Journal of Artificial Intelligence Research 78 (2023) 287-337 Submitted 05/2022; published 10/2023

Embedding Ontologies in the Description Logic $\mathcal{ALC}$ by Axis-Aligned Cones

Özgür Lütfü Özçep OEZCEP@IFIS.UNI-LUEBECK.DE
Mena Leemhuis LEEMHUIS@IFIS.UNI-LUEBECK.DE
University of Lübeck, Germany

Diedrich Wolter DIEDRICH.WOLTER@UNI-BAMBERG.DE
University of Bamberg, Germany

Abstract

This paper is concerned with knowledge graph embedding with background knowledge, taking the formal perspective of logics. In knowledge graph embedding, knowledge—expressed as a set of triples of the form (a R b) ("a is R-related to b")—is embedded into a real-valued vector space. The embedding helps exploiting geometrical regularities of the space in order to tackle typical inductive tasks of machine learning such as link prediction. Recent embedding approaches also consider incorporating background knowledge, in which the intended meanings of the symbols a, R, b are further constrained via axioms of a theory. Of particular interest are theories expressed in a formal language with a neat semantics and a good balance between expressivity and feasibility. In that case, the knowledge graph together with the background can be considered to be an ontology. This paper develops a cone-based theory for embedding in order to advance the expressivity of the ontology: it works (at least) with ontologies expressed in the description logic $\mathcal{ALC}$, which comprises restricted existential and universal quantifiers, as well as concept negation and concept disjunction. In order to align the classical Tarskian Style semantics for $\mathcal{ALC}$ with the sub-symbolic representation of triples, we use the notion of a geometric model of an $\mathcal{ALC}$ ontology and show, as one of our main results, that an $\mathcal{ALC}$ ontology is satisfiable in the classical sense iff it is satisfiable by a geometric model based on cones. The geometric model, if treated as a partial model, can even be chosen to be faithful, i.e., to reflect all and only the knowledge captured by the ontology. We introduce the class of axis-aligned cones and show that modulo simple geometric operations any distributive logic (such as $\mathcal{ALC}$) interpreted over cones employs this class of cones. Cones are also attractive from a machine learning perspective on knowledge graph embeddings since they give rise to applying conic optimization techniques.

1. Introduction

The idea of embedding words or other forms of data into low-dimensional continuous vector spaces has gained much attention as it provides means to connect reasoning with machine learning. Approaches that map individual instances such as words to vector spaces have already been proven useful in various tasks (Goldberg & Levy, 2014; Pennington, Socher, & Manning, 2014; Levy & Goldberg, 2014), yet these approaches cannot grasp the relational—or more generally: the predicate-logical—structure underlying the data. Consequently, the embedding idea was pushed further (see, e.g., Nickel, Tresp, and Kriegel (2011), Bordes, Usunier, García-Durán, Weston, and Yakhnenko (2013) and, for an overview, Wang, Mao, Wang, and Guo (2017)) in order to design embeddings of *knowledge graphs*. As usual, a knowledge graph is defined to be a set of triples of the form $(a\ R\ b)$ for objects a and b and relations R. The idea of embedding knowledge graphs was even pushed further to account also for background knowledge, say an ontology consisting of axioms in some (expressive) logic (Mehran Kazemi & Poole, 2018; Kulmanov, Liu-Wei, Yan, & Hoehndorf, 2019). Embeddings support several reasoning operations and have quickly gained massive attention. Moreover, embeddings may also be regarded as a cognitively justified structure for representing concepts (Gärdenfors, 2000) and reasoning with them.

Many classical approaches of knowledge graph embeddings such as TransE (Bordes et al., 2013) suffer from a lack of full expressivity in the sense laid down by Mehran Kazemi and Poole (2018): given a knowledge graph and a set of triples known to be true (positive set) as well as triples that are known to be false (negative set), the embedding is said to be *fully expressive* if it maps the relations and constants into a space such that $(a\ R\ b)$ holds in the embedding iff it is in the positive set and $(a\ R\ b)$ does not hold iff it is in the negative set. The lack of (full) expressivity of classical approaches to knowledge graph embeddings is due to a decision on how to model relations R: rather than following the classical logic approach, relations are modelled with computationally feasible data structures (e.g., in case of TransE by vector translations) that fit into the general mathematical framework of continuous embeddings. In fact, following the translation approach of TransE one can see that the induced logic is not that of arbitrary relations but that of functional relations only since entities that are said to be related must be located relative to one another with a relation-specific but fixed translation.

These observations motivate research on embeddings that shift the compromise between geometrical models constructible by means of learning and capabilities to capture underlying domain structure towards richer logical structures. This paper proposes a new embedding of ontologies expressed in a concept-centred logic (such as a description logic) into a real-valued vector space. The embeddings works for various such logics, but as a case in example we consider the description logic $\mathcal{ALC}$ (Schmidt-Schauß & Smolka, 1991) which provides restricted forms of existential and universal quantifiers as well as concept negation and concept disjunction and is a common semi-expressive description logic that allows us to go beyond what can be captured with geometric models considered for knowledge graph embeddings so far. The popularity of $\mathcal{ALC}$ makes it worthwhile to study new approaches related to it. To the best of our knowledge, this is the first approach capable of concept negation and disjunction as it occurs naturally in many ontologies. Our main result states that an $\mathcal{ALC}$ ontology is satisfiable in a classical sense iff it is satisfiable by a geometric model that interprets all concept descriptions as cones (concretely: axis-aligned cones). We derive this result by first considering the Boolean part of $\mathcal{ALC}$ ontologies and then generalizing it to full $\mathcal{ALC}$. The geometric models we use are partial and thus allow some uncertainty to be retained, i.e., if x is only known to be a member of the union of two atomic concepts, then our partial model will not commit to saying to which atomic concept x belongs. Put differently, we need partial models in order to represent the knowledge contained in the ontology *faithfully*: exactly those axioms derivable from the ontology should be represented in a partial model. Partial models introduce interesting possibilities into the realm of knowledge graph embeddings when learning from incomplete data. A learner constructing an embedding from data is no longer forced to commit to all feature values of an instance,

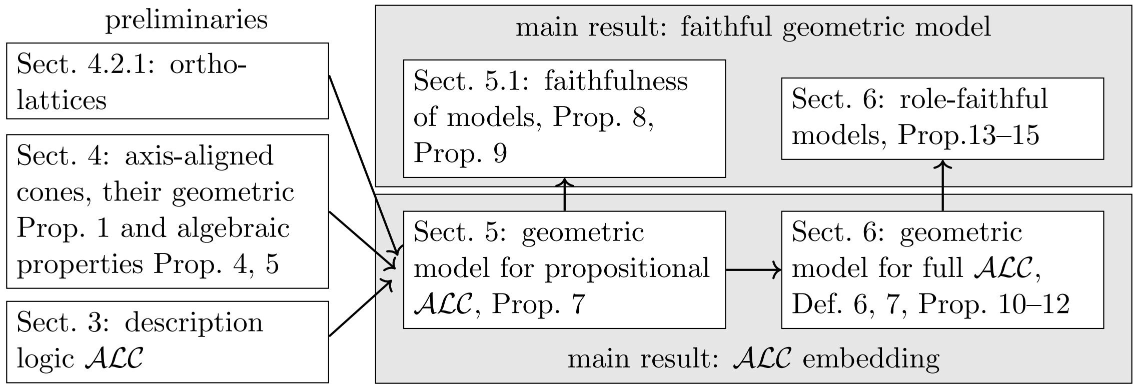


Figure 1: Overview of steps in achieving a geometric embedding for description logic $\mathcal{ALC}$
.

but only to those backed up by the data available and thus is able to incorporate incomplete information in the learning process without being forced to make assumptions about the missing features.

In summary, this paper makes two contributions: First, by detailing an embedding of $\mathcal{ALC}$ we show a way of how logic-level expressivity of knowledge graph embeddings can be advanced to semi-expressive concept languages involving negation. We achieve this result by exploiting the geometric structure of cones and the algebraic-logic structure induced by the operations of classical set intersection and polarity from linear algebra. Second, we present a so-called *geometric model* that is faithful and allows uncertainty in data to be retained in a knowledge graph embedding. Faithful models can be obtained by interpreting the algebraic-logic structure of a specific class of cones.

This paper continues by reviewing related approaches. Section 3 summarizes $\mathcal{ALC}$, Section 4 introduces al-cones and discusses their motivation. In Section 5 we present an embedding for propositional $\mathcal{ALC}$ and show its completeness. Section 6 then considers full $\mathcal{ALC}$. For orientation, the steps towards achieving the geometric model for full $\mathcal{ALC}$ are also summarized in Figure 1. Section 7 sketches how to use our embedding approach for learning. The paper concludes with a brief discussion of the results.

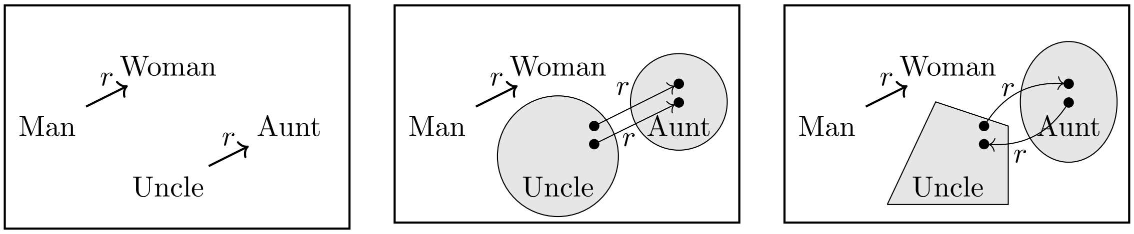
2. Related Work

A class of contemporary research areas in AI is involved with the integration of two or more distinct subfields of AI, sometimes referred to as *hybrid AI*. One of these areas is involved with the integration of knowledge representation (KR) and machine learning (ML) techniques, aiming at various goals such as empowering an agent to reason about knowledge it has learned or to advance machine learning performance by exploiting background knowledge. These questions are particularly relevant to so-called *Knowledge Graph Embeddings* (briefly referred to as embeddings), which connect learning in terms of computing an embedding of concepts in a real-valued vector space to some form of reasoning carried out by geometric operations (see Ji, Pan, Cambria, Marttinen, and Yu (2021), Wang et al. (2017) for an overview). Put differently, the aim of embeddings from a ML perspective is to advance from learning single concepts to learning ontologies, and to exploit ontologies as constraint specification for deriving statistical models from data.

An example for the latter approach in a typical ML task is laid down by Deng, Ding, Jia, Frome, Murphy, Bengio, Li, Neven, and Adam (2014). The main application considered there is object classification on pictures w.r.t. some set of labels (puppy, dog, cat, etc.) The authors consider ontologies of a very simple type called HEX-graph. The HEX-graph consists of nodes for Boolean labels (corresponding to atomic concepts) and edges that stand either for subsumption (leading to hierarchies of concepts, the "h" in "hex", example: puppies are dogs) or for disjointness of concepts (called exclusion, the "ex" in "hex"; example: cats are disjoint from dogs). In Description Logic (DL) terminology as used in this paper to characterize ontologies, the expressivity of HEX-graphs is a fragment of propositional $\mathcal{ALC}$ where only atomic negation is allowed and where conjunction or disjunction are not supported. Deng et al. (2014) then describe a standard model used in the given ML task (concretely: pairwise conditional random fields) and show how to encode the constraints of the hex-graph as additional factors in that model. By presenting a more expressive embedding based on the logic $\mathcal{ALC}$, this paper aims to advance the foundations for integrating machine learning and reasoning.

From a KR perspective, embeddings provide a form of model in the sense of logic, also called *geometrical model* due to the geometric nature of embeddings. Reasoning is then performed by operations on the geometrical model. A particular challenge in grounding logic reasoning in machine learning models is to bridge between semantics underlying learning and reasoning. One example to such neural-symbolic reasoning are Logic Tensor Networks (LTN) (Serafini & d'Avila Garcez, 2016; Badreddine, d'Avila Garcez, Serafini, & Spranger, 2021) that ground a many-valued first-order logic in real values of a deep learning architecture of tensor networks (Socher, Chen, Manning, & Ng, 2013) (thus the name Logic Tensor Networks). The logic of LTN is a Fuzzy Logic based on degrees of satisfiability. Atoms are grounded with degrees of probabilities and predicates as functions over these probabilities. In contrast to the probability-based approach of LTNs, the approach presented in this paper allows concepts to be modeled as geometric objects and predicates as geometric operations. Moreover, though our approach does not rely on fuzzy reasoning, it allows uncertainty to be captured in *partial models*—in particular in so-called *faithful models*. As a last contrasting point to LTNs, let us mention that our contribution to embeddings is to ground the semi-expressive logic $\mathcal{ALC}$ exactly in a real-valued geometry without the need of hyper-parameters that need to be tuned for aligning sub-symbolic and symbolic semantics.

Several approaches to knowledge graph embeddings are derived from the basic idea of the seminal TransE approach (Bordes et al., 2013) that is illustrated in Figure 2 (a). TransE interprets objects, respectively concepts, as vectors in a real-valued vector space. Relations are interpreted as translations of these vectors. For example, Figure 2 (a) shows the words 'man' and 'woman' related by a translation r that could mean 'female form of'. This type of embedding allows analogical reasoning of the following kind: based on the relation between man and woman what would be a possible counterpart X of uncle? In order to answer the question, the translation r that connects 'man' to 'woman' can be applied to 'uncle', obtaining 'aunt'. Though TransE was enormously successful it is also limited with respect to the kinds of background knowledge it can account for. For example, concepts usually have an extension, i.e., have an associated set of objects falling under



(a) (b) (c)

Figure 2: (a) Illustration of TransE; (b) Approach using spheres for $\mathcal{EL}^{++}$ concepts; (c) Quasi-chained Datalog with convex sets

them. Hence, they cannot be represented as single vectors. This problem has motivated the consideration of geometric models (Gutiérrez-Basulto & Schockaert, 2018), i.e. embeddings that are motivated by classical Tarskian style semantics.

One recent approach along this line that is based on the geometric interpretation of knowledge graph embeddings is presented by Kulmanov et al. (2019). Here, the lightweight description logic $\mathcal{EL}^{++}$ (Baader, Brandt, & Lutz, 2005) is considered as a means to express background knowledge. The approach is still committed to the idea of TransE of interpreting relations as translations but invokes the idea of geometric models to interpret concepts as (geometrically shaped) sets of vectors. In the case of the approach of Kulmanov et al. (2019) these geometrically shaped sets for embedding concepts are open n-balls with a fixed radius. Figure 2 (b) illustrates this with the concept 'uncle' and 'aunt' represented as spheres of objects, but still with relations being vector translations (of instances of the concepts). The method of Kulmanov et al. (2019) is limited due to the restricted constructions available in the lightweight description logic $\mathcal{EL}^{++}$, which in particular has no full concept negation—in contrast to the description logic $\mathcal{ALC}$, which is in the focus of this paper.

Gutiérrez-Basulto and Schockaert (2018) define geometric interpretations as classical interpretations of an ontology with two specific constraints: the domain is a Euclidean space of some dimension m and the interpretations of all (n-ary) relations are constrained to convex sets over n-wise Cartesian products of $\mathbb{R}^m$, i.e., $\mathbb{R}^{nm}$. So in contrast to classical knowledge graph embeddings and also to Kulmanov et al. (2019), relations (of arbitrary arity n) are interpreted (classically) by n-wise Cartesian products and (non-classically) requiring them to be convex sets. Figure 2 (c) illustrates the approach: Concepts, i.e., unary relations, such as 'uncle' or 'aunt' are interpreted by arbitrary convex sets (depicted as oval and polygon in the figure). Also relations of arity $n > 1$ are interpreted by convex sets. In our illustration the convex sets for binary relation r would be situated in $\mathbb{R}^4$ since the concepts are embedded in $\mathbb{R}^2$. Pairs contained in relation r are as indicated by the bowed edges not further constrained and not directed in a fixed way. The approach of (Gutiérrez-Basulto & Schockaert, 2018) achieves ontologies expressed as rules in datalog$^{\pm}$ (Calì, Gottlob, & Lukasiewicz, 2009), which admits rules with existentials in the head of the rule (alias tuple-generating dependencies) and integrity constraints, i.e., rules of the form $\forall \vec{x} \psi(\vec{x}) \rightarrow \bot$ with a conjunction of atoms $\psi(\vec{x})$. Their main result is that an ontology consisting of quasi-chained rules is satisfiable classically iff it is satisfiable by a geometric interpretation where all relations are interpreted by convex sets. An existential rule $[B_1 \wedge \cdots \wedge B_n \rightarrow \exists X_1, \ldots, X_j.H_1 \wedge \cdots \wedge H_k$ is called quasi-chained iff each atom B_i shares maximally one variable with the variables of all atoms coming before it, i.e., $(var(B_1) \cup \ldots var(B_{i-1})) \cap B_i \leq 1$.

The use of convex sets for the interpretation of relations can be justified by their importance as, on the one hand, computationally feasible data structures as used in convex optimization (Boyd & Vandenberghe, 2004) and, on the other hand, as a linguistically and cognitively justified structure for representing concepts (Gärdenfors, 2000).

Convexity is preserved under many operations, in particular it is preserved under intersection and projection, which are the main operations expressible in the allowed fragment presented by Gutiérrez-Basulto and Schockaert (2018). Interestingly, due to the use of integrity constraints, these ontologies implicitly use some aspect of negation, namely that of disjointness. And, in fact, Gutiérrez-Basulto and Schockaert (2018) show that quasi-chained datalog$^{\pm}$ covers several well-known logics—some of them from the family of description logics. But neither one can cover full negation: implicitly, negation is allowed to occur on the right hand side as atomic negation, but not as negation on the left hand side of a rule. The latter case negation would allows us, for example, to express coverage aspects which amounts to the use of disjunction. The restriction in their language is not surprising as convexity is neither preserved under set complement nor under set union which are the two algebraic operations they consider.

In our approach, we try to do both, stick to convexity and allow for full negation by identifying a set of geometric structures that can be equipped with suitable algebraic operations. As Gutiérrez-Basulto and Schockaert (2018), we also presume finite satisfiability of the ontology but work dually by constructing concepts on the axes and then placing individuals on these. The reason is that we incorporate an additional structure, namely a scalar product, which in turn induces the negation operator. This one constrains the potential places in which the negations of concepts can be placed. This, in particular, prevents adapting the quasi-chainedness property which was defined for Datalog$^{\pm}$ but not defined for $\mathcal{ALC}$.

Generalizing the notion of full expressiveness (Mehran Kazemi & Poole, 2018) we consider in our approach various notions of faithfulness that express to what extent an embedding represents the knowledge of an ontology. Hohenecker and Lukasiewicz (2020) are also interested in learning embeddings of a knowledge base that are faithful in the following sense: the embedding models exactly the entailments of the knowledge base. But there is a main difference to our approach regarding the entailments: The kind of knowledge base is that of a database with integrity constraints. Hence the kind of reasoning considered by Hohenecker and Lukasiewicz (2020) is that of reasoning with a closed world assumption (actually they also consider local forms of the closed world assumption), and not that of reasoning over genuine ontologies, which adhere to the open world semantics. As a consequence, the kind of negation considered by Hohenecker and Lukasiewicz (2020) is that of negation as failure and not that of full classical negation as in our case.

3. The Description Logic $\mathcal{ALC}$

We are going to work with the description logic $\mathcal{ALC}$ (Schmidt-Schauß & Smolka, 1991; Baader, 2003). We assume that there is a DL vocabulary (signature) given by a set of

Name	Syntax	Semantics
top	$\top$	$\Delta^{\mathcal{I}}$
bottom	$\bot$	$\emptyset$
conjunction	$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
disjunction	$C \sqcup D$	$C^{\mathcal{I}} \cup D^{\mathcal{I}}$
negation	$\neg C$	$\Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$
universal quantifier	$\forall R.C$	$\{x \in \Delta^{\mathcal{I}} \mid \text{For all } y \in \Delta^{\mathcal{I}}\text{: If } (x,y) \in R^{\mathcal{I}} \text{ then } y \in C^{\mathcal{I}}\}$
existential quantifier	$\exists R.C$	$\{x \in \Delta^{\mathcal{I}} \mid \text{There is } y \in \Delta^{\mathcal{I}} \text{ s.t. } (x,y) \in R^{\mathcal{I}} \text{ and } y \in C^{\mathcal{I}}\}$

Table 1: Syntax and semantics for the DL $\mathcal{ALC}$

constants N_c, a set of role names N_R and concept names N_C. The $\mathcal{ALC}$ concepts (concept descriptions) over $N_C \cup N_R$ are described by the grammar

$$C \longrightarrow A \mid \bot \mid \top \mid \neg C \mid C \sqcap C \mid C \sqcup C \mid \exists R.C \mid \forall R.C$$

where $A \in N_C$ is an atomic concept, $R \in N_R$ is a role symbol, and C stands for arbitrary concepts. For readability, $\exists R^{n+1}.C = \exists R.\exists R^n.C$, where $\exists R^1.C = \exists R.C$. A classical $\mathcal{ALC}$ *interpretation* is a pair $(\Delta, (\cdot)^{\mathcal{I}})$ consisting of a set Δ, called the *domain*, and an *interpretation function* $(\cdot)^{\mathcal{I}}$ which maps constants to elements in Δ, concept names to subsets of Δ, and role names to subsets of $\Delta \times \Delta$. The semantics of arbitrary concept descriptions for a given interpretation $\mathcal{I}$ is given in Table 1. An *ontology* $\mathcal{O}$ is defined as a pair $\mathcal{O} = (\mathcal{T}, \mathcal{A})$ of a *terminological box (TBox)* $\mathcal{T}$ and an *assertional box (ABox)* $\mathcal{A}$. A TBox consists of *general inclusion axioms (GCIs)* $C \sqsubseteq D$ ("C is subsumed by D") with concept descriptions C, D. $C \equiv D$ abbreviates $\{C \sqsubseteq D, D \sqsubseteq C\}$. An ABox consists of a finite set of *assertions*, i.e., facts of the form $C(a)$ or of the form $R(a, b)$ for arbitrary concepts C, roles R and $a, b \in N_c$.

An interpretation $\mathcal{I}$ *models* a GCI $C \sqsubseteq D$, for short $\mathcal{I} \models C \sqsubseteq D$, iff $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$. An interpretation $\mathcal{I}$ *models* an ABox axiom $C(a)$, for short $\mathcal{I} \models C(a)$, iff $a^{\mathcal{I}} \in C^{\mathcal{I}}$ and it models an ABox axiom of the form $R(a, b)$ iff $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$. An interpretation is a *model of an ontology* $(\mathcal{T}, \mathcal{A})$ iff it models all axioms appearing in $\mathcal{T} \cup \mathcal{A}$. An ontology $\mathcal{O}$ *entails* a (TBox or ABox) axiom ax, for short $\mathcal{O} \models ax$, iff all models of $\mathcal{O}$ are also models of ax. A concept C is the *most specific concept* (msc) for an ABox-element a and an ontology $\mathcal{O}$, if $\mathcal{O} \vDash C(a)$ and for all concepts C', $\mathcal{O} \vDash C'(a)$ entails $\mathcal{O} \vDash C \sqsubseteq C'$. The *quantifier rank* for arbitrary concepts C is defined as the maximal nesting of quantifiers in it. Formally, for an $\mathcal{ALC}$ concept define the quantifier rank by recursion as $qr(A) = 0$, $qr(\neg C) = qr(C)$, $qr(C_1 \sqcap C_2) = qr(C_1 \sqcup C_2) = max\{qr(C_1), qr(C_2)\}$, and $qr(\exists R.C) = qr(\forall R.C) = qr(C) + 1$.

Each TBox $\mathcal{T}$ generates a Boolean algebra, the so-called *Lindenbaum-Tarski algebra* (Tarski, 1935). The reason why we consider the Lindenbaum-Tarski algebra will become clear later, we just note here that in order to apply the theory of ortholattices (see Sect. 4.2.1 below), on which our treatment of negation rests, we have to algebraize the ontology. The Lindenbaum-Tarski algebra can be defined for any theory in any logic. We show here how it can be defined for an $\mathcal{ALC}$ TBox $\mathcal{T}$. For concepts C, D let $\sim$ be the relation defined by $C \sim D$ iff $\mathcal{T} \models C \sqsubseteq D$ and $\mathcal{T} \models D \sqsubseteq C$. Relation $\sim$ is an equivalence relation inducing for each concept C an equivalence class $[C]$. Define operations $\sqcap$, $\sqcup$, $\neg$ on the equivalence classes by setting $[C] \sqcap [D] = [C \sqcap D]$, $[C] \sqcup [D] = [C \sqcup D]$ and $\neg[C] = [\neg C]$. As the relation $\sim$ is not only an equivalence relation but a congruence relation w.r.t. $\sqcap, \sqcup, \neg$, the given equalities are indeed well-defined and one can show that the equivalence classes fulfil the axioms of a Boolean algebra. Just for completeness, we show here for the case of negation that indeed the logical operations on the equivalence classes are well-defined: $\neg[C]$ is defined by $[\neg C]$. Taken any other representative D of the class $[C]$, i.e. $[D] = [C]$ or, equally, $C \sim D$. By definition this means $\mathcal{T} \models C \equiv D$. But then by definition of the entailment relation and the definition of $\neg$ for concepts we also have $\mathcal{T} \models \neg C \equiv \neg D$, in other words $[\neg C] = [\neg D]$. Hence, indeed $\neg$ on the equivalence classes is well defined. All Lindenbaum-Tarksi algebras that we consider in this paper have a meet operator $\wedge$ and hence induce a lattice with the order $a \leq b$ defined by $a \wedge b = a$. So it makes sense to talk about algebraic atoms (see 4.2.1) in Lindenbaum-Tarski algebras.

4. Al-Cone Models

Our geometric interpretations are based on axis-aligned cones (al-cones) which form a subclass of the class of convex cones. We assume that a Euclidean space E (a vector space over the real numbers $\mathbb{R}$) with a scalar (dot) product $\langle \cdot, \cdot \rangle$ is given. For illustration purposes we will usually assume the real vector space $\mathbb{R}^n$ to be equipped with the scalar product defined as follows: $\langle v, w \rangle := \sum_{i=1}^{n} v_i \cdot w_i = v^T \cdot w$. While we restrict examples to this common scalar product, it is worth noting that our approach only requires $\langle \cdot, \cdot \rangle$ to be a positive semi-definite bilinear form.

A *closed convex cone* X is a non-empty set that fulfils the following property: if $v, w \in X$, then also $\lambda v + \mu w \in X$ for all $\lambda, \mu \geq 0$. The scalar product is necessary to define the polarity operator on cones. The *polar cone* X° of a cone X is defined as follows:

$$X^\circ = \{v \in \mathbb{R}^n \mid \text{ For all } w \in X : \langle v, w \rangle \leq 0\}$$

In case of the usual scalar product $v^T \cdot w$ the polar cone X° contains all vectors that differ in orientation by at least 90 degrees from any vector contained in X. For example, Figure 3 (a) shows three al-cones: X spanning the top-left quadrant, Y along the positive x axis, and Z along the negative y axis. Figure 3 (b) portrays their polar cones, X° spanning the bottom right quadrant, Y° covering the halfplane $\{(x, y) | x < 0\}$ and Z° the halfplane $\{(x, y) | y > 0\}$. Figure 3 (c) includes cone U which is not an al-cone; we later use this example to exemplify special properties of cone structures only exhibited by al-cones.

The scalar product also justifies the choice of closed (and not arbitrary) convex cones. To explain, first note, that the scalar product induces a norm $\|v\| = \sqrt{\langle v, v \rangle}$. This again induces a topology built on open balls $B_\epsilon(v) := \{w \in E \mid \|w - v\| < \epsilon\}$. In turn, this notion induces the open sets $O \subseteq E$ of a topology on the Euclidean space E: O is called open iff for any $v \in O$ there there is $\epsilon > 0$ and a ball $B_\epsilon(v)$ completely contained in O. Closed sets of the topology are defined as set complements $E \setminus O$ of open sets O. The closed convex cones are closed w.r.t. this topology.

The reason for working with (topologically) closed cones is simplicity. Otherwise, at least for the case of al-cones, we would not be able to guarantee that—with set intersection as the conjunction operation—a distributive logic is induced.

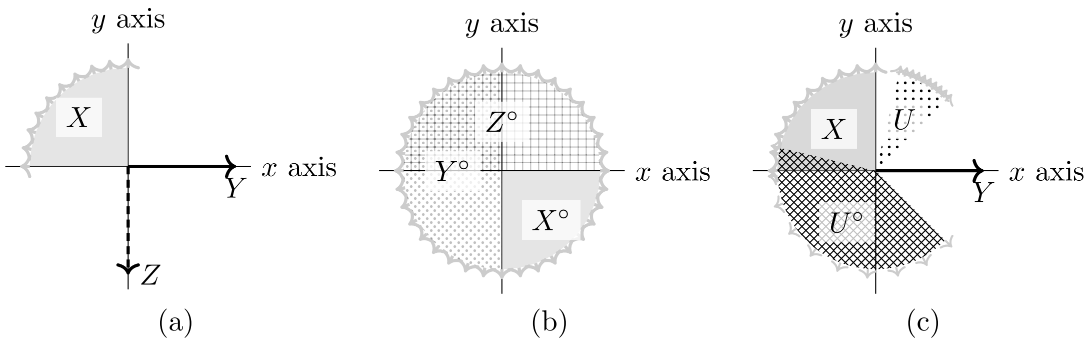


Figure 3: Examples of axis-aligned cones (al-cones) and their polars. Figure (a) shows three al-cones X, Y, Z and (b) their corresponding polar cones that are used to represent negation (in sub-figure (b) for visual clarity). Figure (c) shows al-cones X, Y together with cone U which is not an al-cone and its polar cone U°. We revisit this example to illustrate that al-cones constitute a special family of cones, namely one that exhibits properties of a Boolean algebra.

In the sequel, we use the term "cone" to refer to convex cones. Important cones are the so-called n-orthants or n-hyperoctants. For dimension $n = 2$ these are the 4 quadrants. Using the abbreviations $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$ and $\mathbb{R}_- = \{x \in \mathbb{R} \mid x \leq 0\}$ we can define the 2^n n-dimensional hyperoctants $H_{(b_1,\ldots,b_n)}$ with $b_i \in \{+,-\}$ by $\mathbb{R}_{b_1} \times \cdots \times \mathbb{R}_{b_n}$.

Cones can be constructed by unions of neighboring hyperoctants. Consider next to $+$, $-$, also the value u (for union) and define $\mathbb{R}_u = \mathbb{R}$. Then the axis aligned unions of hyperoctants are defined as sets $\mathbb{R}_{b_1} \times \cdots \times \mathbb{R}_{b_n}$ with $b_i \in \{+,-,u\}$. By further allowing for projections of all the sets mentioned above to the axes of the space we get the *al-cones*. (Just for completeness we note that a projection is a mapping $P : E \longrightarrow X$ of all elements of the space E onto a X where each vector in E is assigned the nearest point in X. But in the following we will not require this notion of projection anymore.)

$$X \text{ is an al-cone} :\Leftrightarrow X = (X_i)_{1\leq i\leq n} = X_1 \times \cdots \times X_n, \text{ where } X_i \in \{\mathbb{R}, \mathbb{R}_+, \mathbb{R}_-, \{0\}\}$$

So, in n dimensions we have 4^n possible al-cones. An illustration of all al-cones in $\mathbb{R}^2$ is given in Figure 4. In that figure we use an abbreviated notation. For example, the right upper quadrant $\mathbb{R}_+ \times \mathbb{R}_+$ is denoted by $(+,+)$. The positive x-axis $\mathbb{R}_+ \times \{0\}$ becomes $(+,0)$, the whole x-axis $(u,0)$ and so on.

The definition immediately entails the fact that al-cones in $\mathbb{R}^n$ are characterized by the sets of half-axes contained in it. So we introduce an operator $\mathit{halfAxes}(\cdot)$ denoting the half-axes of an al-cone X:

$$\begin{aligned} \mathit{halfAxes}(X) \;=\; & \{H \mid H = \{0\} \times \cdots \times \{0\} \times X_i \times \{0\} \times \cdots \times \{0\} \in X \\ & \text{for } X_i \in \{\mathbb{R}_+, \mathbb{R}_-\} \text{ and } i = 1, \ldots, n\} \end{aligned}$$

Let us now see how Boolean operations can be modeled. Aside set-based intersection to represent conjunction, special attention is required. All al-cones are convex cones. Convex cones are preserved under intersection, polarity, and other operations (such as projection), but not under the set-union operator. This applies to al-cones as well, but by interpreting

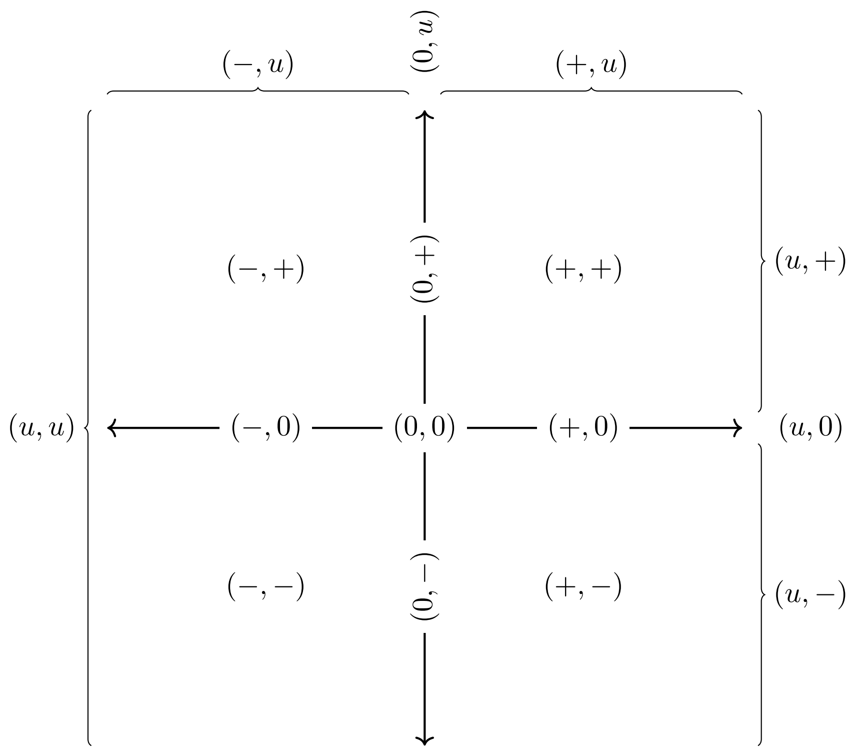


Figure 4: All al-cones in $\mathbb{R}^2$. The positive x-axis $\mathbb{R}_+ \times \{0\}$, e.g., is abbreviated by $(+, 0)$.

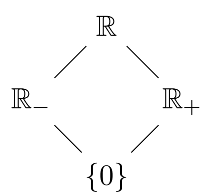


Figure 5: The Boolean Algebra O_4

$\sqcup$ using de Morgan (or, dually, relying on the conic hull operator) we can circumvent that problem.

In fact, for al-cones operations intersection, polarity, and convex union have very simple forms. This is due to the fact that the operations are based on the diamond shaped Boolean algebra $O_4 = \{\{0\}, \mathbb{R}_+, \mathbb{R}_-, \mathbb{R}\}$ with set-inclusion as order: $\{0\} \subseteq \mathbb{R}_+, \mathbb{R}_- \subseteq \mathbb{R}$ (see the illustration in Figure 5). Join $\vee$ on O_4 amounts to set union, the meet operation $\wedge$ to set intersection and the complement $^\perp$ is given as the set complement followed by union with $\{0\}$, i.e., $X^\perp := (\mathbb{R} \setminus X) \cup \{0\}$. We have $(X^\perp)^\perp = X$ and $X \cap Y^\perp = X^\perp \cup Y^\perp$.

Proposition 1 states that intersection of two al-cones becomes componentwise intersection (in O_4), polarity of al-cones becomes componentwise complementation (in O_4) and convex union become componentwise set union (in O_4). The proposition thus teaches us how algebraic computations on al-conces can be performed.

Proposition 1. *Let $X = (X_i)_{1 \le i \le n}, Y = (Y_i)_{1 \le i \le n}$ be two al-cones in $\mathbb{R}^n$. Then:*

1. $X \cap Y = (X_i \cap Y_i)_{1 \leq i \leq n}$
2. $X^\circ = ({X_i}^\perp)_{1 \leq i \leq n}$
3. $(X^\circ \cap Y^\circ)^\circ = (X_i \cup Y_i)_{1 \leq i \leq n}$

Proof.

1. $v = (v_i)_{1 \leq i \leq n} \in X \cap Y$ iff for all i with $1 \leq i \leq n$: $v_i \in X_i$ and $v_i \in Y_i$. This holds in turn iff for all i with $1 \leq i \leq n$: $v_i \in X_i \cap Y_i$. And this holds iff $v \in (X_i \cap Y_i)_{1 \leq i \leq n}$.

2. "$X^\circ \supseteq ({X_i}^\perp)_{1 \leq i \leq n}$": Let $v = (v_i)_{1 \leq i \leq n} \in ({X_i}^\perp)_{1 \leq i \leq n}$. Take any $w = (w_i)_{1 \leq i \leq n} \in X$. Due to the definition of $^\perp$ we have that for all i: $w_i \cdot v_i \leq 0$, so $\langle w, v \rangle \leq 0$ as well. As w was chosen arbitrarily from X we can infer that $v \in X^\circ$.
   "$X^\circ \subseteq ({X_i}^\perp)_{1 \leq i \leq n}$": Let $v = (v_i)_{1 \leq i \leq n} \in X^\circ$, i.e., for all $w \in X$: $\langle v, w \rangle = \Sigma_{1 \leq i \leq n} v_i \cdot w_i \leq 0$. Consider arbitrary i with $1 \leq i \leq n$. We argue by cases: If $X_i = \mathbb{R}$, then w_i can be changed to an arbitrarily large (positively or negatively) w_i', keeping all w_j for $j \neq i$ constant. Even under this change we still have to ensure $\langle v, (w_1, \ldots, w_{i-1}, w_i', w_{i+1}, \ldots, w_n) \rangle \leq 0$. And this is only possible if $v_i = 0$. Hence $v_i \in {X_i}^\perp = \{0\}$. If $X_i = \{0\}$, then ${X_i}^\perp = \mathbb{R}$ and then trivially $v_i \in {X_i}^\perp$. If $X_i = \mathbb{R}_+$, then ${X_i}^\perp = \mathbb{R}_-$. Again, we can make w_i arbitrarily large (but this time only positively). Still the scalar product of the w (changed in its ith component) and of v must be smaller than 0 and hence v can only be in $\mathbb{R}_- = {X_i}^\perp$. If $X_i = \mathbb{R}_-$, then ${X_i}^\perp = \mathbb{R}_+$. The argument works dually: We can make w_i arbitrarily negatively large, hence v_i cannot be negative but $v_i \in \mathbb{R}_+ = {X_i}^\perp$ must hold.

3. Due to 1. and 2. we have

$$\begin{aligned}
(X^\circ \cap Y^\circ)^\circ &= \left(({X_i}^\perp)_{1 \leq i \leq n} \cap ({Y_i}^\perp)_{1 \leq i \leq n}\right)^\circ \\
&= \left(({X_i}^\perp \cap {Y_i}^\perp)_{1 \leq i \leq n}\right)^\circ \\
&= \left(({X_i}^\perp \cap {Y_i}^\perp)^\perp\right)_{1 \leq i \leq n} \\
&= \left(({X_i}^\perp)^\perp \cup ({Y_i}^\perp)^\perp\right)_{1 \leq i \leq n} \\
&= (X_i \cup Y_i)_{1 \leq i \leq n}
\end{aligned}$$

□

As a corollary we get:

Corollary 1. *Al-cones X, Y in $\mathbb{R}^n$ are closed as follows:*

1. *The intersection $X \cap Y$ of X and Y is an al-cone.*
2. *The polar X° is an al-cone.*
3. $(X^\circ)^\circ = X$
4. *$(X^\circ \cap Y^\circ)^\circ$, the convex union of X, Y, is an al-cone.*

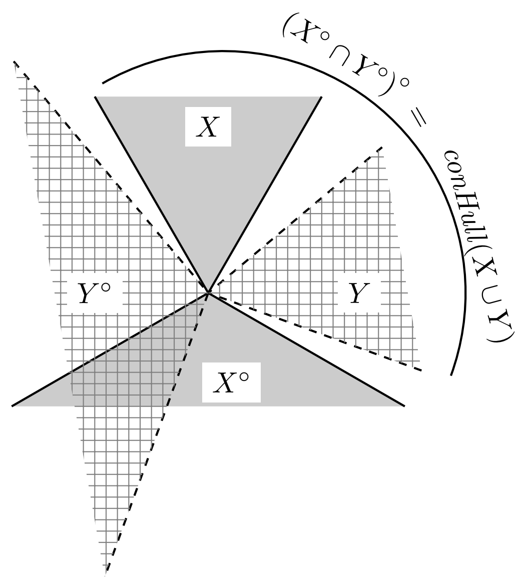


Figure 6: Illustration of the convex union of two cones

The operation of convex union $(X^\circ \cap Y^\circ)^\circ$ in the last item of the corollary is used in our embeddings for representing concept union $\sqcup$. Figure 6 illustrates the construction on arbitrary (not necessarily axis-aligned) cones. Note that the convex union of X and Y is not just the union of the cones but covers the area in between them too.

Within the figure one might already observe the fact that the convex union is an instance of the more general *conic hull* operator $conHull(\cdot)$, which we are now going to introduce. The *conic hull* of a set Y is defined as the smallest closed convex cone containing Y, formally:

$$conHull(Y) = \bigcap\{Z \mid Z \supseteq Y \text{ and } Z \text{ is a closed convex cone}\}$$

This operator is well-defined because the set $\{Z \mid Z \supseteq Y \text{ and } Z \text{ is a closed convex cone}\}$ is not empty (due to the fact that the whole space $E = \mathbb{R}^n$ is a closed convex cone). Moreover $conHull(Y)$ is indeed a closed convex cone (and thus non-empty), because the definition of closed convex cones is a universal sentence and universal sentences are preserved under arbitrary intersections. The following proposition justifies our claim that the convex union is an instance of the conic hull operator. In the proof of the proposition we rely on the antitonicity property of polarity which we state here as a lemma:

Lemma 1. *For all closed convex cones: if $X \subseteq Y$, then $Y^\circ \subseteq X^\circ$.*

Proof. Let $X \subseteq Y$ for closed convex cones X, Y. Then $Y^\circ \subseteq X^\circ$, because if $v \in Y^\circ$, then $\langle v, y\rangle \leq 0$ for all elements $y \in Y$, but then also $\langle v, y\rangle \leq 0$ for all $y \in X \subseteq Y$. □

Proposition 2. *For closed convex cones X, Y:*

$$(X^\circ \cap Y^\circ)^\circ = conHull(X \cup Y)$$

Proof. Clearly $(X^\circ \cap Y^\circ)^\circ$ is a closed cone. We have $X^\circ \cap Y^\circ \subseteq X^\circ$ so by antitonicity (Lemma 1) and elimination of double polarity $X \subseteq (X^\circ \cap Y^\circ)^\circ$. Similarly one infers that $Y \subseteq (X^\circ \cap Y^\circ)^\circ$ and so $X \cup Y \subseteq (X^\circ \cap Y^\circ)^\circ$. But $\mathit{conHull}(X \cup Y)$ is the smallest cone containing $X \cup Y$, so $\mathit{conHull}(X \cup Y) \subseteq (X^\circ \cap Y^\circ)^\circ$.

We are left with showing that $(X^\circ \cap Y^\circ)^\circ$ is also a minimal set containing $X \cup Y$. Let Z be another cone with $Z \supseteq X \cup Y$. We have to show $(X^\circ \cap Y^\circ)^\circ \subseteq Z$, or equivalently $Z^\circ \subseteq (X^\circ \cap Y^\circ)$. We obtain $Z^\circ \subseteq (X \cup Y)^\circ \subseteq X^\circ \cap Y^\circ$. □

The operator $\mathit{halfAxes}(\cdot)$ works smoothly with the operator $\mathit{conHull}(\cdot)$ in that they distribute in the sense explicated in the following proposition.

Proposition 3. *For all al-cones X, Y:*

$$\mathit{halfAxes}(\mathit{conHull}(X \cup Y)) = \mathit{halfAxes}(X) \cup \mathit{halfAxes}(Y)$$

Proof. This is an immediate consequence of item 3 in Proposition 1. □

Before we further embark on technicalities let us motivate the specifics of our approach by discussing two questions: First, why should one use the polarity operator as the interpretation of concept negation? Second, why should we restrict ourselves to axis-aligned cones?

4.1 Motivation of Polarity-Based Negation

The use of the polarity operation for concept negation is motivated by the idea of providing an operator that always maps a concept to a disjoint concept such that the disjoint concept is maximally w.r.t. the underlying similarity relation. Here, the similarity relation is given by the usual scalar product in the Euclidean space: The larger $\langle v, w \rangle$, the more similar are v and w. Usually one also norms the scalar product so that the cosine acts as a similarity measure: $\langle v, w \rangle / (\|v\| \cdot \|w\|)$. Two reasons substantiate our choice.

As for the first reason, one can observe that the idea of considering polarity as a form of negation is related to a general approach of defining negation on the basis of some binary complementary-relation *com*, as explained, e.g., by Dunn (1996) for propositional logics: relation *com* holds between two propositions iff they are complementary in the sense that there is no situation where both propositions are true (but still both may be false). Then, negation of a proposition p can be defined as the disjunction of all those propositions q that are complementary to p. Of course we are interested here in reading subsets C of the embedding space $\mathbb{R}^n$ as concepts and points in $\mathbb{R}^n$ as objects falling into the extension of C. But, drawing a closer connection to the approach of Dunn (1996), one could equally think of the points of $\mathbb{R}^n$ as possible worlds (or states). Then set C becomes a proposition which is exactly true for points contained in it.

As for the second reason, one can consider the following fact, known as Farkas' Lemma (see Figure 7a).

Lemma 2 (Farkas' Lemma, Farkas, 1902)**.** *Let C be the convex cone generated by vectors $v_1, \ldots, v_m \in \mathbb{R}^n$ (i.e., C = smallest convex cone containing all $v_i = \mathit{conHull}(\{v_1, \ldots, v_m\})$) and let $w \in \mathbb{R}^n$. Then either $w \in C$ or there is $z \in \mathbb{R}^n$ such that $\langle z, v_i \rangle \leq 0$ for all i, and $\langle z, w \rangle > 0$.*

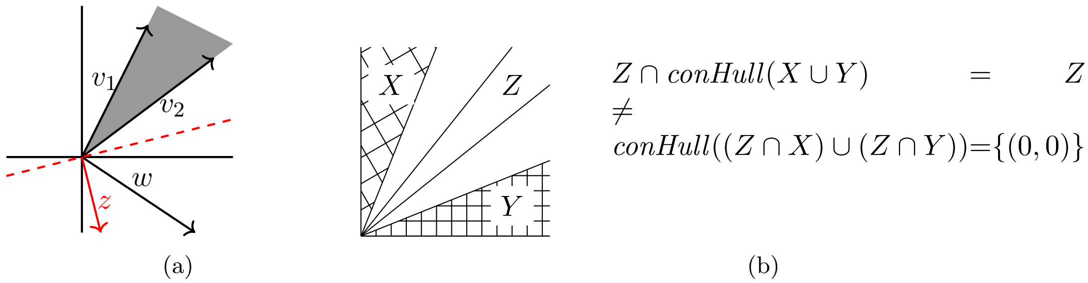


Figure 7: (a) Illustration of Farkas' Lemma; (b) counterexample of distributivity for arbitrary cones

The lemma says that if one considers a vector w that is not in the convex cone, i.e., a vector that is in the complement $\mathbb{R}^n \setminus C$ (the "usual" negation of C, namely set complement), then there is at least a verifier z that is similar to w (namely $\langle z, w \rangle > 0$) and is contained in the polar cone of C. Note that from $\langle z, v_i \rangle \leq 0$ for all i it follows that $\langle z, v \rangle \leq 0$ for all $v \in C$. In the other direction, if one can find a z of the polar cone being similar to w, then w is in the complement of C. In conclusion, polarity can be considered to provide a more intuitive model of negation for scalar-product based similarity structures than set complement.

4.2 Motivation for Considering Al-Cones

Firstly, the usefulness of al-cone models hinges on whether they are able to represent an interesting class of ontologies. In the remainder of this paper we show that al-cone models are indeed complete in the sense that an ontology is satisfiable classically iff it can be embedded into a geometric model based on al-cones.

Secondly, al-cones are motivated by their ability to link to ML. So far we have motivated the use of convex cones as proper structures to handle negation, yet we emphasize that convex cones also are used as computationally feasible data structures in the area of conic optimization (see, e.g., Boyd and Vandenberghe (2004), Section 4.3 on linear optimization problems and Section 4.4.2 on second-order cone programming) and are therefore attractive for ML applications. Here we only regard the case where one aims to learn a statistical model for data that can be characterized by some ontology that has been specified in a logic beforehand, not the case of investigating logics induced by the intersection and polarity operators for arbitrary cones. In fact, if we interpret concept conjunction $\sqcap$ as set intersection and concept disjunction $\sqcup$ as convex union, then the resulting logic cannot be guaranteed to lead to Boolean (and so not to full) $\mathcal{ALC}$ as it would not fulfil the distribution property for $\sqcap$ and $\sqcup$. A simple example is given in Figure 7 (b) for non-al-cones X, Y, Z.

So, in order to handle classical TBoxes such as Boolean $\mathcal{ALC}$ TBoxes, al-cones seem to be an appropriate choice. But are they the only subclass of cones with all possible (closed convex) cones in Euclidean spaces? The example of 7b shows that the class of arbitrary closed convex cones does not satisfy distributivity. The reader is invited to revisit Figure 3 (c) and verify that it presents a configuration similar to Figure 7 (b), only involving a single non-al-cone U. In case of Figure 3 (c), $conHull(X \cup Y) \cap U = U$ is obtained but $X \cap U = \{\vec{0}\}$ and $Y \cap U = \{\vec{0}\}$ holds. In contrast to this, distributivity is preserved for al-cones depicted in Figure 3 (a), (b) where $conHull(X \cup Y) \cap Z = X \cap Z = Y \cap Z = \{\vec{0}\}$ holds. But maybe further distributive subclasses exist? In order to treat this question formally, we shortly describe in the following the basic lattice theoretic structure of arbitrary classes of cones, namely that of ortholattices. Based on this notion and a characterization of Boolean algebras, we then give a characterization result showing that any subclass of cones that fulfils (also) distributivity has the same expressivity as al-cones, meaning they can express the same ontology in a given n-dimensional space. This then paves the way for possible incremental learning algorithms based on al-cones.

4.2.1 Ortholattices

A *lattice* $(L, \leq)$ is a structure with domain L and a partial order $\leq$ such that for any pair of elements $a, b \in L$ there is a smallest upper bound denoted $a \vee b$ and a largest lower bound denoted $a \wedge b$. As usual, $x < y$ means $x \leq y$ and not $y \leq x$. A *bounded lattice* $(L, \leq)$ contains a smallest element $\underline{0}$ and a largest element $\underline{1}$, i.e., elements such that for all $x \in L$ one has $\underline{0} \leq x \leq \underline{1}$.

An *algebraic atom*[1] b in such a lattice is an element that covers $\underline{0}$, i.e., $\underline{0} < b$ holds and for all a with $\underline{0} \leq a \leq b$ either $a = \underline{0}$ or $a = b$. If the lattice is a Lindenbaum-Tarski algebra, we call "algebraic atom" also a representative of the equivalence class that is a algebraic atom in the lattice. The context makes clear whether we mean the equivalence class or a representative. A lattice is called *algebraically atomic* iff each element has an algebraic atom below it. Intuitively, in atomic lattices one excludes the possibility of having an infinitely descending chain whose infimum is $\underline{0}$.

A lattice is called *distributive* iff for all $a, b, c \in L$: $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ (and dually: $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$).

In a lattice an element a^* is called a *complement of* a iff $a \wedge a^* = \underline{0}$ and $a \vee a^* = \underline{1}$. A lattice is said to be *complemented (uniquely complemented)* iff each a has a complement (has exactly one complement). A bounded lattice L is called an *ortholattice* iff it has an orthocomplement $\cdot^\perp$, i.e., a function such that for all $a, b \in L$ the following three conditions hold:

- $a \leq b$ entails $b^\perp \leq a^\perp$ (antitonicity)
- $a^{\perp\perp} = a$ (double complement elimination)
- $\underline{0} = a \wedge a^\perp$ (intuitionistic absurdity)

Any ortholattice satisfies de Morgan's laws, i.e., for any $a, b \in L$ it holds that $(a \wedge b)^\perp = a^\perp \vee b^\perp$ (and dually: $(a \vee b)^\perp = a^\perp \wedge b^\perp$).

A Boolean algebra is an ortholattice in which distributivity holds. Hence algebraic atoms of a Boolean algebra are the algebraic atoms of an ortholattice. As an example of an ortholattice that is not a Boolean algebra consider the hexagon lattice, termed O_6 in the literature shown in Figure 8. It is easy to check that the hexagon is indeed a lattice. It is orthocomplemented with $\cdot^\perp$ defined by $\underline{1}^\perp = \underline{0}$, $\underline{0}^\perp = \underline{1}$, $a^\perp = d$, $d^\perp = a$, $c^\perp = b$, $b^\perp = c$. Indeed, all three properties of an orthocomplement are fulfilled. On the other hand

1. Note that in contrast to the usual wording, we add the specification "algebraic" in order to prevent clashes with the atomic concepts in description logics.

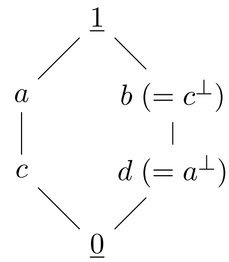


Figure 8: The ortholattice O_6

the lattice is not distributive (even not orthomodular) because we have $c \leq a$ such that $a \wedge (b \vee c) = a \neq c = \underline{0} \vee c = (a \wedge b) \vee c = (a \wedge b) \vee (c \wedge a)$.

There is a more fine-grained characterization by MacNeille (1937) which we use for our characterization result of al-cones. MacNeille's characterization result states that an algebraic structure is a Boolean algebra iff it is an ortholattice and fulfills the following additional axiom:

(*) For all a, b: If for all c: $a \wedge b \leq c$, then $a \leq b^\perp$.

The characterization is such that it dispenses with the special elements $\underline{1}$ and $\underline{0}$ of a lattice. But if the smallest element $\underline{0}$ is allowed, then this axiom can be expressed in the following form:

(wLLJ) For all a, b: If $a \wedge b \leq \underline{0}$, then $a \leq b^\perp$.

We call this rule *weak Johansson's constructive contraposition* as it is a special case of Johansson's constructive contraposition named (LLJ) by Hartonas (2016).

(LLJ) For all a, b, c: If $a \wedge b \leq c$, then $a \wedge c^\perp \leq b^\perp$.

When setting $c = \underline{0}$ in (LLJ) one immediately gets (wLLJ).

4.2.2 Ortholattices generated by Cones

First we note that any class of all closed convex cones (over $\mathbb{R}^n$ for some n) makes up an ortholattice.

Proposition 4. *For any $n \geq 1$ the set of closed convex cones in $\mathbb{R}^n$ is an ortholattice.*

Proof. We first show that the properties of an orthocomplement are fulfilled. Antitonicity: Lemma 1. Double complement elimination: "$(X^\circ)^\circ \subseteq X$": Assume $w \notin X$. Then by Farkas' Lemma there is some z such that $\langle z, v\rangle \leq 0$ for all $v \in X$ and $\langle z, w\rangle > 0$. But the first conjunct says that $z \in X^\circ$ and the second says that z is a falsifier for w being in $(X^\circ)^\circ$. "$(X^\circ)^\circ \supseteq X$": Let $w \in X$. consider an arbitrary $v \in X^\circ$. For all $x \in X$ we have $\langle v, x\rangle \leq 0$, in particular for $x = w$: $\langle v, w\rangle \leq 0$. As v was chosen arbitrarily we have $w \in (X^\circ)^\circ$. Intuitionistic absurdity: Clearly $\{\vec{0}\} \subseteq X \cap X^\circ$ because every closed convex cone contains $\{\vec{0}\}$. On the other hand if $v \in X \cap X^\circ$ then v must in particular fulfill $\langle v, v\rangle \leq 0$ which can be the case only for $v = \{\vec{0}\}$.

Now we are left with showing that the class of closed convex cones is partially ordered by the subset relation $\subseteq$ and has largest lower bounds (meet) and smallest upper bounds (join). For this order there is indeed a meet operation, namely the intersection of two closed convex cones. Since set intersection of two closed convex cones yields a closed convex cone then the set intersection must indeed be the largest lower bound. The join operation can be defined by the conic hull $conHull(X \cup Y)$ over the union of the join arguments, which, necessarily, is the smallest closed convex cone containing X and Y. □

When considering ortholattices of cones we allow us to denote lattice operators with special symbols.

Definition 1. *$\dot{\sqsubseteq}$ denotes the lattice order (amounts to the subset-relation); $\dot{\sqcap}$ denotes the meet operation on cones (amounts to intersection); the symbol $\dot{\sqcup}$ denotes the join operation on cones (corresponds to the convex hull), $^\circ$ denotes orthocomplement (corresponds to polarity); $\dot{\bot}$ stands for the least element (corresponds to $\{\vec{0}\}$) and $\dot{\top}$ for the largest element of the lattice (corresponds to the whole space).*

As we are interested in distributivity, we show with the following characterization result how to restrict geometric models based on closed convex cones to the distributive ones.

Proposition 5. *A subclass of closed convex cones fulfills distributivity iff for each combination of two cones X and Y it holds that $X \dot{\sqcap} Y \dot{\not\sqsubseteq} \dot{\bot}$ or for each $x \in X, y \in Y$, $\langle x, y \rangle \leq 0$.*

Proof. We use MacNeille's axiomatization of Boolean algebras (according to Padmanabhan and Rudeanu (2008), axiom system B67, p. 114):

1. $\forall a \forall b (a \leq b \ \& \ b \leq a \Rightarrow a = b)$
2. $\forall a \forall b \forall c (a \leq b \ \& \ b \leq c \Rightarrow a \leq c)$
3. $\forall a \forall b [(a \wedge b) \leq a \ \& \ (a \wedge b) \leq b \ \& \ \forall c ((c \leq a \ \& \ c \leq b) \Rightarrow c \leq (a \wedge b))]$
4. $\forall a \forall b (a \wedge a^\perp \leq b)$
5. $\forall a \forall b [(\forall c (a \wedge b \leq c)) \Rightarrow a \leq b^\perp]$
6. $\forall a \forall b (a \leq b \Rightarrow b^\perp \leq a^\perp)$

All but one axiom, namely the 5$^{\text{th}}$ axiom, of this axiom system follow trivially from the definition of ortholattice given in Section 4.2.1 and the geometric interpretation of cones and are fulfilled by all cone structures. This reduces the proof to show the validity of the 5$^{\text{th}}$ axiom above. This axiom corresponds to the rule (wLLJ), which said that if $a \wedge b \leq \underline{0}$, then $a \leq b^\perp$. (Note for the correspondence that $a \wedge b$ is smaller than all c iff $a \wedge b$ is smaller than $\underline{0}$).

"→": (wLLJ) can be interpreted on a geometric level as that when X and Y are disjoint, then X must be a subset of Y°. By the definition of negation as polar cone this is possible only if for each $x \in X$ and each $y \in Y$ it holds that $\langle x, y \rangle \leq 0$. When X and Y intersect, then the condition is fulfilled trivially.

"←": Follows from (wLLJ) and the definition of polarity. □

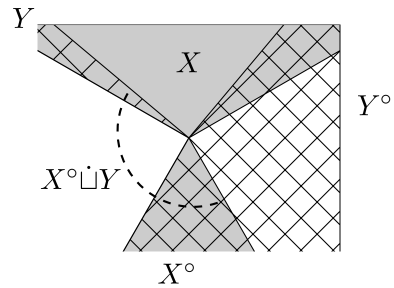


Figure 9: Example of a model of cones that are not al-cones but distributive. Distributive models exclude the configuration shown in Figure 7(b) that some cone is located between other cones and will be covered when building the convex hull.

To give an example, Fig. 9 shows a set of cones which exhibits distributivity, and is not composed out of al-cones. Al-cones satisfy the condition of Proposition 5. In context of introducing al-cones, we presented a counterexample in Fig. 7(b) on page 300.

Revisiting this example, observe that condition of Proposition 5 is not fulfilled as X and U are disjoint, but there are elements $u \in U, x \in X$ with $\langle x, u\rangle > 0$ (thus, $X \not\subseteq U^\circ$). Therefore, distributivity is not satisfied which is easy to verify: Consider $(X \dot{\sqcup} Y) \dot{\sqcap} U$ which evaluates to U, but $(X \dot{\sqcap} U) = (Y \dot{\sqcap} U) = \dot{\bot}$.

Proposition 6. *A distributive cone model with dimension n can be represented with an al-cone model of the same dimension.*

Proof. An al-cone model can be created by placing algebraic atoms on half-axes in $\mathbb{R}^n$ (as explained in more detail, e.g. in Proposition 7).

Consider now a general distributive cone model and the set of algebraic atoms $\mathcal{X}$. For each $X \in \mathcal{X}$: For all $y \in Y^{\mathcal{I}}$ with $Y \in \mathcal{X} \backslash X$ and thus $Y^{\mathcal{I}} \in (X^{\mathcal{I}})^\circ$ and all $x \in X^{\mathcal{I}}$: $\langle x, y\rangle \leq 0$, meaning the angle between x and y has to be at least 90°. Because of distributivity, there must be a $y \in Y^{\mathcal{I}}$ such that $\langle x, y\rangle = 0$, thus the angle between x and y is either 90°, as if that would not be the case, then there is a $y' \in (Y^{\mathcal{I}})^\circ$ with $y' \notin X$ but $\langle x, y'\rangle > 0$ for some $x \in X$, a contradiction to distributivity or the angle between x and all $y \in Y^{\mathcal{I}}$ is 180° with the same argumentation. As this is the case for all $X \in \mathcal{X}$, the greater the cone representing the algebraic atom would be, the less atoms can be placed in one dimension. Therefore, a model requiring as few dimensions as possible models algebraic atoms as rays starting at the point of origin. These algebraic atoms (rays) can be rotated to an axis which then leads to an al-cone model. □

This means that al-cones are the basis for representing $\mathcal{ALC}$ as each non-al-cone distributive model can be converted to an al-cone-model. It is not possible to represent a distributive lattice in a non-al-cone model having a smaller dimension than the al-cone model.

5. Embedding for Propositional $\mathcal{ALC}$

Let us start by considering $\mathcal{ALC}$ ontologies where the TBox language amounts to propositional logic. That is, we first restrict definitions from Section 3 to the concept constructors

$\sqcap$, $\sqcup$, $\neg$ corresponding to the Boolean operations $\wedge, \vee, \neg$. A Boolean $\mathcal{ALC}$ TBox (ABox) consists of GCIs (ABox axioms) using Boolean concepts only.

Let $(\mathcal{T}, \mathcal{A})$ be a Boolean $\mathcal{ALC}$ ontology. We are going to define a geometric interpretation of a special kind that can be a model or an anti-model of a Boolean ontology. It is an ordinary model in the sense that it is a classical predicate logical structure with a domain and interpretations for the atomic concept symbols. It is special in the sense that the domain is of a specific structure, namely a Euclidean space and concept names are interpreted by al-cones and their projections to subspaces. It is non-classical in the sense that some logical constructors are not interpreted by the corresponding set operations, but by operations on al-cones. The interesting aspect is now the following: the fact that Boolean $\mathcal{ALC}$ induces a Lindenbaum-Tarksi Algebra, which is in fact a Boolean Algebra, provides global constraints on the model. These can be satisfied by choosing appropriate positions of the interpretations for all atomic concepts. As a consequence, an ontology is satisfiable classically iff it is satisfied by the geometric interpretation.

Definition 2. *A* Boolean al-cone interpretation $\mathcal{I}$ *is a structure* $(\Delta, (\cdot)^{\mathcal{I}})$ *where* Δ *is* $\mathbb{R}^n$ *for some* $n \in \mathbb{N}$*, and where* $(\cdot)^{\mathcal{I}}$ *maps each concept symbol* A *to some al-cone and each constant* a *to some element in* $\Delta \setminus \{\vec{0}\}$*. An* al-cone interpretation for arbitrary Boolean $\mathcal{ALC}$ concepts *is defined recursively as* $(\top)^{\mathcal{I}} = \Delta$*,* $(\bot)^{\mathcal{I}} = \{\vec{0}\}$*,* $(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$*,* $(\neg C)^{\mathcal{I}} = C^{\circ}$*, and* $(C \sqcup D)^{\mathcal{I}} = (\neg(\neg C \sqcap \neg D))^{\mathcal{I}}$*. The notions of an al-cone being a model and that of entailment are defined as in the classical case (but using al-cone interpretations). In the following, an al-cone interpretation is called a* geometric model*. A geometric model on* $\mathbb{R}^n$ *is a geometric model restricted to al-cones of at most dimension* n*.*

To clarify this definition, in the following, an example of the construction of a geometric model for a simple ontology is presented.

Example 1. *Let us consider a simple example of a Boolean ontology, consisting of an empty TBox and an ABox constructed as follows: Consider the Lindenbaum-Tarski algebra of all Boolean concepts defined on two Boolean names, say* A, B*. Then the generated Boolean algebra of classes of equivalent concepts has* $2^{2^2} = 16$ *elements. Choose for each a "smallest" representative* C_i*,* $i \in \{1, \ldots, 16\}$ *w.r.t., say, the lexicographical ordering of the formulae based on an arbitrary ordering of all propositional symbols and logical symbols. In Figure 10 one can see those 16 (representatives of) concepts. So, e.g.,* $A \sqcap B$ *is the representative chosen for all the infinitely many concepts such as* $A \sqcap A \sqcap B$*,* $A \sqcap B \sqcap (B \sqcup \neg B)$ *etc. that are equivalent to* $A \sqcap B$*.* C_1 *is the bottom concept and let us assume that* C_2 *is the concept* B *and* C_3 *is the concept* $B \sqcap \neg A$*. The ABox has the following form: for each* i *with* $n \geq i \geq 2$ *there is a constant* a_i *and an ABox axiom* $C_i(a_i)$*.*

Thus we represent each of the 15 non-bottom concepts uniquely by some constant. For example, in Figure 10 the constant a_2 *represents the concept* $C_2 = B$*. This ontology is satisfiable in the classical sense. There is an al-cone interpretation that fulfills the ontology in* $\mathbb{R}^2$ *that is constructed as follows: interpret* A *by the left upper quadrant and* B *by the right upper quadrant. This induces uniquely the positions of all other hyperoctants corresponding to the other concepts* C_i*. The interesting point is the localization of constants* a_i*. For each constant* a_i *one localizes the hyperoctant corresponding to* C_i *in the area and positions on it such that it does not coincide with one of the points already associated with a constant.*

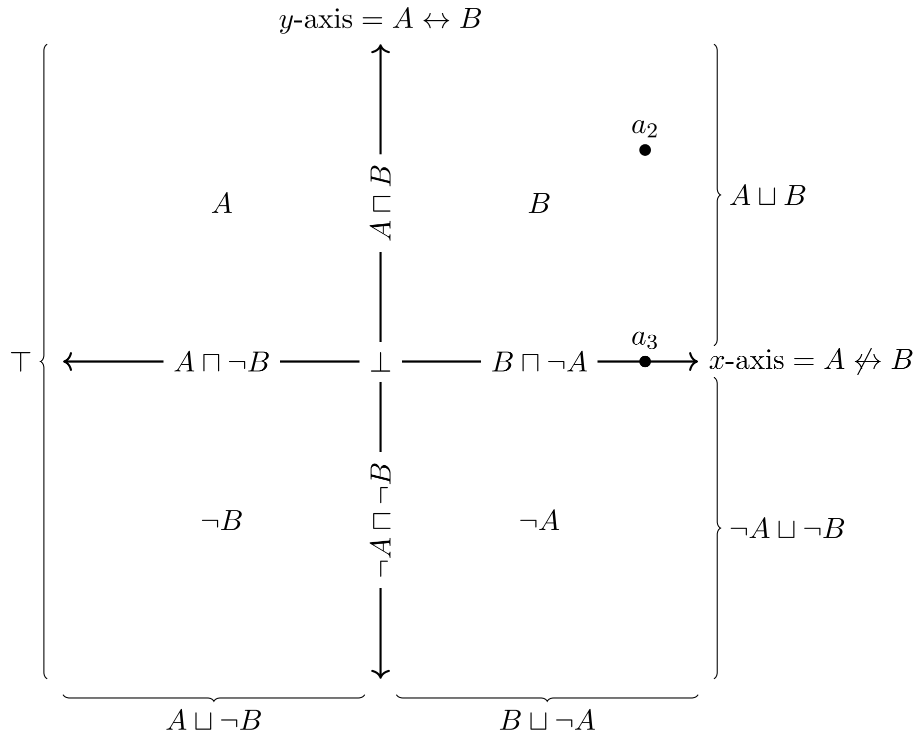


Figure 10: Al-cone model for a simple ABox with empty TBox

Figure 10 illustrates the positions of the concepts: $\bot$ *is by definition the singleton* $\{\vec{0}\}$*,* A *is the left upper quadrant,* $\top$ *is the whole area, etc.*

One can check that the concepts are associated with appropriate al-cones. For example, the negation $\neg A$ *of* A *is indeed the polar cone of the quadrant of* A*. Similarly, consider* $B \sqcap \neg A$*, which is interpreted as the positive* x*-axis* $\mathbb{R}_+ \times \{0\}$*. Its polar cone is the whole left area.*

Now let us turn to an example with a non-empty TBox.

Example 2. *Consider the Boolean algebra induced by the atomic concepts* A, B *under the TBox axiom* $A \sqsubseteq B$*. Then, for example,* $A \sqcap \neg B \sqsubseteq \bot$ *holds and so forth. This gives an embedding with axis-aligned cones as illustrated in Figure 11.*

Using the construction idea of the examples one can show that $\mathcal{ALC}$-ontologies are classically satisfiable iff they are by a geometric model based on al-cones.

Proposition 7. *Boolean* $\mathcal{ALC}$*-ontologies are classically satisfiable iff they are by a geometric model on some finite* $\mathbb{R}^n$ *based on al-cones of the form* $b_1 \times \cdots \times b_n$ *with* $b_i \in \{\{0\}, \mathbb{R}_+, \mathbb{R}_-, \mathbb{R}\}$ *for* $i \in \{1, \ldots, n\}$*.*

Proof. "$\rightarrow$": For the general construction we use the encoding from the examples. Remember that the set of encodings $\{\{0\}, \mathbb{R}_+, \mathbb{R}_-, \mathbb{R}\}$ used in the example forms a diamond shaped Boolean algebra O_4 with $\{0\} \subseteq \mathbb{R}_+, \mathbb{R}_- \subseteq \mathbb{R}$ (see Figure 5). Like the $\mathcal{ALC}$-ontology $\mathcal{O}$, the induced Boolean algebra is finite, say containing no more than 2^k elements, with a finite

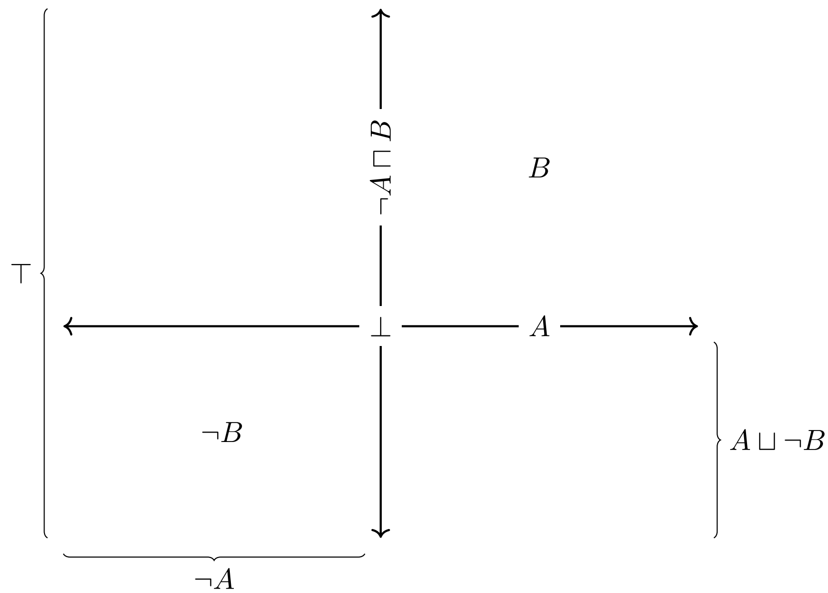


Figure 11: Al-cone model for TBox $\{A \sqsubseteq B\}$

number of atoms m. We choose a dimension n such that $4^n \geq 2^k$ and $n \geq m/2$, the idea being that we have to represent all of the 2^k Boolean concepts by embedding the algebraic atoms of the Boolean algebra directly onto half-axes of $\mathbb{R}$. Now one can see that O_4^n is the n-wise Cartesian product of O_4 where $\sqcap$, $\sqcup$, $\neg$ are defined componentwise based on the $\sqcap$, $\sqcup$, $\neg$ defined on O_4. But in this way $\sqcap$ over O_4 is nothing else than set intersection, $\neg$ is polarity, and $\sqcup$ is defined by de Morgan (see Proposition 1). So, if $\mathcal{O}$ is satisfiable, then we can construct a geometric model.

"$\leftarrow$": Assume that a geometric model $\mathcal{I}_g$ of the ontology $\mathcal{O} = \mathcal{T} \cup \mathcal{A}$ exists. We have to construct a classical model $\mathcal{I} \models \mathcal{T} \cup \mathcal{A}$. For the specification of $\mathcal{I}$ it is sufficient to specify the denotations of the constants occurring in $\mathcal{A}$, denoted by N_c in the following, and for the concept symbols occurring in $\mathcal{O}$, denoted by $N_C = \{A_1, \ldots, A_n\}$ in the following. We may further assume that each ABox axiom is of the form $A_i(c)$. Otherwise, if the axiom is of the form $C(c)$ we would add a new atomic symbol A_{n+j} to N_C and replace $C(x)$ by $A_{n+j}(x), A_{n+j} \equiv C$.

Let the domain $\Delta^{\mathcal{I}}$ be just the set of constants N_c. The constants are interpreted by themselves, i.e. for any constant $c \in N_c$ set $c^{\mathcal{I}} = c$. We are left with specifying the denotations of each $A_i \in N_C$. For each constant $c \in N_c$ let $X \subseteq \Delta^{\mathcal{I}_g}$ be the al-cone representing the most specific concept including $c^{\mathcal{I}_g}$, denoted by concept C. Now consider the following set of algebraic atoms compatible with C, i.e., consider the set

$$atc(c) = \{L = L_1 \sqcap \cdots \sqcap L_n \mid L_i \in \{A_i, \neg A_i\} \text{ and } L^{\mathcal{I}_g} \subseteq X\}$$

Intuitively, $atc(c)$ describes the possible "groundings" of the concept C, as either C is an algebraic atom and thus $atc(c) = \{C\}$ or C is not. If C is not an algebraic atom, then the ability of the geometric model of modeling partial knowledge is used. Thus, e.g., $c^{\mathcal{I}_g} \in (C_1 \sqcup C_2)^{\mathcal{I}_g}$, but neither $c^{\mathcal{I}_g} \in (C_1)^{\mathcal{I}_g}$ nor $c^{\mathcal{I}_g} \in (C_2)^{\mathcal{I}_g}$. However, this ability is not given in a classical interpretation. Thus, an arbitrary element denoted $catc(c) \in atc(c)$ is chosen. Let $catc(c) = L_1 \sqcap \cdots \sqcap L_n$. Now we can define $A_i^{\mathcal{I}} = \{c \in N_c \mid A_i = L_j \text{ for some conjunct } L_j \text{ of } catc(c)\}$. Indeed, $\mathcal{I}$ is a model of the ontology. It makes all ABox axioms $A_i(c)$ true, because $\mathcal{I}_g \models A_i(c)$, hence $catc(c)$ contains A_i and does not contain $\neg A_i$. The constructed model makes also every TBox axiom $C \sqsubseteq D$ true. Because, assume $c \in C^{\mathcal{I}}$. As c is completely specified w.r.t. each symbol A_i we also have $c^{\mathcal{I}_g} \in C^{\mathcal{I}_g}$, hence $c^{\mathcal{I}_g} \in D^{\mathcal{I}_g}$ and so $c \in D^{\mathcal{I}}$. □

5.1 Faithfulness

Geometric models are more specific than classical models in the sense that they impose an underlying structure on the domain given by, firstly, the dimensions representing some latent features that usually are not even mentioned in the ontology and, secondly, a scalar product $\langle \cdot, \cdot \rangle$ over the space. Geometric interpretations are more general than classical interpretations in the sense that they are partial interpretations. Actually, when embedding an ontology into some space one expects the embedding to represent not a single, but all interpretations that make the ontology true. And this is (up to some exceptions, see below) the case in the construction of Proposition 7. This allows partial information to be encoded. Consider, e.g., the difference between a_2 and a_3 in the geometric model of Fig. 10. The individual a_3 is completely identified w.r.t. the given concepts A, B: it lies in the extension of B and in the extension of $\neg A$. For a_2 we "only" know that it must be a B, but we do not know whether it is also an A. Hence, a geometric model correctly reflects knowledge provable in the ontology: one can prove that a_2 is a B but one cannot prove that it is an A nor that it is a $\neg A$.

The geometric model illustrated in Fig. 10 does not reflect all knowledge (not) contained in an ontology in the correct way. For example, it is not possible to represent an object b which is known to be in $A \sqcup B$, but neither known to be an A nor known to be a B. The reason is that by our construction no place is left to reflect this partial knowledge.

The discussion above motivates the definition of *faithful geometric models* of a given ontology. As within ontologies one has to deal with many models there are actually two different adaptations, a weak one and a strong one. Orthogonally to this distinction one may also consider the faithfulness w.r.t. the ABox only (to the data only) or the TBox. An even finer distinction can be given by considering special subclasses of the ABox (which we will do here by considering concept assertions vs. role assertions) and the TBox.

Definition 3. *Let $\mathcal{O}$ be a classically consistent (DL) ontology (or any other representation allowing the distinction between ABox and TBox). For a geometric interpretation $\mathcal{I}$ we have the following notions of being a faithful model of $\mathcal{O}$:*

- *$\mathcal{I}$ is a* strongly concept-faithful *model of $\mathcal{O}$ iff for each concept C and each constant b the following holds: if $b^{\mathcal{I}} \in C^{\mathcal{I}}$, then $\mathcal{O} \models C(b)$;*
- *$\mathcal{I}$ is a* weakly concept-faithful *model of $\mathcal{O}$ iff for each concept C and each constant b the following holds: if $b^{\mathcal{I}} \in C^{\mathcal{I}}$, then $\mathcal{O} \cup \{C(b)\}$ is satisfiable classically;*
- *$\mathcal{I}$ is a* strongly (weakly) ABox-faithful *model of $\mathcal{O}$ iff it is strongly (weakly) concept-faithful and for each role R and constants a, b the following holds: if $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$, then $\mathcal{O} \models R(a, b)$ (resp. $\mathcal{O} \cup \{R(a, b)\}$ is satisfiable classically);*

- *$\mathcal{I}$ is a* strongly (weakly) TBox-faithful *model iff for all TBox axioms $\tau = C \sqsubseteq D$ the following holds: if $\mathcal{I} \models \tau$, then $\mathcal{O} \models \tau$ (resp. $\mathcal{O} \cup \{\tau\}$ is satisfiable) and if there is $x \in \Delta^{\mathcal{I}}$ with $x \in (C \sqcap \neg D)^{\mathcal{I}}$, then $\mathcal{O} \cup \{(C \sqcap \neg D)(a)\}$ is satisfiable for a fresh constant symbol a. (resp. $\mathcal{O} \not\models \tau$)*

When we consider classical geometric models such as those defined by Gutiérrez-Basulto and Schockaert (2018), it will in general not be possible to ensure strong faithfulness. So, the notion of strong faithfulness makes sense only for models which allow for representing partial information.

The notion of faithfulness is a generalization of the notion of full expressivity according to Mehran Kazemi and Poole (2018): Given a knowledge graph and a set of triples known to be true (positive set) as well as triples that are known to be false (negative set), the embedding is said to be *fully expressive* iff it maps the relations and constants into a space such that $(a\ R\ b)$ holds in the embedding iff it is in the positive set and $(a\ R\ b)$ does not hold iff it is in the negative set.

This reading of "fully expressive" is a natural generalization of the separation idea for support vector machines (Burges, 1998): instances are classified as belonging to some concept (class) C by finding a separating maximal margin between the true positive training instances of C and the true negative training instances of C. Hence, the kind of triples $(a\ R\ b)$ considered are such that R is the *has-type*-Relation, a is the object to be classified and b is the class C.

Historically, the expression "fully expressive" has been used rather differently in the knowledge embedding community: The idea was to consider the usual prospects of properties of (binary) relations, say that of being (non)-functional, (a)symmetric, (in)transitive etc. and consider those embeddings to be fully expressive that allow representing relations with those properties. This notion of "fully expressive" depends on the set of of properties considered to be interesting for relations. The notion of full expressivity according to Mehran Kazemi and Poole (2018) has the benefit of being independent of a specific set of properties of a relation. This notion of full expressivity further has the nice property that it provides a good measure to compare different approaches—as illustrated by Gutiérrez-Basulto and Schockaert (2018).

The notion of full expressivity, implicitly, contains a notion of negation: for each triple $(a\ R\ b)$ in the negative set , "its negation" can be considered to belong to the positive set. Unfortunately, such a negative set is usually not given explicitly in the data; rather, the negative triples are randomly generated as part of a procedure usually called "negative sampling" (as done, e.g., by Mehran Kazemi and Poole (2018)). But what would be a fair and justified method to construct those triples $(a\ R\ b)$ in the negative set? Here is a natural suggestion of a method: If one knows for the given relation R that it is disjoint with the relation S, then any for any triple $(a\ S\ b)$ in the positive set there is a justification for considering $(a\ R\ b)$ as part of the negative set. But focusing on just one single relation S that is disjoint from R would lead to skewed data. Ideally, one would like to consider the maximum of all relations known to be disjoint from the given relation R. But, now one may see the analogy to the definition of negation by Dunn (1996) (and our polarity notion) as discussed in Section 4.1—taking disjointness as the complementary-relation *com*. So actually, what one would like to express is that for a relation R one should have a form of negation. On the bottom line these considerations on full expressiveness give further justification to consider negation in ML scenarios.

Now let us come back to our al-cones. We present two ways to gain faithfulness, the first using al-cones, the second using a subclass of al-cones.

Proposition 8. *For classically satisfiable Boolean $\mathcal{ALC}$-ontologies there is a strongly concept-faithful and TBox-faithful geometric model on some finite $\mathbb{R}^{2n}$ based on al-cones of the form $b_1 \times \cdots \times b_{2n}$ with $b_{2i} \in \{0, \mathbb{R}_+, \mathbb{R}_-, \mathbb{R}\}$ and $b_{2i+1} = b_{2i}$. Here n is the number of atomic elements in the Boolean algebra generated by the TBox of the ontology.*

Proof. First a geometric model of the given TBox is generated in size n (using the method described in Prop. 7). By definition, this model is TBox-faithful, thus only strong concept-faithfulness needs to be shown. In the resulting model every dimension is doubled, $(A \sqcup B)^\mathcal{I} = \mathbb{R}_+ \times \mathbb{R}_- \times \mathbb{R}$ for instance becomes $(A \sqcup B)^\mathcal{I} = \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_- \times \mathbb{R}_- \times \mathbb{R} \times \mathbb{R}$. To ensure strong concept-faithfulness it needs to be the case that for any object a in the ABox if $a^\mathcal{I} \in C^\mathcal{I}$, then $O \models C(a)$. This can be achieved by embedding a into its most specific concept (msc) $M^\mathcal{I}$, for which not at the same time $a \in A^\mathcal{I}$ with $A \sqsubset M$ holds, i.e., not embedding it into a concept properly subsumed by $M^\mathcal{I}$. Let M be the msc of some object x. Consider the concepts C_i for $i = 1, \ldots, n$ properly subsumed by M. To ensure concept-faithfulness, $x^\mathcal{I} \in M^\mathcal{I}$ must be valid, however $x^\mathcal{I} \notin C_i^\mathcal{I}$ for all $i = 1, \ldots, n$. Each al-cone $C_i^\mathcal{I}$ must cover at least one dimension d where $(C_i^\mathcal{I})_d \subsetneq (M^\mathcal{I})_d$, as otherwise the subsumption would not be proper. In this dimension d it is either the case that $(M^\mathcal{I})_d = \mathbb{R}$ and $(C_i^\mathcal{I})_d \in \{\mathbb{R}_+, \mathbb{R}_-, \{0\}\}$ or $(M^\mathcal{I})_d = \mathbb{R}_+$ (analog for $\mathbb{R}_-$) and $(C_i^\mathcal{I})_d = \{0\}$. Consider first the second case, there $x^\mathcal{I}$ can be placed in this dimension in $\mathbb{R}_+ \backslash \{0\}$ and thus is not contained in $(C_i^\mathcal{I})_d$. For all other $C_j^\mathcal{I}$ (with $j \neq i$), either $(C_j^\mathcal{I})_d = (C_i^\mathcal{I})_d$ and thus $x^\mathcal{I} \notin C_j^\mathcal{I}$ or $(C_j^\mathcal{I})_d = \mathbb{R}_+$, then there must exist a dimension $k \neq d$ ensuring that $C_j^\mathcal{I} \subsetneq M^\mathcal{I}$. However, in the first case this is not that easy. When considering the model without doubled dimensions, it is possible to place $(x^\mathcal{I})_d$ in $\mathbb{R}_-$, when $(C_i^\mathcal{I})_d = \mathbb{R}_+$, however, it is possible that there is an $C_j^\mathcal{I}$ with $(C_j^\mathcal{I})_d = \mathbb{R}_-$ and $(C_j^\mathcal{I})_k = (M^\mathcal{I})_k$ for $k = 1, \ldots, d-1, d+1, \ldots, n$, thus is a proper subsumption of M only because of the difference in dimension d. Then placing $(x^\mathcal{I})_d$ in $\mathbb{R}_-$ would result in $x^\mathcal{I} \in C_j^\mathcal{I}$, a contradiction. Thus, the doubled dimension is needed. Therefore, $(M^\mathcal{I})_{d,d+1} = \mathbb{R} \times \mathbb{R}$, $(C_i^\mathcal{I})_{d,d+1} = \mathbb{R}_+ \times \mathbb{R}_+$ and $(C_j^\mathcal{I})_{d,d+1} = \mathbb{R}_- \times \mathbb{R}_-$. Then, $(x^\mathcal{I})_{d,d+1}$ can be placed in $\mathbb{R}_+ \times \mathbb{R}_-$, being neither in $C_i^\mathcal{I}$ nor in $C_j^\mathcal{I}$. This follows analogously for the other cases and the other C_i. Thus, x is only contained in the msc and not in any concept properly subsuming it and therefore concept-faithfulness is given. □

When disregarding $\mathbb{R}_+$, $\mathbb{R}_-$ in al-cones we are left with halfspaces, which are sufficient to define faithful models.

Proposition 9. *For classically satisfiable Boolean $\mathcal{ALC}$-ontologies there is a strongly concept-faithful and TBox-faithful geometric model on some finite $\mathbb{R}^n$ using sets of the form $b_1 \times \cdots \times b_n$ with $b_i = \{0\}$ or $b_i = \mathbb{R}$.*

Proof sketch. Assume that an $\mathcal{ALC}$ TBox with n algebraic atoms is given. Arbitrarily assign them to one of the following n spaces $\{0\} \times \{0\} \times \cdots \times \underbrace{\mathbb{R}}_{\text{position } i \in \{1,\ldots,n\}} \times \cdots \times \{0\} \times \{0\}$.

Each of the full axis ax_i, $1 \leq i \leq n$ can be represented by some concept at_{ax_i}. Now, the denotation for each concept symbol A can be found by considering the set I of those i such that $\mathcal{T} \models at_{ax_i} \sqsubseteq A$. Then one can set $A = conHull(\{ax_i \mid i \in I\})$. As we saw earlier (Proposition 2), the closure amounts to applying conic disjunction (defined by de Morgan with intersection and polarity). In turn, conic disjunction for al-cones can be done componentwise according to Proposition 1.

To ensure strong concept-faithfulness it needs to be possible for any object a in the ABox to get embedded into a concept representing its msc, but not at the same time into a concept properly subsuming it. Assume two concepts C and D properly subsuming the msc. As C and D are not subsuming each other, there must be a dimension i in $\mathbb{R}^n$ with $C_i^{\mathcal{I}} = \mathbb{R}$ and $D_i = \{0\}$ and a dimension $j \neq i$ where $D_j^{\mathcal{I}} = \mathbb{R}$ and $C_j = \{0\}$. Thus, an instance a can be placed in a region where $a_i \neq 0$ and $a_j \neq 0$, then it is in the msc but neither in $C^{\mathcal{I}}$ nor in $D^{\mathcal{I}}$. $\square$

Part of the result (the existence of some geometric model) also follows from the the fact that al-cones in $\mathbb{R}^n$ make up a Boolean algebra and that the regions built from full axes make up a Boolean subalgebra.

6. Embedding for Full $\mathcal{ALC}$

Our main task is to define semantics for roles and quantifiers. In contrast to established embedding approaches such as TransE (Bordes et al., 2013) the aim here is not to model the relation $R(a, b)$ between two specific instances a and b directly, but to model it on the conceptual level by proving statements such as $a \in \exists R.C$. This is done by defining an interpretation $(R)^{\mathcal{I}}$ of relation R as matrix $\mathcal{R}$, such that $R(a, b)$ is represented by the statement that $a^{\mathcal{I}}$ is in the al-cone generated by $\mathcal{R}b^{\mathcal{I}}$. Likewise, al-cones can be transformed by matrix multiplication. We will later see that $\mathcal{R}$ can be chosen such that the transformation always yields an al-cone. Then, for an instance $a^{\mathcal{I}}$ the interpretations of the type $(\exists R.C)^{\mathcal{I}}$ with $a^{\mathcal{I}} \in (\exists R.C)^{\mathcal{I}}$ can be determined by applying the matrix $\mathcal{R}$ to the interpretation $(C)^{\mathcal{I}}$ (where C itself can also contain an existential) and checking whether $a^{\mathcal{I}}$ is contained. Therefore, ABox-faithfulness in its orginal form cannot be ensured. However, we show that the notion of concept faithfulness can be extended to roles. Having this extension will enable extending the approach to directly stating concrete relations between instances. In technical terms, modelling (also roles) on the conceptual level amounts to showing that the Lindenbaum-Tarski algebra of an $\mathcal{ALC}$ TBox makes up a Boolean algebra with operators (BAOs) (Jonsson & Tarski, 1951; Jonnson & Tarski, 1952). This in turn amounts to showing that on top of having a Boolean algebra the semantics of $\exists R.(\cdot)$ in a geometric model $\mathcal{I}$ gives an operator with the following properties: $(\exists R.\bot)^{\mathcal{I}} = (\bot)^{\mathcal{I}}$ and $(\exists R.(C \sqcup D))^{\mathcal{I}} = (\exists R.C)^{\mathcal{I}} \dot{\sqcup} (\exists R.D)^{\mathcal{I}}$.

This enables negation to be generalized for concepts that contain roles. So we may express also negated concepts with roles such as $(\neg\exists R.C)(a)$. (Of course we do not model explicitly negation for roles as there is no negation operator for roles in $\mathcal{ALC}$).

The aim is now to adapt the notion of faithfulness to roles and to find a faithful geometric embedding based on cones including roles. Concept-faithfulness can be directly adapted to full $\mathcal{ALC}$, on the one hand restricted up to a specific quantifier rank, on the other hand for an unrestricted ontology.

The next definition adapts geometric faithful models from Definition 3 to full $\mathcal{ALC}$.

Definition 4. *Let $\mathcal{O}$ be a classically consistent (DL) ontology (or any other representation allowing the distinction between ABox and TBox). For a geometric interpretation $\mathcal{I}$ we have the following notions of being a faithful model for $\mathcal{O}$:*

1. *$\mathcal{I}$ is an m-*(quantifier)-rank concept-faithful *model of $\mathcal{O}$ iff for each concept C with rank at most m and each constant b the following holds: if $b^{\mathcal{I}} \in C^{\mathcal{I}}$, then $\mathcal{O} \models C(b)$*
2. *$\mathcal{I}$ is a* concept-faithful *model of $\mathcal{O}$ iff for each concept C and each constant b the following holds: if $b^{\mathcal{I}} \in C^{\mathcal{I}}$, then $\mathcal{O} \models C(b)$.*

First, we consider the problem only on a conceptual level, by which we mean that we do not specify a geometric interpretation of roles but only geometric models of concepts including role information. An idea for the construction would be to consider algebraic atoms including roles and to place them on half axes as for the Boolean case. This is not possible due to the fact that the algebra of $\mathcal{ALC}$ concepts is not an atomic lattice as in the Boolean case.

Example 3. *Consider the example $\mathcal{O}_n = \{\text{loves}(\text{narcis}, \text{narcis}), \text{Vain}(\text{narcis})\}$ described by Baader and Küsters (2006). The instance* narcis *fulfills all concepts of the form $\exists \text{loves}^n.\text{Vain}$, $n = 1, 2, \ldots$ which, connected conjunctively, give a chain of concepts C_i that become infinitely narrower and narrower. The C_is are defined by:*

$$C_0 := \exists R.\text{Vain} \textit{ and } C_i := C_{i-1} \sqcap \exists \text{loves}^{i+1}.\text{Vain}$$

Thus, narcis *would be in a concept of each quantifier rank.*

Even without a cyclic dependency in the ABox, it is possible to have concepts of arbitrary quantifier rank in form of a chain of relations of arbitrary depth, e.g. $R(a_1, a_2)$,$R(a_2, a_3)$, $\ldots$, $R(a_{i-1}, a_i)$ for an arbitrary (unknown) i. Depth i cannot be determined because due to the open world assumption it is possible that relations exist for which the extension is not stated in the ABox. Thus we see that the Lindenbaum-Tarski algebra of $\mathcal{ALC}$ TBoxes (in this case) may not be atomic and an ABox may contain an object for which there is no most specific concept as exemplified in example 3.

Sketch of the solution. Let us sketch our solution to tackle this problem of non-atomicity before we embark on technicalities. We are going to show that we can embed the Lindenbaum-Tarski algebra of an $\mathcal{ALC}$-TBox into an extended algebra with elements that play the role of atomic elements. These elements are constructed as infinite conjunctions of algebraic elements of the Lindenbaum-Tarski Algebra of a special kind (elements of $\mathcal{X}_i$ in Lemma 4 below). So actually we are doing more than embedding just $\mathcal{ALC}$-concepts but we embed carefully chosen infinite versions of $\mathcal{ALC}$ concepts.

To be more concrete, the infinitary logic $\mathcal{ALC}^{inf}$ we refer to is defined as follows: The logical symbols formulae of $\mathcal{ALC}^{inf}$ consist of strings that make up trees that can infinitely branch and can have infinitely long branches. As for finite trees, for these trees we have a notion of a level, level 0 being the level of the root. The branches of the trees are infinite words $\omega : \mathbb{N} \to \Sigma$ where

$$\Sigma = N_R \cup N_C \cup \{\exists, \sqcap, \sqcup, \neg, \bigsqcap, \bigsqcup, (,), .\}$$

is the alphabet consisting of a vocabulary, the logical symbols of $\mathcal{ALC}$ and two additional logical symbols $\bigsqcap, \bigsqcup$. Example words are of the form $\exists R.\exists R. \dots$ or $\exists R.\exists S \sqcap C.\exists R \dots$. Hence, if one wants to define specific subsets of those formulae (as done, e.g., in Lemma 5), one has to rely on coinduction (Rutten, 2005): Operations on infinite branches ω (of a tree) can be defined by defining the operation on the head $\omega(0)$ and on the derivative ω' of ω defined as $\omega'(n) = \omega(n+1)$. The two additional logical symbols of $\mathcal{ALC}^{inf}$ mentioned above, namely $\bigsqcap, \bigsqcup$, are intended to model infinitary conjunction and disjunction (these lead to the infinite branchings). The definition of an $\mathcal{ALC}^{inf}$ concept extends the definition of an $\mathcal{ALC}$ concept by two rules: If X is a (possibly infinite) set of $\mathcal{ALC}^{inf}$ concepts then so are $\bigsqcap X$ and $\bigsqcup X$. The semantics is as expected: $(\bigsqcap X)^{\mathcal{I}} = \bigcap\{C \mid C \in X\}$ and $(\bigsqcup X)^{\mathcal{I}} = \neg \bigcap\{\neg C \mid C \in X\}$.

Intuitively, the extended algebra is a natural limit of arbitrary Lindenbaum-Tarski algebras for $\mathcal{ALC}$ TBoxes. The notion of limit can be made precise by a (new) rank notion. This notion is also necessary to account for the fact that with our construction there might be infinitely many algebraic atoms in our extended algebra—due to the fact that they are based on a possibly infinite number of conjuncts.

We are going to consider first the simple case of an empty TBox in the non-cyclic and in the cyclic case and then treat arbitrary TBoxes.

6.1 Handling the Non-Cyclic Case

Now let us fill the sketch with details. How do the algebraic atoms in the extended algebra look like? They are defined in the Lindenbaum-Tarski algebra of a TBox in $\mathcal{ALC}$ considering the equivalence of $\mathcal{ALC}^{inf}$ concepts.

The idea is to construct them by an adequate choice of its conjuncts by using, e.g., only one positive existential and infinitely many negative ones in it. One first approach (Özçep, Leemhuis, & Wolter, 2020) is to restrict $\mathcal{ALC}$-ontologies to contain concepts up to a specific quantifier rank to get a m-(quantifier)-rank concept-faithful model. But this restricts the expressivity of the model, as faithfulness for higher ranks is lost.

Independent of the used construction, a drawback of a faithful interpretation is that it results in an infinite-dimensional cone model because of its infinitely many algebraic atoms (except for highly restricted ontologies). Because of the infinite-dimensional model, it is not possible to create the model as a whole in a suitable way, therefore, it is necessary to be able to extend the model iteratively whenever necessary without influencing the existing one. This also enables us to stick to the half-axis based construction principle of the Boolean case.

Thus, it is necessary to have a creation principle for algebraic atoms of the extended algebra which is able to handle their (possibly) infinite number. One idea is to use some notion of rank to model concepts of at most a specific rank, and create out of those rank-restricted concepts algebraic atoms based on the extended algebra. Then the modeling known from the Boolean case can be used. Therefore, a notion of rank is searched for, for which rank-restricted algebraic atoms are also algebraic atoms (in the extended algebra) in the non-restricted case and thus enables for iterative extension without having to change the original model.

Therefore, first a definition of rank-restricted interpretations is needed.

Definition 5. *An interpretation* $\mathcal{I}$ *is an* m-rank model of an ontology $(\mathcal{T}, \mathcal{A})$ *with an empty TBox* $\mathcal{T}$ *iff for all constants* a *and concepts* C *such that* $\mathcal{A} \models C(a)$ *and the rank of* C *is at most* m*:* $\mathcal{I} \models C(a)$.

This definition is independent of the type of rank considered. To demonstrate the necessary properties, we consider first the standard (syntactical) notion of quantifier rank as introduced in Section 3 and show its inappropriateness.

Example 4. *Consider a simple ontology with an empty TBox, a concept symbol* A *and a role symbol* R*. For the creation of a* 0*-quantifier-rank model, only the algebraic atom* A *and* $\neg A$ *need to be considered, as the rank* $qr(\exists R.\top) > 0$, $qr(\neg\exists R.\top) > 0$*. Thus, a one-dimensional space is needed. Let* $\mathbb{R}_+$ *represent* $A^{\mathcal{I}}$*, then* $\mathbb{R}_-$ *represents* $(\neg A)^{\mathcal{I}}$ *(or vice versa). A higher-ranked concept, e.g.,* $A \sqcap \exists R.A$ *can be approximated with concept* A*, as* $A \sqcap \exists R.A \sqsubseteq A$*. Therefore, all instances of* $(A \sqcap \exists R.A)^{\mathcal{I}}$ *can be placed in* $\mathbb{R}_+$*. The same holds for instances in* $(A \sqcap \exists R.\neg A)^{\mathcal{I}}$*. Therefore, each* k*-quantifier-rank model (for arbitrary* $k \geq 0$*) can represent all instances of the ABox correctly, however, not necessary as specific as it would be possible using a higher rank.*

Therefore, algebraic atoms of quantifier-rank restricted models are not algebraic atoms of unrestricted models. If the model is extended to an 1*-quantifier-rank model,* A *is not an algebraic atom anymore, as it contains, e.g.,* $A \sqcap \exists R.A$ *and* $A \sqcap \exists R.\neg A$*. Therefore, the instances placed in* $\mathbb{R}_+$ *in the smaller model cannot be placed in* $\mathbb{R}_+ \times \{0\} \times \cdots \times \{0\}$ *in the bigger model as this would contradict* $\exists R.A \sqcap \exists R.\neg A = \bot$*. Therefore, increasing the represented quantifier rank leads to the necessity of constructing a completely new model and thus, the quantifier rank is not suitable for the creation of an iteratively extendable model based on the rank.*

To circumvent the restriction mentioned in the example, it is necessary to define a new notion of rank, which we call the *semantic quantifier rank*. This leads to the possibility of creating a semantic-rank-restricted model which depicts a subspace of the model of the whole ontology, meaning, that if an algebraic atom is modeled on a specific half-axis in the model with rank-restricted concepts, this algebraic atom is also placed on exactly this half-axis in the model of the whole ontology. Thus, algebraic atoms of a smaller rank remain algebraic atoms in models of higher rank and the model is iteratively extendable without influencing the underlying smaller sized model. To reach this, concepts of smaller ranks cannot be interpreted as approximations of concepts of higher rank as done in Example 4. If a concept is either of the represented rank or smaller, then it can be represented correctly and accurate, if it has a higher rank, then it can not be represented or approximated at all (meaning higher ranked concepts are modeled as $\bot$ in a model representing lower ranked concepts). By circumventing the approximation in this way, it is not necessary to change the existing model when adding concepts with a higher rank and thus adding new dimensions to the model.

Intuitively, the semantic quantifier rank describes on which depth it is possible to model a concept having an actual extension (not being $\bot$). It therefore represents the necessary size of the model to be able to model a specific concept. The definition tests in the beginning whether the concept or a part of the concept is equivalent to $\top$ or $\bot$ and then proceeds by induction on the structure of the concept. The semantic rank gives either a natural number or ∞. Hence we assume in the following definition that ∞ is greater than all natural numbers and that $\min, \max, +$ handle ∞ accordingly.

Both, *srank* and *reduce* are defined for existentials only, thus universal quantifiers are interpreted as existentials by setting $\forall R.C = \neg\exists R.\neg C$. $reduce(D)$ is defined inductively as $reduce(D) = D$ for D being an atomic symbol or $reduce(D) = \top$ when D is equivalent to $\top$ or $reduce(D) = \bot$ when D is equivalent to $\bot$. $reduce(C \sqcap D) = reduce(C)$ if $reduce(D)$ is equivalent to $\top$, $reduce(C \sqcap D) = reduce(D)$ if $reduce(C)$ is equivalent to $\top$ and $reduce(C \sqcap D) = reduce(C) \sqcap reduce(D)$ otherwise; similarly for $\sqcup$ and $\neg$. $reduce(\exists R.C) = \bot$ if $reduce(C) = \bot$, else $\exists R.reduce(C)$.

Definition 6. *The* semantic quantifier rank *srank is defined for each $\mathcal{ALC}$ concept $D = reduce(D')$ as follows:*

- $srank(D) = \infty$ *if D is equivalent to* $\bot$
- $srank(D) = 0$ *if D is equivalent to* $\top$
- $srank(D) = 0$ *if D is a concept symbol different from $\bot$ and* $\top$
- $srank(\neg D) = 0$ *if D is a concept symbol different from $\bot$ and* $\top$
- $srank(C \sqcap D) = \max(\{srank(C), srank(D)\})$
- $srank(C \sqcup D) = \min(\{srank(C), srank(D)\})$
- $srank(\exists R.D) = srank(D) + 1$
- $srank(\neg\exists R.D) = 0$
- $srank(\neg(C \sqcap D)) = \min(\{srank(\neg C), srank(\neg D)\})$
- $srank(\neg(C \sqcup D)) = \max(\{srank(\neg C), srank(\neg D)\})$

Note that the case of the universal quantifier is captured by its equivalent description using the existential quantifier: $\forall R.C = \neg\exists R.\neg C$ and thus $srank(\forall R.C) = 0$.

The following example illustrates the calculation of sranks.

Example 5. *Consider concepts $\exists R.C$ and $\exists R^2.C$ for atomic C:*

$$srank(\exists R.C) = 1, srank(\neg\exists R.C) = 0,$$

and

$$srank(\exists R^2.C) = 2, srank(\neg\exists R^2.C) = 0.$$

The definition can be interpreted as follows. Concept symbols can be represented in each model, for these it is not necessary to model roles. Having a conjunction, it is necessary to model both conjuncts, therefore, the conjunct with maximal srank must be considered. Having a union, it is sufficient to model the part with the smaller srank to have a union which is not bottom. The *srank* is increased when an existential is used, as it increases the depth of the path defined by the relations. A negated existential has a *srank* of zero, as there is no relation necessary to model it (only the non-existence of a relation needs to be modeled).

It is necessary to consider concepts and parts of concept D equivalent to $\bot$ and $\top$ separately. For a concept equivalent to $\top$ there exists an extension independent of the srank considered and an arbitrary srank can be used to model the extension. The srank of a concept equivalent to $\bot$ is ∞, as there can't be any concept extension (different from the point of origin) representing the $\bot$-concept. Therefore, concepts being equivalent to $\bot$ or $\top$ are ignored for creating the *srank*. Thus, having a concept for which the *srank* should be determined, it has to be transformed, e.g., by using de Morgan or distributivity to determine all parts representing the $\bot$-concept. After that, for the remaining terms the *srank* can be determined. One can show that the defined rank notion is indeed semantical in the sense that equivalent concepts (based on an empty TBox) have the same quantifier rank. Later on, this approach will be extended to ontologies restricted by a non-empty TBox.

Proposition 10. *If $C \sqsubseteq D$, then $srank(C) \geq srank(D)$.*

Proof. First, it is shown that the *srank* of a concept is independent of its syntactical form (except of subconcepts being equivalent to bottom and top as mentioned in Definition 6). The rules of de Morgan preserve the *srank*, as $srank(\neg(C \sqcap D)) = srank(\neg C \sqcup \neg D) = \min(\{srank(\neg C), srank(\neg D)\}) = srank(\neg C \sqcup \neg D)$. Distributivity is also preserved by *srank* as for example

$$\begin{aligned}
&srank((A \sqcap B) \sqcup C) \\
&= \min\big(\{\max(\{srank(A), srank(B)\}), srank(C)\}\big) \\
&= \max\big(\{\min(\{srank(A), srank(C)\}), \min(\{srank(B), srank(C)\})\}\big) \\
&= srank\big((A \sqcup B) \sqcap (B \sqcup C)\big).
\end{aligned}$$

This follows in the same way for the other cases of de Morgan and distributivity.

Let $C \sqsubseteq D$ and $srank(C) = i$. As $C \sqsubseteq D$, $D = C \sqcup D'$ for some D'. Thus $srank(D) = \min(\{srank(C), srank(D')\}) \leq srank(C) = i$. This can be done, as the *srank* of a concept is independent of its syntactical form as shown above. Thus, $srank(D) \leq srank(C)$. □

The following lemma shows the necessary properties mentioned in the motivation. Let a geometric model M_i be defined as an *i-srank*-model.

Lemma 3. *Model M_i and model M_{i+1} of an ontology $\mathcal{O}$ with an empty TBox have the following properties:*

1. *An algebraic atom C of M_i with $srank(C) = i$ is an algebraic atom of the unrestricted geometric model of $\mathcal{O}$.*
2. *M_i and M_{i+1} can be chosen such that M_i is a subspace of M_{i+1}.*

Proof.

1. Let C be an algebraic atom with $srank(C) = i$. and assume C would not be an algebraic atom of the unrestricted geometric model. Therefore, there exists a concept

D with $D \sqsubseteq C$ and $srank(D) = j > i$ and thus $C = C' \sqcup D$ for some C'. As $srank(D) = j$, it must contain at least one conjunct D' with $srank(D') = j$ and thus $srank(\neg D') = 0$. $C \sqcap \neg D' = (C' \sqcap \neg D') \sqcup (D \sqcap \neg D') \neq \bot$ and thus C is not an algebraic atom. A contradiction.

2. Let M_i be represented in some $\mathbb{R}^n$ where the algebraic atoms are interpreted as half-axis. Let $\mathcal{Y}_i$ be the set of algebraic atoms of M_i and $\mathcal{Y}_{i+1}$ the set of algebraic atoms of M_{i+1}. With 1. it follows that $\mathcal{Y}_i \subseteq \mathcal{Y}_{i+1}$. Algebraic atoms of $\mathcal{Y}_{i+1} \backslash \mathcal{Y}_i$ can thus be represented in some $\mathbb{R}^m$ independently of the $\mathbb{R}^n$ of M_i. Both spaces can be concatenated to a space $\mathbb{R}^{m+n}$ representing M_{i+1}.

□

Example 6 (Example 4 continued)**.** *Consider again the simple ontology with an empty TBox, a concept symbol A and a role symbol R. In the following, it is shown that the srank is suitable for the iterative creation of a geometric model. Consider a 0-srank-model. It is as for the 0-quantifier-rank model a one-dimensional space, but with the difference that not A and $\neg A$ are represented but the algebraic atoms of srank 0, thus, $A \sqcap \neg\exists R.\top$ and $\neg A \sqcap \neg\exists R.\top$ on $\mathbb{R}_+$ resp. $\mathbb{R}_-$. Thus only instances having no relation at all can be represented and instances of $(A \sqcap \exists R.A)^{\mathcal{I}}$ cannot be placed in this model. This circumvents the problem which appeared in Example 4 when extending the dimension. Here, the 1-srank-model consists of the algebraic atoms $A \sqcap \neg\exists R.\top$ and $\neg A \sqcap \neg\exists R.\top$ on the one hand and $X \sqcap \exists R.(Y \sqcap \neg\exists R.\top) \sqcap \neg\exists R.(\exists R.\top)$ for $X, Y \in \{A, \neg A\}$ on the other hand. As there is no algebraic atom in the 1-srank-model which is a proper subsumer of an algebraic atom in the 0-srank-model, it is possible to keep the representation of $(A \sqcap \neg\exists R.\top)^{\mathcal{I}}$ by extending it from $\mathbb{R}_+$ to $\mathbb{R}_+ \times \{0\} \times \{0\}$ (analogously for $(\neg A \sqcap \neg\exists R.\top)^{\mathcal{I}}$). The algebraic atoms of the form $X \sqcap \exists R.(Y \sqcap \neg\exists R.\top) \sqcap \neg\exists R.(\exists R.\top)$ for $X, Y \in \{A, \neg A\}$ can then be placed on the half-axes of the second and third dimension and thus not interfere with the lower dimensional model.*

Having this notion of a semantic rank, it is possible to create a *srank*-restricted model and extend it iteratively. However, this construction treats relations completely on a conceptual level in the sense that there is no geometric operation for the representation of the relation. To mitigate this problem, we observe that it should be possible to describe the changes in the semantic rank caused by a relation in a geometrical way. For example, if $x \in (\exists R^2.\top \sqcap \neg\exists R^3.\top)^{\mathcal{I}}$ and $R(x, y)$ is valid, applying an interpretation of R one time should result in a $y \in (\exists R.\top \sqcap \neg\exists R^2.\top)^{\mathcal{I}}$. Thus, a representation $R^{\mathcal{I}}$ of relation R is needed.

R is represented as an incidence matrix $\mathcal{R}$ that maps each half-axis to arbitrary many half-axes of the model. Interpreting R as incidence matrix allows for extending it iteratively while increasing dimensions.

Definition 7. *An* al-cone interpretation $\mathcal{I}$ *is a Boolean al-cone interpretation* $(\Delta, (\cdot)^{\mathcal{I}})$ *including additionally matrices $\mathcal{R}$ representing relations R. An* al-cone interpretation *of $\mathcal{ALC}$ concepts is defined recursively as for Boolean concepts and defining the concepts of the form $\exists R.C$ as al-cone, as $(\exists R.C)^{\mathcal{I}} = conHull(\{\mathcal{R}^T y \mid y \in C^{\mathcal{I}}\})$ with R interpreted as incidence matrix $\mathcal{R}$. The definition of the all quantifier is given by de Morgan, i.e., $(\forall R.C)^{\mathcal{I}} = (\neg\exists R.\neg C)^{\mathcal{I}}$.*

As $\mathcal{R}$ is an incidence matrix of size $\mathbb{R}^{2n \times 2n}$, where n is the size of the geometric model, the above mentioned definition contains a slight simplification. More concretely, it is based on the incidence vectors $y' \in \mathbb{R}^{2n}$ of $y \in C^{\mathcal{I}}$ where in even dimensions of y' the positive half-axes and in odd dimensions the negative half-axes are represented. Multiplication gives the incidence vector $x' = \mathcal{R}^T y'$ which needs to be transformed to al-cones by splitting it into half-axes. Therefore, $\exists R.C = \mathit{conHull}(\{x \mid \mathit{halfAxes}(x') \text{ with } x' = \mathcal{R}^T y' \text{ with } y' = \text{incidenceVector}(y) \text{ with } y \in C^{\mathcal{I}}\})$. For simplicity this is abbreviated with $(\exists R.C)^{\mathcal{I}} = \mathit{conHull}(\{\mathcal{R}^T y \mid y \in C^{\mathcal{I}}\})$ in the following.

Example 7. *Consider the example of Narcissus mentioned in Example 3. Assume an empty TBox, one concept* Vain *and one role* loves*, for short* V, R.

Let the al-cone interpretation be in $\mathbb{R}^2$, $V^{\mathcal{I}} = \mathbb{R}_+ \times \mathbb{R}_-$ *and*

$$\mathcal{R}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The incidence vector of instance $y^{\mathcal{I}} = [1, -1]^T \in V^{\mathcal{I}}$ *would be* $y' = [1, 0, 0, 1]^T$. $x' = \mathcal{R}^T y' = [1, 0, 1, 0]^T$ *and thus* $x^{\mathcal{I}} = [1, 1]^T$. *This can be split into the half-axes* $\mathbb{R}_+ \times \{0\}$ *and* $\{0\} \times \mathbb{R}_+$. *Thus, when considering all instances of* $V^{\mathcal{I}}$*, then it results in* $(\exists R.V)^{\mathcal{I}} = \mathit{conHull}(\{[1, 0]^T, [0, 1]^T\}) = \mathbb{R}_+ \times \mathbb{R}_+$.

$\mathcal{R}$ can be used to determine possible relations for a specific instance x, e.g., it can be checked, whether $(\mathcal{R}x^{\mathcal{I}}) \dot{\sqcap} C^{\mathcal{I}} \neq \bot^{\mathcal{I}}$, to state that $x^{\mathcal{I}} \in (\exists R.C)^{\mathcal{I}}$. As the concept $\exists R.C$ is created based on the instances of C, to answer such questions, the incidence matrix needs to be used in non-transposed form $\mathcal{R}$ here.

As for the Boolean case, it is also possible in the non-Boolean case to model partial knowledge, e.g., $x \in (\exists R.\top \sqcap (\exists S.\top \sqcup \neg\exists S.\top))^{\mathcal{I}}$ (where the second part does not add any information and is only added for demonstration purposes), meaning it is known that a relation with R exists, however, nothing is known about a relation with S. This can be modeled in the same way as in the Boolean case, by placing instance $x^{\mathcal{I}}$ in $(\exists R.\top)^{\mathcal{I}}$, yet between $(\exists S.\top)^{\mathcal{I}}$ and $(\neg\exists S.\top)^{\mathcal{I}}$. Reasoning about partial knowledge works as follows: Consider the question whether $x^{\mathcal{I}} \in (\exists S.\top)^{\mathcal{I}}$, $\mathcal{S}x^{\mathcal{I}}$ would not result in $\bot^{\mathcal{I}}$, but this does not mean that $x^{\mathcal{I}} \in (\exists S.\top)^{\mathcal{I}}$ is necessarily valid. Therefore, as a second step, it is necessary to check the other way around, namely, to check whether $x^{\mathcal{I}} \in \mathit{conHull}(\{\mathcal{S}^T c \mid c \in \top^{\mathcal{I}}\}))$. The combination of the two steps results in a determining whether x is a positive instance of a concept, or a negative instance or whether only partial knowledge is given. Thus, x has possibly but not necessarily a relation $\mathcal{S}$.

In fact, representation $\mathcal{R}$ of role R behaves as desired with existential quantifiers (and hence also with universal) quantifiers: it makes $\exists R$ indeed to a normal, additive operator f in a Boolean algebra with operators, i.e. f fulfills $f(\underline{0}) = \underline{0}$ (where $\underline{0}$ is the smallest element in the algebra; see Section 4.2.1) and $f(a \vee b) = f(a) \vee f(b)$.

Proposition 11.

1. *Given an al-cone interpretation* $\mathcal{I}$ *and an arbitrary* $\mathcal{R}$*, then* $(\exists R.C)^{\mathcal{I}}$ *yields for arbitrary al-cones* $C^{\mathcal{I}}$ *an al-cone.*

2. $\exists R.X$ *interpreted as* $\mathcal{R}^T X^{\mathcal{I}}$ *maps the cone representing the bottom concept onto itself:* $(\exists R.\bot)^{\mathcal{I}} = \bot^{\mathcal{I}}$ *and distributes over cone disjunction:* $\mathcal{R}^T(X^{\mathcal{I}} \dot{\sqcup} Y^{\mathcal{I}}) = \mathcal{R}^T X^{\mathcal{I}} \dot{\sqcup} \mathcal{R}^T Y^{\mathcal{I}}$, *i.e.,* $\exists R.(C \sqcup D) = \exists R.C \sqcup \exists R.D$.

Proof.

1. The incidence matrix $\mathcal{R}$ only consists of ones and zeros, i.e., $(\mathcal{R})_{i,j} \in \{0, 1\}$. Therefore, multiplication with arbitrary half-axes according to the aforementioned calculation procedure results in a convex hull over half-axes, i.e., an al-cone.

2. $(\exists R.\bot)^{\mathcal{I}} = \bot^{\mathcal{I}}$, as $(\exists R.\bot)^{\mathcal{I}} = \mathit{conHull}(\{\mathcal{R}^T y \mid y \in \{\vec{0}\}\} = \{\vec{0}\} = \bot^{\mathcal{I}}$.

   Due to Proposition 3 for $(C \sqcup D)^{\mathcal{I}}$ there does not exist a half-axis (thus an al-cone) which is not contained in either $C^{\mathcal{I}}$ or $D^{\mathcal{I}}$, yet contained in $(C \sqcup D)^{\mathcal{I}}$. Having this insight and the above mentioned calculation rules for $(\exists R.C)^{\mathcal{I}}$, it follows that

$$\begin{aligned}
(\exists R.(C \sqcup D))^{\mathcal{I}} &= \mathit{conHull}(\{\mathcal{R}^T y \mid y \in (C \sqcup D)^{\mathcal{I}}\} \\
&= \mathit{conHull}(\{\mathcal{R}^T y \mid y \in \mathit{halfAxes}(C \sqcup D)^{\mathcal{I}})\} \\
&= \mathit{conHull}(\{\mathcal{R}^T y, \mathcal{R}^T z \mid y \in \mathit{halfAxes}(C^{\mathcal{I}}), z \in \mathit{halfAxes}(D^{\mathcal{I}})\} \\
&= \mathit{conHull}(\{\mathit{conHull}(\{\mathcal{R}^T y \mid y \in C^{\mathcal{I}}\}), \mathit{conHull}(\{\mathcal{R}^T z \mid z \in D^{\mathcal{I}}\})\}) \\
&= (\exists R.C \sqcup \exists R.D)^{\mathcal{I}}
\end{aligned}$$

□

Now, our first aim is to find a satisfiable geometric model of a given ontology based on the interpretation of relations as incidences matrices on a conceptual level, meaning, if $\mathcal{O} \models C(b)$, then $b^{\mathcal{I}} \in C^{\mathcal{I}}$ for each concept C and each constant b and, if $\mathcal{O} \models R(a, b)$ and $\mathcal{O} \models C(b)$, then $a^{\mathcal{I}} \in (\exists R.C)^{\mathcal{I}}$. In the first step, the focus does not lie on the correct representation of roles, i.e., guaranteeing that if $\mathcal{O} \models R(a, b)$, then $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$. Thus, "role-satisfiability" is considered, i.e. the mapping into correct regions in the sense that if $\mathcal{O} \models R(a, b)$, then $b^{\mathcal{I}} \in \mathcal{R}a^{\mathcal{I}}$.

Proposition 12. *$\mathcal{ALC}$-ontologies are classically satisfiable iff they are satisfiable by a concept faithful geometric model on some (possibly infinite)* $\mathbb{R}^n$ *using sets of the form* $b_1 \times \cdots \times b_n$ *with* $b_i \in \{\{0\}, \mathbb{R}_+, \mathbb{R}_-, \mathbb{R}\}$ *and incidence matrices in* $\mathbb{R}^{2n \times 2n}$.

In order to prove this theorem, we reduce it to the following subproblems. First, concept-faithfulness is proven for an empty TBox with an acyclic ABox. After that, it is extended to an arbitrary TBox with an acyclic ABox. At the bottom line this leads to the case of an arbitrary TBox and a cyclic ABox mentioned in Proposition 12 and thus to the conclusion that it is possible to represent arbitrary $\mathcal{ALC}$-ontologies using al-cones.

Thus, the algebraic atom that a half-axis represents should be determined by the applicability of relation R, thus by exploring the possible paths starting on this half axis by applying the incidence matrix R. The relation matrix is as the geometric model dependent of the *srank* and iteratively extendable.

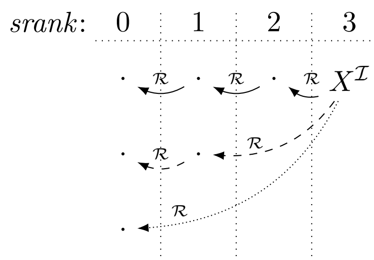


Figure 12: Visualization of the assessment of concepts to the al-cone $X^{\mathcal{I}}$. Starting with X, the arrows depict the application of the relation R (thus incidence matrix $\mathcal{R}$). The numbers indicate the *srank* of the resulting concepts, thus, relation R can be applied until an *srank* of 0 is reached.

Example 8. *An example can be seen in Figure 12. Assume only one relation symbol R and no concept symbols to be given. There are three different connections starting at the half axis $X^{\mathcal{I}}$ with R, the dashed, dotted and solid one. Following the possible paths leads to the fact that $X \sqsubseteq \exists R^3.\neg\exists R.\top \sqcap \exists R^2.\neg\exists R.\top \sqcap \exists R.\neg\exists R.\top$ and additionally to the fact that all other relations are not possible, thus $X \sqsubseteq \neg\exists R^4.\top$, therefore, $X = \exists R^3.\neg\exists R.\top \sqcap \exists R^2.\neg\exists R.\top \sqcap \exists R.\neg\exists R.\top \sqcap \neg\exists R^4.\top$.*

Additionally, it is necessary to define an underlying propositional model to determine the intersection of a concept with the propositional concepts.

This intuition can be represented as incidence matrix.

Definition 8. *Given relations $R, S, T, ...$, the* incidence matrices $\mathcal{R}^*, \mathcal{S}^*, ...$ of $R, S, T, \ldots$ *are defined as follows:*

$$\mathcal{R}^* = \begin{bmatrix} \mathcal{R}'_{0\to 0} & \mathcal{R}'_{1\to 0} & \mathcal{R}'_{2\to 0} & \mathcal{R}'_{3\to 0} & \cdots \\ 0 & 0 & \mathcal{R}'_{2\to 1} & \mathcal{R}'_{3\to 1} & \cdots \\ 0 & 0 & 0 & \mathcal{R}'_{3\to 2} & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{1}$$

with submatrices $\mathcal{R}'_{i\to j}$, where for all $i, j, k \in \mathbb{N}$ with $j < i$ and $k < j$ there exists $m, n, o \in \mathbb{N}$ such that $\mathcal{R}'_{i\to j} \in \mathbb{R}^{m\times n}$ and $\mathcal{R}'_{j\to k} \in \mathbb{R}^{o\times m}$, $\mathcal{R}'_{0\to 0}$ is a zero matrix and the submatrices have a fixed size dependent of the number of relations and concepts given.

Intuitively, the submatrix $\mathcal{R}'_{i\to j}$ depicts the application of relation R to find relations between instances of concepts with *srank* i and *srank* j and thus, submatrix $\mathcal{R}'_{i\to j}$ influences the region of the geometric model containing algebraic atoms of *srank* i.

Example 9. *Consider an empty TBox, with no concept symbol and one relation symbol R. First, it is necessary to determine the number of algebraic atoms added to a i-srank-model when considering an $i+1$-srank-model. Whereas this is straightforward for $i \in \{0, 1\}$ with $(\neg\exists R.\top)^{\mathcal{I}}$ for the 0-srank-model and $\exists R.\neg\exists R.\top \sqcap \neg\exists R^2.\top$ for the 1-srank-model, there are more algebraic atoms for the 2-srank-model: $\exists R^2.(\neg\exists R.\top) \sqcap \exists R.(\neg\exists R.\top) \sqcap \neg\exists R^3.\top$ and*

$\exists R^2.(\neg\exists R.\top) \sqcap \neg\exists R.(\neg\exists R.\top) \sqcap \neg\exists R^3.\top$. *Based on this, the following incidence matrix* $\mathcal{R}^*$ *can be created:*

$$\mathcal{R}^* = \begin{bmatrix} \begin{bmatrix}0\end{bmatrix} & \begin{bmatrix}1\end{bmatrix} & \begin{bmatrix}1 & 0\end{bmatrix} & \dots \\ 0 & 0 & \begin{bmatrix}1 & 1\end{bmatrix} & \dots \\ 0 & 0 & 0\ \ 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

where the submatrices (in brackets) in the first row are from left to right $\mathcal{R}'_{0\to0}, \mathcal{R}'_{1\to0}$ *and* $\mathcal{R}'_{2\to0}$ *and in the second row* $\mathcal{R}'_{2\to1}$. *So, how was this matrix created and how can it be used for determining the concept of some incidence vector x? In the incidence matrix presented here, the relations up to a srank of 2 are depicted. Thus, all algebraic atoms mentioned above need to be representable by the incidence matrix. In the following, it is shown for the half-axes of the geometric model which algebraic atom they depict (where $x^\mathcal{I}, \dots, z^\mathcal{I}$ represent incidence vectors of half-axes). Considering $w^\mathcal{I} = [1\ 0\ \dots]^T$: as $\mathcal{R}^* w^\mathcal{I} = \{\vec{0}\}$, there is no relation starting in w and thus $w \in \neg\exists R.\top$. For $x^\mathcal{I} = [0\ 1\ 0\ \dots]^T$, $\mathcal{R}^* x^\mathcal{I} = [1\ 0\ \dots]^T = w^\mathcal{I}$. Thus, for x there exists a relation R ending in w, such ending in a region where no relation is possible anymore. Thus, $x^\mathcal{I} \in \exists R.(\neg\exists R.\top) \sqcap \neg\exists R^2.\top$. For $y^\mathcal{I} = [0\ 0\ 1\ 0\ \dots]^T$, $\mathcal{R}^* y^\mathcal{I} = \{[0\ 1\ 0\ \dots]^T = x^\mathcal{I}, [1\ 0\ 0\ \dots]^T = w^\mathcal{I}\}$. Hence it is on the one hand possible to apply the relation two times without reaching $\bot$ (from $y^\mathcal{I}$ over $x^\mathcal{I}$ to $w^\mathcal{I}$ and on the other hand to directly reach $w^\mathcal{I}$, i.e. applying the relation only one time. Thus $y^\mathcal{I} \in \exists R.\exists R.(\neg\exists R.\top) \sqcap \exists R.(\neg\exists R.\top) \sqcap \neg\exists R^3.\top$. For $z^\mathcal{I} = [0\ 0\ 0\ 1\ 0\ \dots]^T$, this is not possible, as it is mapped to $x^\mathcal{I}$, however not to $w^\mathcal{I}$ directly. Thus, $z^\mathcal{I} \in \exists R.\exists R.(\neg\exists R.\top) \sqcap \neg\exists R.(\neg\exists R.\top) \sqcap \neg\exists R^3.\top$.*

In each dimension k, the two half-axes of the geometric model represent one algebraic atom each. Interpreting it as incidence vector leads to the case that the positive half-axis in dimension k is represented by a 1 in the incidence vector at position $2k$. The algebraic atom can be determined by considering the matrices $\mathcal{R}^*, \mathcal{S}^*, \dots$ at column $2k$. Thus, to represent all concepts of a specific *srank*, in each submatrix containing the submatrices $\mathcal{R}'_{i\to j}$, where $j = \{0, \dots, i-1\}$, each possible concept of *srank* i has to be created. Thus, before stating the main proposition of reaching concept faithfulness, first, the form and creation of *srank*-algebraic atoms is considered.

Lemma 4. *Let $N_R \cup N_C \cup N_c$ be the signature of the ontology under consideration. For a srank of 0, the set*

$$\mathcal{X}_0 = \left\{ A \sqcap \left(\bigsqcap_{R \in N_R} \neg\exists R.\top \right) \middle| A \in \mathcal{M}_0 \right\},$$

depicts all algebraic atoms of srank 0. Here $\mathcal{M}_0$ depicts the set of algebraic atoms given based only on the concept symbols.

For a given srank $i > 0$, the set of algebraic atoms of srank i can be created based on the algebraic atoms of srank $0, \ldots, i-1$

$$\mathcal{X}_i = \left\{ A \sqcap \left(\bigsqcap_{R \in N_R} \bigsqcap_{B \in \mathcal{X}_{0,\ldots,i-1}} \varphi_{R,B} \right) \sqcap \left(\bigsqcap_{R_j \in N_R \text{ for } j=1,\ldots,i+1} \neg\exists R_1.(\exists R_2.(\ldots \exists R_{i+1}.\top)) \right) \middle| A \in \mathcal{M}_0, \varphi_{R,B} \in \{\exists R.B, \neg\exists R.B\} \right\}$$

for each $X \in \mathcal{X}_i$, $X \sqsubseteq \exists R.B$ for an $R \in N_R$ and a $B \in \mathcal{X}_{i-1}$.

Proof. Each concept $X \in \mathcal{X}_i$ has *srank* i, as $X \sqsubseteq \exists R.B$ for a $R \in N_R$ and $B \in \mathcal{X}_{i-1}$. It has to be shown that it is an algebraic atom. If this would not be the case then there has to be an algebraic atom X' for which $\bot \neq X' \sqcap X$ and $X' \sqsubset X$. Assume $srank(X') = j > i$, thus, $X' \sqsubseteq Y$ where $Y = \exists R_1.(\exists R_2.(\ldots \exists R_j.\top)$ for $R_1, ..., R_j \in N_R$, thus a chain of j positive existentials must exists. Because of the second part of the definition of $\mathcal{X}$, $X \sqcap Y = \bot$ and thus $X \sqcap X' = \bot$, a contradiction. Thus assume $srank(X') \leq i$: By definition, each conjunct containing positive existentials up to *srank* i is either positively or negatively contained in X (because of the first part of the definition of $\mathcal{X}_i$). Thus, it remains to show that there is no conjunct of the type $Z = \neg\exists R.C$ of X' for arbitrary $R \in N_R$ and arbitrary concept C such that $X \sqcap Z \neq \bot$ and $Z \neq X$. Because of the second part of the definition, all Z including one negative and then i or more positive existentials are already included. Thus, assume Z has one negative and $k < i$ positive existentials (with $k \geq 0$), followed by a negative one and arbitrary existentials afterwards. Then, $Z = \neg\exists R.B$, where $B \in \mathcal{X}_{k-1}$ and thus already positively or negatively contained in X, a contradiction. Thus, X is an algebraic atom of *srank* i. As all combinations of algebraic atoms of a lower *srank* are considered, $\mathcal{X}_i$ contains all algebraic atoms of *srank* i. □

To clarify the construction of the algebraic atoms a small example is given.

Example 10. *Consider an ontology with one concept symbol D and one role symbol R and an empty TBox. The algebraic atoms of the Boolean part of the ontology comprise $\mathcal{M}_0 = \{D, \neg D\}$. The construction of $\mathcal{X}_0$ results in $\mathcal{X}_0 = \{D \sqcap \neg\exists R.\top\}$ Then $\mathcal{X}_1$ is constructed as follows:*

$$\begin{aligned} \mathcal{X}_1 &= \left\{ A \sqcap \left(\bigsqcap_{B \in \mathcal{X}_0} \varphi_{R,B} \right) \sqcap \left(\neg\exists R^2.\top \right) \middle| A \in \{D, \neg D\}, \varphi_{R,B} \in \{\exists R.B, \neg\exists R.B\} \right\} \\ &= \{A \sqcap \exists R.(D \sqcap \neg\exists R.\top) \sqcap \exists R.(\neg D \sqcap \neg\exists R.\top) \sqcap \neg\exists R^2.\top, \\ &\quad A \sqcap \exists R.(D \sqcap \neg\exists R.\top) \sqcap \neg\exists R.(\neg D \sqcap \neg\exists R.\top) \sqcap \neg\exists R^2.\top, \\ &\quad A \sqcap \neg\exists R.(D \sqcap \neg\exists R.\top) \sqcap \exists R.(\neg D \sqcap \neg\exists R.\top) \sqcap \neg\exists R^2.\top \mid A \in \{D, \neg D\}\} \end{aligned}$$

The second part, denoting the negative existentials, can be interpreted as all concepts that are not accessible through the relation and, thus, not accessible to the incidence matrix.

Algebraic atoms are restricted so that there must be at least one B of $srank(B) = i-1$ with a positive existential due to the fact that algebraic atoms of *srank* i are considered. Without this restriction, $\mathcal{X}_i$ could contain algebraic atoms of a lower *srank*.

Based on Lemma 4 it is possible to prove the next proposition that states that interpreting the ontology based on the incidence matrices $\mathcal{R}^*$ is a suitable concept-faithful interpretation. The idea is to have an underlying geometric model which only contains information about propositional concepts and having the incidence matrices $\mathcal{R}^*, \mathcal{S}^*, \ldots$ to define the algebraic atoms based on roles.

Proposition 13. *Let $\mathcal{O}$ be an ontology with an empty TBox. Let $\mathcal{I}$ be a geometric interpretation of $\mathcal{O}$ constructed based on an interpretation $\mathcal{M}$ which is $\mathcal{M} = \mathcal{M}_0 \times \mathcal{M}_0 \times \ldots$, thus an infinite direct product of one 0-(quantifier)-rank-concept-faithful geometric interpretation $\mathcal{M}_0$. Furthermore, let each role R be interpreted as $\mathcal{R}^*$ as in Definition 8 with $\mathcal{R}'_{0\to 0}$ in $\mathbb{R}^{m\times m}$ and $m = 2|\mathcal{M}_0|$. Then: It is possible to construct $\mathcal{R}^*$ for each R in a way that $\mathcal{I}$ is concept-faithful (w.r.t. $\mathcal{O}$), and for each constants a, b and relation R, if $\mathcal{O} \models R(a, b)$, then there is a representation $b^{\mathcal{I}}$ and $a^{\mathcal{I}}$ so that $b^{\mathcal{I}} \in \mathcal{R}^* a^{\mathcal{I}}$, if $R(a, b)$ is not part of a cyclic dependency.*

Proof. First, it is shown that the construction is consistent, meaning no contradictions are induced. The construction of the 0-rank-faithful model is consistent as it consists of Boolean $\mathcal{ALC}$. As $\mathcal{R}^*$ only influences positive existential quantification, negation of an existential is defined via polarity. It is by definition not possible that a concept and its negation intersect. As shown in Proposition 11, the relation operator results in an al-cone and fulfills $(\exists R.\bot)^{\mathcal{I}} = \bot^{\mathcal{I}}$ and $\mathcal{R}^T(X^{\mathcal{I}} \dot{\sqcup} Y^{\mathcal{I}}) = \mathcal{R}^T X^{\mathcal{I}} \dot{\sqcup} \mathcal{R}^T Y^{\mathcal{I}}$.

To show concept-faithfulness, it is sufficient to show satisfiability, as the construction principle of Proposition 8 can be used similarly as in the Boolean case to construct a concept-faithful model. Thus, it is shown that all algebraic atoms induced by the TBox can be represented. Therefore, in the following, only half-axes (as representatives of algebraic atoms) are considered.

The proof is done based on induction over the *srank*. Therefore, it is shown that each algebraic atom of a specific *srank* can be represented and therefore, for the geometric interpretation of an arbitrary *srank*, faithfulness is reached. This is done by showing that the construction of $\mathcal{R}^*$, $\mathcal{S}^*, \ldots$ exactly leads to the atomic concepts represented in Lemma 4. First, it is shown that the subspace $\mathbb{R}^m$ of the first m dimensions of the geometric model is a 0-*srank*-model and incorporates all algebraic atoms of *srank* 0. The subspace contains $\mathcal{M}_0$, which is a 0-(quantifier)-rank-faithful geometric model. As $\mathcal{R}'_{0\to 0}$ equals zero for all relations R, each application of the relation operator (in form of $\mathcal{R}^* d$ for a half-axis d in $\mathbb{R}^m$) results in $\{\vec{0}\}$, thus is not possible. Thus, it exactly matches the definition of Lemma 4.

Next, assume that the underlying geometric model combined with the submatrix including all submatrices $\mathcal{R}'$ up to $\mathcal{R}'_{i-1\to i-2}$ represents all atomic concepts up to *srank* $i-1$. It is shown that the submatrices $\mathcal{R}'_{i\to j}$ for $j \in \{0, ..., i-1\}$ for each relation R represent in combination with the geometric model all algebraic atoms of *srank* i and do not lead to any inconsistencies. The submatrix

$$\mathcal{R}_i^* = \begin{bmatrix} \mathcal{R}'_{i\to 0} \\ \mathcal{R}'_{i\to 1} \\ \vdots \\ \mathcal{R}'_{i\to i-1} \\ 0 \\ \vdots \end{bmatrix} \tag{2}$$

for relation R combined with the respective submatrices for the other relations represents in each column one algebraic atom in form of a mapping to lower sranks. Consider, e.g., the incidence vector x' of an algebraic atom X which contains only one non-zero element. $\mathcal{R}^* x'$ can thus be reduced to a $r \cdot 1 = \mathcal{R}^* x'$ where r is one column of $\mathcal{R}^*$.

The exact content of each column r is not considered here, it is adequate to show that each possible algebraic atom could be represented in such a column as each column represents independent of the other one possible half-axis (algebraic atom). $\mathcal{X}_i$ as defined in Lemma 4 is considered. The first conjunct of each algebraic atom, the Boolean concept A is contained, as the geometric model consists of an infinite concatenation of the 0-rank geometric models. Therefore, also in the area influenced by the submatrix considered, there is an intersection with each of the Boolean concepts A possible. Next, it has to be shown that the submatrix $\mathcal{R}'_{i\to j}$ maps an algebraic atom of *srank* i really to one (ore more) algebraic atoms of *srank* j and only to them. This is satisfied by the construction principle of the matrices mentioned in Definition 8. It is not possible that the algebraic atom has a higher *srank*, as the space under the diagonal of $\mathcal{R}^*$ is not populated and therefore, only a reduction of the rank is possible and thus $(\bigsqcap_{R\in N_R} \bigsqcap_{C\in \mathcal{X}_{i+1,i+2,\dots}} \neg\exists R.C)$ is fulfilled. As $\mathcal{R}_i^*$ can contain arbitrary many ones in a column, it is possible to model an existential for arbitrary $B \in \mathcal{X}_{0,\dots,i-1}$ and each combination of relations, as $\mathcal{S}^*, \mathcal{T}^*, \dots$ also can influence the column considered. A column where all relation-matrices $\mathcal{R}'_{i\to i-1}$ have only zero entries is not allowed as it would interfere with the restriction of Lemma 4 for having at least one relation to an algebraic atom with *srank* $i-1$.

When each column-tuple of $\mathcal{R}^*, \mathcal{S}^*, \dots$ is unique, then each algebraic atom is on a unique half-axis. Therefore, for non-cyclic ABoxes, instances $a^\mathcal{I}$ and $b^\mathcal{I}$ have a unique al-cone each. Thus, when $R(a,b)$ is valid, then $b^\mathcal{I} \in R^*(a^\mathcal{I})$. □

6.2 Handling the Cyclic Case

Now we proceed with the problem of handling cyclicity. We observe that cyclicity cannot be represented correctly in the construction above. Assume $R(a,b)$ and $R(b,a)$ is given in the ABox. Then, applying $\mathcal{R}$ on a concept A with $a \in A$ would not reduce its *srank*, as afterwards it is still possible to apply $\mathcal{R}$ infinitely many times. Thus, it needs to be necessary to represent concepts with an infinite *srank*.

The set of possible algebraic atoms has been determined in Lemma 4. The incidence matrix $\mathcal{R}^*$ is created based on the idea of iteratively increasing the *srank* and modeling all possible atomic concepts of this *srank*. Now, it is necessary to consider concepts $X \in \mathcal{X}_\infty$, thus, cyclic dependencies. These concepts are not considered in such an iterative approach. To use the same construction principle of iterative extension it is necessary to define a new notion of rank which is a combination of *srank* and the cycle depth of the concepts. This rank then enables for iterative extension.

This is done based on the idea that each application of a relation is either part of a cycle or not. For each non-cyclic relation, the *srank* can be determined. Out of this, the maximum is chosen. For non-cyclic dependencies, the *srank* does not change. The motivation behind this is that having a cyclic dependency, it is also possible to have non-cyclic behavior in parts, e.g., Narcissus could love himself (a cycle) but could also love a person not loving anyone (a non-cyclic behavior). This has to be determined for each part of the cycle, as in the incidence matrix extended for cyclic dependencies $\mathcal{R}^{**}$, which will be considered in more detail in the proof of Proposition 14 below, it is necessary to place the concept representing a cycle at a *srank* where a mapping of the non-cyclic parts of the concept is possible. On the other hand, the maximum depth of a cycle can be considered. A conjunct with a *srank* of infinity needs to contain a cycle. This means that at some point, applying relation R leads to an algebraic atom which has been visited before. We introduce *cycle_depth* to represent the depth of this cycle. Regarding the Narcissus example, cycle depth would be one.

Definition 9. *The* cyclic semantic rank $srank_c$ $(srank'$,*cycle_depth*) *of a concept* C *given by its defining formula is a pair determined as follows:*

- *If* $C = \bot$, *then* $srank'(C) = srank(C) = \infty$ *and cycle_depth undefined.*
- *Let* C_t *be the (possibly infinite) computation tree that would unfold when computing* $srank(C)$, *i.e.,* $srank(C) = srank(\mathrm{root}(C_t))$.
- *If* $srank(C) = r < \infty$, *then* $srank_c = (r, 0)$;
- *else transform* C_t *to* C'_t *by copying* C_t *and replacing any node* $N \in C_t$ *if a node representing the same concept term as* N *occurs in the subtree rooted at* N. N *is replaced by a new concept symbol* N', *which yields* $srank(N') = 0$ *according to Definition 6; the subtree rooted at* N *is removed. This makes* $srank' = srank(C'_t)$ *evaluate to* $r < \infty$. *Let cycle_depth then be the minimum number of occurrences of* $\exists$ *between re-appearances of* N *in the subtree in* C_t, *maximized over all nodes* N *that were replaced.*

Intuitively, the definition extends *srank* to infinite trees by cutting off re-appearing nodes N and recording the maximum length of cycles cut off.

Example 11. *Consider again an empty TBox and an ontology with a role symbol* R *and a concept symbol* A *as arising in Example 3 (narcis-vain). Consider* $srank_c$ *of concept* C *given by the infinite term* $C = A \sqcap \exists R.\neg A \sqcap \exists R.(A \sqcap \exists R.\neg A \sqcap \exists R.(\dots))$. *First, the computation tree* C_t *as shown in Fig. 13 is created. The calculation of the srank leads to* $srank(C) = \infty$, *as a cycle is included. As marked in the figure with* N, *the node* $\exists R.(A \sqcap \exists R.\neg A \sqcap \exists R.(\dots))$ *occurs several times and its first occurrence is thus replaced with a new symbol* N' *and is given a* $srank(N') = 0$. *Since* $srank(\exists R.\neg A) = 1$, *w have* $srank'(C) = 1$ *and cycle_depth*$(C) = 1$, *as both occurrences of* N *are connected via one existential role quantification, thus every step the origin of the cycle is reached again. Thus,* $srank_c(C) = (1, 1)$.

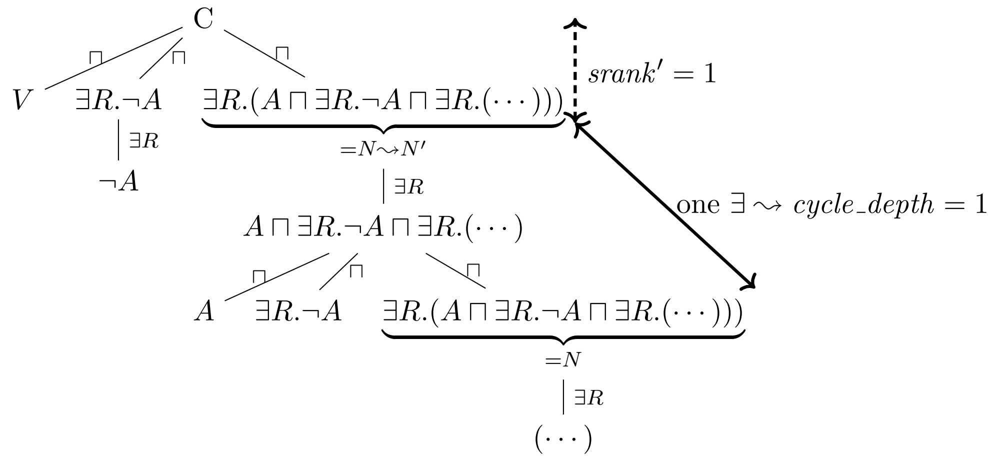


Figure 13: Computation tree C_t of the concept $C = A \sqcap \exists R.\neg A \sqcap \exists R.(A \sqcap \exists R.\neg A \sqcap \exists R.(\dots))$.

Lemma 5. *Let $N_R \cup N_C \cup N_c$ be the signature of the ontology under consideration. Then $\mathcal{X}_{i,0}$ as defined below depicts all algebraic atoms of $srank_c$ (i,0), and $\mathcal{X}_{0,1}$, which is defined by coinduction (Rutten, 2005) below, depicts all algebraic atoms of $srank_c$ (0,1).*

$$\mathcal{X}_{i,0} = \mathcal{X}_i$$

$$\mathcal{X}_{0,1} = \left\{ X^k_{0,1} \;\middle|\; \begin{array}{l} \textit{There is } A \in \mathcal{M}_0, \psi_{R,X^k_{0,1}} \in \{\exists R.(X^k_{0,1})', \neg\exists R.(X^k_{0,1})'\} \textit{ such that} \\ X^k_{0,1} = A \sqcap \left(\bigsqcap_{R \in N_R} \bigsqcap_{B \in \mathcal{X}_0, \mathcal{X}_1, \dots} \neg\exists R.B \right) \\ \sqcap \left(\bigsqcap_{R \in N_R} \psi_{R,(X^k_{0,1})'} \right) \sqcap \left(\bigsqcap_{R \in N_R} \neg\exists R.\neg(X^k_{0,1})' \right) \end{array} \right\}$$

where $\mathcal{M}_0$ depicts the set of algebraic atoms given based only on the concept symbols and for each $X^k_{0,1}$ at least one $\psi^k_{R,X_{0,1}}$ must appear positive. $(X^k_{0,1})'$ is the derivative of $X^k_{0,1}$ required for the coinductive definition.

Proof. Having a cycle-depth of zero, the construction reduces to the construction of $\mathcal{X}_i$, depicted in Lemma 4, thus $\mathcal{X}_{i,0} = \mathcal{X}_i$, as no cycles are contained. Now, consider $X^k_{0,1} \in \mathcal{X}_{0,1}$. $srank'(X^k_{0,1}) = 0$, as all concepts $B \in \mathcal{X}_0, \mathcal{X}_1, \dots$ are only reachable by negated relations. The cycle-depth of $X^k_{0,1}$ is 1, as at least one $\psi^k_{R,X_{0,1}}$ must appear positive, thus a cycle is included. It remains to show that $X^k_{0,1}$ actually is an algebraic atom. Assume this is not the case, thus an algebraic atom $Y \sqsubset X^k_{0,1}$ must exists such that $X^k_{0,1} \sqcap Y \neq \bot$. First, assume $srank'(Y) > 0$. A contradiction, as all concepts of a higher $srank'$ appear negated. Second, assume $srank_c(Y) = (0,0)$. This cannot be the case, as all existentials not part of a cyclic dependency are negated and at least one cycle appears positive in $X^k_{0,1}$ and would interfere with the assumption of a cycle depth of 0. Third, assume $srank_c(Y) = (0,1)$. However, as in $X^k_{0,1}$ all concepts with a cycle-depth greater 1 are negated and each possible cycle of length 1 either appears positive or negative, this is not possible. Fourth, assume $srank_c(Y) = (i, j)$ with $i \geq 0, j > 1$, thus there must exists a cycle of length at least two. However, as $X^k_{0,1}$ contains a conjunction with $\bigsqcap_{R \in N_R} \neg\exists R.\neg X^k_{0,1}$, it is not possible to have a cycle not ending in $X^k_{0,1}$ thus it is not possible to have a cycle depth greater one. Thus, $X^k_{0,1}$ is an algebraic atom with $srank_c(X^k_{0,1}) = (0, 1)$. □

The set of algebraic atoms $\mathcal{X}_{i,j}$ can be derived based on the above two but are omitted here for reason of readability.

Now, based on $srank_c$, an iterative creation of the geometric model is possible. For each of these tuples, a submatrix, thus, a specific region in the geometric model can be created to depict the concepts having this configuration. The matrix where all submatrices representing a cycle depth greater than zero (thus containing a cycle) are set to zero contains exactly the submatrices contained in $\mathcal{R}^*$ and has its behavior (except including more zero matrices). The basic idea for the submatrix $\mathcal{R}'_{i,j\to j}$ depicting a specific $srank_c$ i and a cycle-depth j is to place it at position $\mathcal{R}^{**}_{k\ldots l,k\ldots l}$ for some k and l in the new incidence matrix $\mathcal{R}^{**}$, thus, to enable to model connections trough relations between arbitrary columns of the matrix, thus, creating arbitrary circles. In the same column, it is also possible to place some $\mathcal{R}'_{i,j\to h}$ for $h < j$ and j is the *srank*' of $srank_c$ i.

Thus, Proposition 13 can be extended to a proposition covering cyclic ABoxes.

Proposition 14. *Let $\mathcal{O}$ be an ontology with an empty TBox. Let $\mathcal{I}$ be a geometric interpretation of $\mathcal{O}$ constructed based on an interpretation $\mathcal{M}$ which is $\mathcal{M} = \mathcal{M}_0 \times \mathcal{M}_0 \times \ldots$, thus an infinite direct product of one 0-(quantifier)-rank-concept-faithful geometric interpretation $\mathcal{M}_0$. Let further $\mathcal{I}$ interpret each role R as some incidence matrix. Then: $\mathcal{I}$ is concept faithful (w.r.t. $\mathcal{O}$), and for each constants a, b and relation R, if $\mathcal{O} \models R(a, b)$, then $b^{\mathcal{I}} \in \mathcal{R}^{**}a^{\mathcal{I}}$.*

In the proposition above we rely on the usual "direct product" operation on structures that is known in model theory and that can be applied to a possibly infinite number of input structures (Chang & Keisler, 1990, p. 224): the domain is the Cartesian product of the domains of the input structures and the interpretations of the non-logical symbols are given component-wise.

Proof. In the following, it is shown how concepts having a cycle depth greater zero and a *srank*' of zero and concepts having a *srank*' greater zero and a cycle-depth of one can be modeled in an incidence matrix $\mathcal{R}^{**}$. The proof of the other concepts follows analogously.

Assume a cycle depth of i and an *srank*' of zero. It is shown that a submatrix $\mathcal{R}'_{i,j\to j}$ of $\mathcal{R}^{**}_{k\ldots l,k\ldots l}$ can be chosen which models these dependencies. As the $srank'$ is zero, all other elements in columns $k, \ldots, l$ are zero. Thus, all possible combinations of conjunctions of zero $srank'$ concepts with relations have to be considered. These are only finitely many and could be represented in this submatrix.

The incidence matrix when considering only elements having a cycle depth of zero or one has the following form:

$$\mathcal{R}^{**} = \begin{bmatrix} \mathcal{R}'_{0,0\to 0} & 0 & \mathcal{R}'_{2,1\to 0} & \mathcal{R}'_{3,1\to 0} & \cdots \\ 0 & \mathcal{R}'_{1,1\to 1} & \mathcal{R}'_{2,1,1\to 0} & 0 & \cdots \\ 0 & 0 & \mathcal{R}'_{2,1\to 1} & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{3}$$

where the first column with submatrices depicts $srank_c$ of $(0,0)$, the second column with submatrices the $srank_c$ $(0,1)$, the third $(1,1)$ and the fourth $(0,1)$. $\mathcal{R}^{**}$ contains the submatrices of $\mathcal{R}^*$ as submatrices, as there are concepts without any cyclic dependency possible, which are represented by interaction with the submatrices of $\mathcal{R}^*$. The cyclic dependency is represented by the additional submatrices $\mathcal{R}'_{j,i\to i}$ which are on the diagonal and enable to cover arbitrary connections between algebraic atoms. To model cycles of length one, we have to ensure that when the submatrix is one at position $\{k,l\}$ it needs to be one at position $\{l,k\}$. As for each $srank_c$ there are only finitely many combinations, it is possible to model the incidence matrix iteratively in this way. Extension to more than one relation and a higher $srank_c$ is achieved analogously. □

6.3 Non-Empty TBoxes

As in general it cannot be assumed and is not suitable to have an empty TBox, in the following it is shown how the geometric model and the incidence matrix can be restricted to fulfill the TBox axioms.

The same iterative modeling approach from above is going to be used here. But now, the information of each TBox-axiom has to be incorporated, even if the model is restricted to an srank smaller than the srank of a part of the axiom. Therefore, it is necessary to know the relevance of specific TBox-axioms to the srank-models. It is not possible to use the definition of a *srank* for the empty TBox without change, as, e.g., having the axiom $A \sqsubseteq \exists R.C$, then the *srank* would be $srank(A) = 0$, $srank(\exists R.C) = 1$. However, A is known to be a subconcept of $\exists R.C$, this means, each for each $a \in A$, there must exists a $b \in C$ with $R(a,b)$. Therefore, it is not possible to place A in a 0-*srank*-model, as then, it would be possible that an element in A does not have any relation, what conflicts with the axiom. Having a concept, e.g., $\exists R.C$ with $srank(\exists R.C) = 3$, this influences also concepts based on the *srank* of this concept, e.g., $\exists R^2.C$, thus, this concept would have $srank(\exists R^2.C) = 4$, as applying a relation increases the rank by one.

To model this, an extension of the semantic rank accounting for TBoxes is defined. Therefore, the notion of *circular relationship* is introduced, stating that a concept is part of a cyclic dependency. This can be defined as a slight adaption of (Baader & Nutt, 2003, p. 56): For two atomic concepts A and B, A *directly uses* B if it appears on the right hand side of the definition of A. A concept is part of a cyclic dependency if it uses itself (where *uses* is the transitive closure of *directly uses*).

Definition 10. *The* TBox-specific semantic rank $srank_{\mathcal{T}}$ (TBox-srank for short) *is defined as follows:* $srank_{\mathcal{T}}(C) = srank(C) = \infty$ *if* C *is part of a circular relationship and else is* *defined based on the same defining rules as for the srank (see Definition 6), with one added rule:*

$$srank_{\mathcal{T}}(C) = max(\{srank(C)\} \cup \{srank(D) \mid \mathcal{T} \models C \sqsubseteq D\})$$

where the other rules are changed to

$$\begin{aligned} srank(C \sqcap D) &= \max(\{srank_{\mathcal{T}}(C), srank_{\mathcal{T}}(D)\}) \\ srank(C \sqcup D) &= \min(\{srank_{\mathcal{T}}(C), srank_{\mathcal{T}}(D)\}) \\ srank(\neg(C \sqcap D)) &= \min(\{srank_{\mathcal{T}}(\neg C), srank_{\mathcal{T}}(\neg D)\}) \\ srank(\neg(C \sqcup D)) &= \max(\{srank_{\mathcal{T}}(\neg C), srank_{\mathcal{T}}(\neg D)\}) \end{aligned}$$

We extend the arithmetics for ∞ in the usual way, setting $\infty + 1 = \infty$.

Thus, the calculation of $srank_{\mathcal{T}}$ is influenced on the one hand by the *srank* of the concept but on the other hand on its subsumption-relation to other concepts.

Lemma 6. *If a TBox contains a concept $\neg\exists R.C$ for arbitrary role R and concept C and $srank_{\mathcal{T}}(\neg\exists R.C) > 0$ or contains a concept D with $srank_{\mathcal{T}}(D) = \infty$, then the TBox contains a cyclic dependency.*

Proof. The lemma follows trivially for infinite *srank*. Assume a concept $\neg\exists R.C$ for arbitrary R and a propositional concept C and a TBox-axiom $\neg\exists R.C \sqsubseteq \exists S^i.D$ for arbitrary S, i and propositional concept D. Then, we get the equation $srank_{\mathcal{T}}(\neg\exists R.C) = i$. $\exists R.C$ has a $srank_{\mathcal{T}}(\exists R.C) = 1$, as C is a propositional concept. Thus, $\exists R.C$ can be rewritten to $\exists R.(C \sqcap \neg\exists R.\top \sqcap \neg\exists S.\top)$. This is a contradiction, as $\neg\exists R.\top \sqcap \neg\exists S.\top = \bot$. Therefore, an extension is necessary and thus $\exists R.C = \exists R.(C \sqcap \neg\exists R.\top \sqcap \exists S^i.D)$. This can be done in the same way for D and thus leads to an infinite extension of the concept and thus to an infinite $srank_{\mathcal{T}}$. Therefore, a cyclic dependency exists. This can trivially be adapted to general axioms. □

Example 12 (Example 5 continued)**.** *Consider a TBox with $\exists R.C \equiv \exists R^2.C$ (and therefore $\neg\exists R.\exists R.C \equiv \neg\exists R.C$). Based on the sranks calculated above*

$$srank(\exists R.C) = 1, srank(\neg\exists R.C) = 0, srank(\exists R^2.C) = 2, srank(\neg\exists R.\exists R.C) = 0,$$

it is possible to calculate the TBox-sranks:

$$srank_{\mathcal{T}}(\neg\exists R.\exists R.C) = \max(\{srank(\neg\exists R.\exists R.C), srank(\neg\exists R.C)\}) = 0$$

For the first axiom, the calculation is more complex:

$$srank_{\mathcal{T}}(\exists R.C) = \max(\{srank(\exists R.C)\} \cup \{srank(\exists R^2.C)\}) = \max(\{1\}, \{srank_{\mathcal{T}}(\exists R.C) + 1\})$$

Therefore, $srank_{\mathcal{T}}(\exists R.C) = srank_{\mathcal{T}}(\exists R^2.C) = \infty$.

We first focus on acyclic TBoxes and ABoxes, this means that each concept has a finite TBox-srank and each concept of the form $\neg\exists R.C$ for arbitrary R, C has an TBox-srank of 0. Then, the creation of a geometric model can be done as described in Definition 8 and Proposition 11, except that the axioms have to be considered. Thus, the first appearance of a concept is at its $srank_{\mathcal{T}}$, before that, it appears only negative. Therefore, the 0-quantifier-rank concept-faithful model for the representation of the 0-*srank*-model is only created with the propositional concepts which have a TBox-srank of 0 and is extended for each TBox-srank with the propositional concepts having this rank. Submatrices $R'_{i \to j}$ are not allowed to perform any mappings that contradict the axioms.

Definition 11. *A geometric interpretation for an arbitrary non-cyclic TBox and non-cyclic ABox is given as*

- *an incidence matrix $\mathcal{R}^*$ as defined in Definition 8 for each relation R;*
- *each $\mathcal{R}'_{i \to 0,\dots,i-1}$ is based on the part of the geometric model representing a product of arbitrary many 0-quantifier-rank geometric models of the propositional concepts having a TBox-srank of at least i.*

Proposition 15. *Let $\mathcal{O}$ be an ontology with an arbitrary non-cyclic TBox and non-cyclic ABox. Let $\mathcal{I}$ be a geometric interpretation of $\mathcal{O}$ as defined in Definition 11. Then: $\mathcal{I}$ is concept-faithful (w.r.t $\mathcal{O}$), and for each constants a, b and relation R, if $\mathcal{O} \models R(a, b)$, then there is a representation $b^{\mathcal{I}}$ and $a^{\mathcal{I}}$ so that $b^{\mathcal{I}} \in \mathcal{R}^* a^{\mathcal{I}}$, if $R(a, b)$ is not part of a cyclic dependency.*

Proof. As the TBox is acyclic, for each concept C, either $srank_{\mathcal{T}}(C) = 0$ or $srank_{\mathcal{T}}(\neg C) = 0$. Thus, it is possible to model a k-*srank*-model for arbitrary $k \geq 0$ without having both C and $\neg C$ to be contradictory. As a concept C is only different from $\bot^{\mathcal{I}}$ in a k-$srank_{\mathcal{T}}$-model, if $srank_{\mathcal{T}}(C) \leq k$, it has to be ensured that a concept does not appear at a lower *srank* and that only non-contradictory algebraic atoms are modeled. Non-contradiction is ensured for propositional concepts because of Definition 11.

As proven in Proposition 13, each algebraic atom can be represented. Based on this, only columns are considered which represent the desired algebraic atoms and thus, the restrictions are satisfied. □

This leads to the proof of Proposition 12 stating that it is possible to model a faithful geometric model for ontologies over full $\mathcal{ALC}$.

Proof of Proposition 12. Considering TBoxes with finite ranks, the proof follows directly from Proposition 15.

Now consider TBoxes with infinite ranks, e.g., a TBox containing the axiom $A = \exists R.A$. Then $srank_{\mathcal{T}}(A) = \infty$ and $srank_{\mathcal{T}}(\neg A) = \infty$ or a TBox containing a concept $\neg \exists R.C$ with TBox-srank greater than 0, as described in Lemma 6. Therefore, it is necessary to consider the construction introduced in Proposition 14. There, it is allowed to model cycles in the ABox, and therefore, infinite sranks. Based on this, the construction principle depicted in Propositions 14 and 15 can be used. □

The approach presented above enables us to create a faithful geometric model of a given ontology. It opens up the possibility to restrict the model to a given srank without affecting the expressivity of the given concepts. However, even for a restricted rank, the faithful geometric model could grow exponentially (depending on the TBox) and is therefore possibly not practical because of its size. Therefore, on the one hand, it is possible to extend the TBox with axioms which model a known bias in the data to circumvent that this bias is learned. This incorporates helpful information and reduces the size of the model. On the other hand, it is possible to focus on specific subparts of the model and model them faithfully. Thus, being able to model an ontology faithfully is not only helpful when full faithfulness is needed but also helpful as it is then known that each desired subproblem can be modeled correctly.

Example 13. *Consider again the example of Narcissus first mentioned in Example 3. Assume an empty TBox, one concept* Vain *and one role* loves*, for short V, R.*

It is possible to create a faithful representation, yet quite complex and of infinite dimension. Therefore, it is necessary to make suitable restrictions on the faithfulness. Here, the focus lies on reflexivity of Narcissus*. Therefore, reflexivity is to be modeled faithfully. This still results in an infinite model, however, of a simpler structure.*

One example for modeling is

$$\begin{aligned} V^{\mathcal{I}} &= \mathbb{R}_+ \times \mathbb{R}_+ \times \dots \\ (\neg V)^{\mathcal{I}} &= \mathbb{R}_- \times \mathbb{R}_- \times \dots . \end{aligned}$$

Then the relation R can be modeled as

$$\mathcal{R} = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & \dots \\ & & 1 & 0 & & & & & & & \\ & & 0 & 1 & & & & & & & \\ & & 0 & 0 & & & & & & & \\ & & \vdots & \vdots & & & & & & & \end{bmatrix},$$

where the empty regions contain zeros.

The geometric model is based on only one concept. Thus, each odd column of $\mathcal{R}$ represents a conjunction with $V^{\mathcal{I}}$, each even column a conjunction with $(\neg V)^{\mathcal{I}}$. The first two columns represent the area of the geometric model where no relation is possible. Therefore, a person Charlie represented by the incidence vector $c^{\mathcal{I}} = [1, 0, 0, \dots]^T$ would be vain (as $c^{\mathcal{I}} \in V^{\mathcal{I}}$) but would not love any person (as $\mathcal{R}c^{\mathcal{I}} = \{\vec{0}\} = \bot^{\mathcal{I}}$).

The fifth to the tenth column represent the different algebraic atoms of srank 1*. A person Bob represented by $b^{\mathcal{I}} = [0, 0, 0, 0, 1, 0, \dots]^T$ is vain, as $b^{\mathcal{I}} \in V^{\mathcal{I}}$. $\mathcal{R}b^{\mathcal{I}}$ leads to the incidence vectors $[1, 0, \dots]^T$ and $[0, 1, 0, \dots]^T$, thus, $b \in \exists R.(V \sqcap \neg\exists R.\top)$ and $b \in \exists R.(\neg V \sqcap \neg\exists R.\top)$. Thus, Bob is vain and loves one person being vain and one person not being vain, whereas both the beloved persons do not love anyone.*

The third and fourth column represent the reflexivity and thus the property relevant for expressing the narcissism of Narcissus. As Narcissus is vain, it has to be modeled in an odd dimension of the model. Thus, $n^{\mathcal{I}} = [0, 0, 1, 0, \dots]^T$. $\mathcal{R}n^{\mathcal{I}}$ leads to $[0, 0, 1, 0, \dots]^T$, thus it is mapped onto itself. Therefore, $n \in \exists R.(V \sqcap \exists R.(V \sqcap \dots))$ and reflexivity is modeled. Additionally, it leads to $[1, 0, \dots]^T$ and $n \in \exists R.(V \sqcap \neg\exists R.\top)$. Thus, Narcissus loves himself and he loves a vain person which does not love anyone.

7. Using Cone-based Embeddings for Learning

In this short section we sketch a general method of using cone-based embedding approaches for accomplishing classical learning tasks. First, a possible learning approach is outlined. After that, the advantages of a cone-based embedding over an embedding with TransE (and related approaches) are pointed out by considering a simple example. Development of a practical learning algorithm and its experimental evaluation are outside the scope of this paper.

Shortly after Özçep et al. (2020) presented the idea of cone-based geometric models for the first time, learning methods using cones appeared in NeurIPS papers (Zhang, Wang, Jiajun, Shuiwang, & Feng, 2021; Bai, Ying, Ren, & Leskovec, 2021). However, these are not aligned with logic methods. For an example of a multi-label learning scenario, but restricted to the case of Boolean $\mathcal{ALC}$ ontologies, i.e., neglecting roles, we point the interested reader to approaches by Leemhuis, Özçep, and Wolter (2020, 2022).

As a general learning scenario for cone models without roles, it is possible to consider multi-label learning: Given a set X of instances, with $(x, y) \in X$ where $x \in \mathbb{R}^n$ represents the features of instance (x, y) and y gives one or several labels assigned to the instance. The task is then to predict y for an (unseen) instance based on features x. It is possible to interpret label information as ABox (features get asserted to an instance). Additionally, it is possible that some background information in form of a TBox is given. Having this information, a mapping function can be learned which maps the instances from the feature space into the space of the geometric model. This could be done using a loss function which on the one hand forces the instances to lie in the cone corresponding to their ABox-assignments, and on the other hand creates a geometric model by forcing the cones to fulfill the TBox-axioms. A sufficient dimension to accommodate the geometric model would need to be chosen beforehand. When a restriction of the model to $\mathcal{ALC}$ is desired, then it is necessary to extend the loss function by incorporating either the restrictions mentioned in Proposition 5 or stronger restrictions enforcing the cones to be axis-aligned. In case of roles, a generalized multi-label learning scenario can be considered, where labels are not necessarily Boolean concepts but also role-based concepts or the ABox contains even role-assignments. Learning can then be accomplished by extending the loss function by terms such that for each role assignment $R(a, b)$, applying R to a would end up in b.

To point out advantages of a cone-based embedding vs an embedding with TransE, let us consider the following example:

Example 14. *An ontology $\mathcal{O}$ is given, with concept symbols M (male), F (female) and role R (has_parent). Let $\mathcal{T}$ be the TBox $\mathcal{T} = \{\exists R.M \equiv \top, \exists R.F \equiv \top, M \equiv \neg F\}$, thus each person has a mother and a father. ABox $\mathcal{A}$ states that Alice (a) has two parents, Bob (b) and Charlotte (c) and it is known that Bob has a parent called Robin (r) whose sex is unknown. $\mathcal{A} = \{R(a, b), R(a, c), R(b, r), M(b), F(c), F(a)\}$.*

Using TransE, the TBox-information cannot be handled (directly) and has to be ignored. A possible embedding can be seen in Figure 14 on the left. We manually constructed this embedding, yet it is a plausible outcome of learning, as the ABox-information is correctly expressed. It is not possible to model the concept information, even a relation is_of_type *cannot be modeled as Alice and Charlotte would have the same type. It is not possible to locate Alice, Charlotte, and their type (e.g., Human) such that Alice and Human as well as*

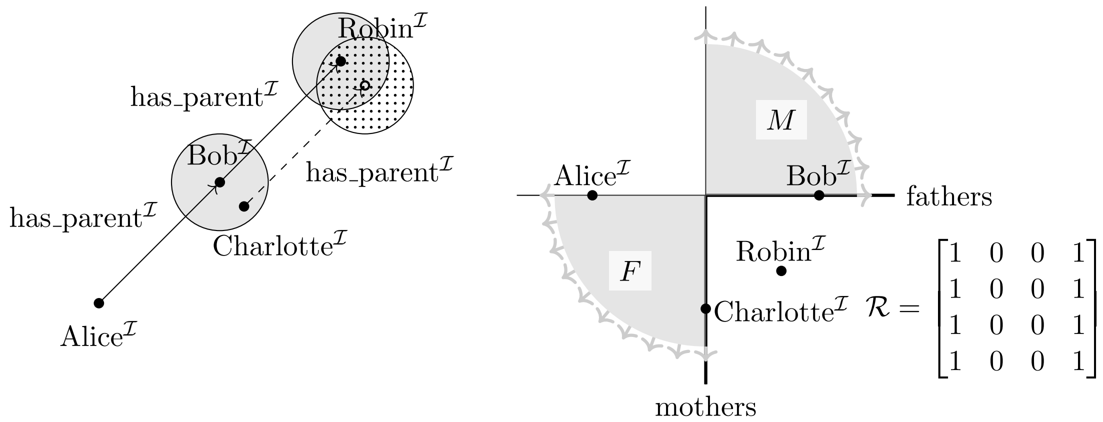


Figure 14: On the left, a possible embedding of the TBox mentioned in the text can be seen using TransE. The has_parent-relation is modeled as translation, the instances as points and the circle depicts the area of uncertainty, thus each instance in the circle is the endpoint of the relation. On the right, the respective al-cone embedding can be seen, where $\mathcal{R}$ depicts the incidence matrix for the relation *has_parent*.

Charlotte and Human are connected by the same translation. Neither is it possible to model that Alice has two (disjoint) parents (Bob, Charlotte) since translations are functional. In our example Bob and Charlotte are embedded to locations close to one another in order to apply some thresholding when interpreting relations (as is typically done): Bob and Charlotte are both close to the endpoint of the has_parent *translation. However, this leads to the unintended situation that Bob and Charlotte have (one ore more) parents in common, in this case Robin. Another drawback is that it is only possible to derive information for known instances. In the scenario, Robin would not have any parents at all, which contradicts the TBox.*

These drawbacks can be circumvented when using a cone-based embedding. A possible embedding can be seen in Figure 14 on the right. There, it is possible to place all instances in regions depending on their concepts. Robin can be modeled in between concepts male (M) and female (F) as the sex is unknown. The relation matrix $\mathcal{R}$ states that each instance of the model needs to have a father and a mother, even when they are not given in the ABox. Thus, for Robin it can also be represented that Robin has parents, albeit unknown ones. All TBox-axioms are modeled correctly and are thus respected. In a higher-dimensional model it would even be possible to exclude symmetric parent-relationships, taking full advantage of the expressivity of the TBox. Our cone model does not state concrete relations between the instances, in contrast to TransE. However, it can represent non-functional relationships, TBox-information, partial information and allows us to infer facts beyond what is covered by the ABox.

8. Conclusion and Outlook

Starting from an interpretation of negation as a polarity operator we presented embeddings of $\mathcal{ALC}$ ontologies that interpret all concepts as axis-aligned cones. This result adds an interesting alternative to embeddings considered so far as it advances the logical structure that can be captured by an embedding. In particular, the proposed approach is able to handle concept negation and disjunction. The model of axis-aligned cones (al-cones) investigated in this paper is shown to be universal in the sense that all embeddings of a disjunctive logic based on cones must employ, modulo simple geometric operations, axis-aligned cones.

As a side product of defining negation by polarity we obtain partial models, i.e. models with individuals for which one does not *know* whether they belong to a concept or not. This is different from approaches considered in classical embedding scenarios and can lead to interesting applications where one does not want a learnt model to perform certain generalizations. This is also different from partial models that have been investigated in the context of general logic programs, notably in the context of well-founded semantics (Van Gelder, Ross, & Schlipf, 1991). These differ from our partial models since those partial models are meant to treat $p \leftarrow \neg q$ and its contraposition $q \leftarrow \neg p$ differently, whereas in our case the contraposition rule (and double elimination) holds. In fact, as Van Gelder et al. (1991) note, there is a 3-valued logic interpretation for partial models in the well-founded semantics whereas for our partial models this kind of extensional semantics based on truth values is not possible: Consider the truth value of $p \vee q$. Assume that the ontology does not entail p (so p gets assigned the third truth value different from true and false) and q is not entailed (so q gets assigned the third truth value different from true and false). These two assignments do not determine the truth value of $p \vee q$: because the truth value of $p \vee q$ depends on whether the ontology entails the disjunction or not.

So a fine-grained treatment of the kind of uncertainty in our partial models must rely on intensional semantics and could proceed by considering *partial models* as done by Hartonas (2016) or by providing an epistemic operator $\Box$ (Donini, Lenzerini, Nardi, Nutt, & Schaerf, 1998, for example), i.e. a special modal logic operator where the accessibility relation expresses a kind of accessibility between epistemic states. Such an operator would allow us to distinguish between instances known to be in a specific concept C, i.e., those in $\Box C$ and instances which could be in the concept or its negation. How a translation could work for arbitrary orthologics (i.e. logics tailored towards ortholattices) is given in a small result of Goldblatt (1974).

Acknowledgment

The authors would like to thank the anonymous reviewers for their constructive and detailed feedback that helped us to improve the paper. This article is an extension of a paper by Özçep et al. (2020).

References

Baader, F. (2003). Description logic terminology. In Baader, F., Calvanese, D., McGuinness, D., Nardi, D., & Patel-Schneider, P. (Eds.), *The Description Logic Handbook*, pp. 485–495. Cambridge University Press.

Baader, F., Brandt, S., & Lutz, C. (2005). Pushing the el envelope. In *IJCAI'05: Proceedings of the 19th international joint conference on Artificial intelligence*, pp. 364–369, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Baader, F., & Küsters, R. (2006). Nonstandard inferences in description logics: The story so far. In Gabbay, D., Goncharov, S., & Zakharyaschev, M. (Eds.), *Mathematical Problems from Applied Logic I*, Vol. 4 of *International Mathematical Series*, pp. 1–75. Springer-Verlag.

Baader, F., & Nutt, W. (2003). Basic description logics. In Baader, F., Calvanese, D., McGuinness, D., Nardi, D., & Patel-Schneider, P. (Eds.), *The Description Logic Handbook*, pp. 43–95. Cambridge University Press.

Badreddine, S., d'Avila Garcez, A., Serafini, L., & Spranger, M. (2021). Logic tensor networks. *Artificial Intelligence*, *303*, 103649.

Bai, Y., Ying, R., Ren, H., & Leskovec, J. (2021). Modeling Heterogeneous Hierarchies with Relation-specific Hyperbolic Cones. In *Proc. 35th Annual Conference on Neural Information Processing Systems (NeurIPS 2021)*, p. arXiv:2110.14923.

Bordes, A., Usunier, N., García-Durán, A., Weston, J., & Yakhnenko, O. (2013). Translating embeddings for modeling multi-relational data. In Burges, C. J. C., Bottou, L., Ghahramani, Z., & Weinberger, K. Q. (Eds.), *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States.*, pp. 2787–2795.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, *2*(2), 121–167.

Calì, A., Gottlob, G., & Lukasiewicz, T. (2009). Datalog+-: A unified approach to ontologies and integrity constraints. In *Proceedings of the 12th International Conference on Database Theory*, pp. 14–30. ACM Press.

Chang, C., & Keisler, H. (1990). *Model Theory*. Studies in Logic and the Foundations of Mathematics. Elsevier Science.

Deng, J., Ding, N., Jia, Y., Frome, A., Murphy, K., Bengio, S., Li, Y., Neven, H., & Adam, H. (2014). Large-scale object classification using label relation graphs. In Fleet, D., Pajdla, T., Schiele, B., & Tuytelaars, T. (Eds.), *Computer Vision — ECCV 2014*, Vol. 8689 of *Lecture Notes in Computer Science*, pp. 48–64. Springer International Publishing.

Donini, F. M., Lenzerini, M., Nardi, D., Nutt, W., & Schaerf, A. (1998). An epistemic operator for description logics. *Artificial Intelligence*, *100*(1), 225–274.

Dunn, J. M. (1996). Generalized ortho-negation. In Wansing, H. (Ed.), *Negation—A Notion in Focus*, Perspektiven der Analytischen Philosophie / Perspectives in Analytical Philosophy 7, pp. 3–26. De Gruyter.

Farkas, J. (1902). Theorie der einfachen ungleichungen.. *Journal für die reine und angewandte Mathematik*, *124*, 1–27.

Gärdenfors, P. (2000). *Conceptual Spaces: The Geometry of Thought*. The MIT Press, Cambridge, Massachusetts.

Goldberg, Y., & Levy, O. (2014). word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method. *arXiv e-prints*, *0*(0), 0.

Goldblatt, R. I. (1974). Semantic analysis of orthologic. *Journal of Philosophical Logic*, *3*(1), 19–35.

Gutiérrez-Basulto, V., & Schockaert, S. (2018). From knowledge graph embedding to ontology embedding? An analysis of the compatibility between vector space representations and rules. In Thielscher, M., Toni, F., & Wolter, F. (Eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference, KR 2018, Tempe, Arizona, 30 October – 2 November 2018.*, pp. 379–388. AAAI Press.

Hartonas, C. (2016). Reasoning with incomplete information in generalized galois logics without distribution: The case of negation and modal operators. In Bimbó, K. (Ed.), *J. Michael Dunn on Information Based Logics*, pp. 279–312. Springer International Publishing, Cham.

Hohenecker, P., & Lukasiewicz, T. (2020). Ontology reasoning with deep neural networks. *J. Artif. Intell. Res.*, *68*, 503–540.

Ji, S., Pan, S., Cambria, E., Marttinen, P., & Yu, P. S. (2021). A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(2), 494–514.

Jonnson, B., & Tarski, A. (1952). Boolean algebras with operators. *American Journal of Mathematics*, *74*(1), 127–162.

Jonsson, B., & Tarski, A. (1951). Boolean algebras with operators. part i. *American Journal of Mathematics*, *73*(4), 891–939.

Kulmanov, M., Liu-Wei, W., Yan, Y., & Hoehndorf, R. (2019). 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)*.

Leemhuis, M., Özçep, Ö. L., & Wolter, D. (2020). Multi-label learning with a cone-based geometric model. In *Proceedings of the 25th International Conference on Conceptual Structures (ICCS 2020)*.

Leemhuis, M., Özçep, Ö. L., & Wolter, D. (2022). Learning with cone-based geometric models and orthologics. *Annals of Mathematics and Artificial Intelligence*, *90*, 1159–1195.

Levy, O., & Goldberg, Y. (2014). Neural word embedding as implicit matrix factorization. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D., & Weinberger, K. Q. (Eds.), *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*, pp. 2177–2185.

MacNeille, H. M. (1937). Partially ordered sets. *Trans. Amer. Math. Soc.*, *42*(3), 416–460.

Mehran Kazemi, S., & Poole, D. (2018). SimplE Embedding for Link Prediction in Knowledge Graphs. *arXiv e-prints*, *1*(1), 1. Available at https://ui.adsabs.harvard.edu/abs/2018arXiv180204868M.

Nickel, M., Tresp, V., & Kriegel, H.-P. (2011). A three-way model for collective learning on multi-relational data. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML'11, pp. 809–816, USA. Omnipress.

Özçep, Ö. L., Leemhuis, M., & Wolter, D. (2020). Cone semantics for logics with negation. In Bessiere, C. (Ed.), *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, pp. 1820–1826. ijcai.org.

Padmanabhan, R., & Rudeanu, S. (2008). *Axioms for Lattices and Boolean Algebras*. World Scientifc Press.

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation.. In *EMNLP*, Vol. 14, pp. 1532–1543.

Rutten, J. J. M. M. (2005). A coinductive calculus of streams. *Mathematical. Structures in Comp. Sci.*, *15*(1), 93–147.

Schmidt-Schauß, M., & Smolka, G. (1991). Attributive concept descriptions with complements. *Artificial Intelligence*, *48*, 1–26.

Serafini, L., & d'Avila Garcez, A. S. (2016). Learning and reasoning with logic tensor networks. In *Proceedings of the International Conference of the Italian Association for Artificial Intelligence (AI*IA)*, pp. 334–348.

Socher, R., Chen, D., Manning, C. D., & Ng, A. (2013). Reasoning with neural tensor networks for knowledge base completion. In *Advances in Neural Information Processing Systems (NIPS)*, pp. 926–934.

Tarski, A. (1935). Grundzüge des Systemenkalküls. Erster Teil. *In: Fundamenta Mathematicae*, *25,*, 503–526.

Van Gelder, A., Ross, K. A., & Schlipf, J. S. (1991). The well-founded semantics for general logic programs. *J. ACM*, *38*(3), 619–649.

Wang, Q., Mao, Z., Wang, B., & Guo, L. (2017). Knowledge graph embedding: A survey of approaches and applications. *IEEE Transactions on Knowledge and Data Engineering*, *29*(12), 2724–2743.

Zhang, Z., Wang, J., Jiajun, C., Shuiwang, J., & Feng, W. (2021). Cone: Cone embeddings for multi-hop reasoning over knowledge graphs. In *Advances in Neural Information Processing Systems*.