

Introducing Multiboxes: Increasing the Trustworthiness of Ontology Embeddings

Giuseppe Federico^{1,2,*}, Mena Leemhuis³ and Oliver Kutz¹

¹KRDB Research Centre for Knowledge-Based Artificial Intelligence, Free University of Bozen-Bolzano, Bolzano, Italy

³Institute for Symbolic AI, Johannes Kepler University, Linz, Austria

²Sony AI, Barcelona

Abstract

In the current AI landscape, the fast increase in performance and result quality of various AI systems is accompanied by a need and growing interest in trustworthy and interpretable AI approaches. One way of increasing the trustworthiness of subsymbolic learning approaches is to incorporate background knowledge information in the form of ontologies. It is, however, necessary to bridge the gap between the set-based, Tarskian semantics of ontologies, and the subsymbolic viewpoint of learning approaches. In this context, ontology embeddings come into play by representing concepts geometrically in a vector space. These representations are often based on simple geometric structures such as boxes or spheres. In this paper, it is argued that an embedding in the form of a geometric representation of ontologies may enforce representations of concepts that are not conforming to those simple geometric structures. Namely, embedded concepts may not necessarily be convex or they may even be unconnected. We argue that it is necessary to allow such non-conforming structures in order to guarantee trustworthiness, both based on the intensional representation of concepts and based on their learnability. To address these issues, we thus introduce a class of representations that extend the simple box semantics of embeddings, namely so-called multiboxes. We show that the framework of multiboxes is both sufficiently expressive to regain trustworthiness, and at the same time provides a computationally feasible embedding regime. Our main technical contribution is that every $\mathcal{ELHO}(\circ)^+$ -ontology is satisfiable if and only if it has a multibox interpretation.

Keywords

Ontology embeddings, Box embeddings, Convexity, Explainability, Neurosymbolic AI

1. Introduction

Trustworthiness of learning approaches is a topic of increasing importance, particularly in the context of subsymbolic learning approaches. In this context, ontologies come in handy as a way to represent background knowledge information in a structured fashion. Including these into the learning process increases the trustworthiness by allowing for interpretable and knowledge grounded outcomes. However, there is a gap between the subsymbolic, distance-based representation and the structured, Tarskian-style semantics of ontologies. Ontologies have an extensional semantics, thus interpretations of ontologies are focusing on individuals and their concept membership from a set-theoretic viewpoint. Thus, any topological or geometric information, e.g., on the shape or similarity of concepts or individuals is not modeled. Here, other views turn out useful, e.g., prototypes [1] or the *Conceptual Spaces* by Gärdenfors [2]. Gärdenfors proposes that concepts are represented as convex objects in some space that is based on quality dimensions. In this context, for each concept and each individual, an intensional representation exists describing the attributes of the concept resp. of the individual. A similar view has been adopted in the area of *Knowledge Base Embeddings* (KBEs) [3]. There, ontologies are embedded into a vector space by representing concepts as specific types of convex objects, e.g., spheres [4] or boxes [5]. Logical operators between concepts or individuals are represented as geometric operations between the respective representations. In contrast to conceptual spaces, in KBE, the dimensions do not have a meaning. Thus, there is an implicit intensional representation of the concepts and individuals

Proceedings of the Joint Ontology Workshops (JOWO) - Episode XII: The Tropical Spring of Ontology, co-located with the 16th International Conference on Formal Ontology in Information Systems (FOIS 2026), September 21-22, 2026, Vitória (ES), Brazil

*Corresponding author.

✉ Giuseppe.Federico@student.unibz.it (G. Federico); mena.leemhuis@jku.at (M. Leemhuis); Oliver.Kutz@unibz.it (O. Kutz)

id 0009-0006-6487-1544 (G. Federico); 0000-0003-1017-8921 (M. Leemhuis); 0000-0003-1517-7354 (O. Kutz)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

learned. Here, although sometimes contested as e.g. in [6], the entity similarity assumption is assumed to hold, which states that similarities of concepts or individuals are reflected as short distance between them in the embedding space. This fact can be used both for inferring new information based on, e.g., similar individuals, and in general for having a strong connection between symbolic ontologies and subsymbolic information.

To achieve a trustworthy approach, it is, however, necessary to satisfy several further desiderata. Two key aspects are:

(1) An ontology should be representable geometrically in a Tarskian-style manner such that an ontology is satisfiable if and only if it has a geometric representation. Thus, the semantics of the embedding should be as expected by a human user.

(2) The geometric representation should actually be able to allow for an embedding to satisfy the entity similarity assumption, thus realising an embedding that does not introduce strong biases.

Existing embedding approaches show limits in both of these aspects. In the following, we are particularly focusing on the choice of the concept representation and its influence on the trustworthiness. This issue is discussed in detail in this paper, proposing a specific representation of concepts that is based on simple topological structures and at the same time provides good learnability and a sufficient expressivity. These structures are the so-called *multiboxes*.

More in detail, we show that for each $\mathcal{ELHO}(\circ)^\perp$ ontology there exists an embedding based on multiboxes modeling this ontology. Thus, they are sufficiently general. However, they still allow for implementations with low computational cost and thus support good learnability.

2. Preliminaries

In this section, we introduce the formal frame in which our discussion takes place. We first introduce the description logic $\mathcal{ELHO}(\circ)^\perp$, whose choice is motivated by its computational advantages, as subsumption is polynomial, and by its employment in prominent ontologies like SNOMED [7] for clinical documentation and the Gene Ontology [8] for modeling genes and their interactions. Secondly, we define boxes and box embeddings of ontologies as the basis for the multibox embeddings introduced in Section 4.

2.1. The Description Logic $\mathcal{ELHO}(\circ)^\perp$

We consider the $\mathcal{ELHO}(\circ)^\perp$ -fragment of the well-known description logic $\mathcal{EL}^{++}$ [9]. Although this logic is relatively expressive, it exhibits some fundamental limitations: it neither allows for negation nor universal quantification. For general details on syntax and semantics of description logics, see [10].

A DL vocabulary is given by a set of individual names $\mathbf{I}$, a set of role names $\mathbf{R}$ and concept names $\mathbf{C}$. The $\mathcal{ELHO}(\circ)^\perp$ concepts over $\mathbf{C} \cup \mathbf{R} \cup \mathbf{I}$ are written according to the syntax rules

$$C \longrightarrow A \mid \{a\} \mid \perp \mid \top \mid C \sqcap C \mid \exists R.C$$

where $A \in \mathbf{C}$, $a \in \mathbf{I}$, $\{a\}$ denotes a nominal concept, $R \in \mathbf{R}$ and C stands for arbitrary concepts. A terminological box (Tbox) $\mathcal{T}$ consists of general inclusion axioms (GCIs) $C \sqsubseteq D$ with C, D being concept descriptions. An assertional box (Abox) $\mathcal{A}$ consists of individual assertions, namely facts of the form $a : C$ or $(a, b) : R$ with $a, b \in \mathbf{I}$, $R \in \mathbf{R}$ and C being a concept description. The Rbox $\mathcal{R}$ consists of *role inclusions* $R_1 \circ \dots \circ R_k \sqsubseteq R$ and *role hierarchies* $R_1 \sqsubseteq R$. An ontology $\mathcal{O}$ is a triple of the form $(\mathcal{T}, \mathcal{R}, \mathcal{A})$.

The semantics of arbitrary concept descriptions for a given interpretation $\mathcal{I}$ is given in Table 1 (details can be found in [9]). In this paper, the focus lies on finite ontologies.

Name	Syntax	Semantics
top	$\top$	Δ
bottom	$\perp$	$\emptyset$
nominal	$\{a\}$	$\{a^{\mathcal{I}}\}$
conjunction	$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
existential restriction	$\exists R.C$	$\{x \in \Delta \mid \exists y \in \Delta : (x, y) \in R^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\}$
role concatenation	$(R_1 \circ R_2)^{\mathcal{I}}$	$\{(a, c) \mid \exists b \in \Delta : (a, b) \in R_1^{\mathcal{I}}, (b, c) \in R_2^{\mathcal{I}}\}$

Table 1

Syntax and semantics of $\mathcal{ELHO}(\circ)^{\perp}$ [9]

2.2. Box Embeddings of Ontologies

We introduce boxes and their geometrical properties, and explain how they can serve as a basis for an interpretation of ontologies. Box embeddings are widely studied in the context of knowledge base embeddings, see, e.g., [11]. However, also outside of this area, boxes turned out useful as basis for concept representations due to their computational ease and therefore good learnability. The aim of this work is to introduce multibox interpretations as an extension of box interpretations with increased representative abilities.

Definition 1. A box in some $\mathbb{R}^n$, for $n \in \mathbb{N}$, is defined as a Cartesian product of intervals $I_1 \times \dots \times I_n$, where each $I_i \subseteq \mathbb{R}$ is a closed interval or a half line or the whole $\mathbb{R}$.

One main feature of axis-aligned boxes is their closure under intersection. The intersection of two axis-aligned boxes A and B is defined as follows:

$$C := A \cap B = I_1^A \cap I_1^B \times \dots \times I_n^A \cap I_n^B.$$

In Figure 1 (a), an example of such an intersection of two boxes can be seen. Note here that arbitrary dimensional boxes are considered and that the examples are restricted to the two-dimensional case for representation reasons only.

Let $A \subseteq \mathbb{R}^n$ for some n , we define $\text{BoxHull}(A)$ to be the smallest box containing all elements of a set A , this can be explicitly defined as

$$\text{BoxHull}(A) = \{(x_1, \dots, x_n)^T \mid x_i \in \text{ConvHull}(\{a_i \mid a \in A\}) \text{ for } 1 \leq i \leq n\}$$

where $\text{ConvHull}(X)$ is the convex hull of X and a_i is the value at the i -th dimension of the vector a . Let the set $\mathcal{B}^n = \{\text{BoxHull}(X) \mid X \subseteq \mathbb{R}^n\}$ be the set of all boxes in $\mathbb{R}^n$ including the whole space $\mathbb{R}^n$ and the empty set.

Now, assume that an $\mathcal{ELHO}(\circ)^{\perp}$ -ontology $\mathcal{O} = (\mathcal{T}, \mathcal{R}, \mathcal{A})$ is given. Its box interpretation is based on a (classical) model of this ontology under some restrictions. Concepts are represented as boxes in the domain $\mathbb{R}^n$. Individuals are represented as points, and nominals as special type of boxes that include only one point. The relations are defined as general as possible. They are, however, not allowed to be arbitrary: as each concept needs to be represented as a convex object, also each pre-image of a box needs to result in a box (e.g., the representation of $\exists R.C$).

The following definition follows in principle the one in [11, Def.1], however, focuses on the fact that box interpretations are a special case of classical interpretations.

Definition 2. A classical interpretation $\mathcal{I} = (\Delta, \cdot^{\mathcal{I}})$ is a box interpretation if

- $\Delta \in \mathcal{B}^n$ for some $n \in \mathbb{N}$,
- $\cdot^{\mathcal{I}}$ maps
 - each concept name $A \in \mathcal{C}$ to some box in $\mathcal{B}^n$
 - each nominal concept $\{c\}$ to the box $\{c^{\mathcal{I}}\}$

- each role $R \in \mathbf{R}$ to a subset $R^{\mathcal{I}} \subseteq \Delta \times \Delta$ such that for every $B \in \mathcal{B}^n$: $R^{\mathcal{I}^{-1}}(B) \in \mathcal{B}^n$.

For ease of discrimination, a box interpretation is denoted with ξ on the structure $(\Xi, \cdot^\xi)$.

It is important to note that this definition of box embeddings is one focused on its relation to classical interpretations. For application settings, thus for the definition of actual embedding approaches, more flexible (but less classical) definitions are used. To avoid overfitting and to ease learning, it is often not the case that arbitrary relations (as defined in Definition 2) are modeled but restricted ones that can be, for instance, modeled as translations. Whereas these would still be interpretable as classical interpretations, there are also many embedding approaches that change basic principles of interpretations, e.g., modeling a conjunction not as an intersection.

3. Suitable Concept Representations

One main design question for KBE approaches, which is equally raised for any approach that considers embedding ontological information in general, is the question how to choose a suitable concept representation framework.

In this context, two different aspects need to be considered. The first, and from a practical viewpoint more urgent point, is the computational ease and avoidance of overfitting afforded by the representation framework. This aspect is discussed in Section 3.1. The second aspect considers the intensional semantics of concepts, thus their inherent meaning and structure, an issue which is partly independent of the representation issues referred to in the first point. This aspect is discussed in Section 3.2.

3.1. Computational Aspects of Concept Modelling

The aim of KBE approaches is two-fold: first, an embedding should be able to sufficiently represent the ontological information. Thus, the minimal requirement would be soundness in the sense that a geometric interpretation only exists if the ontology is satisfiable. This prevents, for instance, to draw inferences based on inconsistent information. Ideally, also completeness should be obtained, meaning in this context that for each satisfiable ontology a geometric interpretation exists that is, ideally, also findable by an embedding approach. However, as discussed by Bourgaux et al. [12], this is not sufficient. It is furthermore necessary to resemble classical Tarskian semantics as much as possible. This is a necessary building block of a trustworthy system: if an ontology is represented, then this representation should resemble a “natural” human-compatible semantics. This, however, leads to a tradeoff between computational cost and representational power.

The most obvious solution would be to learn an embedding in the style of the classical interpretation, thus modeling concepts as arbitrary sets in some geometric space. This, however, would interfere with the second goal of embedding approaches, namely the learnability. To be able to infer (ideally) correct new facts, it is necessary to have a concept representation that is both simple enough to be computable and allows for inferences. An interpretation that models exactly the information already stated in the ontology would be perfectly trustworthy in the sense that no incorrect information is inferred, however, it is not useful, as nothing is inferred at all that is not already deducible from the ontology.

However, trustworthiness covers a wider spectrum of aspects of such an approach (see, e.g., the work of Li et al. [13] on Trustworthy AI). Therefore, we are omitting a definition of trustworthiness here and concentrate on some key aspects, namely not only the reliability of the result but also the performance quality and robustness against noise. A geometric interpretation should not only model exactly the information stated in the ontology but also need to allow for new inferences, e.g., based on the similarity of individuals. Additionally, it should be still interpretable and semantically and logically grounded.

To obtain this, a simpler representation of concepts is needed that (1) is computationally feasible (thus has a sufficiently low amount of variables to learn per dimension of the embedding space and also the embedding space has a restricted dimension), (2) uses data regularities for learning and (3) avoids overfitting thus does not only models the given data but allows for inferences.

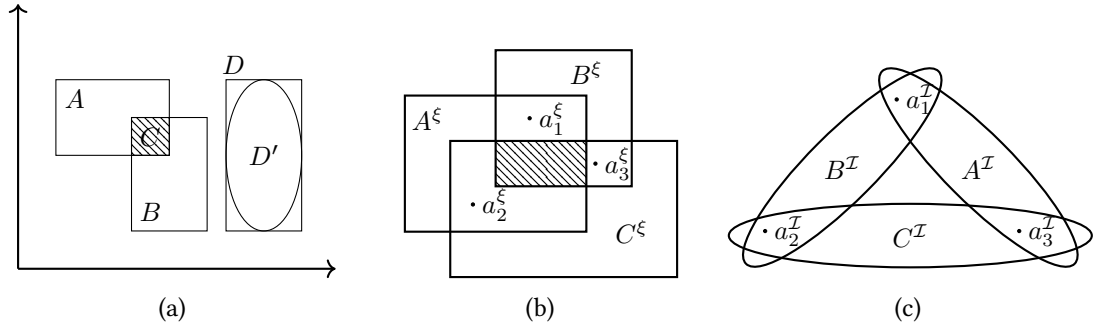


Figure 1: (a) A, B, C and D are examples for axis-aligned boxes. The intersection of two axis-aligned boxes (e.g., A and B) results always in an axis-aligned box (in this case C). $D = \text{BoxHull}(D')$, thus D is the box hull of D' ; (b) a box interpretation exemplifying the Helly Property; (c) a classical model of the ontology given in Example 4.

The first step is to consider convex concepts. These are advantageous, as they allow for representing concepts based on their defining points. It allows for making inferences, as if two individuals are member of a concept, then the individual represented in between these two is also in that concept. This, of course, severely reduces the reliability of the system: There is no guarantee that the inference is actually correct, however, it increases the performance and robustness. The robustness increases, as outliers, e.g., erroneous data can be corrected in this model due to geometric regularities. This leads to a clear tradeoff between reliability on the one side and performance and robustness on the other side.

However, even when having convex concept representations, these could still be too complex. An additional necessary property is therefore to have a bounded number of variables defining the convex set. This is, e.g., the case for axis-aligned boxes, where there is an upper and a lower bound per dimension, meaning there are $2n$ variables needed for describing a n -dimensional box. This avoids overfitting and makes learning approaches feasible. Such a simplification, however, comes always to the cost of a reduced trustworthiness. Indeed, it is easy to show (as proved in [11]) that there is an $\mathcal{ELHO}(\circ)^\perp$ ontology $\mathcal{O}$ that does not admit any box interpretation. The reason behind this is a combinatorial property of boxes, the so called “Helly Property”. When using boxes, ontologies interfering with this property can not be modeled.

Definition 3 (Helly Property, adapted from [14]). *A family B fulfills the Helly Property if it is the case that: $\bigcap_{b \in B} b \neq \emptyset$ if and only if for all $b_1, b_2 \in B$: $b_1 \cap b_2 \neq \emptyset$.*

Using the fact that boxes fulfill the Helly Property [14], it has been shown in [11] that there exists a classically satisfiable $\mathcal{ELHO}(\circ)^\perp$ ontology $\mathcal{O}$ such that no box interpretation in $\mathbb{R}^n$, for arbitrary $n \in \mathbb{N}$, satisfies $\mathcal{O}$. An example of such an ontology is given in the following.

Example 4. Let $\mathcal{O} = (\mathcal{T}, \mathcal{R}, \mathcal{A})$ with $\mathcal{T} = \{A \sqcap B \sqcap C \sqsubseteq \perp\}$, $\mathcal{R} = \{\}$ and $\mathcal{A} = \{a_1 : A \sqcap B, a_2 : B \sqcap C, a_3 : A \sqcap C\}$. This ontology is clearly satisfiable (see Figure 1 (c)). However, when trying to construct a box interpretation either one individual is incorrectly assigned or the Tbox axiom is not fulfilled (see Figure 1 (b) for a box interpretation not in line with the ontology). This resembles the Helly Property.

Therefore, the choice of a sufficient concept representation influences the trustworthiness of the approach. Not each representation is able to be used as basis for representing each ontology. This problem can be tackled by using more complex geometric representations (e.g., cones [15] instead of spheres or boxes). It is also possible to relax the underlying notions, e.g., of convexity, as done for the conceptual orthospaces [16]. Another option is to consider weaker ontology languages such as $\mathcal{ELH}$. Lacerda et al. [17] present an embedding based on convex concept representations that is able to represent each $\mathcal{ELH}$ ontology. In this context, they have also considered non-convex representations. Another example by Lourenço et al. [18] considers box embeddings for DL-Lite^H, also omitting above mentioned problems, as DL-Lite^H does not allow for disjointness of concepts. TransBox [19] allows for modeling concept conjunction as an approximated intersection, circumventing this problem, however, to the cost of loosing a close connection to classical semantics.

We are considering here a different variant: we want to stick to boxes due to their computational feasibility, to keep the expressive description logic $\mathcal{ELHO}(\circ)^+$, and be as close to classical semantics as possible. On the other hand we want to circumvent restrictions, e.g., based on the Helly Property. Therefore, we are using multiboxes, thus unions of several boxes representing one concept. Doing this, we circumvent the restrictions due to the Helly Property. Multiboxes are introduced in detail in Section 4. These multiboxes are expected to show a good tradeoff between different aspects of trustworthiness: they are, as will be shown in Theorem 3, able to represent every satisfiable $\mathcal{ELHO}(\circ)^+$ -ontology geometrically. However, they still are not arbitrary sets but based on the simple structure of boxes and therefore should have sufficiently good learnability qualities. Thus, a trustworthy embedding does not only exist but is also plausibly findable.

3.2. Semantic Aspects of Concept Modeling: Convexity and Connectedness

In particular based on the entity similarity hypothesis and also cognitive principles such as the *universal law of generalization* [20], the assumption and goal of embedding methods is to model individual or concept similarity as distance between the resp. representations. Thus, it is not sufficient to consider whether an ontology is adequately representable. Additionally, it is necessary to justify whether this hypothesis should hold. This of course needs to be tested empirically, however, also based on the general construction of concepts some conclusions can be drawn about whether an adequate semantics is represented. Based on the computational ease but also motivated from a cognitive viewpoint, it is widely argued that concept representations should be convex, for example by Gärdenfors [2]. However, he also proposes weaker representations, such as connectedness and starshapedness. Starshapedness is influenced by the prototypical view on concepts: there is one prototypical representation that needs to have a convex connection to each other element of the concept, however, convexity between arbitrary elements of the concept does not need to hold. Convexity is often considered as a too strong restriction, Hernández-Conde argues in [21] that “none of the arguments provided by Gärdenfors for the convexity constraint constitutes a compelling reason in favor of that requirement”. This strong position is motivated by several reasons, he underlines that most of the assumptions made by Gärdenfors are problematic, that the notion of convexity depends on the chosen metric and the Euclidean one is not obviously preferable. He also criticizes how the notion of “natural concept” is used from Gärdenfors only as an extreme defense against the revelation of non-convex concepts like the “bleen” concept of Goodman [22]. In [21, ch.8], the author presents his strongest arguments against convexity, the first one is the more interesting for us, it is simple and states that “there are non convex physical properties”. He introduces this issue in the attempt to represent the concept of “apple”. Since its geometrical shape is not convex how could it be captured in a convex manner?

However, problems also arise when considering the color dimension. An apple can be green, red, or yellow, but not every color in between, for instance, not blue (see Figure 2 (a) for an illustration). Here, it is necessary to note that the geometric interpretations we are considering are not based on quality dimensions and thus are not conceptual spaces but a relaxed version of it without introducing meaningful dimensions. Thus, the apple-example is not directly applicable. The general line of argument, however, still holds: if for some concepts, there is a convex connection between “red” and “green” and for others, it is not, then the representation could be problematic to learn, even without quality dimensions. This is due to the fact that all concept representations are based on the same notion of convexity. Building on this observation, we pose in discussion even the connectedness assumption, by asking, are there non connected physical concepts?

As a starting point, we can again discuss the example of the color of an apple. There, it can be plausibly argued that the apple concept, consisting of differently colored apples is not only not convex but also unconnected (see again Figure 2 (a)). One solution strategy would be to consider an implicit representation of the color attribute in combination with other attributes. This is possible, as we do not consider quality dimensions. Then, there could be a two-colored apple in between the two, instead of an, e.g., orange one. This, however, depends heavily on the use case and could interfere with other concept representations in the same domain that rely on other notions of betweenness, thus of convexity. This

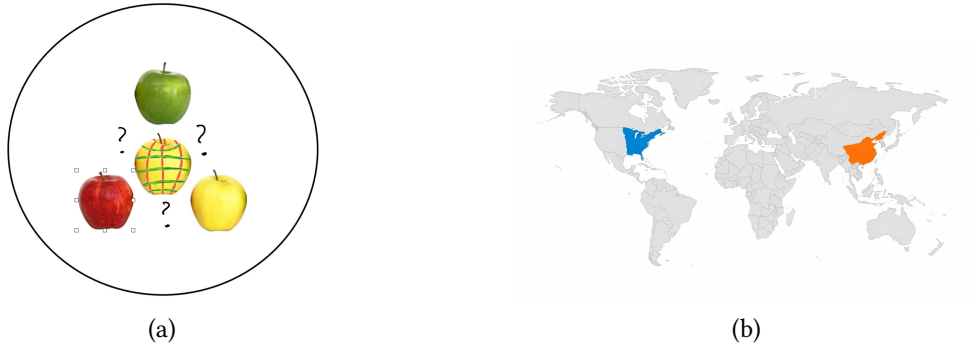


Figure 2: (a) The conceptual representation of apples, depicting the non-existence of in-between concepts; (b) habitat of alligator, depicting an unconnected concept.

apple example can be straightforwardly adapted to other concepts defined based on an equivalence relation, thus concepts consisting of a partition of different subconcepts, thus are a disjunction of disjoint subconcepts.

The implications of such unconnectedness and other occurrences of it are discussed in the following two examples, one based on a spatial argument considering the concept ‘habitat of’, the other based on the difference between visible and genetic features.

Example 5. *One prominent example is the habitat of animals. The geographical areas of an animal’s habitat are often disconnected. Let the concept ‘habitat of x ’ be represented as the area covered by the countries in which species x can be found today. A spatial representation can be seen in Figure 2 (b). Note here again, that a geometric interpretation of this fact would not have these two clearly defined spatial dimensions. In this context, alligators can be considered. The genus *Alligator* has only two living species: the American alligator (*Alligator mississippiensis*) and the Chinese alligator (*Alligator sinensis*). Thus, in our trivial representation the concept ‘habitat of alligator’ is represented by two disconnected regions.*

Even when representing these spatial properties only implicitly, there is an inherent unconnectedness, as the two habitats are not only differing spatially, but also based on many other properties such as climate, flora, and fauna.

It is important to note that the use of formal ontologies is widespread and promising in disciplines such as evolutionary biology and biological geography. Notable examples include ontologies such as ENVO (Environmental Ontology) and OntoBio (see [23] and [24]). Therefore, these ontologies represent ideal candidates for testing the embedding method proposed in Section 4. Additionally, spatial considerations in general are a clear problematic aspect of convex concept representations, independent of the use case.

In the previous example, we saw how some simple concepts can have a “natural” geometrical representation that is not connected. The question is now, is it possible to have a good convex representation of an unconnected concept? The answer is negative and the reason is topological: the number of connected components of a set is a topological invariant. If a spatial concept is intrinsically disconnected, then any representation preserving its topological structure must also be disconnected. The reason is that connectedness is a topological invariant preserved by homeomorphisms; therefore, any homeomorphic representation has the same number of connected components. In this sense, connectedness is a stronger constraint than convexity: non-convexity can often be eliminated by a suitable transformation, whereas disconnectedness cannot without loss of topological information.

Thus, at least for geometrical concepts, it is important to introduce some embedding technique that allows for non convex and unconnected representation. In the following, we give another, non-spatial example representing a scenario in which the representation of a set of entities tries to capture simultaneously totally different aspects of these entities.

Example 6. *We are considering phylogeny, thus the diversification of a species, and evolutionary convergence. In particular, we want to introduce the concepts of phylogenetic tree and evolutionary convergence, then we discuss the role of convexity and connection in this specific setting. A phylogenetic tree is a diagram*

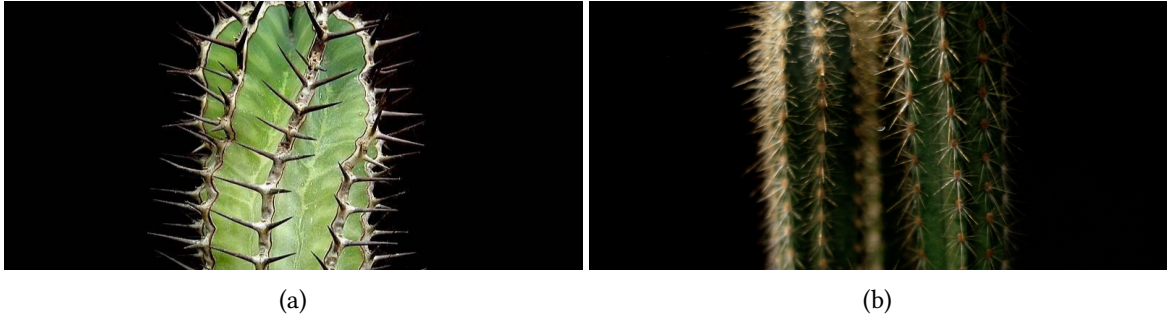


Figure 3: An euphorbia virosa (a) and a cactus (b); showcasing the notion of evolutionary convergence discussed in Example 6

that shows the evolutionary relationships between species. Each node represents a species. When node A is an ancestor of node B , it means species B evolved from species A . It is often the case, in biological disciplines, that for a single concept there are several possible non equivalent definitions and that of evolutionary convergence makes no exception. A rich discussion regarding the definition of “evolutionary convergence” is contained in [25], for our purposes it is enough to consider the definition given in [26]: “Evolutionary convergence is the evolution of similar features in independent evolutionary lineages.” In this context, some plants can be considered, namely Cactus (Cactaceae, Figure 3 (a)) and Euphorbia virosa (Euphorbiaceae, Figure 3 (b)), two species with a completely distinct phylogenetical tree. These plants live in different continents and have a totally different internal functions. However, under a morphological point of view they look very similar; both have columnar structure, spines, and a succulent stem. What we are describing is a situation in which two entities are really similar regarding some features but totally different regarding others. So, suppose now that we want to give a geometrical representation of the plant kingdom in way that (1) distances between species in the representation space reflect the distances in the phylogenetical tree; (2) the representation criteria is also based on the features of the species. Namely, we want that our representation is able to distinguish species in the phylogenetical sense but it is also able to capture their similarities. For instance, let C be the concept “Plant with a columnar shape and spines that lives in desertic areas” and consider possible representations of C . In this example, C contains two sub-concepts A (cactus) and B (euphorbia virosa) that must be at a huge distance in the representation space. Moreover, there are no other species with the properties described by C , so C needs to have at least two components, one containing the representation of A and the other containing the representation of B . Then, it would be possible to model some form of context-based similarity by representing the two as concepts that are quite similar in some dimensions, however, unconnected in others. When, e.g., looking at multibox embeddings, it would be possible that two boxes are in some dimensions similar or even overlapping, in others, however, clearly disjoint.

4. MultiBoxes as a Basis for a Trustworthy Embedding Approach

Our aim is now to combine the two needs, discussed in Section 3.1 and Section 3.2 to gain an embedding approach that is trustworthy in the sense that it (1) allows for a computationally efficient embedding technique able to represent the behavior of classical interpretations and (2) is able to model the semantics of the concepts such that both the similarity hypothesis for concepts and individuals can be easier reached and the shape of the concepts have a clearer semantic justification. This is done in the following with the help of multiboxes. The basic idea of multiboxes is to combine the computational properties of boxes (thus particularly the small number of variables for training and the convexity for inference) with the flexibility of unconnected representations. Therefore, one concept is not represented by one box as for box interpretations but by a union of a finite family of boxes.

Definition 7. A multibox in some $\mathbb{R}^n$ is the union of a finite family of boxes. We denote by $\mathcal{MB}^n$ the sets of all multiboxes contained in $\mathbb{R}^n$.

The definition of multiboxes allows for a flexible representation of concepts due to two facts: (1) it is possible to represent unconnected concepts as disjoint boxes. This particularly allows for modeling multibox interpretations that behave locally like box interpretations, however, allow for representing one concept in different domains, e.g., based on the habitat example one domain regarding the American world, the other domain the Asian world. This viewpoint particularly enables to model concepts in a controlled way based on a restricted amount of disjoint boxes with a particular focus on concepts occurring in different/disconnected domains.

However, multiboxes allow also for another viewpoint (2), as the boxes a multibox consists of do not need to be disjoint. Therefore, it is possible to use a multibox representation as an approximation of an arbitrary geometric shape. Here, the special advantage is that the size resp. the number of boxes allowed as basis for a multibox determine the granularity of the approximation. A large number of boxes as basis leads to a good approximation, however, increases computational need and reduces learnability and inference power.

An important feature of boxes is their closure under intersection. This property, that is essential for representing conjunction, is satisfied also by multiboxes.

Proposition 1. *If M_1, M_2 are elements of $\mathcal{MB}^n$ then $M_1 \cap M_2$ is an element of $\mathcal{MB}^n$.*

Proof. Let $M_1 = \bigcup_{i=1}^k B_1^i$ and $M_2 = \bigcup_{j=1}^l B_2^j$ for $B_{1,2}^i \in \mathcal{B}^n$. The intersection of two boxes results always in a box, hence for all $i \in k, j \in l$ we have that $B_1^i \cap B_2^j$ is a box. We have that $M_1 \cap M_2 = \bigcup_{i \leq k, j \leq l} B_1^i \cap B_2^j$ and so it is a finite union of boxes. $\square$

Notice that this closure property requires that we consider multiboxes composed by an arbitrarily large number of boxes. For example, the intersection of two multiboxes made up of two boxes can be composed of four different boxes, as shown in Figure 4 (a). In this case, the computational advantages of boxes, namely the small number of defining variables, would be lost. However, even a multibox consisting of only a few boxes allows for modeling different contexts and also allows for circumventing problems with the Helly Property (as will be discussed in Example 9). Additionally, a multibox consisting of a large number of boxes tends to overfit and thus is not useful in practice. Therefore, it seems reasonable to restrict the concept representations to multiboxes consisting of a sufficiently small number of boxes. Results on the connection between representative abilities and the number of boxes per multibox are subject to future work.

In the following definition, we consider the most general case, allowing concepts to be represented as multiboxes that are unions of an arbitrary family of boxes, thus without any restriction on connectedness or the number of boxes. This can be used as starting point, however, the discussion above shows that different restrictions of this definition are of interest, particularly also regarding the number of boxes as basis for the multibox allowed.

Definition 8. *A classical interpretation $\mathcal{I} = (\Delta, \cdot^{\mathcal{I}})$ is a multibox interpretation if*

- $\Delta \in \mathcal{B}^n$ for some $n \in \mathbb{N}$,
- $\cdot^{\mathcal{I}}$ maps
 - each concept name $A \in \mathbf{C}$ to some multibox in $\mathcal{MB}^n$
 - each role $R \in \mathbf{R}$ to a subset $R^{\mathcal{I}} \subseteq \Delta \times \Delta$ such that for every $B \in \mathcal{B}^n$: $(R^{\mathcal{I}})^{-1}(B) \in \mathcal{MB}^n$.

To easily distinguish classical, box and multibox interpretations, a multibox interpretation is denoted as $\mathcal{M} = (M, \cdot^{\mathcal{M}})$ in the following. We stress that, since it is a classical interpretation, its semantics on general $\mathcal{ELHO}(\circ)^+$ concepts is the one given in Table 1. By looking at Definition 2 it is clear that any box interpretation is also a multibox interpretation, in this sense multibox interpretations are an extension of box interpretations. To exemplify the ability of the multiboxes to define a multibox interpretation for each satisfiable ontology, in the following an example is given how an ontology not in line with the Helly Property, thus not representable with a box interpretation can be modeled with the help of multiboxes.

Example 9. Let $\mathcal{O}$ be the ontology of example 4, namely $\mathcal{O} = (\mathcal{T}, \mathcal{R}, \mathcal{A})$ with $\mathcal{T} = \{A \sqcap B \sqcap C \sqsubseteq \perp\}$, $\mathcal{R} = \{\}$ and $\mathcal{A} = \{a_1 : A \sqcap B, a_2 : B \sqcap C, a_3 : A \sqcap C\}$. As we have already discussed, this satisfiable $\mathcal{ELHO}(\circ)^\perp$ -ontology does not admit a box interpretation. However, we can write a multibox interpretation $\mathcal{M} = (\mathbb{R}^n, \cdot^\mathcal{M})$ such that $\mathcal{M} \models \mathcal{O}$. In Figure 4 (b), an example for such a multibox interpretation $\mathcal{M}$ can be seen, exemplifying the ability of multibox interpretations to circumvent the representational restrictions of box embeddings.

Next, we prove the closure result of multiboxes with respect to $\mathcal{ELHO}(\circ)^\perp$, namely we show that the interpretation of a general $\mathcal{ELHO}(\circ)^\perp$ concept is a multibox.

Proposition 2. Let $(M, \cdot^\mathcal{M})$ be a multibox interpretation with $M \in \mathcal{B}^n$. The interpretation of each $\mathcal{ELHO}(\circ)^\perp$ concept is an element of $\mathcal{MB}^n$.

Proof. We proceed by induction on concepts. By definition for all concept names or nominal concepts $A \in \mathbf{C}$ we have that $A^\mathcal{M}$ is a multibox and if $C \in \{\perp, \top\}$ then $C^\mathcal{M}$ is a multibox.

If $C := D \sqcap E$ for some concepts D and E and $D^\mathcal{M}, E^\mathcal{M}$ are elements of $\mathcal{MB}^n$ then by definition $C^\mathcal{M} = D^\mathcal{M} \cap E^\mathcal{M}$. Thus, Proposition 1 implies that $C^\mathcal{M} \in \mathcal{MB}^n$.

If $C := \exists R.D$ for a relation $R \in \mathbf{R}$ and a concept D with $D^\mathcal{M} \in \mathcal{MB}$ we have that $D^\mathcal{M} = \bigcup_{i=1}^k B_i$ for some $B_i \in \mathcal{B}^n$. Moreover

$$(\exists R.D)^\mathcal{M} = (R^\mathcal{M})^{-1}(D^\mathcal{M}) = \bigcup_{i=1}^k (R^\mathcal{M})^{-1}(B_i);$$

since we have $(R^\mathcal{M})^{-1}(B) \in \mathcal{MB}^n$ for every $B \in \mathcal{B}^n$ we conclude that $(\exists R.D)^\mathcal{M} \in \mathcal{MB}^n$. $\square$

We proved that multibox interpretations are closed with respect to $\mathcal{ELHO}(\circ)^\perp$ concepts. It is natural to ask whether this closure property extends to more expressive description logics. In particular, how can multibox interpretations be defined for concepts involving additional operations such as negation or disjunction? For disjunction, the situation is straightforward: since the union of multiboxes is again a multibox, the standard semantic clause can be retained:

$$(A \sqcup B)^\mathcal{M} = A^\mathcal{M} \cup B^\mathcal{M},$$

for any concepts A and B , and any interpretation $\mathcal{M}$. However, when considering negation a problem appears. Under our current definition, a multibox is a closed subset of a Euclidean space, whereas its complement is generally open and therefore not a multibox. Consequently, the standard semantic clause

$$(\neg A)^\mathcal{M} = M \setminus A^\mathcal{M}$$

does not preserve the class of multiboxes. A possible solution involves the use of open boxes instead of closed boxes or consider a special definition of the subconcept relation. Exploring such extensions is left for future work. For the purposes of this paper, it is sufficient to restrict attention to the description logic $\mathcal{ELHO}(\circ)^\perp$.

We are now ready to prove a completeness theorem for these representations, namely, we will show that an $\mathcal{ELHO}(\circ)^\perp$ ontology is satisfiable if and only if it has a model that is a multibox interpretation.

Theorem 3. For all $n \in \mathbb{N}$ and each $\mathcal{ELHO}(\circ)^\perp$ ontology $\mathcal{O}$ the following is true:

$$\mathcal{O} \text{ is satisfiable} \iff \mathcal{M} \models \mathcal{O} \text{ for some multibox interpretation } \mathcal{M} \text{ with domain } \mathbb{R}^n.$$

Proof. Let $\mathcal{O}$ be an $\mathcal{ELHO}(\circ)^\perp$ ontology and $\mathcal{I} = (\{d_i\}_{i \in J}, \mathcal{I})$ be a finite model of $\mathcal{O}$, namely J is a finite set of indices and $\mathcal{I} \models \mathcal{O}$. Fix $n \in \mathbb{N}$. It is straightforward to define a multibox interpretation $\mathcal{M}$ with $\mathbb{R}^n$ as its domain such that $\mathcal{M} \models \mathcal{O}$. Notice that for each $i \in J$ select a point $e_i \in \mathbb{R}^n$ in way that $d_i = d_j$ iff $e_i = e_j$ for each $i, j \in J$. Then we impose that $a^I = \{d_i\}$ iff $a^\mathcal{M} = \{e_i\}$ for each $a \in \mathbf{I}$. Each subset of $\{e_i\}_{i \in J}$ is in $\mathcal{MB}^n$. We let

$$\{a\}^\mathcal{M} = \{a^\mathcal{M}\} \text{ for each nominal concept } \{a\}, \quad (1)$$

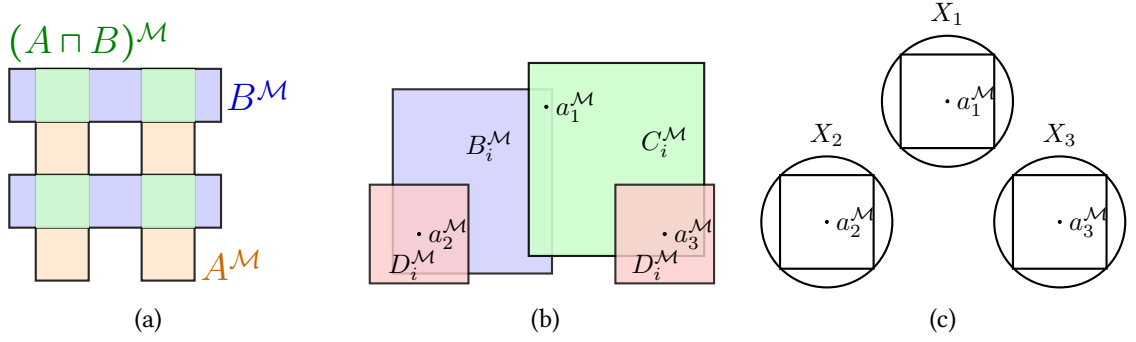


Figure 4: (a) the intersection of two multiboxes made up by two boxes that results in a multibox made up by four boxes; (b) a multibox model of the ontology in Example 4; (c) a sketch of the construction presented in Example 10.

$$C^{\mathcal{I}} = \{d_i\}_{i \in J} \iff C^{\mathcal{M}} = \{e_i\}_{i \in J} \text{ for each } C \in \mathbf{C}, \quad (2)$$

$$(e_i, e_j) \in R^{\mathcal{M}} \iff (d_i, d_j) \in R^{\mathcal{I}} \text{ for each } R \in \mathbf{R}. \quad (3)$$

By induction on concepts, it is immediate to establish that second condition above holds also for general concepts C . Therefore, it is immediate to verify that if φ is an individual assertion, a general axiom inclusion, or a role hierarchy statement expressible in $\mathcal{ELHO}(\circ)^\perp$ the following holds

$$\mathcal{I} \models \varphi \iff \mathcal{M} \models \varphi. \quad (4)$$

□

The previous proof gives us a simple method to obtain from any interpretation an $\mathcal{ELHO}(\circ)^\perp$ equivalent multibox interpretation. However, from a practical point of view, the multibox interpretation that comes out from that construction is poor of geometrical meaning. Each concept contains just a finite number of points, this means that each concept has zero volume (it is a set of measure zero with the Lebesgue measure). In the following example, we describe how to modify it in order to have a model with a more interesting geometry.

Example 10. Assume that for an $\mathcal{ELHO}(\circ)^\perp$ ontology $\mathcal{O}$ in some $\mathbb{R}^n$ a multibox interpretation based on the construction principle of Theorem 3 is constructed. There the volume of the interpretation of each concept is zero. Since $\mathbb{R}^n$ is an Hausdorff space we can always find balls X_i with $e_i \in X_i$ (e_i as defined in the proof of Theorem 3, such that $X_i \cap X_j = \emptyset$ whenever $i \neq j$). Then for each of these balls X_i we take a box X'_i inscribed in it. Finally we use these to define a new interpretation $\mathcal{M}'$. We let $a_i^{\mathcal{M}'} = a_i^{\mathcal{M}}$ for each i . Then for all $C \in \mathbf{C}$, we let

$$C^{\mathcal{M}} = \{e_i : i \in I\} \iff C^{\mathcal{M}'} = \bigcup_{i \in I} X'_i.$$

Also relations need to be considered, however define their interpretations is not immediate and it goes beyond the purposes of this example.

A sketch of this construction is given in Figure 4(c). The construction just described, together with the proof of Theorem 3, could be a first step toward the building of an initialization of a multibox embedding starting from a standard model. Another possible step involves reducing the number of unconnected components in the interpretation by joining some boxes if possible. In Figure 4(b) we see a multibox model of $\mathcal{O}$ in which only one concept has a non connected interpretation. Note here that this task is normally solved by the embedding framework based on loss functions. These considerations are therefore a starting point either for designing loss functions or for initializing the model. Constructing a model based on those construction principles is normally not advisable, as it is missing the geometric regularities.

Theorem 3 states that it is actually possible to find for each satisfiable ontology a multibox interpretation. This is a vital property, as it allows for (at least theoretically) using this approach for every

$\mathcal{ELHO}(\circ)^\perp$ ontology in a practical setting without needing to be concerned about potential inconsistent outcomes. One additional advantage of such multibox models (that is also a property of the box models mentioned in Definition 2) is that they allow for a faithful representation in the sense that not only the ontology but also the consequences of the ontology are modeled in the geometric interpretation. This seems to be a trivial property, as it is trivially fulfilled by classical interpretations. It is, however, not the case as pointed out in [12]. Due to the completeness proved in Theorem 3, we are even able to state a stronger statement: If one general GCI, role inclusion axiom or individual assertion is a consequence of each multibox interpretation of an ontology $\mathcal{O}$, then it is actually a consequence of the ontology. This particularly states some sort of bias-freeness of the multibox approach: There aren't any GCIs, role inclusion axioms or individual assertions following solely due to some construction principles out of the multibox interpretation. This sort of faithfulness of the multibox family is the matter of next proposition.

Proposition 4. (*semantic faithfulness of multiboxes*) *Let $\mathcal{O}$ be an $\mathcal{ELHO}(\circ)^\perp$ ontology $\mathcal{O}$ and φ be a general GCI, role inclusion axiom or individual assertion. We have that*

$$\mathcal{O} \models \varphi \iff \text{for each multibox interpretation } \mathcal{M}, \mathcal{M} \models \mathcal{O} \text{ implies } \mathcal{M} \models \varphi$$

Proof. ($\Rightarrow$) Let $\mathcal{M}$ be a multibox interpretation of $\mathcal{O}$. By definition, $\mathcal{M} \models \mathcal{O}$. By Theorem 3, this always exists. Since $\mathcal{M}$ is also a standard interpretation, from $\mathcal{M} \models \mathcal{O}$ and $\mathcal{O} \models \varphi$ we get that $\mathcal{M} \models \varphi$.

($\Leftarrow$) Assume that $\mathcal{O} \not\models \varphi$, it has to be shown that then $\mathcal{M} \not\models \varphi$ for all multibox interpretations $\mathcal{M}$. Let $\mathcal{I}$ be a classical interpretation such that $\mathcal{I} \models \mathcal{O}$ and $\mathcal{I} \not\models \varphi$. This exists, as otherwise $\mathcal{O} \models \varphi$ would be the case. It follows from the proof of Theorem 3, in particular from statement (4), that we can pick a multibox model $\mathcal{M}$ such that $\mathcal{M} \models \mathcal{O}$ and $\mathcal{M} \not\models \varphi$. $\square$

5. Conclusions

This paper has introduced multiboxes as a promising framework to increase the trustworthiness of embedding approaches along two axes: on the one hand, it is possible to increase the representation strength of the framework, improving the robustness and performance aspects of trustworthiness. On the other hand, the alignment with the semantics of concepts is sharpened, thus improving the trustworthiness regarding reliability and plausibility of the inferred results.

The work reported in this paper with the introduction of the multibox framework acts only as a starting point and needs to be extended in several important directions.

First and foremost, it is necessary to extend this view from a purely theoretical investigation to an experimental one, namely to check whether multibox interpretations (1) are actually increasing the abilities to represent concepts in a semantically meaningful way; (2) have suitable properties regarding learnability, specifically, to establish that the interpretation that exists due to Theorem 3 can actually be found; (3) support that such learnable multibox representations actually increase inferential quality.

Beside these practical and experimental issues, it is also relevant to consider modelling scenarios and the different use cases of families of disjoint boxes and families of joint boxes (or mixtures of both). Particularly relevant is to examine in detail whether it is possible to give upper bounds both for the dimensionality needed (for a particular ontology embedding) or for the number of boxes making up a multibox concept representation. For instance, it would be expected that a sufficiently low number of boxes per multibox is of particular importance for learning approaches in order to avoid overfitting and to allow for low enough computational costs.

In conclusion, with the introduction of the multibox approach, we have provided a representation framework for ontology embeddings that provides some key advantages over other approaches with direct impact on the trustworthiness debate, namely being able to be both 'more precise' (representation strength) and 'more adequate' (fit to intended semantics). We therefore hope that future work will establish the multibox framework as a valuable choice for trustworthy ontology embeddings.

Acknowledgments

The research of Giuseppe Federico was funded in part by Sony CSL/Sony AI. The research of Mena Leemhuis was funded in part by the Austrian Science Fund (FWF) 10.55776/COE12. Oliver Kutz acknowledges the financial support through the ‘Abstractron’ project funded by the Autonome Provinz Bozen - Südtirol (Autonomous Province of Bolzano/Bozen) through the Research Südtirol/Alto Adige 2022 Call.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

- [1] E. Rosch, Cognitive representations of semantic categories., *Journal of Experimental Psychology: General* 104 (1975) 192–233. doi:10.1037/0096-3445.104.3.192.
- [2] P. Gärdenfors, *Conceptual Spaces: The Geometry of Thought*, The MIT Press, Cambridge, Massachusetts, 2000. doi:10.7551/mitpress/2076.001.0001.
- [3] J. Chen, O. Mashkova, F. Zhapa-Camacho, R. Hoehndorf, Y. He, I. Horrocks, *Ontology Embedding: A Survey of Methods, Applications and Resources*, *IEEE Transactions on Knowledge and Data Engineering* (2025) 1–20. doi:10.1109/tkde.2025.3559023.
- [4] M. Kulmanov, W. Liu-Wei, Y. Yan, R. Hoehndorf, *EL Embeddings: Geometric Construction of Models for the Description Logic EL++*, in: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19, International Joint Conferences on Artificial Intelligence Organization*, 2019, pp. 6103–6109. doi:10.24963/ijcai.2019/845.
- [5] B. Xiong, N. Potyka, T.-K. Tran, M. Nayyeri, S. Staab, *Faithful Embeddings for $\mathcal{EL}^{++}$ Knowledge Bases*, in: *The Semantic Web – ISWC 2022*, Springer International Publishing, 2022, pp. 22–38. doi:10.1007/978-3-031-19433-7_2.
- [6] H. Paulheim, N. Hubert, J. Portisch, *What do knowledge graph embeddings learn to represent?*, in: *Handbook on Neurosymbolic AI and Knowledge Graphs*, IOS Press, 2025. doi:10.3233/faia250206.
- [7] K. Donnelly, *SNOMED-CT: The Advanced Terminology and Coding System for EHealth*, *Studies in Health Technology and Informatics* 121 (2006) 279–290.
- [8] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock, *Gene Ontology: tool for the unification of biology*, *Nature Genetics* 25 (2000) 25–29. doi:10.1038/75556.
- [9] F. Baader, S. Brandt, C. Lutz, *Pushing the el envelope*, in: *IJCAI’05: Proceedings of the 19th International Joint Conference on Artificial Intelligence*, 2005, pp. 364–369. doi:10.25368/2022.144.
- [10] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, P. F. Patel-Schneider (Eds.), *The Description Logic Handbook: Theory, Implementation and Applications*, second ed., Cambridge University Press, 2007. doi:10.1017/CBO9780511711787.
- [11] M. Leemhuis, O. Kutz, *Understanding the expressive capabilities of knowledge base embeddings under box semantics*, in: L. H. Gilpin, E. Giunchiglia, P. Hitzler, E. van Krieken (Eds.), *Proceedings of The 19th International Conference on Neurosymbolic Learning and Reasoning*, volume 284 of *Proceedings of Machine Learning Research*, PMLR, 2025, pp. 303–321.
- [12] C. Bourgaux, R. Guimarães, R. Koudijs, V. Lacerda, A. Ozaki, *Knowledge Base Embeddings: Semantics and Theoretical Properties*, in: *Proceedings of the TwentyFirst International Conference on Principles of Knowledge Representation and Reasoning, KR-2024, International Joint Conferences on Artificial Intelligence Organization*, 2024, pp. 823–833. doi:10.24963/kr.2024/77.

- [13] B. Li, P. Qi, B. Liu, S. Di, J. Liu, J. Pei, J. Yi, B. Zhou, Trustworthy AI: from principles to practices, *ACM Comput. Surv.* 55 (2023) 177:1–177:46. doi:10.1145/3555803.
- [14] J. Eckhoff, Intersection properties of boxes. Part I: An upper-bound theorem, *Israel Journal of Mathematics* 62 (1988) 283–301. doi:10.1007/bf02783298.
- [15] Ö. Özçep, M. Leemhuis, D. Wolter, Cone Semantics for Logics with Negation, in: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, International Joint Conferences on Artificial Intelligence Organization, 2020, pp. 1820–1826. doi:10.24963/ijcai.2020/252.
- [16] M. Leemhuis, Ö. Özçep, Conceptual orthospaces—convexity meets negation, *International Journal of Approximate Reasoning* 162 (2023) 109013. doi:10.1016/j.ijar.2023.109013.
- [17] V. Lacerda, A. Ozaki, R. Guimarães, Strong Faithfulness for ELH Ontology Embeddings, *Transactions on Graph Data and Knowledge* 2 (2024) 2:1–2:29. doi:10.4230/TGDK.2.3.2.
- [18] B. F. Lourenço, H. Morgan, A. Ozaki, A. Pavlović, E. Sallinger, BoxLitE: A Faithful Knowledge Base Embedding Based on Convex Optimization, in: *Proceedings of the 23rd International Conference on Principles of Knowledge Representation and Reasoning*, 2026, pp. 1006–1017. doi:10.24963/kr.2026/94.
- [19] H. Yang, J. Chen, U. Sattler, TransBox: $\mathcal{EL}^{++}$ -closed Ontology Embedding, in: *Proceedings of the ACM on Web Conference 2025, WWW '25*, ACM, 2025, p. 22–34. doi:10.1145/3696410.3714672.
- [20] R. N. Shepard, Toward a universal law of generalization for psychological science, *Science* 237 (1987) 1317–1323. doi:10.1126/science.3629243.
- [21] J. V. Hernández-Conde, A case against convexity in conceptual spaces, *Synthese* 194 (2016) 4011 – 4037.
- [22] N. Goodman, *The Structure of Appearance*, volume 107 of *Boston Studies in the Philosophy of Science*, 3rd ed., Springer Netherlands, 1977. doi:10.1007/978-94-010-1184-6.
- [23] P. L. Buttigieg, N. Morrison, B. Smith, C. J. Mungall, S. E. Lewis, The environment ontology: Contextualising biological and biomedical entities, *Journal of Biomedical Semantics* 4 (2013) 43. doi:10.1186/2041-1480-4-43.
- [24] A. C. Albuquerque, J. L. Campos dos Santos, A. N. d. Castro, Ontobio: A biodiversity domain ontology for amazonian biological collected objects, in: *2015 48th Hawaii International Conference on System Sciences*, 2015, pp. 3770–3779. doi:10.1109/HICSS.2015.453.
- [25] C. T. Stayton, The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence, *Evolution* 69 (2015) 2140–2153. doi:10.1111/evo.12729.
- [26] J. B. Reece, L. A. Urry, M. L. Cain, S. A. Wasserman, P. V. Minorsky, R. B. Jackson, et al., *Campbell biology*, volume 9, Pearson Boston, 2014.