Logics for Reasoning with Comparative Distances

Mehmet Giritli[1]

Arbeitsgruppe Grundlagen der Künstlichen Intelligenz, Institut für Informatik, Albert-Ludwigs-Universität, Georges-Köhler-Allee 52, D-79110 Freiburg im Breisgau, Germany

Abstract

We study two logical formalisms of different expressive strengths for dealing with qualitative distance information. We utilise Theodore de Laguna's concept of 'can-connect' as a framework for handling this kind of information. This also allows us to seamlessly handle distance information using a more cognitively plausible, point-free ontological basis. The first-order logic we studied have very nice expressive capabilities and it is finitely axiomatisable, but as expected it is undecidable. We also explored a computationally more feasible alternative for qualitative reasoning about distances in the form of a modal logic formalism. Naturally, this logic is much less expressive than its first-order counterpart and it is decidable with an NP-complete satisfiability problem. We also provide a complete finite axiomatisation of the modal logic formalism.

Keywords: First-order logic, Modal logic, Knowledge representation, Spatial reasoning

1. Introduction

In this paper we deal with the revitalization of Theodore de Laguna's notion of 'can-connect'. Our purpose is to develop first-order and modal logic formalisms that have the ability to represent and reason with *comparative* distance information. Laguna's original idea appears in an article [1] which he regards as an appendix to his "revisit to the basic elements

Email address: mehmet@giritli.eu (Mehmet Giritli)

[1]Present address: Prestij Sit. No:32, G.A.U Yolu, Karaoglanoglu, Kyrenia, Cyprus. Tel: +90 533 8604364

of mathematical geometry from the window of actual human experience [2, 3]". This framework serves as a perfect basis for a qualitative formalism that can be used for reasoning about distances.

At the heart of Laguna's work lies his ontology built on the notion of 'solids'[2]. Especially from a philosophical point of view, the discussion regarding the choice of suitable ontologies for spatial formalisms played an important role in the field's research [4, 5, 6, 7, 8, 9]. As Simons [10] says, the problem is that, "nobody has ever perceived a 'point' or ever will do so, whereas people have perceived individuals of finite extent". In other words, to the researchers of the field, using an ontological basis like the one in question within formalisms that aim to represent and reason with our physical surrounding space means an attractive harmony between these formalisms and what they claim to be talking about. Unfortunately, choosing such an approach over point-based representations generally implies abandoning the comfort provided by using well-established standard mathematical models.

The intended semantics of the can-connect notion can be described as follows: A solid a can-connect two other solids, say b and c, whenever a can be moved into simultaneous contact with solids b and c, while all the solids a, b and c remain deformation-free during this process. By using this notion, very simple but effective distance measurements between spatial entities can be introduced in a very natural way: If a can-connect solids b and c, but it can not d and e, then this implies that solid b is nearer to c, than d is to e. From here, "equidistant objects" can also be easily defined.

This method of dealing with distances has four main advantages: Firstly, we are able to handle the distance information between non-atomic solids (instead of points) in a natural and cognitively plausible way. Secondly, we do not need to incorporate any numeric parameters or values into formalisms which are built around can-connect, allowing mathematically much simpler theories to be formed. Thirdly, one can easily compare two distances with each other within formalisms utilising can-connect; in other words, we can make statements of the form "the distance between a and b is greater than the distance between c and d". This is important be-

[2]There are various versions of this terminology used across qualitative spatial reasoning literature including 'individuals', 'regions' and 'volumes'. They have essentially the same meaning.

cause as the work of Wolter and Zakharyaschev on quantitative distance logics shows, without the use of a notion like can-connect, comparison of distances within formalisms is actually very difficult, if not impossible [11]. Finally, as our results on the modal logic formalism will show, computational feasibility of using this framework compared to other logical approaches for reasoning about distances is great.

Given the fundamental importance of distance data across various areas of computer science, there are already many studies which developed formalisms in various expressive levels to talk about distance information [12, 13, 14, 15, 16, 17, 11]. Even non-Euclidean distances are of interest in computer science: In the development of 'logics of similarity' in the field of approximate reasoning [18, 19], similarity measures are used to classify various sets of objects [20] and require reasoning in metric spaces that are non-Euclidean. In spatial reasoning, distance information is especially important as it deals with the physical space and distance information allows one of the most basic types (besides more abstract topological information) of relationships between spatial entities to be established. Moreover, with detailed distance information one can even represent and reason about the size and shape aspects of spatial entities [21].

However, the investigation of the theoretical properties of reasoning with distances came only recently with the "logics of metric spaces" [17] combining elements from modal and description logics to form formalisms that can reason with the standard mathematical model of distances, i.e., metric spaces. Results only show that, computationally most feasible logical fragments of these formalisms have an NEXPTIME upper bound. An extensive study along the same line of work combining topology and metric offers an EXPTIME-complete formalism [11]. Unfortunately these results draw a rather gloomy picture about the computational feasibility of reasoning about distance information and shows the need for a more qualitative approach accompanied with theoretical investigations of computational properties.

We construct two types of languages in which we embed can-connect relation: First, we embed the ternary can-connect relation into a first-order language and interpret it using standard metric spaces. We provide a finite axiomatisation of the resulting first-order logic and prove the corresponding representation theorems. Our axiomatisation provides 'mereology' as a subtheory, just like axiomatisations of topological 'connection' relation in the literature bring mereology as a subtheory [22]. Moreover, our first-order logic allows construction of new solids from the old ones, e.g., given a and b, the sum of a and b, the product of a and b and also their complements are parts of the theory as well. This naturally implies that one can make expressions of the form "while neither a nor b can-connect c and d, $a + b$ can-connect c and d". Unfortunately, we also establish that this logic has an undecidable satisfiability problem.

In the second part of the paper, we introduce a polyadic multi-modal language with the usual Kripkean semantics built around can-connect. This language mainly consists of a polyadic modality of the form $\langle \mathbf{CC} \rangle(\varphi, \psi)$, with the intended semantics that "here can connect somewhere φ and somewhere ψ". We introduce a "relational semantics version" of the metric spaces suitable to interpret can-connect primitive together with an equivalence result. Using relational semantics, finite model property and decidability results follow. Finally we establish that the satisfiability problem is NP-complete. In terms of logical properties, we provide a complete finite axiomatisation of this modal logic and prove the necessary representation theorems.

2. First-Order Comparative Distance Logic

The goal of this section is to develop a first-order logic which can talk about distance information in a qualitative and cognitively plausible manner, parallel to the main scheme of the paper. In this section we will also introduce the core semantic structure which we will be working with throughout the entire paper. The main characteristic of the structure in question is that it embeds the notion of 'individuals' (in contrast to points) inside metric spaces. Distance information between individuals will be handled with the help of Laguna's notion of can-connect. Unfortunately, one of the main results of this section will be that despite its simplicity, reasoning about qualitative distances via such structures in a first-order setting is computationally infeasible.

2.1. Language and Semantics

We begin by introducing the first-order language $\mathcal{L}^1$, which has the usual properties that can be expected from any first-order language. $\mathcal{L}^1$ contains denumerably many variable symbols, which we generally denote by $x, y, z, \ldots$ etc. and denumerably many constant symbols, which we generally denote by $c_1, c_2, c_3, \ldots$ etc. A term in the language $\mathcal{L}^1$ is either a variable or a constant.

Naturally, $\mathcal{L}^1$ contains the standard basic boolean operators $\vee, \neg$ and the constant verum $\top$ besides the first-order existential quantifier $\exists$. The operators of $\wedge, \rightarrow, \leftrightarrow, \bot$ and $\forall$ represent the usual duals and shorthands for the aforementioned basic operators. Finally and most importantly, $\mathcal{L}^1$ contains a non-logical, ternary primitive relation symbol CC.

Atomic formulas of $\mathcal{L}^1$ are expressions in the form $t_1 = t_2$ or $\mathsf{CC}(t_1, t_2, t_3)$, where t_1, t_2 and t_3 are terms. Arbitrary formulas of $\mathcal{L}^1$ are generated in the usual recursive manner using the basic operators of the language.

The language $\mathcal{L}^1$ is interpreted over models based on structures which consist of a metric space and a set of 'individuals'. Hence, we call them 'metric structures with individuals'. More precisely, we will be dealing with structures in the following form:

$$\mathfrak{F} = \left\langle \mathsf{W}, d, \mathbb{I} \right\rangle$$

where $\langle \mathsf{W}, d \rangle$ is a metric space and $\mathbb{I}$ is a set each member of which is called an 'individual' and satisfy the following constraints:

(**CNT**1) $\mathbb{I} \subseteq 2^{\mathsf{W}}$ and $\mathsf{W} \in \mathbb{I}$,

(**CNT**2) $\forall x \in \mathbb{I}[x \neq \emptyset]$,

(**CNT**3) $\forall p \in \mathsf{W}[\{p\} \in \mathbb{I}]$,

(**CNT**4) $\forall x \in \mathbb{I}[x \neq \mathsf{W} \Rightarrow \sim x \in \mathbb{I}]$,

(**CNT**5) $\forall x \in \mathbb{I} \forall y \in \mathbb{I}[x \cap y \neq \emptyset \Rightarrow x \cap y \in \mathbb{I}]$,

(**CNT**6) $\forall x \in \mathbb{I} \forall y \in \mathbb{I}[x \cup y \in \mathbb{I}]$.

Now, our models are pairs in the form:

$$\mathfrak{M} = \left\langle \mathfrak{F}, \mathsf{C} \right\rangle$$

where $\mathfrak{F}$ is a metric structure with individuals and C is a function interpreting the constants symbols of $\mathcal{L}^1$ as individuals from $\mathbb{I}$.

Let α, β be two formulas, t_1, t_2 be two terms and let $\mathfrak{a}$ be an assignment function mapping free occurring variables to the elements of $\mathbb{I}$. The interpretation of arbitrary $\mathcal{L}^1$ formulas is achieved in the usual inductive manner by defining a relation of truth $\models_{\mathfrak{a}}$ as follows:

- $\mathfrak{M} \models_{\mathfrak{a}} \top$,
- $\mathfrak{M} \models_{\mathfrak{a}} \alpha \wedge \beta$ iff $\mathfrak{M} \models_{\mathfrak{a}} \alpha$ and $\mathfrak{M} \models_{\mathfrak{a}} \beta$,
- $\mathfrak{M} \models_{\mathfrak{a}} \neg\alpha$ iff $\mathfrak{M} \not\models_{\mathfrak{a}} \alpha$,
- $\mathfrak{M} \models_{\mathfrak{a}} t_1 = t_2$ iff $\mathfrak{a}(t_1) = \mathfrak{a}(t_2)$,
- $\mathfrak{M} \models_{\mathfrak{a}} \mathsf{CC}(x, y, z)$ iff

$$\exists p_1 \exists p_2 \exists p_3 \exists p_4 \Big[p_1 \in \mathfrak{a}(y) \wedge p_2 \in \mathfrak{a}(z) \wedge p_3 \in \mathfrak{a}(x) \wedge p_4 \in \mathfrak{a}(x) \wedge d(p_1, p_2) \leq d(p_3, p_4) \Big],$$

- $\mathfrak{M} \models_{\mathfrak{a}} \exists x \alpha$ iff $\mathfrak{M} \models_{\mathfrak{b}} \alpha$ where $\mathfrak{b}$ is an assignment which differs from $\mathfrak{a}$, if at all, only on x.

The class of all metric models with individuals is denoted by M. As usual, validity (of a formula α) in every metric model with individuals under any assignment is denoted by writing $\mathsf{M} \models \alpha$.

2.2. Axiomatisation

Combining the axioms and inference rules for first-order logic with the axioms intended to capture the necessary properties of comparative distance logic, which are given below through **AXM**1 to **AXM**10, results with the formation of a proof system, its theory which we will denote by **Ax**CD_1 and denote its 'relation of proof' by $\vdash$. A proof in this proof system is a usual sequence of sentences of $\mathcal{L}^1$ such that each sentence is either an axiom of the system or derivable from the previous elements of the sequence using modus ponens or universal generalisation, in which case we write **Ax**$\text{CD}_1 \vdash \alpha$, where α is the formula proved. Short comments about definitions and axioms follows.

(**AXM**1) $\forall x \forall y [\mathsf{CC}(x, y, y)]$,

(**AXM**2) $\forall x \forall y \forall z [\mathsf{CC}(x, y, z) \rightarrow \mathsf{CC}(x, z, y)]$,

(**AXM**3) $\forall x \forall y \forall z \forall p \forall q [\mathsf{CC}(x, y, z) \wedge \neg\mathsf{CC}(x, p, q) \rightarrow \neg\exists r [\mathsf{CC}(r, p, q) \wedge \neg\mathsf{CC}(r, y, z)]]$,

(**DEF**) $\mathsf{I}(x, y) \equiv_{\text{def}} \forall z [\mathsf{CC}(z, x, y)]$, (**AXM**4) $\forall x \forall y[\forall z[\mathsf{I}(x,z) \leftrightarrow \mathsf{I}(y,z)] \rightarrow x = y]$,

(**AXM**5) $\exists x \forall y[\mathsf{I}(x,y)]$,

(**DEF**) $\mathsf{P}(x,y) \equiv_{\text{def}} \forall z[\mathsf{I}(x,z) \rightarrow \mathsf{I}(y,z)]$,

(**AXM**6) $\forall x \forall y \exists z \forall p[\mathsf{I}(z,p) \leftrightarrow [\mathsf{I}(x,p) \vee \mathsf{I}(y,p)]]$,

(**AXM**7) $\forall x \forall y[\mathsf{I}(x,y) \rightarrow \exists z \forall p[\mathsf{I}(z,p) \leftrightarrow \exists q[\mathsf{P}(q,x) \wedge \mathsf{P}(q,y) \wedge \mathsf{I}(p,q)]]]$,

(**AXM**8) $\forall x[\exists y \neg \mathsf{I}(x,y) \rightarrow \exists z \forall p[\mathsf{I}(z,p) \leftrightarrow \exists q[\neg \mathsf{I}(q,x) \wedge \mathsf{I}(q,p)]]]$,

(**DEF**) $\mathsf{A}(x) \equiv_{\text{def}} \forall y[\mathsf{P}(y,x) \rightarrow x = y]$,

(**AXM**9) $\forall x \exists y[\mathsf{A}(y) \wedge \mathsf{P}(y,x)]$,

(**DEF**) $\mathsf{AP}(x,y) \equiv_{\text{def}} \mathsf{A}(x) \wedge \mathsf{P}(x,y)$,

(**DEF**) $(x,y) \leq (z,p) \equiv_{\text{def}} \forall q[\mathsf{CC}(q,z,p) \rightarrow \mathsf{CC}(q,x,y)]$,

(**DEF**) $(x,y) = (z,p) \equiv_{\text{def}} (x,y) \leq (z,p) \wedge (z,p) \leq (x,y)$,

(**DEF**) $(x,y) < (z,p) \equiv_{\text{def}} (x,y) \leq (z,p) \wedge \neg(x,y) = (z,p)$,

(**AXM**10) $\forall x \forall y \forall z[\mathsf{CC}(x,y,z) \leftrightarrow \exists p \exists q \exists r \exists s[\mathsf{AP}(p,x) \wedge \mathsf{AP}(q,x) \wedge \mathsf{AP}(r,y) \wedge \mathsf{AP}(s,z) \wedge (r,s) \leq (p,q)]]$.

The first three axioms are very intuitive, they capture the essential properties of the notion of can-connect: Axiom **AXM**1 states that any entity can-connect any other entity with itself. With the axiom **AXM**2 symmetric nature of can-connect notion is captured: If an entity can-connect other two entities y and z, then it also can-connect z and y. **AXM**3 is the following property of can-connect: If an entity can-connect a pair of entities but it can-NOT-connect some other pair, then there could be no entity which can-connect the latter pair and yet can-NOT-connect the former pair. Axiom **AXM**4 is the identity axiom. It allows us to establish the identity of entities based on our primitive notion of can-connect.

The definition for the predicate $\mathsf{I}(x,y)$ captures the non-empty intersection of x and y, i.e., it is true whenever two solids share a common sub-solid. This is proved in Theorem 2.2. Based on the definition of the intersection predicate, we go ahead and define the parthood predicate $\mathsf{P}(x,y)$ from which all mereological relations can easily be defined with the help of the first-order expressiveness. A proof that parthood predicate actually corresponds to the subset relation is proved in Theorem 2.2. The other important definition worth mentioning is that of an atom. Intuitively, $\mathsf{A}(x)$ stands for 'x is a point'. An atom-solid is a solid which contains only itself as a part; it has no other parts. In other words, with our axiomatisation we rediscover points from solids.

Axioms **AXM**5, **AXM**6, **AXM**7 and **AXM**8 create new entities from the old ones with the help of the identity axiom. More precisely, **AXM**5 entails the existence of a unique universe which we shall denote by **U**. Given two entities x and y, while **AXM**6 entails the existence of a unique entity $x + y$, **AXM**7 entails the existence of a unique entity $x * y$, whenever we have $\mathsf{I}(x, y)$. Finally, **AXM**8 entails the existence of a unique entity $-x$, whenever we have $\exists y \neg \mathsf{I}(x, y)$.

Finally, axiom **AXM**9 states that every entity contains an atomic entity (entities whose only sub-part is itself) and axiom **AXM**10 manifests the interaction between atomic entities and can-connect primitive.

Many similar axiomatisations can be found in the literature of spatial logics. For example, Asher and Vieu present a successful axiomatisation of mereotopology ("geometry of common sense") [22]. However, there are quite a number of problematic first-order axiomatisation attempts with regard to their basic logical properties as well [23]. Such problems are often in the form of inconsistent axiom systems [5] or semantically incomplete systems [24]. There are even studies with the pursuit of achieving an absolutely complete (in contrast to semantic completeness which is what we deal with in this paper) first-order axiom systems [25], which is an impossible task given the likely undecidability of such logics [26] and the fact that every absolutely complete and recursively enumerable theory must be decidable.

The following lemma will be used in the forthcoming proofs:

Lemma 2.1. *The following formulas are theorems of* $\mathbf{Ax}CD_1$*:*

- $\neg[(c_1, c_2) \leq (c_3, c_4) \wedge (c_3, c_4) < (c_1, c_2)]$,
- $[(c_1, c_2) \leq (c_3, c_4) \wedge (c_3, c_4) = (c_5, c_6)] \rightarrow (c_1, c_2) \leq (c_5, c_6)$.

Proof. In order to see through the first claim, note that $(c_1, c_2) \leq (c_3, c_4) \wedge (c_3, c_4) < (c_1, c_2)$ is, by definition, equivalent to $(c_1, c_2) \leq (c_3, c_4) \wedge (c_3, c_4) \leq (c_1, c_2) \wedge \neg (c_3, c_4) = (c_1, c_2)$, which is again by definition equivalent to $(c_3, c_4) = (c_1, c_2) \wedge \neg (c_3, c_4) = (c_1, c_2)$, which is a contradiction.

For the second claim, note that by definition $(c_1, c_2) \leq (c_3, c_4) \wedge (c_3, c_4) = (c_5, c_6)$ implies that $(c_1, c_2) \leq (c_3, c_4) \wedge (c_3, c_4) \leq (c_5, c_6)$. This obviously implies that $(c_1, c_2) \leq (c_5, c_6)$. □

2.3. Soundness and Completeness Theorems

Establishing a semantical foundation for any kind of spatial logic is essential. The lack of such investigations in the study of several spatial logics has been subject of righteous criticism from within the field [23]. This is because of the fact that any spatial logic study which lacks necessary semantical investigation, bears the risk of not being able to capture the type of reasoning it promises to achieve. In other words, an established semantical foundation guarantees that a logic is able to represent and reason with the structures that are of interest.

In this section, we will establish that the first-order comparative distance logic is sound and semantically complete with respect to the class of all metric structures with individuals M. While the soundness has a completely standard proof, completeness proof employs a Henkin-style argument which consists of more interesting model construction procedures.

Theorem 2.2 (Soundness)**.** *Let φ be a formula. We have that* $\mathbf{Ax}CD_1 \vdash \varphi \Rightarrow \mathsf{M} \models \varphi$*.*

Proof. The proof is by induction on the complexity of φ. It is sufficient to establish the base case, which amounts to show that all of the axioms **AXM**1-**AXM**10 are valid on any metric model with individuals. So, first let,

$$\mathfrak{M} = \left\langle \mathfrak{F}, \mathsf{C} \right\rangle$$

be an arbitrary model where,

$$\mathfrak{F} = \left\langle \mathsf{W}, d, \mathbb{I} \right\rangle$$

is a metric structure with individuals.

First, let us establish the case of **AXM**1, i.e., that $\mathfrak{M} \models \forall x \forall y [\mathsf{CC}(x, y, y)]$. It is sufficient to show that for every $x \in \mathbb{I}$ and for every $y \in \mathbb{I}$ there are $p_1 \in x, p_2 \in x, p_3 \in y$ and $p_4 \in y$ such that $d(p_3, p_4) \leq d(p_1, p_2)$. Since $x \neq \emptyset \neq y$ from **CNT**2, we can simply pick some arbitrary $p_1 = p_2 \in x$ and $p_3 = p_4 \in y$ giving us $d(p_1, p_2) = d(p_3, p_4) = 0$, which is what we want.

Now we focus on the case of **AXM**2. So we have to show that $\mathfrak{M} \models \forall x \forall y \forall z[\mathsf{CC}(x, y, z) \rightarrow \mathsf{CC}(x, z, y)]$. We will proceed as follows: Suppose that for some x, y and z in $\mathbb{I}$, there are $p_1 \in x$, $p_2 \in x$, $p_3 \in y$ and $p_4 \in z$ such that $d(p_3, p_4) \leq d(p_1, p_2)$. But since d is symmetric, it follows that $d(p_4, p_3) \leq d(p_1, p_2)$, which gives us what we want.

Let us now consider axiom **AXM**3. This is slightly more complicated than the previous cases. Assume that for some x, y, z, p and q in $\mathbb{I}$, there are $p_1 \in x$, $p_2 \in x$, $p_3 \in y$ and $p_4 \in z$ such that $d(p_3, p_4) \leq d(p_1, p_2)$ and on the other hand, for every $p'_1 \in x$, $p'_2 \in x$, $p'_3 \in p$ and $p'_4 \in q$ we have that $d(p'_1, p'_2) < d(p'_3, p'_4)$.

Now, for the sake of a contradiction suppose that there is a $r \in \mathbb{I}$ such that there are $p_5, p_6 \in r$, $p_7 \in p$ and $p_8 \in q$ such that $d(p_7, p_8) \leq d(p_5, p_6)$ while for every $p'_1 \in r$, $p'_2 \in r$, $p'_3 \in y$ and $p'_4 \in z$, we have that $d(p'_1, p'_2) < d(p'_3, p'_4)$.

Combining the information we have so far, it easily follows that we have $d(p_5, p_6) < d(p_3, p_4)$ and $d(p_1, p_2) < d(p_7, p_8)$. On the other hand, since $d(p_3, p_4) \leq d(p_1, p_2)$, we conclude that $d(p_5, p_6) < d(p_7, p_8)$, which contradicts with the fact that we have $d(p_7, p_8) \leq d(p_5, p_6)$. This ends the case of axiom **AXM**3.

Before we continue any further, let us establish the fact that for any assignment $\mathfrak{a}$, we have that $\mathfrak{M} \models_{\mathfrak{a}} \mathsf{I}(x, y)$ iff $\mathfrak{a}(x) \cap \mathfrak{a}(y) \neq \emptyset$. To see this from left to right, assume that $\mathfrak{M} \models_{\mathfrak{a}} \forall z[\mathsf{CC}(z, x, y)]$. This means that, for every $z \in \mathbb{I}$, there are $p_1 \in z$, $p_2 \in z$, $p_3 \in \mathfrak{a}(x)$ and $p_4 \in \mathfrak{a}(y)$ such that $d(p_3, p_4) \leq d(p_1, p_2)$. Now suppose that $z = \{p_5\}$ for some $p_5 \in \mathsf{W}$. But this implies that $p_1 = p_2 = p_5$ and moreover that $d(p_1, p_2) = 0$. Hence, we must have $d(p_3, p_4) = 0$. So, $p_3 = p_4$. Thus, $\mathfrak{a}(x) \cap \mathfrak{a}(y) \neq \emptyset$. The opposite direction can be easily established by using a similar argument.

Similarly, we will show that for any assignment $\mathfrak{a}$, we have that $\mathfrak{M} \models_{\mathfrak{a}} \mathsf{P}(x, y)$ iff $\mathfrak{a}(x) \subseteq \mathfrak{a}(y)$. For the direction from left to right, assume that we have $\mathfrak{M} \models_{\mathfrak{a}} \mathsf{P}(x, y)$. Together with the previous paragraph this means that for every $z \in \mathbb{I}$, $\mathfrak{a}(x) \cap z \neq \emptyset \Rightarrow \mathfrak{a}(y) \cap z \neq \emptyset$. For the sake of a contradiction, suppose we have $\mathfrak{a}(x) \supset \mathfrak{a}(y)$. This implies that there is a $p_1 \in \mathfrak{a}(x)$ such that $p_1 \notin \mathfrak{a}(y)$. Since we have $\{p_1\} \in \mathbb{I}$ due to **CNT**3, it follows that we must also have $\mathfrak{a}(y) \cap \{p_1\} \neq \emptyset$. This is a contradiction.

For the opposite direction, suppose we $\mathfrak{a}(x) \subseteq \mathfrak{a}(y)$. We will again proceed with a contraposition argument. So assume that $\mathfrak{M} \models_{\mathfrak{a}} \exists z[\mathsf{I}(x, z) \wedge \neg \mathsf{I}(y, z)]$. This implies that for some assignment $\mathfrak{b}$ that differs from $\mathfrak{a}$ only on z, we have $\mathfrak{b}(x) \cap \mathfrak{b}(z) \neq \emptyset \wedge \mathfrak{b}(y) \cap \mathfrak{b}(z) = \emptyset$. Since $\mathfrak{b}(x) \subseteq \mathfrak{b}(y)$ by the assumption, it follows that $\mathfrak{b}(x) \cap \mathfrak{b}(z) \subseteq \mathfrak{b}(y)$. This implies that $\mathfrak{b}(y) \cap \mathfrak{b}(z) \neq \emptyset$, a contradiction.

Let us continue with the case of axiom **AXM**4. We will show that for every $x \in \mathbb{I}$ and $y \in \mathbb{I}$ we have that $\forall z[x \cap z \neq \emptyset \Rightarrow y \cap z \neq \emptyset] \Rightarrow x \subseteq y$. If $y = \mathsf{W}$ then we are through. So assume that $y \neq \mathsf{W}$. Then from **CNT**4, it follows that $\sim y \in \mathbb{I}$. Suppose that for every $x \in \mathbb{I}$ and $y \in \mathbb{I}$ we have that $\forall z[x \cap z \neq \emptyset \Rightarrow y \cap z \neq \emptyset]$ and for sake of a contradiction, assume that $x \nsubseteq y$. The latter implies that $x \cap \sim y \neq \emptyset$. But from the hypothesis, this implies that $y \cap \sim y \neq \emptyset$, which is a contradiction.

The case of **AXM**6 is trivial once we observe that for every $x \in \mathbb{I}$ and $y \in \mathbb{I}$ we have that $x \cup y \in \mathbb{I}$ from **CNT**6 and thus, we can always select this as an assignment for $x + y$. Now all that remains to be done is to show that for every $x \in \mathbb{I}$ and $y \in \mathbb{I}$ we have that $\forall z[(x \cup y) \cap z \neq \emptyset \Leftrightarrow [(x \cap z) \neq \emptyset \vee (y \cup z) \neq \emptyset]]$, which is a well known fact itself. A similar proof for the validity of axioms **AXM**7 and **AXM**8 can be easily generated by using the constraints **CNT**5 and **CNT**4, respectively. The case of axiom **AXM**5 is absolutely trivial since $\mathsf{W} \in \mathbb{I}$.

Now, first notice that we have $\mathfrak{M} \models_{\mathfrak{a}} \mathsf{A}(x)$ iff $\mathfrak{a}(x)$ is a singleton. Now, to see the case of axiom **AXM**9, let $x \in \mathbb{I}$. From **CNT**2, it follows that $x \neq \emptyset$. Pick $p_1 \in x$. Now from **CNT**3, it follows that $\{p_1\} \in \mathbb{I}$. Hence, we have found a singleton $\{p_1\}$ such that $\{p_1\} \subseteq x$.

Before we move into the case of **AXM**10, assume that $\mathfrak{M} \models_{\mathfrak{a}} (x, y) \leq (z, p)$. By definition, we get that $\mathfrak{M} \models_{\mathfrak{a}} \forall q[\mathsf{CC}(q, z, p) \rightarrow \mathsf{CC}(q, x, y)]$. We will show that this means $\inf\{d(p_1, p_2) \mid p_1 \in \mathfrak{a}(x), p_2 \in \mathfrak{a}(y)\} \leq \inf\{d(p_1, p_2) \mid p_1 \in \mathfrak{a}(z), p_2 \in \mathfrak{a}(p)\}$. For the sake of a contradiction, assume not. Then, $\exists p_1 \in \mathfrak{a}(z)$ and $\exists p_2 \in \mathfrak{a}(p)$ such that $\forall p_3 \in \mathfrak{a}(x)$ and $\forall p_4 \in \mathfrak{a}(y)$ we have $d(p_1, p_2) < d(p_3, p_4)$. Note that $\{p_1, p_2\} \in \mathbb{I}$. From here, it follows that we have $\mathfrak{M} \models_{\mathfrak{b}} \mathsf{CC}(q, z, p)$ and $\mathfrak{M} \not\models_{\mathfrak{b}} \mathsf{CC}(q, x, y)]$ where $\mathfrak{b}$ is an assignment which differs from $\mathfrak{a}$ only on q such that $\mathfrak{b}(q) = \{p_1, p_2\}$. This is a contradiction.

Now, finally to see the case of **AXM**10, assume that for some assignment $\mathfrak{a}$ we have $\mathfrak{M} \models_{\mathfrak{a}} \mathsf{CC}(x, y, z)$. Then there are $p_1 \in \mathfrak{a}(x)$, $p_2 \in \mathfrak{a}(x)$, $p_3 \in \mathfrak{a}(y)$ and $p_4 \in \mathfrak{a}(z)$ such that $d(p_3, p_4) \leq d(p_1, p_2)$. First notice that from **CNT**3 it follows that $\{p_1\}, \{p_2\}, \{p_3\}$ and $\{p_4\}$ are all in $\mathbb{I}$. Now, together with the above paragraph it follows that $\mathfrak{M} \models_{\mathfrak{b}} \mathsf{AP}(p, x) \wedge \mathsf{AP}(q, x) \wedge \mathsf{AP}(r, y) \wedge \mathsf{AP}(s, z) \wedge (r, s) \leq (p, q)$ where $\mathfrak{b}$ is an assignment which differs from $\mathfrak{a}$, if at all, on p, q, r and s such that $\mathfrak{b}(p) = \{p_1\}$, $\mathfrak{b}(q) = \{p_2\}$, $\mathfrak{b}(r) = \{p_3\}$ and $\mathfrak{b}(s) = \{p_4\}$.

□

We now turn our attention to the completeness of the axiomatic system

$\mathbf{Ax}\mathrm{CD}_1$ with respect to the class of all metric models with individuals by employing the well-known Henkin-argumentation in our proof. We begin by recalling one of the standard lemmas which constitutes an important part of the Henkinian proof method. This lemma has a well-known proof and there is no need to provide one here:

Lemma 2.3 (Witness (or Saturation) Lemma)**.** *Every $\mathbf{Ax}CD_1$-consistent set of sentences Σ can be extended to a saturated set Σ' in the extension of $\mathcal{L}^1$, $\mathcal{L}^1(c_1, c_2, \ldots)$, such that $\Sigma' \vdash \exists x\varphi \rightarrow \varphi[c_k/x]$, for every formula with one free variable φ and c_k is a witness for x.*

Next, we present the Henkin-argument at the core of our completeness proof, conveniently known as the "Henkin Lemma". Unlike the previous one, Henkin Lemma must be provided with a proof. Since the proof is quite long, it is split into multiple shorter lemmata.

Lemma 2.4 (Henkin Lemma)**.** *Every $\mathbf{Ax}CD_1$-consistent, maximal and saturated set of sentences Γ yields a metric model with individuals $\mathfrak{M}$ such that for any formula φ, we have that $\mathfrak{M} \models \varphi$ iff $\varphi \in \Gamma$.*

Proof. Let γ be a set of $\mathbf{Ax}\mathrm{CD}_1$-consistent set of sentences. Our main task is to build an appropriate metric model with individuals. It follows from the Lindenbaum's Lemma and from the Saturation Lemma (Lemma 2.3) that, we can extend γ to a $\mathbf{Ax}\mathrm{CD}_1$-consistent, maximal and saturated set of sentences Γ. Now let $\mathfrak{C}$ denote the collection of constants occurring in Γ. We will utilise equivalence classes over constant symbols to create the 'individuals' of our model. In order to achieve this, we first define the relation $\equiv$ such that for every $c_1 \in \mathfrak{C}$ and $c_2 \in \mathfrak{C}$ we have that:

$$c_1 \equiv c_2 \text{ iff } \Gamma \vdash c_1 = c_2.$$

Clearly, $\equiv$ is an equivalence relation over $\mathfrak{C}$. So let us define the equivalence classes induced by the relation $\equiv$ as follows:

$$|c_1| = \{c_2 \in \mathfrak{C} \mid c_1 \equiv c_2\}.$$

We have now constructed the basic elements of our model. However, in our models individuals are represented as sets of points. So far, we have only created the elements to stand for individuals. So we now need to "fill" these individuals with suitable points.

First, we define the universe -the set of all points- where our individuals will inherit their points from. We will denote the universe by W and define it as follows:

$$\mathsf{W} = \{c \in \mathfrak{C} \mid \Gamma \vdash \mathsf{A}(c)\}.$$

In other words, points are simply derived from the constants which are "syntactically points according to Γ". Now we have to assign points to the corresponding individuals. This can be achieved quite easily as follows:

$$\mathcal{P}(|c_1|) = \{c_2 \in \mathsf{W} \mid \Gamma \vdash \mathsf{P}(c_2, c_1)\}.$$

Now we arrived at the most complicated and important part of the proof: inducing a metric space over W. More specifically, we have to define a metric function $d\colon \mathsf{W} \times \mathsf{W} \to \mathbb{R}^+ \cup \{0\}$ such that the existing metric information within Γ is represented via d.

We will devote an inductive construction procedure which will take points from W as the input and at the end of the procedure, it will return a function d satisfying all the constraints we have mentioned in the above paragraph. It is very important to note that W is a countable set, since it is a subset of $\mathfrak{C}$ which is in turn a subset of the denumerable language $\mathcal{L}^1(c_1, c_2, \dots)$ of the Witness Lemma 2.3. *This observation is critical since without the guarantee of countability, an inductive construction procedure can not be used for higher cardinalities.*

Before giving the procedure in detail, we will define some shorthands for simplifying the specification of the procedure. First, we set up some relations on $\mathsf{W} \times \mathsf{W}$. Let c_1, c_2, c_3 and c_4 in W. We set:

- $(c_1, c_2) \sqsubseteq (c_3, c_4)$ iff $\Gamma \vdash (c_1, c_2) \leq (c_3, c_4)$,
- $(c_1, c_2) \sqsubset (c_3, c_4)$ iff $\Gamma \vdash (c_1, c_2) \leq (c_3, c_4) \wedge \neg(c_3, c_4) \leq (c_1, c_2)$,
- $(c_1, c_2) \sqcap (c_3, c_4)$ iff $\Gamma \vdash (c_1, c_2) \leq (c_3, c_4) \wedge (c_3, c_4) \leq (c_1, c_2)$.

The procedure given below works by considering *every different triple* of points from W, one triple in each iteration, until all the combinations of all the points from W are handled. Given an arbitrary triple of points, say c_1, c_2 and c_3, there are three values (one for each of the pairs (c_1, c_2), (c_2, c_3) and (c_1, c_3)) to be assigned by the procedure to the function d. In order to keep a track of the pairs whose value has been already assigned (note that same pairs will most likely occur within many different triples), they are added into a set as soon as their processing is finished by the procedure. This tracking set is denoted AV_n (Assigned Values), where n is the number of iterations. Another relevant notation which we will frequently use is the following:

$$d(\mathrm{AV}_n) = \{d(c_1, c_2) \mid \{c_1, c_2\} \in \mathrm{AV}_n\}.$$

In plain words, $d(\mathrm{AV}_n)$ denotes the set of all values which are so far (until nth iteration) assigned by the procedure.

The assignment of values to pairs is done in a certain order. Namely, before any processing, the procedure orders the pairs based on the constraints inherited from Γ. For example, if we have that $(c_1, c_2) \leq (c_2, c_3) \leq (c_1, c_3)$, then the procedure begins by dealing with the pair $\{c_1, c_2\}$ first, then deals with the pair $\{c_1, c_3\}$ in the second order and finally finishes assigning all three pairs by processing the pair $\{c_2, c_3\}$.

Before we give the procedure in detail, let us finally analyse the underlying strategy used by the procedure in assigning the values to d. The procedure has to achieve two main goals: First of all, the metric constraints concerning individuals inherited from Γ must be satisfied. Second of all, d must satisfy the necessary constraints in order to qualify as a metric.

For the first goal, we rely on the fact that $\mathbb{R}^+$ is dense: the procedure is always guaranteed to find appropriate values from $\mathbb{R}^+$ to assign for d such that the constrains inherited from Γ are satisfied.

In order to achieve the second goal, procedure ensures that in each iteration, the value picked for the maximal pair is less than twice the value picked for the minimal pair. This guarantees that the function d we end up with satisfies the triangle inequality and becomes a metric.

Technically speaking, the strategy in question is implemented by using yet another tracking set of values which we will denote by MPV_n (Maximal Pair Values), where n is the number of iterations. It works as follows: In each iteration, half of the value assigned for the maximal pair is added into MPV_n. In the iterations that follow, the value to be assigned for the minimal pair is chosen such that it is greater than all of the elements in MPV_n.

Let us now give the construction procedure in detail:

Construction 2.1 (Metric Construction)**.** The procedure consists of two main parts: The initial part in step 1 and the inductive step 2.

1. Assume that the first input to the procedure is the triple c_1, c_2 and c_3 from W such that $(c_1, c_2) \leq (c_2, c_3) \leq (c_1, c_3)$. Pick three arbitrary

elements r_1, r_2 and r_3 from $\mathbb{R}^+$ such that the appropriate ones of the following constraints are satisfied:
First pick r_1 and r_2:

- If $(c_1, c_2) \sqsubset (c_2, c_3)$ then $0 < r_1 < r_2 < 2 \cdot r_1$ or,
- if $(c_1, c_2) \square (c_2, c_3)$ then $0 < r_1 = r_2$.

Now pick r_3:

- If $(c_2, c_3) \sqsubset (c_1, c_3)$ then $0 < r_2 < r_3 < 2 \cdot r_1$ or,
- if $(c_2, c_3) \square (c_1, c_3)$ then $0 < r_2 = r_3$.

Now make the assignments for the function d as follows:

- $d(c_1, c_2) = d(c_2, c_1) = r_1$,
- $d(c_2, c_3) = d(c_3, c_2) = r_2$,
- $d(c_1, c_3) = d(c_3, c_1) = r_3$.

Set $\mathrm{AV}_1 = \{\{c_1, c_2\}, \{c_2, c_3\}, \{c_1, c_3\}\}$ and $\mathrm{MPV}_1 = \{\frac{d(c_1,c_3)}{2}\}$.

2. Assume that $n-1$th step has already been executed. If $\{\{c_1, c_2\} \mid c_1, c_2 \in \mathsf{W}\} = \mathrm{AV}_{n-1}$, then quit the procedure. Otherwise, start executing the nth step as follows: Pick a triple from W, say c_1, c_2 and c_3, such that at least one of the three pairs is not an element of AV_{n-1} -otherwise there is nothing to do. Suppose that we have the following order among the pairs: $(c_1, c_2) \leq (c_2, c_3) \leq (c_1, c_3)$.
   (a) *Firstly, assign a value for d on the minimal pair* (c_1, c_2)*:*
   If the pair is already processed by an earlier iteration of the procedure, i.e., if $\{c_1, c_2\} \in \mathrm{AV}_{n-1}$, then skip this step and continue with step 2b. Otherwise:
      i. If $\forall \{x, y\} \in \mathrm{AV}_{n-1}[(c_1, c_2) \sqsubset (x, y)]$ then:
         - pick $r \in \mathbb{R}^+$ such that $\max \mathrm{MPV}_{n-1} < r < \min d(\mathrm{AV}_{n-1})$ and assign $d(c_1, c_2) = d(c_2, c_1) = r$.

      ii. If $\forall \{x, y\} \in \mathrm{AV}_{n-1}[(x, y) \sqsubset (c_1, c_2)]$ then:
         - pick $r \in \mathbb{R}^+$ such that $\max d(\mathrm{AV}_{n-1}) < r < 2 \cdot \min d(\mathrm{AV}_{n-1})$ and assign $d(c_1, c_2) = d(c_2, c_1) = r$.

      iii. If $\exists \{x, y\} \in \mathrm{AV}_{n-1}[(c_1, c_2) \square (x, y)]$ then:
         - assign $d(c_1, c_2) = d(c_2, c_1) = d(x, y)$.

      iv. If none of the above is the case then:

- pick $r \in \mathbb{R}^+$ such that $\max\{d(x,y) \mid \{x,y\} \in \mathrm{AV}_{n-1} \wedge (x,y) \sqsubset (c_1,c_2)\} < r < \min\{d(x,y) \mid \{x,y\} \in \mathrm{AV}_{n-1} \wedge (c_1,c_2) \sqsubset (x,y)\}$ and assign $d(c_1,c_2) = d(c_2,c_1) = r$.

v. Finally set $\mathrm{AV}_n = \mathrm{AV}_{n-1} \cup \{\{c_1,c_2\}\}$.

(b) *Secondly, assign a value for d on the maximal pair* (c_1,c_3)*:*
If the pair is already processed by an earlier iteration of the procedure, i.e., if $\{c_1,c_3\} \in \mathrm{AV}_{n-1}$, then skip this step and continue with step 2c. Otherwise:

i. If $\forall(x,y) \in \mathrm{AV}_n[(x,y) \sqsubset (c_1,c_3)]$ then:
   - pick $r \in \mathbb{R}^+$ such that $\max d(\mathrm{AV}_n) < r < \min\{2 \cdot \min d(\mathrm{AV}_n), 2 \cdot d(c_1,c_2)\}$ and assign $d(c_1,c_3) = d(c_3,c_1) = r$.

ii. If $\exists(x,y) \in \mathrm{AV}_n[(c_1,c_3) \square (x,y)]$ then:
   - assign $d(c_1,c_3) = d(c_3,c_1) = d(x,y)$.

iii. If none of the above is the case then:
   - pick $r \in \mathbb{R}^+$ such that $\max\{d(x,y) \mid \{x,y\} \in \mathrm{AV}_n \wedge (x,y) \sqsubset (c_1,c_3)\} < r < \min\{\min\{d(x,y) \mid \{x,y\} \in \mathrm{AV}_n \wedge (c_1,c_3) \sqsubset (x,y)\}, 2 \cdot d(c_1,c_2)\}$ and assign $d(c_1,c_3) = d(c_3,c_1) = r$.

iv. Finally set $\mathrm{AV}_n = \mathrm{AV}_n \cup \{\{c_1,c_3\}\}$ and $\mathrm{MPV}_n = \mathrm{MPV}_{n-1} \cup \{\frac{d(c_1,c_3)}{2}\}$.

(c) *Thirdly, assign a value for d on the final pair* (c_2,c_3)*:*
If the pair is already processed by an earlier iteration of the procedure, i.e., if $\{c_2,c_3\} \in \mathrm{AV}_{n-1}$, then skip this step and end the current iteration. Otherwise:

i. If $\exists\{x,y\} \in \mathrm{AV}_n[(c_2,c_3) \square (x,y)]$ then:
   - assign $d(c_2,c_3) = d(c_2,c_3) = d(x,y)$.

ii. Otherwise:
   - pick $r \in \mathbb{R}^+$ such that $\max\{d(x,y) \mid \{x,y\} \in \mathrm{AV}_n \wedge (x,y) \sqsubset (c_2,c_3)\} < r < \min\{d(x,y) \mid \{x,y\} \in \mathrm{AV}_n \wedge (c_2,c_3) \sqsubset (x,y)\}$ and assign $d(c_2,c_3) = d(c_3,c_2) = r$.

iii. Finally set $\mathrm{AV}_n = \mathrm{AV}_n \cup \{\{c_2,c_3\}\}$.

Now it only remains to make the finishing touch: For every $c \in \mathsf{W}$, we set:

$$d(c,c) = 0.$$

This ends the construction procedure.

Let us make some final remarks about the procedure before proceeding any further. First of all, note that all maximums and minimums of the sets that are used in the procedure, e.g., $\min d(\mathrm{AV}_{n-1})$, $\max \mathrm{MPV}_n$ etc., are guaranteed to always exist. This can be easily shown by induction on n.

The fact that we have $\max \mathrm{MPV}_{n-1} < \min d(\mathrm{AV}_{n-1})$, which was used in step 2(a)i, can be seen through Lemma 2.6 below. On the other hand, the fact that $\max d(\mathrm{AV}_{n-1}) < 2 \cdot \min d(\mathrm{AV}_{n-1})$ can also be proved in a similar way to Lemma 2.6 by induction on n: In order to see the induction through, it is sufficient to examine the elements added in step 2b, where the maximal element of each step is added into AV_n and to notice that they are always $< 2 \cdot \min d(\mathrm{AV}_n)$.

In order to see that $\max\{d(x,y) \mid \{x,y\} \in \mathrm{AV}_{n-1} \wedge (x,y) \sqsubset (c_1,c_2)\} < \min\{d(x,y) \mid \{x,y\} \in \mathrm{AV}_{n-1} \wedge (c_1,c_2) \sqsubset (x,y)\}$ in step 2(a)iv, Lemma 2.5 is sufficient. Validity of all of the other assumptions made in the rest of the procedure can be verified in similar ways.

Now, we have to establish that the function d constructed by the procedure above is actually a metric. First of all, note that from Construction 2.1 it is obvious that d satisfies the following two constraints: $\forall c_1 \in \mathsf{W}$ and $\forall c_2 \in \mathsf{W}$ and $\forall r \in \mathbb{R}^+ \cup \{0\}$:

- $d(c_1,c_2) = 0$ iff $c_1 = c_2$ and,
- $d(c_1,c_2) = r$ iff $d(c_2,c_1) = r$.

In other words, it only remains to establish that d satisfies the triangle inequality. In order to establish that, we first need to prove the following two lemmas:

Lemma 2.5. *For every c_1, c_2, c_3 and c_4 in W, we have that $(c_1,c_2) \sqsubseteq (c_3,c_4)$ iff $d(c_1,c_2) \leq d(c_3,c_4)$.*

Proof. Let $c_1 \in \mathsf{W}$, $c_2 \in \mathsf{W}$, $c_3 \in \mathsf{W}$ and $c_4 \in \mathsf{W}$ and consider the procedure of Construction 2.1 by which d is defined. Clearly, we have that $\{c_1,c_2\} \in \mathrm{AV}_m$ and $\{c_3,c_4\} \in \mathrm{AV}_m$ for some m. So, if we could show for every n and any $\{c_1,c_2\} \in \mathrm{AV}_n$ and $\{c_3,c_4\} \in \mathrm{AV}_n$ that we have $(c_1,c_2) \sqsubseteq (c_3,c_4)$ iff $d(c_1,c_2) \leq d(c_3,c_4)$, i.e., at any stage of the construction procedure the claim holds for all the pairs processed so far, then we will have the proof we are looking for. The proof of this claim is by induction on n.

The base case for $n = 1$ is immediate from step 1 of Construction 2.1. For the inductive step, assume that for every $\{c_1,c_2\} \in \mathrm{AV}_{n-1}$ and $\{c_3,c_4\} \in \mathrm{AV}_{n-1}$

we have $(c_1, c_2) \sqsubseteq (c_3, c_4)$ iff $d(c_1, c_2) \leq d(c_3, c_4)$. Now pick $\{c_1, c_2\} \in \mathrm{AV}_n$ and $\{c_3, c_4\} \in \mathrm{AV}_n$.

If both $\{c_1, c_2\} \in \mathrm{AV}_{n-1}$ and $\{c_3, c_4\} \in \mathrm{AV}_{n-1}$, then we are immediately through by the induction hypothesis. So, suppose that we have $\{c_1, c_2\} \in \mathrm{AV}_{n-1}$ and $\{c_3, c_4\} \in \mathrm{AV}_n \setminus \mathrm{AV}_{n-1}$.

This means that the pair $\{c_3, c_4\}$ is added into AV_n via either of the steps 2a, 2b or 2c. Firstly, suppose that the pair $\{c_3, c_4\}$ is added into AV_n via step 2a of the procedure. In this step there are four exclusive cases to be considered corresponding to each of the sub-steps:

In the case of 2(a)i, in order to see through the claim from left to right direction assume that $(c_1, c_2) \sqsubseteq (c_3, c_4)$. However, this assumption contradicts with step 2(a)i's premise that $\forall\{x, y\} \in \mathrm{AV}_{n-1}[(c_3, c_4) \sqsubset (x, y)]$, since we have $(c_1, c_2) \sqsubseteq (c_3, c_4)$ and $\{c_1, c_2\} \in \mathrm{AV}_{n-1}$ by the previous assumption. This means that such a pair can not be actually added in step 2(a)i.

Conversely, assume that $d(c_1, c_2) \leq d(c_3, c_4)$. Since $\{c_1, c_2\} \in \mathrm{AV}_{n-1}$, it follows that $d(c_3, c_4) < d(c_1, c_2)$, which is a contradiction. So, just like in the opposite direction, it is impossible that a value for $d(c_3, c_4)$ is assigned in step 2(a)i under current assumptions.

In the case of 2(a)ii, we have that $\max d(\mathrm{AV}_{n-1}) < d(c_3, c_4)$. Therefore, we have $d(c_1, c_2) < d(c_3, c_4)$ since $\{c_1, c_2\} \in \mathrm{AV}_{n-1}$. There is nothing to show in the opposite direction since we already have $(c_1, c_2) \sqsubset (c_3, c_4)$ by the premise of step 2(a)ii, which implies that $(c_1, c_2) \sqsubseteq (c_3, c_4)$.

In the case of 2(a)iii, in order to see through the claim from left to right direction assume that $(c_1, c_2) \sqsubseteq (c_3, c_4)$. We have that $(c_3, c_4) \square (c_5, c_6)$ for some $\{c_5, c_6\} \in \mathrm{AV}_{n-1}$ and $d(c_5, c_6) = d(c_3, c_4)$. On the other hand, since we have $(c_1, c_2) \sqsubseteq (c_3, c_4)$ it follows from Lemma 2.1 that we also have $(c_1, c_2) \sqsubseteq (c_5, c_6)$. Now using the induction hypothesis it follows that $d(c_1, c_2) \leq d(c_5, c_6) = d(c_3, c_4)$. The opposite direction follows in a similar way using the induction hypothesis.

In the case of 2(a)iv, in order to see through the claim from left to right direction assume that $(c_1, c_2) \sqsubseteq (c_3, c_4)$. We have that $\max\{d(x, y) \mid \{x, y\} \in \mathrm{AV}_{n-1} \wedge (x, y) \sqsubset (c_3, c_4)\} < d(c_3, c_4)$. Since we have that $(c_1, c_2) \sqsubseteq (c_3, c_4)$ and $\neg(c_1, c_2) \square (c_3, c_4)$ (otherwise step 2a would have been finalised by the case 2(a)iii), it means that we have $(c_1, c_2) \sqsubset (c_3, c_4)$. Since $\{c_1, c_2\} \in \mathrm{AV}_{n-1}$, it follows that $d(c_1, c_2) < d(c_3, c_4)$. The opposite direction is obvious.

Now, suppose that the pair $\{c_3, c_4\}$ is added into AV_n via step 2b of the procedure. For this step, there are three exclusive cases to be considered corresponding to the sub-steps:

In the case of 2(b)i, since we have that $\{c_1, c_2\} \in \mathrm{AV}_{n-1} \subseteq \mathrm{AV}_n$ by the assumption, it follows from the very premise that we already have $(c_1, c_2) \sqsubset (c_3, c_4)$. Moreover, we get that $d(c_1, c_2) < d(c_3, c_4)$ since $\max d(\mathrm{AV}_n) < d(c_3, c_4)$. So, there is nothing to show for this case.

In the case of 2(b)ii first assume that $(c_1, c_2) \sqsubseteq (c_3, c_4)$. From the premise, we have that $(c_3, c_4) \sqcap (c_5, c_6)$ for some $\{c_5, c_6\} \in \mathrm{AV}_n$ and $d(c_3, c_4) = d(c_5, c_6)$. From Lemma 2.1 it follows that $(c_1, c_2) \sqsubseteq (c_5, c_6)$. Since $\mathrm{AV}_{n-1} \subseteq \mathrm{AV}_n$, there are two possibilities: Either $\{c_5, c_6\} \in \mathrm{AV}_{n-1}$ or $\{c_5, c_6\} \in \mathrm{AV}_n \setminus \mathrm{AV}_{n-1}$. In the former case, it follows from the induction hypothesis that $d(c_1, c_2) \leq d(c_5, c_6) = d(c_3, c_4)$. In the latter case, first notice that this implies that $\{c_5, c_6\}$ must have been added into AV_n via step 2a. From what is already established in the first part of this very proof regarding the case of step 2a, it follows that $d(c_1, c_2) \leq d(c_5, c_6)$, which gives us what we want since $d(c_3, c_4) = d(c_5, c_6)$. The opposite direction is very similar.

Now consider case 2(b)iii. We have that $\max\{d(x, y) \mid \{x, y\} \in \mathrm{AV}_n \wedge (x, y) \sqsubset (c_3, c_4)\} < d(c_3, c_4)$. Since $(c_1, c_2) \sqsubseteq (c_3, c_4)$ and $\neg(c_1, c_2) \sqcap (c_3, c_4)$ (otherwise step 2b would have been finalised by the sub-step 2(b)ii), we have $(c_1, c_2) \sqsubset (c_3, c_4)$. On the other hand, since we have $\{c_1, c_2\} \in \mathrm{AV}_{n-1} \subseteq \mathrm{AV}_n$, it follows that $d(c_1, c_2) < d(c_3, c_4)$ as desired.

For the opposite direction, assume that $d(c_1, c_2) \leq d(c_3, c_4)$. It is sufficient to show that we do not have $(c_3, c_4) \sqsubset (c_1, c_2)$. For sake of a contradiction, suppose we have $(c_3, c_4) \sqsubset (c_1, c_2)$. Since $\{c_1, c_2\} \in \mathrm{AV}_{n-1} \subseteq \mathrm{AV}_n$ and since from the premise of step 2(b)iii we have $d(c_3, c_4) < \min\{d(x, y) \mid \{x, y\} \in \mathrm{AV}_n \wedge (c_3, c_4) \sqsubset (x, y)\}$, it follows that $d(c_3, c_4) < d(c_1, c_2)$. A contradiction.

Thirdly and finally, suppose that the pair $\{c_3, c_4\}$ is added into AV_n via step 2c of the procedure. In this step, there are only two exclusive cases to be considered corresponding to the sub-steps and the necessary proofs are very similar to the previous cases.

□

Lemma 2.6. *For every $n \in \mathbb{N}$, $x \in MPV_n$ and $\{c_1, c_2\} \in AV_n$, we have that $x < d(c_1, c_2)$.*

Proof. The proof is by induction on n. The base case for $n = 1$ is obvious from step 1 of Construction 2.1. For the inductive step, assume that for every $x \in \mathrm{MPV}_{n-1}$ and for every $\{c_1, c_2\} \in \mathrm{AV}_{n-1}$, we have that $x < d(c_1, c_2)$. Now let $x \in \mathrm{MPV}_n$ and $\{c_1, c_2\} \in \mathrm{AV}_n$. Suppose that $\{c_1, c_2\} \in \mathrm{AV}_n \setminus \mathrm{AV}_{n-1}$.

Firstly, suppose that the pair $\{c_1, c_2\}$ is added into AV_n via step 2a of the procedure. In this step, there are four sub-cases based on which a value for $d(c_1, c_2)$ is assigned:

In case 2(a)i we have that $\max \mathrm{MPV}_{n-1} < d(c_1, c_2)$. So, if $x \in \mathrm{MPV}_{n-1}$, then we are easily through.

Alternatively, suppose that $x \in \mathrm{MPV}_n \setminus \mathrm{MPV}_{n-1}$. First, note that the set MPV_{n-1} is extended to MPV_n in step 2b. If x was added into MPV_n either with step 2(b)i or with step 2(b)iii, one can easily derive from the assignments made in these steps that we have $x < d(c_1, c_2)$ in both cases. Now suppose that x was added into MPV_n with case 2(b)ii. Then, we have that $x = \frac{d(c_3,c_4)}{2}$ for some $\{c_3, c_4\}$ such that either $\{c_3, c_4\} \in \mathrm{AV}_{n-1}$ or $\{c_3, c_4\} \in \mathrm{AV}_n \setminus \mathrm{AV}_{n-1}$. Former case implies that we have $\max \mathrm{MPV}_n = \max \mathrm{MPV}_{n-1}$. But now we have that $x \leq \max \mathrm{MPV}_n = \max \mathrm{MPV}_{n-1} < d(c_1, c_2)$, which is what we want. The latter case implies that $\{c_3, c_4\}$ was added into AV_n in step 2a. But this means that the pair $\{c_3, c_4\}$ is actually the pair $\{c_1, c_2\}$. In other words, we have that $x = \frac{d(c_1,c_2)}{2} < d(c_1, c_2)$.

In case 2(a)ii, we have that $\max d(\mathrm{AV}_{n-1}) < d(c_1, c_2)$. Now, if $x \in \mathrm{MPV}_{n-1}$, then from the induction hypothesis it follows that $x < \max d(\mathrm{AV}_{n-1}) < d(c_1, c_2)$. The case of $x \in \mathrm{MPV}_n \setminus \mathrm{MPV}_{n-1}$ has an almost identical proof to the corresponding part of case 2(a)i in the above paragraph.

In case 2(a)iii, we have that $d(c_1, c_2) = d(c_3, c_4)$ for some $\{c_3, c_4\} \in \mathrm{AV}_{n-1}$. If $x \in \mathrm{MPV}_{n-1}$, then from the induction hypothesis it follows that $x < d(c_3, c_4) = d(c_1, c_2)$. We again leave the proof of case $x \in \mathrm{MPV}_n \setminus \mathrm{MPV}_{n-1}$ since it can be easily derived from the case of 2(a)i.

Finally, in case 2(a)iv, we have that $\max\{d(x, y) \mid \{x, y\} \in \mathrm{AV}_{n-1} \wedge (x, y) \sqsubset (c_1, c_2)\} < d(c_1, c_2)$. If $x \in \mathrm{MPV}_{n-1}$, then from the induction hypothesis it follows that $x < \max\{d(x, y) \mid \{x, y\} \in \mathrm{AV}_{n-1} \wedge (x, y) \sqsubset (c_1, c_2)\} < d(c_1, c_2)$ and we are through. Case $x \in \mathrm{MPV}_n \setminus \mathrm{MPV}_{n-1}$ can be derived from the above corresponding case of 2(a)i.

If the pair $\{c_1, c_2\}$ is added into AV_n either via step 2b or step 2c, then it suffices to notice that there is a pair $\{c_3, c_4\} \in \mathrm{AV}_n \setminus \mathrm{AV}_{n-1}$ added in step 2a such that $(c_3, c_4) \sqsubseteq (c_1, c_2)$ and as we established in the above, $x < d(c_3, c_4)$. Using Lemma 2.5, from here it follows that $x < d(c_1, c_2)$ as desired. This completes the proof. □

Let us now show that d satisfies the triangle inequality. Let c_1, c_2 and c_3 in W. It is sufficient to establish that we have $d(c_1, c_3) \leq d(c_1, c_2) + d(c_2, c_3)$.

Consider the ordering among the pairs $\{c_1, c_3\}$, $\{c_1, c_2\}$ and $\{c_2, c_3\}$. We consider two possibilities: Firstly, suppose that the pair $\{c_1, c_3\}$ is not the maximal pair. This means that we have either $(c_1, c_3) \sqsubseteq (c_1, c_2)$ or $(c_1, c_3) \sqsubseteq (c_2, c_3)$. But then from Lemma 2.5, it follows that we have either $d(c_1, c_3) \leq d(c_1, c_2)$ or $d(c_1, c_3) \leq d(c_2, c_3)$. In either case, we get that $d(c_1, c_3) \leq d(c_1, c_2) + d(c_2, c_3)$ as desired.

Secondly, suppose that (c_1, c_3) is the maximal pair. It is sufficient to show that $\frac{d(c_1,c_3)}{2} \leq d(c_1, c_2)$ and $\frac{d(c_1,c_3)}{2} \leq d(c_2, c_3)$. Since (c_1, c_3) is the maximal pair, it follows from step 2b of Construction 2.1 that for some n we have that $\frac{d(c_1,c_3)}{2} \in \mathrm{MPV}_n$ and $\{\{c_1, c_2\}, \{c_2, c_3\}\} \subseteq \mathrm{AV}_n$. Now the desired result follows immediately from Lemma 2.6 and this shows that d satisfies the triangle equality. As we have already mentioned preceding Lemma 2.5, d possesses all the other necessary properties and we conclude that the pair $\langle \mathsf{W}, d\rangle$ is a metric space.

Finally, we are almost ready to put together our "Henkin model," except that we yet to define the set of individuals. But this can be done quite easily by setting:

$$\mathbb{I} = \{\mathcal{P}(|c|) \mid c \in \mathfrak{C}\}.$$

We first set

$$\mathfrak{F} = \left\langle \mathsf{W}, d, \mathbb{I} \right\rangle.$$

Now we can give our constructed model as follows:

$$\mathfrak{M} = \left\langle \mathfrak{F}, \mathsf{C} \right\rangle,$$

where C is a function interpreting the constant symbols such that for every $c \in \mathfrak{C}$,

$$\mathsf{C}(c) = \mathcal{P}(|c|).$$

In order to complete the proof of the Henkin Lemma, we provide the following two lemmata:

Lemma 2.7. *Let φ be a formula. Then we have that $\mathfrak{M} \models \varphi$ iff $\varphi \in \Gamma$.*

Proof. The proof is by induction on the complexity of φ. It is sufficient to establish the base case alone, since the rest of the inductive cases are highly routine. This amounts to prove that we have $\mathfrak{M} \models \mathsf{CC}(c_1, c_2, c_3)$ iff $\mathsf{CC}(c_1, c_2, c_3) \in \Gamma$.

In order to prove the claim in the direction from right to left, assume that $\mathsf{CC}(c_1, c_2, c_3) \in \Gamma$. From axiom **AXM**10 we have $\exists c_1' \exists c_1'' \exists c_2' \exists c_3' [\mathsf{AP}(c_1', c_1) \wedge$

$\mathsf{AP}(c_1'', c_1) \wedge \mathsf{AP}(c_2', c_2) \wedge \mathsf{AP}(c_3', c_3) \wedge (c_2', c_3') \leq (c_1', c_1'')]$. Now, by the very definition of $\mathcal{P}(|c_1|)$ it can be easily seen that $c_1' \in \mathcal{P}(|c_1|)$, $c_1'' \in \mathcal{P}(|c_1|)$, $c_2' \in \mathcal{P}(|c_2|)$ and $c_3' \in \mathcal{P}(|c_3|)$. Moreover, from Lemma 2.5 we get that $d(c_2', c_3') \leq d(c_1', c_1'')$. In other words, we have $\mathfrak{M} \models \mathsf{CC}(c_1, c_2, c_3)$ as desired.

Conversely, suppose that $\mathfrak{M} \models \mathsf{CC}(c_1, c_2, c_3)$. Then there exists $c_1' \in \mathcal{P}(|c_1|)$, $c_1'' \in \mathcal{P}(|c_1|)$, $c_2' \in \mathcal{P}(|c_2|)$ and $c_3' \in \mathcal{P}(|c_3|)$ such that $d(c_2', c_3') \leq d(c_1', c_1'')$. On the other hand, it follows from the construction that $\mathsf{AP}(c_2', c_2)$, $\mathsf{AP}(c_1', c_1)$, $\mathsf{AP}(c_1'', c_1)$ and $\mathsf{AP}(c_3', c_3)$. Moreover, from Lemma 2.5, we get that $(c_2', c_3') \sqsubseteq (c_1', c_1'')$. Now it follows from axiom **AXM**10 and the maximal consistency of Γ that we have $\mathsf{CC}(c_1, c_2, c_3) \in \Gamma$ as desired. This completes the proof. □

Lemma 2.8. *$\mathfrak{F}$ satisfies all constraints **CNT**1- **CNT**6, i.e., $\mathfrak{F}$ is a metric structure with individuals.*

Proof. Let us begin by establishing that **CNT**1 is satisfied over $\mathfrak{F}$. By definition, we have that $\mathcal{P}(|c|) \subseteq \mathsf{W}$ for every $c \in \mathfrak{C}$. So, we clearly have that $\mathbb{I} \subseteq 2^{\mathsf{W}}$. On the other hand, from axiom **AXM**5 it follows that $\mathbf{U} \in \mathfrak{C}$ and $\forall c \in \mathfrak{C}$ we have that $\Gamma \vdash \mathsf{P}(c, \mathbf{U})$. By definition, this entails that $\mathcal{P}(|\mathbf{U}|) = \mathsf{W}$. Hence, $\mathsf{W} \in \mathbb{I}$ as desired.

Now lets show the case of **CNT**2. Let some arbitrary $c_1 \in \mathfrak{C}$. We will show that $\mathcal{P}(|c_1|) \neq \emptyset$. However, from axiom **AXM**9 we immediately get that $\exists c_2[\mathsf{A}(c_2) \wedge \mathsf{P}(c_2, c_1)]$. So it follows that $c_2 \in \mathcal{P}(|c_1|)$. This proves **CNT**2.

To see the case of **CNT**3, let $c_1 \in \mathsf{W}$. Then by definition we have that $\Gamma \vdash \mathsf{A}(c_1)$. In other words, $\Gamma \vdash \forall x[\mathsf{P}(x, c_1) \rightarrow c_1 = x]$. However, this means that $\mathcal{P}(|c_1|) = \{c_1\}$. Since $\mathcal{P}(|c_1|) \in \mathbb{I}$, the desired result follows.

We will finally establish that **CNT**6 is satisfied by $\mathfrak{F}$. The cases of **CNT**4 and **CNT**5 follow in a very similar way. Let $x \in \mathbb{I}$ and $y \in \mathbb{I}$. By definition, it follows that there are $c_1 \in \mathfrak{C}$ and $c_2 \in \mathfrak{C}$ such that $x = \mathcal{P}(|c_1|)$ and $y = \mathcal{P}(|c_2|)$. On the other hand, from axiom **AXM**6 we derive that there exists $c_3 \in \mathfrak{C}$ such that $\forall p[\mathsf{I}(c_3, p) \leftrightarrow [\mathsf{I}(c_1, p) \vee \mathsf{I}(c_2, p)]]$.

Let $z \in \mathcal{P}(|c_1|) \cup \mathcal{P}(|c_2|)$. We will show that $z \in \mathcal{P}(|c_3|)$. First assume that $z \in \mathcal{P}(|c_1|)$ (the case of $z \in \mathcal{P}(|c_2|)$ can be established in a similar way). Then by definition we get that $\mathsf{AP}(z, c_1)$. This implies that $\mathsf{I}(c_1, z)$ from the definition of predicate P. So it follows that $\mathsf{I}(c_3, z)$. However, since $\mathsf{A}(z)$, it follows from axiom **AXM**7 and the definition of predicate A that $\mathsf{P}(z, c_3)$. So we finally get that $z \in \mathcal{P}(|c_3|)$ as desired.

Conversely assume that $z \in \mathcal{P}(|c_3|)$. So we have that $\mathsf{AP}(z, c_3)$ and from here that $\mathsf{I}(z, c_3)$. Therefore, $\mathsf{I}(c_1, z) \vee \mathsf{I}(c_2, z)$. Since $\mathsf{A}(z)$, it follows that

$\mathsf{P}(z, c_1) \vee \mathsf{P}(z, c_2)$. In other words, we have either $z \in \mathcal{P}(|c_1|)$ or $z \in \mathcal{P}(|c_2|)$, i.e., $z \in \mathcal{P}(|c_1|) \cup \mathcal{P}(|c_2|)$. □

With Lemma 2.8, we also complete the proof of the Henkin Lemma. □

Now we finally conclude that,

Theorem 2.9 (Completeness)**.** *Let φ be a formula. Then we have that* $\mathsf{M} \models \varphi \Rightarrow \mathbf{Ax}CD_1 \vdash \varphi$.

Proof. Follows directly from Lindenbaum's Lemma, Saturation Lemma (Lemma 2.3) and the Henkin Lemma (Lemma 2.4). □

2.4. Undecidability

Now we briefly examine the expressive strength of the first-order comparative distance logic:

Theorem 2.10. *The satisfiability problem of $\mathcal{L}^1$ formulas in* M *is undecidable, i.e., the first-order comparative distance logic is undecidable.*

Proof. It follows that the same proof used to show the undecidability of quantitative first-order distance logic [17] can be used to obtain the undecidability of first-order comparative distance logic as well. The proof uses a reduction from the theory of graphs, which is known to be hereditarily undecidable [27]. This means that not only the graph theory itself is undecidable, but so is every subtheory of it. Recall that graphs are structures in the form $\langle \mathsf{W}, R \rangle$ where R is a reflexive and symmetric binary relation on $\mathsf{W} \times \mathsf{W}$. Theory of graphs is then the logic arising by interpreting the first-order language of a single binary relation symbol using graphs. For more details see Rabin's work [27].

We will begin by defining a metric structure with individuals for any given graph. Let $G = \langle \mathsf{W}, R \rangle$ be a graph. For every $c_1, c_2 \in \mathsf{W}$ we set:

$$d^R(c_1, c_2) = \begin{cases} 0, & \text{if } c_1 = c_2, \\ 1, & \text{if } c_1 \neq c_2 \wedge R(c_1, c_2), \\ 2, & \text{if } \neg R(c_1, c_2). \end{cases}$$

It is easy to show that $\langle \mathsf{W}, d^R \rangle$ is a metric space and that $\mathfrak{F}(G) = \langle \mathsf{W}, d^R, \mathbb{I} \rangle$, where $\mathbb{I} = 2^{\mathsf{W}} - \emptyset$, satisfies all the constraints **CNT**1-**CNT**6, i.e., $\mathfrak{F}(G)$ is a metric structure with individuals.

Next, we define our syntactic translation. For every formula φ of the language of graph theory, let $\varphi^\dagger$ to stand for the translation of φ into $\mathcal{L}^1$ by replacing every occurrence of the atom $R(x, y)$ by $\forall z[\mathsf{CC}(z, x, y)]$. Note that $\varphi^\dagger$ is now a formula of $\mathcal{L}^1$ and we have that $G \models \varphi$ iff $\mathfrak{F}(G) \models \varphi^\dagger$. To see this, it is sufficient to establish that for any $c_1, c_2 \in \mathsf{W}$:

$$R(c_1, c_2) \text{ iff } \mathfrak{F}(G) \models_{\mathfrak{a}} \forall z[\mathsf{CC}(z, x, y)], \tag{1}$$

where $\mathfrak{a}(x) = \{c_3 \in \mathsf{W} \mid d(c_1, c_3) \leq 1\}$ and $\mathfrak{a}(y) = \{c_2\}$.

To prove (1), first suppose that $\mathfrak{a}$ is an assignment as in the previous paragraph. To see the validity of the claim from left to right, assume that we have $R(c_1, c_2)$ for some $c_1, c_2 \in \mathsf{W}$. Then by the definition of d^R it follows that either $d^R(c_1, c_2) = 0$ or $d^R(c_1, c_2) = 1$. In the former case, we immediately derive that $\mathfrak{a}(x) \cap \mathfrak{a}(y) \neq \emptyset$. On the other hand, if $d^R(c_1, c_2) = 1$, then we have that $c_2 \in \mathfrak{a}(x)$. So we get $\mathfrak{a}(x) \cap \mathfrak{a}(y) \neq \emptyset$ again. It is obvious that $\mathfrak{a}(x) \cap \mathfrak{a}(y) \neq \emptyset$ implies $\mathfrak{F} \models_{\mathfrak{a}} \forall z[\mathsf{CC}(z, x, y)]$.

Conversely suppose that $\mathfrak{F} \models_{\mathfrak{a}} \forall z[\mathsf{CC}(z, x, y)]$. Since every metric structure with individuals must contain at least one singleton-individual by **CNT**3, this implies that $\mathfrak{a}(x) \cap \mathfrak{a}(y) \neq \emptyset$ and hence, $d^R(c_1, c_2) \leq 1$. Hence, we have either $d^R(c_1, c_2) = 0$ or $d(c_1, c_2) = 1$. In either case we have $R(c_1, c_2)$ as desired. This completes the proof of (1).

Consider the set of all formulas φ of the language of theory of graphs such that $\varphi^\dagger$ is true in every comparative distance model. From above, it follows that this set is in fact a subtheory of the theory of graphs, which is hereditarily undecidable, i.e., not only itself but every subtheory of graph theory is undecidable. □

3. Modal Comparative Distance Logic

This section is devoted to the development of a modal logic formalism which can talk about distance information in a comparative and qualitative manner, just like the first-order logic of the previous section. The biggest difference between this section and the previous one is naturally the use of a much less expressive -hence, computationally much more feasible-language to talk about essentially identical semantic structures, i.e., metric structures with individuals.

3.1. Language and Semantics

We will use a modal language containing denumerably many proposition letters the set of which will be denoted by $\mathscr{P}$ and its elements by

$p, q, r, \ldots$, and the usual basic boolean operators $\vee$ and $\neg$ together with the standard proposition constants $\top$ and $\bot$. The main component of the language is the polyadic modal operator $\langle \mathbf{CC} \rangle(\alpha, \beta)$ ('here can-connect somewhere at which α and somewhere at which β'). We denote this language by $\mathcal{L}$. As we will demonstrate shortly, the standard **S5** modal operator $\exists$ ('global modality') can be easily defined within the modal comparative distance logic. Duals of the modalities $\langle \mathbf{CC} \rangle$ and $\exists$ are denoted by $[\mathbf{CC}]$ and $\forall$, respectively.

Despite of the fact that the modal operator $\exists$ can be obtained from the language as outlined above, in some cases we will be better off with a language which explicitly contains an individual **S5** operator. Lemma 3.20 of the modal completeness theorem is the occasion in which we will make use of explicitly defining this operator. This language, extending the language $\mathcal{L}$ merely with the **S5** operator $\exists$, will be denoted by $\mathcal{L}[\langle \mathbf{CC} \rangle, \exists]$.

Naturally, we would like to interpret formulas of $\mathcal{L}[\langle \mathbf{CC} \rangle, \exists]$ over 'metric models with individuals' (see Section 2.1). Or more precisely, with structures of the form:

$$\mathfrak{N} = \left\langle \mathsf{W}, d, \mathbb{I}, \mathsf{V} \right\rangle,$$

where the pair $\langle \mathsf{W}, d \rangle$ is a metric space, $\mathbb{I} \subseteq 2^{\mathsf{W}}$ ('individuals') and $\mathsf{V} \colon \mathscr{P} \to 2^{\mathbb{I}}$ is a valuation function. Arbitrary formulas are then interpreted as follows: For all formulas α, β, every $p \in \mathscr{P}$ and $w \in \mathbb{I}$ we have:

- $\mathfrak{N}, w \models p$ iff $w \in \mathsf{V}(p)$,
- $\mathfrak{N}, w \models \alpha \wedge \beta$ iff $\mathfrak{N}, w \models \alpha$ and $\mathfrak{N}, w \models \beta$,
- $\mathfrak{N}, w \models \neg\alpha$ iff $\mathfrak{N}, w \not\models \alpha$,
- $\mathfrak{N}, w \models \exists\alpha$ iff $\exists u[\mathfrak{N}, u \models \alpha]$,
- $\mathfrak{N}, w \models \langle \mathbf{CC} \rangle(\alpha, \beta)$ iff

$$\exists u \exists v [\exists p_1 \in u \exists p_2 \in v \exists p_3 \in w \exists p_4 \in w \; [d(p_1, p_2) \leq d(p_3, p_4)] \wedge \mathfrak{N}, u \models \alpha \text{ and } \mathfrak{N}, v \models \beta].$$

However, we would like to work with relational semantics and have our formulas interpreted in the usual Kripkean way. The following section deals with this switch and establishes the necessary equivalence theorem between metric models and their relational representations.

3.1.1. Relational Representation of Metric Models

In this section, we will introduce relational semantics for the modal comparative distance logic, so that the formulas of $\mathcal{L}[\langle\mathbf{CC}\rangle, \exists]$ will be interpreted in the usual Kripkean way. Let us begin by introducing our targeted relational structures.

A comparative distance frame is a structure which can be given by the pair,

$$\mathfrak{F} = \left\langle \mathsf{W}, \mathsf{CC} \right\rangle$$

where W is the domain set the elements of which ('states') represent 'individuals' and CC is a ternary accessibility relation over $\mathsf{W} \times \mathsf{W} \times \mathsf{W}$ which will be used to interpret the binary $\langle\mathbf{CC}\rangle$ modality. A comparative distance frame $\mathfrak{F}$ satisfies the following constraints:

(**CNT**1) $\forall w \forall u [\mathsf{CC}(w, u, u)]$,

(**CNT**2) $\forall w \forall u \forall v [\mathsf{CC}(w, u, v) \Rightarrow \mathsf{CC}(w, v, u)]$,

(**CNT**3) $\forall w \forall u \forall v \forall y \forall z [\mathsf{CC}(w, u, v) \wedge \neg\mathsf{CC}(w, y, z) \Rightarrow \neg\exists t [\mathsf{CC}(t, y, z) \wedge \neg\mathsf{CC}(t, u, v)]]$.

Therefore, a comparative distance model based on a comparative distance frame $\mathfrak{F}$ is as usual a pair,

$$\mathfrak{M} = \left\langle \mathfrak{F}, \mathsf{V} \right\rangle$$

where V is a valuation function such that $\mathsf{V}\colon \mathscr{P} \to 2^{\mathsf{W}}$, mapping proposition letters to sets of states. Now, we are ready to give the relational interpretation of $\mathcal{L}[\langle\mathbf{CC}\rangle, \exists]$ formulas by defining a relation of truth in the usual inductive way. For all formulas α, β, every $w \in \mathsf{W}$ and $p \in \mathscr{P}$ we have:

- $\mathfrak{M}, w \models p$ iff $w \in \mathsf{V}(p)$,
- $\mathfrak{M}, w \models \alpha \wedge \beta$ iff $\mathfrak{M}, w \models \alpha$ and $\mathfrak{M}, w \models \beta$,
- $\mathfrak{M}, w \models \neg\alpha$ iff $\mathfrak{M}, w \not\models \alpha$,
- $\mathfrak{M}, w \models \exists\alpha$ iff $\exists u [\mathfrak{M}, u \models \alpha]$,
- $\mathfrak{M}, w \models \langle\mathbf{CC}\rangle(\alpha, \beta)$ iff $\exists u \exists v [\mathsf{CC}(w, u, v)$ and $\mathfrak{M}, u \models \alpha$ and $\mathfrak{M}, v \models \beta]$.

We denote the class of all comparative distance frames by F and the class of all comparative distance models by M. We will write $\mathsf{M} \models \varphi$ to denote the validity of the formula φ over every comparative distance model. Modal comparative distance logic is the set of all formulas of the language $\mathcal{L}$ that are valid on every comparative distance model. Similarly, modal comparative distance logic with global modality is the set of all formulas of the language $\mathcal{L}[\langle \mathbf{CC} \rangle, \exists]$ that are valid on every comparative distance model.

Note that, under the given semantics, and in particular constraint **CNT**1, we can obtain the 'global modality' as follows:

$$\exists \varphi \coloneqq \langle \mathbf{CC} \rangle (\varphi, \varphi).$$

The following theorem shows that the relational semantics we introduced above is a "true representation" of the original metric models.

Theorem 3.1. *Let φ be a formula. There is a metric model with individuals satisfying φ iff there is a comparative distance model satisfying φ.*

Proof. Follows from Lemmas 3.2 and 3.3 below. □

The following lemma establishes the proof of Theorem 3.1 in the direction from left to right:

Lemma 3.2. *For every formula φ, if there is a metric model with individuals satisfying φ, then there is a comparative distance model satisfying φ.*

Proof. Let,

$$\mathfrak{M} = \left\langle \mathsf{W}, d, \mathbb{I}, \mathsf{V} \right\rangle$$

be a metric model with individuals, i.e., $\langle \mathsf{W}, d \rangle$ is a metric space, $\mathbb{I} \subseteq 2^{\mathsf{W}}$ and V is a valuation function mapping proposition letters to the subsets of $\mathbb{I}$. A comparative distance model $\mathfrak{N}$ can be easily obtained from the metric model $\mathfrak{M}$ by defining a ternary accessibility relation in the obvious way as follows: For every w, u and v in $\mathbb{I}$:

$$\mathsf{CC}(w, u, v) \text{ iff } \exists p_1 \in w \exists p_2 \in w \exists p_3 \in u \exists p_4 \in v [d(p_3, p_4) \leq d(p_1, p_2)].$$

Now set,

$$\mathfrak{N} = \left\langle \mathbb{I}, \mathsf{CC}, \mathsf{V} \right\rangle.$$

Let φ be a formula. We will first establish that for every $w \in \mathbb{I}$, we have that $\mathfrak{M}, w \models \varphi$ iff $\mathfrak{N}, w \models \varphi$.

Let $w \in \mathbb{I}$. The proof is by induction on the complexity of φ. Base case and the boolean cases are standard. So consider the case when $\varphi = \langle \mathbf{CC} \rangle(\alpha, \beta)$ for some formulas α, β. We have that, $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(\alpha, \beta)$ iff $\exists u \exists v[\exists p_1 \in w \exists p_2 \in w \exists p_3 \in u \exists p_4 \in v[d(p_3, p_4) \leq d(p_1, p_2)]$ and $\mathfrak{M}, u \models \alpha$ and $\mathfrak{M}, v \models \beta]$ iff (by the definition of CC above) $\mathsf{CC}(w, u, v) \wedge \mathfrak{M}, u \models \alpha$ and $\mathfrak{M}, v \models \beta$ iff (induction hypothesis) $\mathsf{CC}(w, u, v) \wedge \mathfrak{N}, u \models \alpha$ and $\mathfrak{N}, v \models \beta$ iff $\mathfrak{N}, w \models \langle \mathbf{CC} \rangle(\alpha, \beta)$.

Now, it only remains to establish that $\mathfrak{N}$ is a comparative distance model. In order to see that $\mathfrak{N}$ satisfies constraint **CNT**1, let $w \in \mathbb{I}$ and $u \in \mathbb{I}$. It is sufficient to pick two arbitrary points $p_1 \in w$ and $p_2 \in u$ to see that $0 = d(p_2, p_2) \leq d(p_1, p_1) = 0$. Hence, we have that $\mathsf{CC}(w, u, u)$.

Next, we show that $\mathfrak{N}$ satisfies constraint **CNT**2. Let $w \in \mathbb{I}$, $u \in \mathbb{I}$ and $v \in \mathbb{I}$ such that $\mathsf{CC}(w, u, v)$. It follows that $\exists p_1 \in w \exists p_2 \in w \exists p_3 \in u \exists p_4 \in v[d(p_3, p_4) \leq d(p_1, p_2)]$. Since d is a metric, it follows that $d(p_4, p_3) = d(p_3, p_4) \leq d(p_1, p_2)$. Thus, it follows that $\mathsf{CC}(w, v, u)$ as desired.

Finally, we establish that **CNT**3 is satisfied. Let w, u, v, y, z and t from $\mathbb{I}$ such that $\mathsf{CC}(w, u, v) \wedge \neg\mathsf{CC}(w, y, z)$. Now, for the sake of a contradiction suppose that we also have that $\neg\mathsf{CC}(t, u, v) \wedge \mathsf{CC}(t, y, z)$. Then, by the definition of CC, the following hold:

$$\exists p_1 \in w \exists p_2 \in w \exists p_3 \in u \exists p_4 \in v[d(p_3, p_4) \leq d(p_1, p_2)] \tag{2}$$

$$\exists p_5 \in t \exists p_6 \in t \exists p_7 \in y \exists p_8 \in z[d(p_7, p_8) \leq d(p_5, p_6)] \tag{3}$$

$$\forall p'_1 \in w \forall p'_2 \in w \forall p'_7 \in y \forall p'_8 \in z[d(p'_1, p'_2) < d(p'_7, p'_8)] \tag{4}$$

$$\forall p'_5 \in t \forall p'_6 \in t \forall p'_3 \in u \forall p'_4 \in v[d(p'_5, p'_6) < d(p'_3, p'_4)] \tag{5}$$

From (2), (3) and (5), it follows that $d(p_5, p_6) < d(p_3, p_4) \leq d(p_1, p_2)$. On the other hand, from (2), (3) and (4), we get that $d(p_1, p_2) < d(p_7, p_8) \leq d(p_5, p_6)$, which is a contradiction. This completes the proof of the lemma. □

The proof of Theorem 3.1 in the direction from right to left is given by the following lemma, which has a more complicated proof compared to Lemma 3.2.

Lemma 3.3. *For every formula φ, if there is a comparative distance model satisfying φ, then there is a metric model with individuals satisfying φ.*

Proof. Let,

$$\mathfrak{M} = \left\langle \mathsf{W}, \mathsf{CC}, \mathsf{V} \right\rangle$$

be a comparative distance model. As it will be established in Section 3.2, modal comparative distance logic has the finite model property. Moreover, an examination of the proofs of Section 3.2 will also show that the finite model property result is established independently from the results of this section. This means that, for our purposes, we can safely assume that $\mathfrak{M}$ is a finite model and build up rest of the proof on this assumption.

The first step in our proof is to construct a metric model with individuals by taking the states of $\mathfrak{M}$ as our domain of points and defining an appropriate metric on it. However, we have already established in Construction 2.1 that, given a set of comparative distance constraints over pairs of individuals $\mathsf{W}^2 \times \mathsf{W}^2$ (up to a size of $\aleph_0$), one can always find a metric space $\langle \mathsf{W}, d \rangle$ in which the metric d embodies those constraints. The relevant constraints in the current context together with their shorthands can be given as follows:

- $(w, u) \sqsubseteq (v, y)$ iff $\forall W[\mathsf{CC}(W, v, y) \Rightarrow \mathsf{CC}(W, w, u)]$,
- $(w, u) \sqsubset (v, y)$ iff $(w, u) \sqsubseteq (v, y) \land \neg(v, y) \sqsubseteq (w, u)$,
- $(w, u) \sqcap (v, y)$ iff $(w, u) \sqsubseteq (v, y) \land (v, y) \sqsubseteq (w, u)$.

Moreover, Lemma 2.5 establishes that these constraints are passed on to the constructed metric space, i.e., for every w, u, v and y in W we have that:

$$(w, u) \sqsubseteq (v, y) \text{ iff } d(w, u) \leq d(v, y).$$

Finally and most importantly, we define the basic concept which will provide the cornerstone of the following construction procedure.

Definition 3.1 (Diameter and Extension of a Point)**.** Given a *finite* metric space $\langle \mathsf{W}, d \rangle$ and a comparative distance frame $\langle \mathsf{W}, \mathsf{CC} \rangle$, a function $D \colon \mathsf{W} \to \mathbb{R}^+$ is called a 'diameter function' iff for every $w \in \mathsf{W}$ we have that:

$$\max \left\{ d(u, v) \mid u, v \in \mathsf{W} \land \mathsf{CC}(w, u, v) \right\} < D(w) < \min \left\{ d(u, v) \mid u, v \in \mathsf{W} \land \neg\mathsf{CC}(w, u, v) \right\}.$$

Now, for every $w \in \mathsf{W}$, the 'extension of w', denoted by $E(w)$, is a point which can be added to extend the given metric space to a new one, say $\langle \mathsf{W}', d' \rangle$, such that $\mathsf{W}' = \mathsf{W} \cup \{E(w)\}$ and:

$$d'(w, E(w)) = D(w).$$

Now, we are ready to extend the metric space $\langle \mathsf{W}, d \rangle$ to a metric model with individuals using the following construction procedure.

Construction 3.1. The procedure consists of two main parts. In the first part, we extend the given metric space $\langle \mathsf{W}, d \rangle$ to a metric space where every point in W is coupled by its extension. In the second part, we only put the pieces together to construct the model we are after.

1. In this part, we will recursively extend the given metric space $\langle \mathsf{W}, d \rangle$ by adding the extension of one point in each step. So, initially set $\mathsf{W}' = \mathsf{W}$ and $d' = d$, and repeat the following procedure until we have $E(w) \in \mathsf{W}'$ for every $w \in \mathsf{W}$:

    - Pick some $w \in \mathsf{W}$ such that $E(w) \notin \mathsf{W}'$.
    - Extend the set W' to W'' as follows: $\mathsf{W}'' = \mathsf{W}' \cup \{E(w)\}$.
    - Extend the function d' to a function d'' on W'' as follows: For every $u, v \in \mathsf{W}''$:

    $$d''(u, v) = \begin{cases} 0, & \text{if } u = v = E(w); \\ D(w), & \text{if } u = E(w) \wedge v = w \text{ (or vice versa);} \\ d'(w, v) + D(w), & \text{if } u = E(w) \wedge v \neq w \text{ (or vice versa);} \\ d'(u, v), & \text{otherwise.} \end{cases}$$

    - Set $\mathsf{W}' = \mathsf{W}''$, $d' = d''$ and start over unless we have $E(w) \in \mathsf{W}'$ for every $w \in \mathsf{W}$.

2. From the previous step, we have a metric space $\langle \mathsf{W}', d' \rangle$ extending $\langle \mathsf{W}, d \rangle$ such that every point in W is coupled with its extension in W'. Using this fact, we now create the set of individuals:

    - For every $w \in \mathsf{W}$, define $\mathbb{I}(w) = \{w, E(w)\}$. We now set the set of individuals as follows:

    $$\mathbb{I} = \{\mathbb{I}(w) \mid w \in \mathsf{W}\}.$$

- Finally, we set the valuation function as follows: For every $p \in \mathscr{P}$:
$$\mathsf{V}'(p) = \{\mathbb{I}(w) \mid w \in \mathsf{V}(p)\}.$$

Now, our final metric model with individuals can be obtained by setting:
$$\mathfrak{N} = \left\langle \mathsf{W}', d', \mathbb{I}, \mathsf{V}' \right\rangle.$$

Following proofs establish that the metric model with individuals generated by the above construction is indeed what we are looking for.

Fact 3.1. Let $w, u \in \mathsf{W}$. Then the following statements hold:

1. $d(w, u) = d'(w, u)$,
2. if $w \neq u$, then $d'(E(w), E(u)) = d(w, u) + D(w) + D(u)$.

Proof. (1) is immediate from the construction. In order to see (2), let $w, u \in \mathsf{W}$ such that $w \neq u$. Suppose that the procedure of Construction 3.1 added $E(w)$ into W' before it added $E(u)$. By construction, we have that $d'(E(w), u) = d'(w, u) + D(w)$. In one of the following iterations, when the procedure extends W' by $E(u)$, it sets $d'(E(u), E(w)) = d'(u, E(w)) + D(u) = d'(w, u) + D(w) + D(u) = d(w, u) + D(w) + D(u)$, which is exactly what we wanted to find. □

Lemma 3.4. *$\langle \mathsf{W}', d' \rangle$ is a metric space.*

Proof. We first show that the structure $\langle \mathsf{W}', d' \rangle$ satisfies the identity axiom, i.e., $\forall w \in \mathsf{W}' \forall u \in \mathsf{W}'[d'(w, u) = 0 \text{ iff } w = u]$. Let $w, u \in \mathsf{W}'$. Since $\langle \mathsf{W}', d' \rangle$ is an extension of the metric space $\langle \mathsf{W}, d \rangle$, we are only interested in the case when at least one of w and u is an 'extension' of some other point.

It is easy to see from Construction 3.1 that, if $w = u = E(v)$ for some $v \in \mathsf{W}$, then $d'(w, u) = 0$. Conversely, suppose that $d'(w, u) = 0$. Since the 'diameter' of any point must be greater than 0 by Definition 3.1, it follows from Construction 3.1 that the only way for this to happen is when $w = u = E(v)$ for some $v \in \mathsf{W}$.

It is completely trivial to see that the function d' is symmetric. So, let us now move on to show that it satisfies the triangle inequality. Picking three arbitrary points from W' to establish triangle inequality gives us three possibilities to consider for the combination of extension and non-extension points, each with multiple sub-possibilities. First, let $w, u, v \in \mathsf{W}$.

From Fact 3.1 and Construction 3.1 and the fact that d is a metric, the following can be easily seen:

1. The following two configurations of two extension points and one non-extension point:
   - $d'(E(w), E(u)) + d'(E(u), v) = d(w, u) + D(w) + D(u) + d(u, v) + D(u) \geq d(w, v) + D(w) = d'(E(w), v)$;
   - $d'(E(w), u) + d'(u, E(v)) = d(w, u) + D(w) + d(u, v) + D(v) \geq d(w, v) + D(w) + D(v) = d'(E(w), E(v))$.
2. The following two configurations of one extension point and two non-extension points:
   - $d'(E(w), u) + d'(u, v) = d(w, u) + D(w) + d(u, v) \geq d(w, v) + D(w) = d'(E(w), v)$;
   - $d'(w, E(u)) + d'(E(u), v) = d(w, u) + D(u) + d(u, v) + D(u) \geq d(w, v) = d'(w, v)$.
3. And finally, the following configuration for three extension points:
   - $d'(E(w), E(u)) + d'(E(u), E(v)) = d(w, u) + D(w) + D(u) + d(u, v) + D(u) + D(v) \geq d(w, v) + D(w) + D(v) = d'(E(w), E(v))$.

Therefore, d' satisfies the triangle inequality. Hence, the structure $\langle \mathsf{W}', d' \rangle$ is a metric space. □

Lemma 3.5. *For every formula φ, we have that $\mathfrak{M}, w \models \varphi$ iff $\mathfrak{N}, \mathbb{I}(w) \models \varphi$.*

Proof. The proof is by induction on the complexity of φ. Base case and the boolean cases are straightforward by Construction 3.1. So, let us now consider the case of $\varphi = \langle \mathbf{CC} \rangle(\alpha, \beta)$.

In the direction from left to right, suppose that $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(\alpha, \beta)$. Then, $\exists u \exists v[\mathsf{CC}(w, u, v)$ and $\mathfrak{M}, u \models \alpha$ and $\mathfrak{M}, v \models \beta]$. By the definition of diameter and extension (Definition 3.1), we have that $d'(w, E(w)) = D(w) > d(u, v) = d'(u, v)$. On the other hand, from the induction hypothesis we have that $\mathfrak{N}, \mathbb{I}(u) \models \alpha$ and $\mathfrak{N}, \mathbb{I}(v) \models \beta$. Obviously, $u \in \mathbb{I}(u)$, $v \in \mathbb{I}(v)$ and $w, E(w) \in \mathbb{I}(w)$. Thus, $\mathfrak{N}, \mathbb{I}(w) \models \langle \mathbf{CC} \rangle(\alpha, \beta)$.

Conversely, suppose that $\mathfrak{N}, \mathbb{I}(w) \models \langle \mathbf{CC} \rangle(\alpha, \beta)$. Then, $\exists u \in \mathsf{W}$ and $\exists v \in \mathsf{W}$ and there are points $p_1, p_2 \in \mathbb{I}(w)$, $p_3 \in \mathbb{I}(u)$ and $p_4 \in \mathbb{I}(v)$ such that $\mathfrak{N}, \mathbb{I}(u) \models \alpha$ and $\mathfrak{N}, \mathbb{I}(v) \models \beta$ and $d'(p_1, p_2) \geq d'(p_3, p_4)$. Moreover, from the induction hypothesis we get that $\mathfrak{M}, u \models \alpha$ and $\mathfrak{M}, v \models \beta$.

Now, observe that we must have $d'(p_1, p_2) = D(w)$ and that $d'(p_3, p_4) \geq d(u, v)$. So, it follows that $D(w) \geq d(u, v)$. By Construction 3.1, this implies that $\mathsf{CC}(w, u, v)$. Hence, we get that $\mathfrak{M}, w \models \langle\mathbf{CC}\rangle(\alpha, \beta)$ as desired. This completes the proof of the lemma. □

Now the proof of Lemma 3.3 follows from Construction 3.1, Lemma 3.4 and Lemma 3.5. □

3.2. Finite Model Property and Decidability

In this section, we will show that the modal comparative distance logic enjoys strong finite model property with respect to the class of all comparative distance models. In Section 3.1 we defined two modal languages, $\mathcal{L}$ and $\mathcal{L}[\langle\mathbf{CC}\rangle, \exists]$, the latter of which is a simple extension of the former merely with the 'global modality'. It is a well known fact that, whenever a decidability result for a modal logic is obtained by using a filtration-based technique to establish strong finite model property, the same can be done with that modal logic's extension by the global modality (Blackburn et al. [28], pg. 418, par. 4 and also pg. 422, Theorem 7.8). Since $\mathcal{L}[\langle\mathbf{CC}\rangle, \exists]$ is merely an extension of $\mathcal{L}$ by the global modality $\exists$, it is sufficient to establish that the modal comparative distance logic has the strong finite model property in order to conclude that the modal comparative distance logic with global modality also has the strong finite model property.

The proof that modal comparative distance logic has the strong finite model property is a standard one based on the filtration technique. Given a model $\mathfrak{M}$ and a formula φ, we will provide a procedure which constructs a finite model $\mathfrak{M}^{\mathrm{Fin}}$ such that φ is satisfied in $\mathfrak{M}$ iff φ is satisfied in $\mathfrak{M}^{\mathrm{Fin}}$. Let us now start giving the details of this construction procedure.

Construction 3.2. We say that a set of formulas Σ is 'symmetry-closed' whenever we have that:

$$\langle\mathbf{CC}\rangle(\alpha, \beta) \in \Sigma \text{ iff } \langle\mathbf{CC}\rangle(\beta, \alpha) \in \Sigma.$$

Let Σ be a finite, symmetry and subformula closed set of formulas and $\mathfrak{M} = \langle \mathsf{W}, \mathsf{CC}, \mathsf{V} \rangle$ be a comparative distance model. We begin by defining a relation over $\mathsf{W} \times \mathsf{W}$ which we will denote by $\equiv_\Sigma$. For every $w, u \in \mathsf{W}$ set:

$$w \equiv_\Sigma u \text{ iff } \forall \varphi \in \Sigma[\mathfrak{M}, w \models \varphi \Leftrightarrow \mathfrak{M}, u \models \varphi].$$

In plain words, $\equiv_\Sigma$ is the modal equivalence relation with respect to the set of modal formulas Σ. It is obvious that $\equiv_\Sigma$ is an equivalence relation. We denote the equivalence class of a $w \in \mathsf{W}$ induced by this relation with $|w|$. We will now define the model $\mathfrak{M}^{\mathrm{Fin}}$ by taking the 'filtration of $\mathfrak{M}$ through Σ'. This can be done by defining the following:

- $\mathsf{W}^{\mathrm{Fin}} = \{|w| \mid w \in \mathsf{W}\}$;
- $\mathsf{CC}^{\mathrm{Fin}}(|w|, |u|, |v|)$ iff for every $\langle \mathbf{CC} \rangle(\varphi, \psi) \in \Sigma$
$$[\mathfrak{M}, u \models \varphi \text{ and } \mathfrak{M}, v \models \psi \Rightarrow \mathfrak{M}, w \models \langle \mathbf{CC} \rangle(\varphi, \psi)];$$
- For every $p \in \mathscr{P}$ such that $p \in \Sigma$, $\mathsf{V}^{\mathrm{Fin}}(p) = \{|w| \mid \mathfrak{M}, w \models p\}$.

Now finally set,
$$\mathfrak{M}^{\mathrm{Fin}} = \left\langle \mathsf{W}^{\mathrm{Fin}}, \mathsf{CC}^{\mathrm{Fin}}, \mathsf{V}^{\mathrm{Fin}} \right\rangle,$$
as the filtration of $\mathfrak{M}$ through Σ.

Let Σ be a finite, symmetry and subformula closed set of formulas and let $\mathfrak{M} = \langle \mathsf{W}, \mathsf{CC}, \mathsf{V} \rangle$ be a comparative distance model. If $\mathfrak{M}^{\mathrm{Fin}} = \langle \mathsf{W}^{\mathrm{Fin}}, \mathsf{CC}^{\mathrm{Fin}}, \mathsf{V}^{\mathrm{Fin}} \rangle$ is the filtration of $\mathfrak{M}$ through Σ, then we have the following three lemmata:

Lemma 3.6. *For every formula $\varphi \in \Sigma$ and every state $w \in \mathsf{W}$, we have that $\mathfrak{M}, w \models \varphi$ iff $\mathfrak{M}^{Fin}, |w| \models \varphi$.*

Proof. The proof is by induction on the complexity of φ. The base case is trivial from Construction 3.2 and the boolean cases are straightforward. So, it only remains to establish the modal case when $\varphi = \langle \mathbf{CC} \rangle(\alpha, \beta)$.

To see it from left to right, assume that $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(\alpha, \beta)$. Then we have that $\exists u \exists v[\mathsf{CC}(w, u, v) \wedge u \models \alpha \text{ and } v \models \beta]$. First, from here it follows that we have $\mathsf{CC}^{\mathrm{Fin}}(|w|, |u|, |v|)$. Moreover, from the induction hypothesis it also follows that $\mathfrak{M}^{\mathrm{Fin}}, |u| \models \alpha$ and $\mathfrak{M}^{\mathrm{Fin}}, |v| \models \beta$. Hence, we get that $\mathfrak{M}^{\mathrm{Fin}}, |w| \models \langle \mathbf{CC} \rangle(\alpha, \beta)$ as desired.

Now in order to see it in the opposite direction, assume that we have $\mathfrak{M}^{\mathrm{Fin}}, |w| \models \langle \mathbf{CC} \rangle(\alpha, \beta)$. From here we get that $\exists |u| \exists |v| [\mathsf{CC}^{\mathrm{Fin}}(|w|, |u|, |v|) \wedge \mathfrak{M}^{\mathrm{Fin}}, |u| \models \alpha \text{ and } \mathfrak{M}^{\mathrm{Fin}}, |v| \models \beta]$. Note that we have $\langle \mathbf{CC} \rangle(\alpha, \beta) \in \Sigma$ since $\varphi \in \Sigma$ and $\langle \mathbf{CC} \rangle(\alpha, \beta)$ is a subformula of φ and Σ is subformula-closed. Moreover, from the induction hypothesis we get that $\mathfrak{M}, u \models \alpha$ and $\mathfrak{M}, v \models \beta$. But now, it follows from Construction 3.2 that $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(\alpha, \beta)$. This completes the proof of the lemma. □

Lemma 3.7. *$\mathfrak{M}^{Fin}$ is a comparative distance model.*

Proof. It is sufficient to establish that $\mathfrak{F}^{\text{Fin}} = \langle \mathsf{W}^{\text{Fin}}, \mathsf{CC}^{\text{Fin}} \rangle$ is a comparative distance frame, which amounts to show that the frame constraints **CNT**1-**CNT**3 (see Section 3.1.1) hold over $\mathfrak{F}^{\text{Fin}}$.

Let us first establish that **CNT**1 is satisfied over $\mathfrak{F}^{\text{Fin}}$. Let $|w|, |u| \in \mathsf{W}^{\text{Fin}}$ and pick some $\langle \mathbf{CC} \rangle(\alpha, \beta) \in \Sigma$. Suppose that $\mathfrak{M}, u \models \alpha$ and $\mathfrak{M}, u \models \beta$. Since $\mathfrak{F}$ satisfies the frame constraint **CNT**1, it follows that we have $\mathsf{CC}(w, u, u)$. Hence, $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(\alpha, \beta)$. From Construction 3.2, we derive that $\mathsf{CC}^{\text{Fin}}(|w|, |u|, |u|)$ as desired.

Let us now consider the case of **CNT**2. Let $|w|, |u|, |v| \in \mathsf{W}^{\text{Fin}}$ and suppose that $\mathsf{CC}^{\text{Fin}}(|w|, |u|, |v|)$. In order to see that we have $\mathsf{CC}^{\text{Fin}}(|w|, |v|, |u|)$, pick some $\langle \mathbf{CC} \rangle(\alpha, \beta) \in \Sigma$ and assume that $\mathfrak{M}, v \models \alpha$ and $\mathfrak{M}, u \models \beta$. Since Σ is symmetry-closed, we have that $\langle \mathbf{CC} \rangle(\beta, \alpha) \in \Sigma$. By the assumption and Construction 3.2, it follows that we have $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(\beta, \alpha)$. Since $\mathfrak{F}$ satisfies **CNT**2, it is easy to see that this implies $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(\alpha, \beta)$. Hence, from Construction 3.2, we get that $\mathsf{CC}^{\text{Fin}}(|w|, |v|, |u|)$.

Finally we address the case of **CNT**3. Let $|w|, |u|, |v|, |y|, |z| \in \mathsf{W}^{\text{Fin}}$ and suppose that we have $\mathsf{CC}^{\text{Fin}}(|w|, |u|, |v|) \wedge \neg \mathsf{CC}^{\text{Fin}}(|w|, |y|, |z|)$. For sake of a contradiction, suppose that $\exists |t| \in \mathsf{W}^{\text{Fin}}$ such that $\neg \mathsf{CC}^{\text{Fin}}(|t|, |u|, |v|) \wedge \mathsf{CC}^{\text{Fin}}(|t|, |y|, |z|)$.

From here, it follows that there is a $\langle \mathbf{CC} \rangle(\alpha, \beta) \in \Sigma$ such that $\mathfrak{M}, y \models \alpha$ and $\mathfrak{M}, z \models \beta$ and $\mathfrak{M}, w \not\models \langle \mathbf{CC} \rangle(\alpha, \beta)$. Since we also have that $\mathsf{CC}^{\text{Fin}}(|t|, |y|, |z|)$ by the assumption, it follows that $\mathfrak{M}, t \models \langle \mathbf{CC} \rangle(\alpha, \beta)$. In the very same way, it follows that there is a formula $\langle \mathbf{CC} \rangle(\gamma, \delta) \in \Sigma$ such that $\mathfrak{M}, u \models \gamma$ and $\mathfrak{M}, v \models \delta$ and $\mathfrak{M}, t \not\models \langle \mathbf{CC} \rangle(\gamma, \delta)$ and since $\mathsf{CC}^{\text{Fin}}(|w|, |u|, |v|)$, it also follows that $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(\gamma, \delta)$.

To summarise, we have that $\mathfrak{M}, t \models \langle \mathbf{CC} \rangle(\alpha, \beta) \wedge \neg \langle \mathbf{CC} \rangle(\gamma, \delta)$ and $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(\gamma, \delta) \wedge \neg \langle \mathbf{CC} \rangle(\alpha, \beta)$. Now it is easy to see that this contradicts with the fact that $\mathfrak{F}$ satisfies **CNT**3. This completes the proof of the lemma. □

Lemma 3.8. *The size of $\mathfrak{M}^{Fin}$ is exponential in the size of Σ, i.e., $|\mathsf{W}^{Fin}| \leq 2^{|\Sigma|}$.*

Proof. Define a function $f \colon \mathsf{W}^{\text{Fin}} \to 2^{\Sigma}$ such that for every $|w| \in \mathsf{W}^{\text{Fin}}$ we have,

$$f(|w|) = \{\varphi \in \Sigma \mid \mathfrak{M}^{\text{Fin}}, w \models \varphi\}.$$

It is sufficient to show that f is a well-defined and injective function. To see that f is well-defined, let $|w|, |u| \in \mathsf{W}^{\text{Fin}}$ and suppose that $|w| = |u|$. By definition, this means that w and u are modally equivalent with respect to Σ. From here it immediately follows that $f(|w|) = f(|u|)$.

To see that f is also injective, suppose that $f(|w|) = f(|u|)$ for some $|w|, |u| \in \mathsf{W}^{\mathrm{Fin}}$. By the definition of f, this means that w and u are modally equivalent with respect to Σ. In other words, $w \equiv_\Sigma u$. Hence, $|w| = |u|$ as desired. □

Now it only remains to put the pieces together, which gives us the main result:

Theorem 3.9 (Strong Finite Model Property)**.** *Let φ be a formula. If φ is satisfiable over a comparative distance model, then it is satisfiable over a finite comparative distance model of size at most $2^{|\varphi|}$. In other words, modal comparative distance logic has the strong finite model property with respect to* M*, the class of all comparative distance models.*

Corollary 3.10 (Strong Finite Model Property)**.** *Modal comparative distance logic with global modality has the strong finite model property with respect to* M*.*

Moreover, we have that:

Lemma 3.11. *The class of all finite comparative distance models is recursive.*

Finally, we present our main result which follows directly from Theorem 3.9, Corollary 3.10 and Lemma 3.11:

Theorem 3.12. *Modal comparative distance logic and modal comparative distance logic with global modality have decidable satisfiability problems.*

3.3. Computational Complexity

In this section we show that the modal comparative distance logic has an NP-complete satisfiability problem. We adapt a proof method which relies on the fact that modal comparative distance logic has the finite model property, which was established in the previous section. In fact, the proof will go a little bit further and establish that the modal comparative distance logic has the polysize model property.

Core of the proof consists of Construction 3.3 below: Given a formula φ and a finite model $\mathfrak{M}^{\mathrm{Fin}}$, this construction procedure generates a new model $\mathfrak{M}^{\varphi}$ by appropriately selecting states from $\mathfrak{M}^{\mathrm{Fin}}$ such that by the end of the procedure, size of $\mathfrak{M}^{\varphi}$ is only polynomial in the size of φ (in contrast to the exponential model generated in the finite model property proof above) and the satisfiability of φ can be preserved in $\mathfrak{M}^{\varphi}$. Now let us continue with the details.

Construction 3.3. Let φ be a formula and,

$$\mathfrak{M}^{\text{Fin}} = \left\langle \mathsf{W}^{\text{Fin}}, \mathsf{CC}^{\text{Fin}}, \mathsf{V}^{\text{Fin}} \right\rangle$$

be a *finite* comparative distance model such that $\mathfrak{M}^{\text{Fin}}, W \models \varphi$ for some $W \in \mathsf{W}^{\text{Fin}}$. We will select suitable states from $\mathfrak{M}^{\text{Fin}}$ to construct a new model $\mathfrak{M}^{\varphi}$ such that the size of $\mathfrak{M}^{\varphi}$ is polynomial in the size of φ.

First, let $\langle \mathbf{CC} \rangle(\alpha_1, \beta_1), \ldots, \langle \mathbf{CC} \rangle(\alpha_n, \beta_n)$ be an enumeration of all of the subformulas of φ in the form of $\langle \mathbf{CC} \rangle(\cdot, \cdot)$ and which are are satisfiable in $\mathfrak{M}^{\text{Fin}}$. For each pair of formulas α_k and β_k, where $1 \leq k \leq n$, choose a pair of states w_k and u_k from W^{Fin} such that w_k and u_k is a pair with minimal distance in between satisfying formulas α_k and β_k, respectively. More precisely, we choose a pair of states w_k and u_k from W^{Fin} such that:

$$\begin{aligned} &[\mathfrak{M}^{\text{Fin}}, w_k \models \alpha_k \text{ and } \mathfrak{M}^{\text{Fin}}, u_k \models \beta_k] \wedge \\ &\qquad \forall v[\forall y \forall z[\mathsf{CC}^{\text{Fin}}(v, y, z) \text{ and } \mathfrak{M}^{\text{Fin}}, y \models \alpha_k \text{ and } \mathfrak{M}^{\text{Fin}}, z \models \beta_k] \Rightarrow \\ &\qquad \mathsf{CC}^{\text{Fin}}(v, w_k, u_k)] \end{aligned} \tag{6}$$

Now, the critical question is whether such a pair of states can always be found. However, since every formula $\langle \mathbf{CC} \rangle(\alpha_k, \beta_k)$ is satisfied in $\mathfrak{M}^{\text{Fin}}$ by the assumption and $\mathfrak{M}^{\text{Fin}}$ is a finite model, it is easy to see that such a pair of states (w_k and u_k) must exist. So, we can now finally set:

- $\mathsf{W}^{\varphi} = \{W\} \cup \bigcup_{k=1}^{n} \{w_k, u_k\}$,
- $\mathsf{CC}^{\varphi} = \mathsf{CC}^{\text{Fin}} \restriction \mathsf{W}^{\varphi}$,
- $\mathsf{V}^{\varphi} = \mathsf{V}^{\text{Fin}} \restriction \mathsf{W}^{\varphi}$.

And finally set,

$$\mathfrak{M}^{\varphi} = \left\langle \mathsf{W}^{\varphi}, \mathsf{CC}^{\varphi}, \mathsf{V}^{\varphi} \right\rangle.$$

Lemma 3.13. *$\mathfrak{M}^{\varphi}$ is a comparative distance model.*

Proof. Since $\mathfrak{M}^{\varphi}$ is a restriction of $\mathfrak{M}^{\text{Fin}}$ and $\mathfrak{M}^{\text{Fin}}$ is a comparative distance model, it follows straightforwardly that $\mathfrak{M}^{\varphi}$ satisfies constraints **CNT**1-**CNT**3. □

Lemma 3.14. *For every subformula ψ of φ and every state $w \in \mathsf{W}^\varphi$, we have that $\mathfrak{M}^{Fin}, w \models \psi$ iff $\mathfrak{M}^\varphi, w \models \psi$.*

Proof. Let ψ be a subformula of φ. The proof is naturally by induction on the complexity of ψ. Let $w \in \mathsf{W}^\varphi$. Since $\mathfrak{M}^\varphi$ is simply a restriction of $\mathfrak{M}^{\mathrm{Fin}}$, base case follows trivially.

Now suppose $\psi = \neg\alpha$. Then we have that $\mathfrak{M}^{\mathrm{Fin}}, w \models \neg\alpha$ iff $\mathfrak{M}^{\mathrm{Fin}}, w \not\models \alpha$ iff (by the induction hypothesis) $\mathfrak{M}^\varphi, w \not\models \alpha$ iff $\mathfrak{M}^\varphi, w \models \neg\alpha$.

Alternatively suppose that $\psi = \alpha\wedge\beta$. Then, $\mathfrak{M}^{\mathrm{Fin}}, w \models \alpha\wedge\beta$ iff $\mathfrak{M}^{\mathrm{Fin}}, w \models \alpha$ and $\mathfrak{M}^{\mathrm{Fin}}, w \models \beta$ iff (by the induction hypothesis) $\mathfrak{M}^\varphi, w \models \alpha$ and $\mathfrak{M}^\varphi, w \models \beta$ iff $\mathfrak{M}^\varphi, w \models \alpha \wedge \beta$.

Now we address the case of $\psi = \langle\mathbf{CC}\rangle(\alpha, \beta)$. In the direction from left to right, suppose that we have $\mathfrak{M}^{\mathrm{Fin}}, w \models \langle\mathbf{CC}\rangle(\alpha, \beta)$. Since ψ is a subformula of φ, it follows from (6) of Construction 3.3 that there must be a pair of states u_α and v_β in W^φ such that they are the closest pair having $\mathfrak{M}^{\mathrm{Fin}}, u_\alpha \models \alpha$ and $\mathfrak{M}^{\mathrm{Fin}}, v_\beta \models \beta$. Now from the induction hypothesis we get that $\mathfrak{M}^\varphi, u_\alpha \models \alpha$ and $\mathfrak{M}^\varphi, v_\beta \models \beta$.

On the other hand, since we have $\mathfrak{M}^{\mathrm{Fin}}, w \models \langle\mathbf{CC}\rangle(\alpha, \beta)$ by the assumption, we get $\exists u \exists v[\mathsf{CC}^{\mathrm{Fin}}(w, u, v) \wedge \mathfrak{M}^{\mathrm{Fin}}, u \models \alpha$ and $\mathfrak{M}^{\mathrm{Fin}}, v \models \beta]$. But now it follows as a consequence of (6) that we must have $\mathsf{CC}^{\mathrm{Fin}}(w, u_\alpha, v_\beta)$ and thus, $\mathsf{CC}^\varphi(w, u_\alpha, v_\beta)$ by the construction. This gives the desired result.

In the opposite direction, suppose that $\mathfrak{M}^\varphi, w \models \langle\mathbf{CC}\rangle(\alpha, \beta)$. Then we have that $\exists u \exists v[\mathsf{CC}^\varphi(w, u, v) \wedge \mathfrak{M}^\varphi, u \models \alpha$ and $\mathfrak{M}^\varphi, v \models \beta]$. Since $\mathfrak{M}^\varphi$ is a restriction of $\mathfrak{M}^{\mathrm{Fin}}$, it follows from here that $\mathsf{CC}^{\mathrm{Fin}}(w, u, v)$. On the other hand, from the induction hypothesis we get that $\mathfrak{M}^{\mathrm{Fin}}, u \models \alpha$ and $\mathfrak{M}^{\mathrm{Fin}}, v \models \beta$. This obviously implies that $\mathfrak{M}^{\mathrm{Fin}}, w \models \langle\mathbf{CC}\rangle(\alpha, \beta)$ as desired.

□

Lemma 3.15. *Modal comparative distance logic has the polysize model property.*

Proof. A quick examination of Construction 3.3 reveals that the size of $\mathfrak{M}^\varphi$ is only polynomial in the size of input formula φ. More precisely, the size of $\mathfrak{M}^\varphi$ is equal to twice the number of modalities in the input formula plus 1 at the maximum. Thus, from Lemmas 3.14 and 3.13 we conclude that the modal comparative distance logic has the polysize model property. □

Theorem 3.16. *The satisfiability problem of modal comparative distance logic is NP-complete.*

Proof. Since the class of all comparative distance frames can be defined by a first-order sentence, it follows that the membership problem of the class of all comparative distance models is polynomial ([28], pg. 376, Lemma 6.36). Now, from lemma 3.15 it follows that the satisfiability problem of modal comparative distance logic is NP-complete ([28], pg. 375, Lemma 6.35). □

3.4. Soundness and Completeness Theorems

In this section, we provide an axiomatic system for syntactic reasoning with comparative distances. We will introduce an axiomatic system and it will be shown again in this section that the introduced system is sound and complete with respect to the class of all comparative distance frames. Both the soundness and the completeness proofs follow a standard methodology.

3.4.1. Axiomatic System

We begin by constructing an axiomatic system which we will denote with $\mathbf{Ax}\mathrm{CD}_{\diamond}$. Naturally, $\mathbf{Ax}\mathrm{CD}_{\diamond}$ consists of axioms for propositional logic, the standard axioms of minimal modal logic **K** for each modal operator we use and the axioms which capture the essential nature of comparative distance reasoning. In addition to these, it contains the standard inference rules of uniform substitution, generalization and of course, modus ponens. This results with the following axiom schemata for $\mathbf{Ax}\mathrm{CD}_{\diamond}$:

(**AXM**1) $[\mathbf{CC}](p \to q, r) \to [[\mathbf{CC}](p, r) \to [\mathbf{CC}](q, r)]$,

(**AXM**2) $[\mathbf{CC}](p, q \to r) \to [[\mathbf{CC}](p, q) \to [\mathbf{CC}](p, r)]$,

(**AXM**3) $\forall(p \to q) \to [\forall p \to \forall q]$,

(**AXM**4) $\exists\exists p \to \exists p$,

(**AXM**5) $p \to \exists p$,

(**AXM**6) $p \to \forall\exists p$,

(**AXM**7) $\langle\mathbf{CC}\rangle(p, q) \to \exists p \wedge \exists q$,

(**AXM**8) $\exists(p \wedge q) \to \langle\mathbf{CC}\rangle(p, q)$,

(**AXM**9) $\langle\mathbf{CC}\rangle(p, q) \to \langle\mathbf{CC}\rangle(q, p)$, (**AXM**10) $\langle \mathbf{CC} \rangle(p,q) \land \neg\langle \mathbf{CC} \rangle(r,s) \rightarrow \forall[\langle \mathbf{CC} \rangle(r,s) \rightarrow \langle \mathbf{CC} \rangle(p,q)]$.

Axioms **AXM**1, **AXM**2 and **AXM**3 are those corresponding to the axioms of the minimal modal logic **K** (**AXM**1 and **AXM**2 in the polyadic form) making the logic generated by **Ax**CD$_\diamond$ a 'normal modal logic.' This is a property which will be necessary in the application of some of the theorems that are fundamental to our argumentation.

Axioms **AXM**4, **AXM**5 and **AXM**6 are the axioms more commonly known by the names **4** (of transitive frames), **T** (of reflexive frames) and **B** (of symmetric frames), respectively. They constitute (together with the **K**-axioms) the axiom system of the modal logic **S5** and in the current context, they define the behaviour of our $\exists$ operator, which is obviously intended as a **S5** modality. Axiom **AXM**7 is called the inclusion axiom and it defines the interaction between the modal operators $\langle \mathbf{CC} \rangle$ and $\exists$.

Finally, axioms **AXM**8, **AXM**9 and **AXM**10 aim to syntactically capture the nature of the comparative distance frame conditions **CNT**1, **CNT**2 and **CNT**3 (see Section 3.1.1), respectively.

We will denote deduction in **Ax**CD$_\diamond$ by using the notation $\vdash_{\mathbf{Ax}\mathrm{CD}_\diamond}$. So, for any formula φ which is deductible in **Ax**CD$_\diamond$, we write $\vdash_{\mathbf{Ax}\mathrm{CD}_\diamond} \varphi$ to denote that φ is a theorem of the logic arising from system **Ax**CD$_\diamond$. The following theorem establishes that all theorems of the axiomatic system **Ax**CD$_\diamond$ are tautologies for the class of all comparative distance frames F.

Theorem 3.17 (Soundness)**.** *For every formula φ, we have that $\vdash_{\mathbf{Ax}CD_\diamond} \varphi \Rightarrow \mathsf{F} \models \varphi$.*

Proof. It is sufficient to establish the validity of axioms **AXM**1 - **AXM**10 over arbitrary frames from F. So, let $\mathfrak{F} = \langle \mathsf{W}, \mathsf{CC} \rangle \in \mathsf{F}$ and set $\mathfrak{M} = \langle \mathfrak{F}, \mathsf{V} \rangle$ for some arbitrary valuation V.

While axioms **AXM**1, **AXM**2 and **AXM**3 are the axioms of minimal modal logic **K**, axioms **AXM**4, **AXM**5, **AXM**6 and **AXM**7 are the well-known axioms of **S5**. So, we will skip the well-known proofs for the soundness of these axioms, which are obvious to the mind of the experienced reader. Let us focus on the more interesting axioms **AXM**8, **AXM**9 and **AXM**10.

Let $w \in \mathsf{W}$. In order to establish the soundness of **AXM**8, assume that $\mathfrak{M}, w \models \exists(p \land q)$. So, there is a $u \in \mathsf{W}$ such that $\mathfrak{M}, u \models p \land q$. On the other hand, since $\mathfrak{F}$ is a comparative distance frame, it satisfies frame condition

CNT1. Hence, we derive that $\mathsf{CC}(w, u, u)$. Thus, we get $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(p, q)$, which is what we want.

For the case of **AXM**9, assume that we have $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(p, q)$. We will show that this implies $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(q, p)$. From the hypothesis, it follows that $\exists u \exists v[\mathsf{CC}(w, u, v) \wedge \mathfrak{M}, u \models p$ and $\mathfrak{M}, v \models q]$. Since $\mathfrak{F}$ satisfies **CNT**2, we have that $\mathsf{CC}(w, v, u)$. Hence, $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(q, p)$ as desired.

Finally, to address the case of axiom **AXM**10, first suppose that we have $\mathfrak{M}, w \models \langle \mathbf{CC} \rangle(p, q) \wedge \neg \langle \mathbf{CC} \rangle(r, s)$. From here, it follows that $\exists u \exists v[\mathsf{CC}(w, u, v) \wedge \mathfrak{M}, u \models p$ and $\mathfrak{M}, v \models q]$.

For sake of a contradiction, assume that we have $\mathfrak{M}, w \models \exists \neg [\langle \mathbf{CC} \rangle(r, s) \rightarrow \langle \mathbf{CC} \rangle(p, q)]$. This means that $\exists y[\mathfrak{M}, y \models \langle \mathbf{CC} \rangle(r, s)$ and $\mathfrak{M}, y \models \neg \langle \mathbf{CC} \rangle(p, q)]$. So, it follows that $\exists z \exists t[\mathsf{CC}(y, z, t) \wedge \mathfrak{M}, z \models r$ and $\mathfrak{M}, t \models s]$.

Now let us put the pieces together: Since we have $\mathfrak{M}, w \models \neg \langle \mathbf{CC} \rangle(r, s)$, it follows that $\neg \mathsf{CC}(w, z, t)$. Similarly, since $\mathfrak{M}, y \models \neg \langle \mathbf{CC} \rangle(p, q)$, we conclude that $\neg \mathsf{CC}(y, u, v)$. But then we have that $\mathsf{CC}(w, u, v) \wedge \neg \mathsf{CC}(w, z, t)$ while on the other hand that $\mathsf{CC}(y, z, t) \wedge \neg \mathsf{CC}(y, u, v)$. This clearly violates the frame condition **CNT**3 and it contradicts with the fact that $\mathfrak{F}$ is a comparative distance frame. □

We now turn our attention to the semantic completeness of axiomatic system **Ax**CD$_\diamond$ with respect to the class of all comparative distance frames. Completeness proof exploits the canonical model method based on maximal consistent sets of the logic.

We begin with the construction of the canonical model. Note that, since **Ax**CD$_\diamond$ is a normal modal logic, it must be strongly complete with respect to its canonical model ([28], pg. 199, Theorem 4.22), which can be defined by the following construction.

Construction 3.4. This construction simply builds a model by using the set of all maximal consistent sets in the following well-known way:

- $\mathsf{W} = \{w \mid w$ is a maximal **Ax**CD$_\diamond$-consistent set$\}$;
- For every $w, u, v \in \mathsf{W}$ set:

$$\mathsf{CC}(w, u, v) \text{ iff } \forall \alpha \forall \beta[\alpha \in u \wedge \beta \in v \Rightarrow \langle \mathbf{CC} \rangle(\alpha, \beta) \in w];$$

- For every $p \in \mathscr{P}$ set:

$$\mathsf{V}(p) = \{w \in \mathsf{W} \mid p \in w\}.$$

Finally, canonical frame and canonical model are set as follows:

$$\mathfrak{F} = \left\langle \mathsf{W}, \mathsf{CC} \right\rangle \quad \text{and} \quad \mathfrak{M} = \left\langle \mathfrak{F}, \mathsf{V} \right\rangle.$$

We mention the following lemma and its theorem which will act as the cornerstone of our completeness proof. The proofs are very well-known and standard and can be found in Blackburn et al. [28]. A proof for the Truth Lemma can be found on page 199 (Lemma 4.21) and a proof for the Canonical Model Theorem is again on the same page (Theorem 4.22).

Lemma 3.18 (Truth Lemma)**.** *For every formula φ and state $w \in \mathsf{W}$, we have that $\mathfrak{M}, w \models \varphi \Leftrightarrow \varphi \in w$.*

Combined with the Lindenbaum Lemma, Lemma 3.18 immediately gives the following result:

Theorem 3.19 (Canonical Model Theorem)**.** **Ax***$CD_\diamond$ is strongly complete with respect to $\mathfrak{M}$.*

In order to establish strong completeness of **Ax**$\mathrm{CD}_\diamond$ with respect to the class of all comparative distance frames, all that is needed to be demonstrated is that the canonical frame $\mathfrak{F}$ satisfies the frame conditions **CNT**1 - **CNT**3. The following lemma deals with this issue.

We should also note that this lemma is the reason for using the language $\mathcal{L}[\langle \mathbf{CC} \rangle, \exists]$ instead of simpler $\mathcal{L}$.

Lemma 3.20. *$\mathfrak{F}$ is a comparative distance frame, i.e., it satisfies frame conditions **CNT**1 - **CNT**3.*

Proof. We begin by showing that constraint **CNT**1 is satisfied by the canonical frame $\mathfrak{F}$. Let $w, u \in \mathsf{W}$ and also let φ and ψ be any two formulas such that $\varphi, \psi \in u$. From Lemma 3.18, it follows that $\mathfrak{M}, u \models \varphi \wedge \psi$. Using the well-known canonicity of **S5** we have that $\mathfrak{M}, w \models \exists(\varphi \wedge \psi)$[3]. So, again from Lemma 3.18, it follows that $\exists(\varphi \wedge \psi) \in w$. On the other hand, since w is a maximal consistent set, it must contain the formula $\exists(\varphi \wedge \psi) \rightarrow \langle \mathbf{CC} \rangle(\varphi, \psi)$, which is an instance of **AXM**8. Since maximal consistent sets are closed under modus ponens, it follows that $\langle \mathbf{CC} \rangle(\varphi, \psi) \in w$. From Construction 3.4, we conclude that $\mathsf{CC}(w, u, u)$ as desired.

[3]Without the globality operator $\exists$, this step would not happen

Now let us establish that constraint **CNT**2 is satisfied by $\mathfrak{F}$. Let $w, u, v \in \mathsf{W}$ and assume that $\mathsf{CC}(w, u, v)$. This means that, for all formulas φ' and ψ' we have $\varphi' \in u \wedge \psi' \in v \Rightarrow \langle \mathbf{CC} \rangle(\varphi', \psi') \in w$. Now let $\varphi \in v$ and $\psi \in u$. From the hypothesis, it follows that $\langle \mathbf{CC} \rangle(\psi, \varphi) \in w$. As the formula $\langle \mathbf{CC} \rangle(\psi, \varphi) \rightarrow \langle \mathbf{CC} \rangle(\varphi, \psi)$ is an instance of **AXM**9, it must be contained in w. Using modus ponens, it follows that $\langle \mathbf{CC} \rangle(\varphi, \psi) \in w$. Thus, from Construction 3.4 we get that $\mathsf{CC}(w, v, u)$.

Finally, we address the more interesting case of **CNT**3. Let $w, u, v, y, z \in \mathsf{W}$ and assume that we have $\mathsf{CC}(w, u, v) \wedge \neg\mathsf{CC}(w, y, z)$. From here and from Construction 3.4 it follows that for all formulas φ and ψ, we have that $\varphi \in u \wedge \psi \in v \Rightarrow \langle \mathbf{CC} \rangle(\varphi, \psi) \in w$. On the other hand, we derive that there are formulas $\alpha \in y$ and $\beta \in z$ such that $\langle \mathbf{CC} \rangle(\alpha, \beta) \notin w$ or equivalently, that $\neg\langle \mathbf{CC} \rangle(\alpha, \beta) \in w$ as w is maximal consistent.

For sake of a contradiction, assume that there exists a $t \in \mathsf{W}$ such that $\neg\mathsf{CC}(t, u, v) \wedge \mathsf{CC}(t, y, z)$. So, we have that for all formulas φ and ψ, $\varphi \in y \wedge \psi \in z \Rightarrow \langle \mathbf{CC} \rangle(\varphi, \psi) \in t$. Moreover, it follows that there are formulas $\gamma \in u$ and $\delta \in v$ such that $\neg\langle \mathbf{CC} \rangle(\gamma, \delta) \in t$.

Combining all the information we have gathered so far, first we have that $\langle \mathbf{CC} \rangle(\gamma, \delta) \in w$ and $\neg\langle \mathbf{CC} \rangle(\alpha, \beta) \in w$, which entails that $\langle \mathbf{CC} \rangle(\gamma, \delta) \wedge \neg\langle \mathbf{CC} \rangle(\alpha, \beta) \in w$ since w is maximal consistent. Moreover, since the formula $\langle \mathbf{CC} \rangle(\gamma, \delta) \wedge \neg\langle \mathbf{CC} \rangle(\alpha, \beta) \rightarrow \forall[\langle \mathbf{CC} \rangle(\alpha, \beta) \rightarrow \langle \mathbf{CC} \rangle(\gamma, \delta)]$ is an instance of **AXM**10, using modus ponens we derive that $\forall[\langle \mathbf{CC} \rangle(\alpha, \beta) \rightarrow \langle \mathbf{CC} \rangle(\gamma, \delta)] \in w$. Using Lemma 3.18, it is easy to see that $\langle \mathbf{CC} \rangle(\alpha, \beta) \rightarrow \langle \mathbf{CC} \rangle(\gamma, \delta) \in t$. Since we also have $\langle \mathbf{CC} \rangle(\alpha, \beta) \in t$ from the above, it follows that $\langle \mathbf{CC} \rangle(\gamma, \delta) \in t$, which is a contradiction since $\neg\langle \mathbf{CC} \rangle(\gamma, \delta) \in t$ and t is consistent. This completes the proof of the lemma. □

We summarize our achievement with the following completeness theorem:

Theorem 3.21 (Strong Completeness)**.** **Ax**$CD_\diamond$ *is strongly complete with respect to the class of all comparative distance frames, i.e., for every formula* φ *we have that* $\mathsf{F} \models \varphi \Rightarrow \vdash_{\mathbf{Ax}CD_\diamond} \varphi$.

Proof. Follows directly from Theorem 3.19 and Lemma 3.20. □

4. Conclusion, Related Work & Future Research

Our goal was to develop *computationally feasible* logical formalisms which can talk about distance information and investigate their logical and computational properties. We chose to base this study on Theodore de Laguna's qualitative notion of 'can-connect' [1] and it included a first-order logic and a modal logic formalism which utilised Laguna's framework to talk about comparative distances between spatial entities or, as Laguna himself put it, 'solids'.

The expressive strength of the modal language is naturally limited. But this was a goal rather than a deficiency as our main target was to create a computationally feasible formalism. As a matter of fact, this simplistic framework not only benefited us in terms of computational efficiency, but it also enabled design of languages which do not need any embedded numerical parameters or similar additional syntax to talk about distances. Moreover, it helped us to work with distance information using the cognitively plausible notion of 'spatial entities' in contrast to the theoretically motivated 'points'.

Our theoretical investigations have established that, while the first-order comparative distance logic is finitely axiomatisable but undecidable, modal comparative distance logic is finitely axiomatisable, enjoys the finite model property and decidable. Moreover, we established that the satisfiability problem of modal comparative distance logic is NP-complete.

In comparison to the 'logics of metric spaces' [17] which have only some decidable fragments whose computational complexity have NEXPTIME upper bounds (lower bounds remain an open problem), this work presents a much less expressive modal formalism which is computationally much more feasible. Wolter and Zakharyaschev's 'logic of metric and topology' [11], which targets reasoning with the induced topologies of metric spaces to create a formalism where qualitative and quantitative approaches coexist, was shown to be EXPTIME-complete and it is thus computationally also much more costly. But surprisingly enough, modal comparative distance logic can manage something any of mentioned relatively much more expressive logics can not: It is quite easy to compare the distance between two spatial entities with modal comparative distance logic, whereas in the case of aforementioned logics this is not possible at all unless the language is modified in a non-trivial way [11]. We can simply write the formula to state that 'a is closer to b than c is to d':

$$\forall[\langle \mathbf{CC} \rangle(a, b) \rightarrow \langle \mathbf{CC} \rangle(c, d)].$$

Looking at the future research potential, the first thing that comes to mind when a modal logic is to be expressively strengthened, is to consider its hybrid extension. One of the immediate results of such an extension would be that, we will be able to express when two solids overlap with the following formula:

$$\forall\langle \mathbf{CC}\rangle(\mathsf{i},\mathsf{j})$$

where i and j are nominal letters. It is hard to go any further than this using modal satisfiability. For example, defining the subset relation between two solids is beyond the limits of the modal language. However, modal logic gives us the concept of validity,the dual of satisfiability, with which we can get a better grasp of the frame level relationships. Using validity we can now check if a is a subset of b as follows:

$$\forall\langle \mathbf{CC}\rangle(a,\mathsf{i}) \rightarrow \forall\langle \mathbf{CC}\rangle(b,\mathsf{j})$$

where i and j are nominal letters and, a and b are constant symbols. The questions that follow are the expected: is such an extension going to be decidable and if so, what its computational complexity going to be?

Finally, spatial reasoning is widely known for its use of mereotopological relations. Despite of a past attempt to create a modal logic for such relationships which ended with an undecidable modal logic [29], recently it was shown that a decidable modal logic of such relations can be actually formed [30]. Keeping in mind that there is an obvious limitation for talking about distance information using mere topological relationships, the question regarding future research potential here is whether can-connect primitive can be contained together with mereotopological relations within a decidable modal logic or not. And if so, what would be the computational cost of mixing them up?

References

[1] T. de Laguna, Point, line and surface as sets of points, The Journal of Philosophy 19 (1922) 449–461.

[2] T. de Laguna, The nature of space -part one, The Journal of Philosophy 19 (1922) 393–407.

[3] T. de Laguna, The nature of space -part two, The Journal of Philosophy 19 (1922) 421–440.

[4] A. Tarski, Foundations of the geometry of solids, in: J. H. Woodger (Ed.), Logic, Semantics, Metamathematics, Oxford Clarendon Press, 1956, pp. 24–29.

[5] B. L. Clarke, Individuals and points, Notre Dame Journal of Formal Logic 26 (1985) 61–75.

[6] N. M. Gotts, J. M. Gooday, A. G. Cohn, A connection based approach to common-sense topological description and reasoning, The Monist 79 (1996) 51–75.

[7] I. Pratt, O. Lemon, Ontologies for plane, polygonal mereotopology, Notre Dame Journal of Formal Logic 38 (1997) 225–245.

[8] A. G. Cohn, B. Bennett, J. M. Gooday, N. M. Gotts, Representing and reasoning with qualitative spatial relations about regions, in: O. Stock (Ed.), Temporal and spatial reasoning, Kluwer, 1997.

[9] A. G. Cohn, B. Bennett, J. M. Gooday, N. Gotts, RCC: a calculus for region based qualitative spatial reasoning, GeoInformatica 1 (1997) 275–316.

[10] P. Simons, Parts: A Study in Ontology, Clarendon Press, Oxford, 1987.

[11] F. Wolter, M. Zakharyaschev, A logic of metric and topology, Journal of Symbolic Logic (2005).

[12] N. Rescher, J. Garson, Topological logic, Journal of Symbolic Logic 33 (1968) 537–548.

[13] G. H. von Wright, A modal logic of place, in: E. Sosa (Ed.), The Philosophy of Nicholas Rescher, Reidel, Dordrecht, 1974, pp. 65–73.

[14] K. Segerberg, A note on the logic of elsewhere, Theoria 46 (1980) 183–187.

[15] R. Jansana, Some logics related to von wright's logic of place, Notre Dame Journal of Formal Logic 35 (1994) 88–98.

[16] O. Lemon, I. Pratt, On the incompleteness of modal logics of space: Advancing complete modal logics of place, in: M. Kracht, M. D. Rijke, H. Wansing, M. Z. (eds.), M. Zakharyaschev, O. Lemon, I. Pratt (Eds.), Advances in Modal Logic '96.

[17] O. Kutz, H. Sturm, N. Suzuki, F. Wolter, M. Zakharyaschev, Logics of metric spaces, ACM Transactions on Computational Logic 4 (2003) 260–294.

[18] D. Dubois, H. Prade, F. Esteva, P. Garcia, L. Godo, A logical approach to interpolation based on similarity relations, International Journal of Approximate Reasoning 17 (1997) 1–36.

[19] F. Esteva, P. Garcia, L. Godo, R. Rodriguez, A modal account of similarity-based reasoning, International Journal of Approximate Reasoning 16 (1997) 235–260.

[20] P. Clote, R. Backofen, Computational Molecular Biology, John Wiley & Sons, Chichester, 2000.

[21] A. G. Cohn, S. M. Hazarika, Qualitative spatial representation and reasoning: An overview, Fundamenta Informaticae 46 (2001) 1–29.

[22] N. Asher, L. Vieu, Toward a geometry for common sense: A semantics and a complete axiomatization for mereotopology, in: C. Mellish (Ed.), Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, Morgan Kaufmann, San Francisco, 1995, pp. 846–852.

[23] O. J. Lemon, Semantical foundations of spatial logics, in: L. C. Aiello, J. Doyle, S. Shapiro (Eds.), Principles of Knowledge Representation and Reasoning (KR96), Morgan Kaufmann, San Francisco, CA., 1996, pp. 212–219.

[24] D. A. Randell, Z. Cui, A. G. Cohn, A spatial logic based on regions and connection, in: B. Nebel, C. Rich, W. Swartout (Eds.), Principles of Knowledge Representation and Reasoning (KR92), Morgan Kaufmann, San Mateo, 1992, pp. 165–176.

[25] B. Bennett, Carving up space: steps towards construction of an absolutely complete theory of spatial regions, in: L. P. J.J. Alfres, E. Orlowska (Eds.), Proceedings of JELIA'96, pp. 337–353.

[26] A. Grzegorczyk, Undecidability of some topological theories, Fundamenta Mathematicae 38 (1951) 137–152.

[27] M. O. Rabin, A simple method for undecidability proofs and some applications., in: Y. Bar-Hillel (Ed.), Logic and Methodology of Sciences, North Holland, 1965, pp. 58–68.

[28] P. Blackburn, M. de Rijke, Y. Venema, Modal Logic, Cambridge Tracts in Theoretical Computer Science, Cambridge University Press, Cambridge, 2002.

[29] C. Lutz, F. Wolter, Modal logics of topological relations, Logical Methods in Computer Science 2 (2006).

[30] Y. Nenov, D. Vakarelov, Modal logics for mereotopological relations, in: C. Areces, R. Goldblatt (Eds.), Advances in Modal Logic, College Publications, 2008, pp. 249–272.