 can be accessed or understood in W_2. Formally,*

$$(\Sigma_1 \cap \Sigma_2 \neq \emptyset) \rightarrow R(W_1, W_2).$$

Each theorem refines our understanding of the relationship between accessible relations and the overlap of non-logical symbols between different worlds. The accessible relation among possible worlds reveals the commonalities between them, a feature that can be employed to bridge understanding gaps and facilitate information exchange between different systems. As each possible world shares aspects of the conceptualization, this overlapping knowledge serves as a foundation for semantic integration, aiding in achieving interoperability between the OVs.

The three theorems presented above lay the groundwork for understanding the dynamics of semantic integration in decentralized systems. They also further enrich our understanding of how semantic equivalence operates within these OVs.

6 FORMAL DEFINITIONS

This section extends the foundational concepts provided by Adhnouss (2023), with an emphasis on understanding semantic integration within decentralized information systems.

6.1 Conceptualization

The conceptualization (C) for a domain D is represented by the triple $C =< D, W, \Re >$, where D represents the domain, W is a set of maximal extensional structures within the domain, and $\Re$ stands for a set of intensional structures over the domain space $< D, W >$. Intensional relations $\rho_n : W \rightarrow 2^{Dn}$ are defined over this domain space, and the set of admissible extensions of ρ, $E_\rho = \rho(w) | w \in W$, is also defined here.

The intended extensional structure of w as per C is symbolized as $SwC =< D, R_{wC} >$, with $R_{wC} = \rho(w) | \rho \in \Re$, denoting the set of extensions (relative to w) of the elements of $\Re$. The set SC is defined as $SwC | w \in W$, representing all the intended world structures of C.

The structure of the domain in its extensional form is expressed as $C_{ex} =< D, R >= SwC$, which models the structure of the Conceptualized Extension (CE).

6.2 Intensional Semantic Level (Epistemology)

The intensional semantic level corresponds to the concept of epistemology described in (Adhnouss et al., 2023). We denote this as $\Theta =< C, \Im >$, where C stands for the conceptualization of a unique perspective of the domain and $\Im$ is an intensional interpretation function that maps elements of D to the constant symbols of V and elements of $\Re$ to predicate symbols of V. This is expressed as $\Im : V \rightarrow D \cup \Re$.

To align Θ with a specific domain, an extensional interpretation function I and a set of axioms are introduced.

6.3 Extensional Semantic Level (OV)

Given the intensional semantic level Θ and an extensional interpretation I, a model $M =< Sw, I >$ of L (extensional semantic Φ) is considered compatible with Θ if:

- The subset SwC is included in SC.
- For every $c \in V$, $\Im(c)$ is equal to $I(c)$.
- For every predicate $p \in V$, if $\Im(p) = \rho$ then there exists a world $w \in W$ such that $\rho(w) = I(p)$.

We denote the set of all extensions (models) of L that are compatible with Θ as the set of intended extensions $I_\Theta(L)$ of L in accordance with Θ.

The extensional semantic level (Φ) of an (Ov) is a specification of C defined by a language L, an extensional interpretation I, and a set of axioms that align the intensional interpretation with the intended extensions $I_\Theta(L)$.

In conclusion, we note that:

- The Extensional Semantic Level (Φ) commits to a Conceptualization C if it has been designed to represent C and approximate the reality D through its extensions.
- A language L commits to Θ if it adheres to the Conceptualization C such that Φ is consistent with C.
- L commits to Φ for a given Θ such that $I_\Theta(L)$ is incorporated in the models for Φ.

This rigorous framework lays the groundwork for semantic integration discussions, promoting interoperability within decentralized information systems.

7 OV MAPPING POSSIBILITIES

Understanding the relationship between conceptualization, the intended model, and the OV is crucial in a decentralized environment. This correlation is illustrated in Figures 1,2 and 3. These illustrations delineate three distinct scenarios that encapsulate the potential for mapping between two independent OVs, which are conceptualized as possible worlds (PW):

1. **Unified Intended Model:** This scenario features two independent OVs reflecting an identical intended model, signifying that they embody the same possible world. Condition:

$$(|W| = 1) \text{ and } (|D| = 1)$$

   This unity of the intended model implies that these OVs coexist within an isolated system, where they share the exact same conceptualization of the domain and illustrate a shared perspective. In such an environment, the overlap of these OVs encapsulates the mutual understanding of the domain, paving the way for an effective system-to-system mapping. Deciphering the unified intended model and its connotations is imperative for realizing semantic integration and interoperability.

2. **Intersecting Intended Models:** This scenario occurs when two independent OVs have overlapping intended models, implying they depict two unique and independent possible worlds. Condition:

$$(|W| > 1 \text{ and } |D| = 1) \text{ and } W \cap D \neq \emptyset$$

   This scenario is the primary interest of our research. In this situation, several OVs intersect, contributing to a single conceptualization of the Domain. This situation emerges in a decentralized environment, which is the central focus of our investigation. Even though the OVs share certain elements through the intersection of their intended models, they also portray diverse perspectives of the same domain. The accessibility relations among the possible worlds are instrumental in forming connections and facilitating mappings between the overlapping sections of the intended models. These relations ensure that the mapping between the OVs is exhaustive and dependable, enabling effective integration.

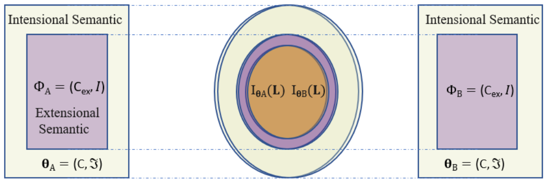

Figure 1: Scenario (1).

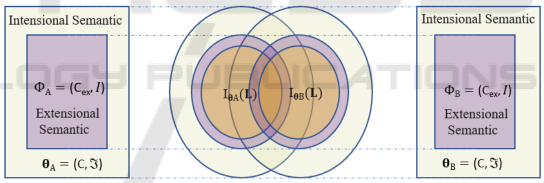

Figure 2: Scenario (2).

3. **Distinct Intended Models:** In the final scenario, two independent OVs exhibit no overlap in their intended models, suggesting the absence of a shared domain. Condition:

$$(|W| > 1 \text{ and } |D| > 1)$$

   This scenario transpires when there are multiple PWs and multiple Domains. Each OV might independently map to a different Domain, with no obligatory overlap or shared components. Consequently, mapping one OV onto the other becomes unfeasible in this scenario. Without an intersection of intended models, there is no mutual understanding or shared ground that can aid the integration process. This absence of overlap presents significant challenges for attaining semantic integration and interoperability between the two OVs.

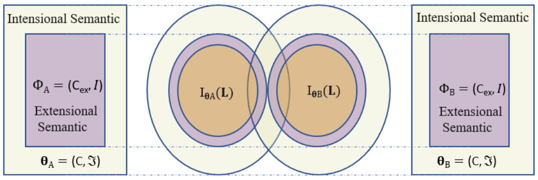

Figure 3: Scenario (3).

These scenarios depict varying degrees of mapping possibilities between OVs in decentralized environments, where each OV represents a possible world (PW). The crux of our study lies within the ambit of mapping possibilities (scenario 2), where two independent OVs intersect in their intended models. We propose a method to tackle this scenario, aiming to ensure the thoroughness and applicability of the mapped OV within the target information systems.

Theorem 4. *For a given conceptualization $C_{in} = \langle D, W, \Im \rangle$, a language L with non-logical symbols V, and an intensional semantics $\Theta = \langle C_{in}, \Im \rangle$, there exists a unique set of intended models of L according to Θ.*

Proof. Let's assume the existence of two sets of intended models $I_\Theta(L)_1$ and $I_\Theta(L)_2$, which are different. This would imply the existence of a model Φ compatible with Θ, such that $\Phi \in I_\Theta(L)_1$ but $\Phi \notin I_\Theta(L)_2$. However, by definition, if $\Phi \in I_\Theta(L)_1$, it is compatible with Θ, and similarly, Φ should belong to $I_\Theta(L)_2$. Therefore, such a model Φ cannot exist, leading to the conclusion that $I_\Theta(L)_1 = I_\Theta(L)_2$. □

Theorem 5. *Consider two distinct intensional conceptualizations $C_{in1} = \langle D_1, W_1, \Re_1 \rangle$ and $C_{in2} = \langle D_2, W_2, \Re_2 \rangle$, a language L with non-logical symbols V, and two intensional semantics $\Theta_1 = \langle C_{in1}, \Im_1 \rangle$ and $\Theta_2 = \langle C_{in2}, \Im_2 \rangle$, if the two sets of intended models for C_{in1} and C_{in2} overlap, the overlapped part consists of shared concepts and properties.*

Proof. Let's define the following: $D_1 = \{e_{1i} \mid 1 \leq i \leq n\}$ $D_2 = \{e_{2j} \mid 1 \leq j \leq m\}$ $I_{\Theta_1}(L) = \{\Phi_{1i} \mid 1 \leq i \leq k\}$ $I_{\Theta_2}(L) = \{\Phi_{2j} \mid 1 \leq j \leq l\}$

For each $1 \leq i \leq k$, let $\Phi_{1i} = \langle Cex_{1i}, I_{1i} \rangle$, where $Cex_{1i} = \langle D_1, R_{1i} \rangle$. Similarly, for each $1 \leq j \leq l$, let $\Phi_{2j} = \langle Cex_{2j}, I_{2j} \rangle$, where $Cex_{2j} = \langle D_2, R_{2j} \rangle$.

If $I_{\Theta_1}(L) \cap I_{\Theta_2}(L) \neq \emptyset$, then the overlapped part can be represented as $\langle D_1 \cap D_2, R_{1i} \cap R_{2j} \rangle$ for some $1 \leq i \leq k$ and $1 \leq j \leq l$.

The non-empty intersection $D_1 \cap D_2$ implies the presence of common concepts in the two conceptualizations. Furthermore, considering $R_{1i} = \{\rho(w) \mid w \in W_1\}$ and $R_{2j} = \{\rho(w) \mid w \in W_2\}$, where ρ represents the conceptual relation "Has", the non-empty intersection $R_{1i} \cap R_{2j}$ suggests the existence of relations $\{(e_1, e_2) \mid e_1 \in D_1 \wedge e_1 \in D_2 \wedge e_2 \in D_1 \wedge e_2 \in D_2\}$, where each e_2 represents a shared property of the shared concept e_1. □

These theorems establish a formal understanding of the mappings between different conceptualizations in a decentralized information system. Specifically, Theorem 1 confirms the uniqueness of the set of intended models for a language L with respect to a specific intensional semantics Θ. Meanwhile, Theorem 2 delves into scenarios where the intended models of two different conceptualizations overlap, demonstrating that the overlapped part is precisely defined by the shared concepts and properties in the conceptualizations.

These foundational theorems lay the groundwork for developing a theoretical framework supporting semantic interoperability in decentralized information systems. Their implications extend far beyond theoretical boundaries, enabling the identification of shared concepts and properties among diverse independent OVs and facilitating data and information mapping between disparate systems.

It is essential to acknowledge that while these theorems provide a solid foundation for semantic interoperability, their practical implementation will require further research to develop effective algorithms and protocols in real-world decentralized systems. Addressing situations where the intended models of independent OVs do not overlap or only partially overlap will pose a significant challenge, necessitating continuous investigation and exploration.

8 AN ILLUSTRATION OF SEMANTIC INTEGRATION: A HEALTHCARE SCENARIO

We consider a medical domain wherein we have entities represented by the set E:

- P represents the Patient
- N stands for Name
- M for Medical History
- C for Current Condition
- A for Address
- H for Healthcare Provider

Within this domain, we establish a binary relation r such that $r(P, X)$ signifies that a Patient P has property X.

We opt for English as the language to model this domain, defining a vocabulary of non-logical symbols V as follows:

- **Patient**
- **Name**
- **MedicalHistory**
- **CurrentCondition**
- **Address**
- **HealthcareProvider**
- **has** (as the predicate symbol)

All elements, except **has**, are constant symbols.

With the domain $D = E$ and the single conceptual relation $\rho = r$, our conceptual schema $C_{ex} = \langle D, R \rangle$ encompasses D and the set of relations $R = \{\rho\}$. In this context, the possible world w posits that a patient can possess the aforementioned properties.

We proceed to construct a model of the language, denoted as $\Phi_1 = \langle C_{ex}, I_1 \rangle$. Here, the relational structure R, which arises from the application of ρ to w, yields $R = \{(P,N),(P,M),(P,C),(P,A),(P,H)\}$.

Let's assume that I_1 is an interpretation function assigning elements of R to predicate symbols in V. We define $I_1(\textbf{has}) = \{(P,N),(P,M),(P,C)\}$. This definition implies that our model prioritizes the clinical aspects of a patient.

Similarly, we can establish another model $\Phi_2 = \langle C_{ex}, I_2 \rangle$, where I_2 is defined as $I_2(\textbf{has}) = \{(P,N),(P,A),(P,H)\}$. This model underscores the logistical and administrative aspects of patient care.

Assuming that Φ_1 and Φ_2 adhere to intensional semantics Θ, they are the intended models of the language L, given Θ. Hence, $I_\Theta(L) = \{\Phi_1, \Phi_2\}$.

In this instance, we can observe that if two information systems stick to identical conceptualizations and use the same vocabulary, they can achieve mutual agreement as the vocabulary symbols are consistently interpreted.

Let's consider two distinct intensional conceptualizations C_{in1} and C_{in2}:

- $C_{in1} = \langle D_1, W_1, \Re_1 \rangle$, where:
  - $D_1 = \{P, N, M, C\}$,
  - W_1 corresponds to "Patient has property Name, Medical History, and Current Condition",
  - $W_1 = \{W_1\}$,
  - $\Re_1 = \{r\}$.
- $C_{in2} = \langle D_2, W_2, \Re_2 \rangle$, where:
  - $D_2 = \{P, N, A, H\}$,
  - W_2 corresponds to "Patient has property Name, Address, and Healthcare Provider",
  - $W_2 = \{W_2\}$,
  - $\Re_2 = \{r\}$.

Given the same language L and vocabulary V, we construct intensional semantics $\Theta_1 = \langle C_{in1}, \Im_1 \rangle$ and $\Theta_2 = \langle C_{in2}, \Im_2 \rangle$, where:

- $\Im_1(\textbf{Patient}) = P$, $\Im_1(\textbf{Name}) = N$,
  $\Im_1(\textbf{MedicalHistory}) = M$,
  $\Im_1(\textbf{CurrentCondition}) = C$,
  $\Im_1(\textbf{has}) = r$.
- $\Im_2(\textbf{Patient}) = P$,
  $\Im_2(\textbf{Name}) = N$,
  $\Im_2(\textbf{Address}) = A$,
  $\Im_2(\textbf{HealthcareProvider}) = H$,
  $\Im_2(\textbf{has}) = r$.

We then define models $\Phi_1 = \langle C_{ex1}, I_1 \rangle$ and $\Phi_2 = \langle C_{ex2}, I_2 \rangle$, corresponding to these intensional semantics. When we examine $I_{\Theta_1}(L) \cap I_{\Theta_2}(L)$, we notice a convergence on the concept of Patient and its property Name.

In such a scenario, when two information systems commit to different conceptualizations but employ the same vocabulary, partial agreement can be achieved. Symbols related to shared concepts and properties are consistently interpreted. In contrast, symbols associated with distinct concepts and properties might induce misunderstandings and conflicts.

This discussion illustrates that semantic interoperability between two information systems hinges not only on syntactic compatibility (i.e., the vocabulary used) but also on semantic congruity (i.e., the intended models of the language per the intensional semantics).

Therefore, to ensure efficient semantic integration and semantic interoperability, systems must consider both the shared meaning of the vocabulary, specifically non-logical symbols, and the underlying conceptualizations and intensional semantics. A common understanding of the domain's concepts and relations is essential to achieve effective interoperability.

9 CONCLUSIONS

In this paper, we have embarked on an exploration of semantic integration within decentralized systems, with a particular focus on OVs and their semantics. We highlighted the importance of semantics, and more specifically, referential semantics, in achieving semantic integration in the absence of a global view. The concept of semantic equivalence between OVs, both at the syntactic and the semantic level, was discussed in detail, emphasizing its significance in the alignment of different OVs.

We also proposed three theorems that examine the relationship between accessibility relations and the overlap of non-logical symbols between different worlds. These theorems offer a novel perspective on semantic integration, revealing the potential of accessibility relations and overlapping non-logical symbols as critical factors in achieving semantic consistency and integration.

Despite the insights this paper presents, several areas of study demand future research. In particular, our approach could be extended and enriched by further exploring the roles of axioms and their semantics in the OVs. As axioms are often the source of intensional semantics, understanding their function could offer a deeper insight into semantic integration. Moreover, the exploration of further ways to achieve semantic integration, such as the development of ontology matching techniques or the use of machine learning algorithms, could also enhance this work.

Finally, a critical area for future investigation is the application of our theoretical framework to real-world scenarios. Practical implementations in diverse fields such as healthcare, finance, and e-commerce could serve to validate our approach, providing essential feedback to refine our understanding of semantic integration within decentralized systems. This is an area we are currently investigating, and we hope our findings will stimulate further research and discussion in this domain.

REFERENCES

Adhnouss, F. M. A., El-Asfour, H. M. A., McIsaac, K., and El-Feghi, I. (2023). A hybrid approach to representing shared conceptualization in decentralized ai systems: Integrating epistemology, ontology, and epistemic logic. *AppliedMath*, 3(3):601–624.

Adhnouss, F. M. A., El-Asfour, H. M. A., McIsaac, K. A., El-Feghia, I., Aburukba, R., and Wahaishi, A. (2022). Ontological view-driven intensional semantic integration for information systems in a decentralized environment. In *Proceedings of the 14th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2022, Volume 2: KEOD, Valletta, Malta, October 24-26, 2022*, pages 109–116. SCITEPRESS.

Ali, I. and McIsaac, K. A. (2020). Intensional model for data integration system in open environment. In *KEOD*, pages 189–196.

Cao, Q., Beden, S., and Beckmann, A. (2022). A core reference ontology for steelmaking process knowledge modelling and information management. *Computers in Industry*, 135:103574.

Chandrashekar, D., Fazal, S. M., and Suneetha, K. (2023). 73 enterprise application development using semantic web languages for semantic data modeling. In *Data Science with Semantic Technologies: Deployment and Exploration*, pages 73–88. CRC Press.

de Mello, B. H., Rigo, S. J., da Costa, C. A., da Rosa Righi, R., Donida, B., Bez, M. R., and Schunke, L. C. (2022). Semantic interoperability in health records standards: a systematic literature review. *Health and Technology*, 12(2):255–272.

Dhakal, S., Zhang, L., and Lv, X. (2022). Ontology-based semantic modelling to support knowledge-based document classification on disaster-resilient construction practices. *International Journal of Construction Management*, 22(11):2059–2078.

Du, Y., Shi, L., Zhai, X., Gong, H., and Zhang, Z. (2022). Knowledge extract and ontology construction method of assembly process text. In *MATEC Web of Conferences*, volume 355, page 02029. EDP Sciences.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International journal of human-computer studies*, 43(5-6):907–928.

Guarino, N., Oberle, D., and Staab, S. (2009). What is an ontology? In Staab, S. and Studer, R., editors, *Handbook on ontologies*, pages 1–17. Springer.

Majkić, Z. and Prasad, B. (2018). Intensional fol for reasoning about probabilities and probabilistic logic programming. *International Journal of Intelligent Information and Database Systems*, 11(1):79–96.

Melluso, N., Grangel-González, I., and Fantoni, G. (2022). Enhancing industry 4.0 standards interoperability via knowledge graphs with natural language processing. *Computers in Industry*, 140:103676.

Polenghi, A., Roda, I., Macchi, M., Pozzetti, A., and Panetto, H. (2022). Knowledge reuse for ontology modelling in maintenance and industrial asset management. *Journal of Industrial Information Integration*, 27:100298.

Wang, Y. D., Ghenniwa, H., and Shen, W. (2009). Ontological view based semantic transformation for distributed systems. In *2009 IEEE International Conference on Systems, Man and Cybernetics*, pages 4007–4012. IEEE.

Xue, Y., Ghenniwa, H. H., and Shen, W. (2012). Frame-based ontological view for semantic integration. *Journal of Network and Computer Applications*, 35(1):121–131.

Zhang, J., Mytkowicz, T., Kaufman, M., Piskac, R., and Lahiri, S. K. (2022). Using pre-trained language models to resolve textual and semantic merge conflicts (experience paper). In *Proceedings of the 31st ACM SIGSOFT International Symposium on Software Testing and Analysis*, pages 77–88.