

# DISTANCES BETWEEN FORMAL THEORIES

MICHÈLE FRIEND, MOHAMED KHALED, KOEN LEFEVER AND GERGELY SZÉKELY

ABSTRACT. In the literature, there have been several methods and definitions for working out if two theories are “equivalent” (essentially the same) or not. In this article, we do something subtler. We provide means to measure distances (and explore connections) between formal theories. We introduce two main notions for such distances. The first one is that of *axiomatic distance*, but we argue that it might be of limited interest. The more interesting and widely applicable notion is that of *conceptual distance* which measures the minimum number of concepts that distinguish two theories. For instance, we use conceptual distance to show that relativistic and classical kinematics are distinguished by one concept only. We also develop further notions of distance, and we include a number of suggestions for applying and extending our project.

## 1. INTRODUCTION

It is well known that the theory of strict partial orders and the theory of partial orders are “equivalent”, i.e., they have the same essential content. To capture this vague idea, defining a precise equivalence between theories, several formal definitions have been made, e.g., *logical equivalence*, *definitional equivalence*, *categorical equivalence*, etc. Which theory is equivalent to which other theory depends on the point of view from which one decides to explore the equivalence between the theories in question.

In the last few decades, the concept of equivalence between theories (henceforth: “theory-equivalence”) has become important for studying the connections between formal theories. Many interesting results have been derived from investigating such equivalence, e.g., [Andréka et al., 2005], [Barrett and Halvorson, 2016], [Japaridze and Jongh, 1998], [Pinter, 1978] and [Visser, 2006]. We can also look at the question starting from non-equivalence. Given two non-equivalent theories (according to any chosen definition of theory-equivalence), some natural questions arise: (1) Can these theories be modified into equivalent theories (in a non-trivial way)? (2) If this can be done, can we do it in finitely many steps? In other words, what is the degree of their non-equivalence?

In this article, we lay down the first steps of a research programme to answer these questions. In order to investigate some ways to measure how far two theories are from each other; we introduce a framework that can give a qualitative and quantitative analysis of the connections between formal theories. We focus on formal theories that are formulated in any of the following logical systems: sentential logic, ordinary first order logic (FOL), finite variables fragments of FOL and/or infinitary versions of FOL. We develop several notions for distances between theories, we discuss these notions and we make comparisons between them.

---

2010 *Mathematics Subject Classification*. Primary 03B99, 03C07, 03A10. Secondary 03G99, 03B80.

*Key words and phrases*. Network of theories, degrees of non-equivalence, conceptual distance, relativistic and classical kinematics.

The idea is very simple: based on a symmetric relation capturing a notion of minimal change, we introduce a general way to define a distance on any class of objects (not just theories) equipped with an equivalence relation. The idea is a generalization of the distance between any two nodes in the same graph, in graph theory. After, we give particular examples when the given class is a class of theories and the equivalence relation is a fixed notion of theory-equivalence.

The first particular example, is that of logical equivalence. As a measure for the degree of logical non-equivalence, we introduce the concept of *axiomatic distance*. The idea is to count the minimum number of axioms that are needed to be added or “removed” to get from one theory to the other.<sup>1</sup> Since any finite number of axioms can be concatenated by conjunction resulting in only one axiom, one may think that the axiomatic distance, if it is finite, between two given theories  $T$  and  $T'$  must be  $\leq 2$ , i.e., we need at most two steps to get  $T$  from  $T'$ : one step for axiom addition and another one for axiom removal. This is why we have the intuition that axiomatic distance may not be very interesting, cf., Problem 1 and Theorem 3.11.

Then we turn to definitional equivalence. Two theories are definitionally equivalent if they cannot be distinguished by a concept (a formula defining some notion). As a measure for the degree of definitional non-equivalence, we define *conceptual distance*. This distance counts the minimum number of *concepts* that separate two theories. We find that this distance is of special interest in the study of logic. We give examples and we count conceptual distance between some specific theories, see, e.g., Theorem 4.7 and Theorem 4.10. We also explore a connection between conceptual distance and *spectrum of theories* which is a central topic in model theory, cf., Theorem 4.11.

In algebraic logic, Lindenbaum-Tarski algebras of logical theories (sometimes these are called concept algebras) are often introduced as the algebras of different concepts of the corresponding theories. Thus, counting concepts amounts to counting elements of Lindenbaum-Tarski algebras. In fact, our definition of conceptual distance herein is a careful translation of an algebraic distance between Lindenbaum-Tarski algebras. Such an algebraic distance allows us to define conceptual distance between theories in any algebraizable logic, e.g., modal logic and intuitionistic logic. This general algebraic distance (and its application on concept algebras) is planned to be investigated in details in a forthcoming algebra oriented paper.

Furthermore, we investigate the possible application of conceptual distance in the logical foundation of physical theories in ordinary first order logic. We prove that conceptual distance between classical and relativistic kinematics is one. In other words, only one concept distinguishes classical and relativistic kinematics: the existence of a class of observers who are at absolute rest. This is indeed an interesting result in its own right, not only for logicians but also for physicists. Such a result opens several similar questions about how many concepts (and what are they) differentiate two physical theories when their phenomena can be described in FOL.

In philosophy of physics, this might be important because, on the one hand it is clear that we are not presently converging towards one unified theory of physics in the sense of converging to one set of laws from which all the phenomena of physics can be derived. On the other hand, we can give logical foundation to several physical theories: Newtonian mechanics, relativity theories and some parts of quantum theory. Given these logical representations, we would like to know the exact relationship between physical theories. If we know this, then we can form an impression of how far we are from such a philosophical dream – the dream of the unity of physics. Or, we can adjust

---

<sup>1</sup>By “removing an axiom” here we only mean the trivial converse of adding an axiom in the following sense:  $T$  is a theory resulting from “removing” one axiom from  $T'$  if  $T'$  can be reached from  $T$  by adding one axiom.

our hopes and expectations, and rest content with a unity of science at a more general level: as a network of logical theories with precise relations between them.

With definitions and metrics on distance developed here, we have maps of the network of logical theories. When we draw such maps of networks, the topology may suggest very interesting and fruitful questions. For instance: if there is a distance other than zero or one, then is there already a known theory in between? or if not, we can ask what are the limitative properties of that theory and what is its philosophical significance? By engaging in such studies, we see the “edge” of the limitative results, and by examining this edge we more precisely understand the rapport between meta-logical limitative results and physical phenomena.

In the present paper, we assume familiarity with the basic notions of set theory. For instance, what is a set, a class, a relation, etc. The only difference is that in several occurrences in this paper, we decided not to distinguish different kinds of infinities. Therefore, together with the standard notion of cardinality, we are going to speak about the *size of set*  $X$ , defined as follows:

$$\|X\| \stackrel{\text{def}}{=} \begin{cases} k & X \text{ is finite and has exactly } k\text{-many elements,} \\ \infty & \text{if } X \text{ is an infinite set.} \end{cases}$$

We also make use of von Neumann ordinals. For example,  $\omega$  is the smallest infinite ordinal, sometimes we denote  $\omega$  by  $\mathbb{N}$  to indicate that it is the set of natural numbers (non-negative integers).

## 2. NOTIONS OF LOGIC

In the course of this paper, let  $\alpha$  and  $\beta \leq \alpha + 1$  be two fixed ordinals. We consider a natural generalisation of ordinary first order logic, we denote it by  $\mathbf{L}_\alpha^\beta$ , which is inspired from the definitions and the discussions in [Henkin et al., 1985, section 4.3]. Roughly, the formulas of  $\mathbf{L}_\alpha^\beta$  uses a fixed set of individual variables  $\{v_i : i \in \alpha\}$  and relation symbols of rank strictly less than  $\beta$ . For simplicity, we assume that our languages do not contain any function symbols and/or constant symbols.

In particular,  $\mathbf{L}_0^1$  is sentential (propositional) logic, while  $\mathbf{L}_\omega^\omega$  is ordinary first order logic. The so-called finite variables fragments of first order logic are the logics  $\mathbf{L}_n^{n+1}$ , for finite ordinals  $n$ 's. When  $\alpha$  and  $\beta$  are infinite,  $\mathbf{L}_\alpha^\beta$  is called infinitary logic. Throughout, since  $\alpha$  and  $\beta$  are fixed, languages, theories, etc., are understood to be languages for  $\mathbf{L}_\alpha^\beta$ , theories in  $\mathbf{L}_\alpha^\beta$ , etc.

**2.1. The syntax of  $\mathbf{L}_\alpha^\beta$ .** More precisely, A *language*  $\mathcal{L}$  for  $\mathbf{L}_\alpha^\beta$  is a set of relation symbols such that each relation symbol  $R \in \mathcal{L}$  is assigned a rank  $\text{rank}(R) < \beta$ . Relation symbols of rank 0 are called *sentential constants*. To construct the formulas of language  $\mathcal{L}$ , we also need some other symbols: equality “=” (we deal with quantifier logics with identity), brackets “(” and “)”, conjunction “ $\wedge$ ”, negation “ $\neg$ ” and the existential quantifier “ $\exists$ ”. We also use the necessary symbols to write sequences of variables  $(v_{i_m} : m \in I)$ , for any indexing set  $I \subseteq \alpha$ . *The set of formulas*  $\text{Fm}$  of  $\mathcal{L}$  is the smallest set that satisfies:

- (a)  $\text{Fm}$  contains each *basic formula* of  $\mathcal{L}$ , where the basic formulas are the following two types of formulas:
  - (i) The equalities  $v_i = v_j$ , for any  $i, j \in \alpha$ .
  - (ii)  $R(v_{i_m} : m < \text{rank}(R))$ , for any relation symbol  $R$ .
- (b)  $\text{Fm}$  contains  $\varphi \wedge \psi$ ,  $\neg\varphi$  and  $\exists v_i \varphi$ , for each  $\varphi, \psi \in \text{Fm}$ .

As usual, we use the following abbreviations.

- If  $P$  is a sentential constant, then we just write  $P$  instead of  $P()$ .
- If  $R$  is a relation symbol of finite positive rank, say  $k$ , then we write  $R(v_{i_0}, \dots, v_{i_{k-1}})$  instead of  $R(v_{i_m} : m < k)$ .
- We use disjunction, implication, equivalence and universal quantifier:

$$\begin{aligned} \varphi \vee \psi &\stackrel{\text{def}}{=} \neg(\neg\varphi \wedge \neg\psi) & \varphi \rightarrow \psi &\stackrel{\text{def}}{=} \neg(\varphi \wedge \neg\psi) \\ \varphi \leftrightarrow \psi &\stackrel{\text{def}}{=} (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi) & \forall v_i \varphi &\stackrel{\text{def}}{=} \neg(\exists v_i \neg\varphi) \end{aligned}$$

- We also use grouped conjunction and disjunction: Empty disjunction is defined to be  $\varphi \wedge \neg\varphi$  and empty conjunction is defined to be  $\varphi \vee \neg\varphi$  (for any arbitrary but fixed formula  $\varphi \in \text{Fm}$ ). Let  $\varphi_0, \dots, \varphi_m \in \text{Fm}$ , then

$$\bigvee_{0 \leq i \leq m} \varphi_i \stackrel{\text{def}}{=} \varphi_0 \vee \dots \vee \varphi_m \quad \text{and} \quad \bigwedge_{0 \leq i \leq m} \varphi_i \stackrel{\text{def}}{=} \varphi_0 \wedge \dots \wedge \varphi_m.$$

**2.2. The semantics of  $L_\alpha^\beta$ .** A *model*  $\mathfrak{M}$  for language  $\mathcal{L}$  is a non-empty set  $M$  enriched with operations  $R^{\mathfrak{M}} \subseteq M^{\text{rank}(R)}$ , for each  $R \in \mathcal{L}$  (for a sentential constant  $P$ ,  $P^{\mathfrak{M}} \subseteq M^0 = \{\emptyset\}$ ).<sup>2</sup> An *assignment* in  $\mathfrak{M}$  is a function  $\tau$  that assigns for each variable an element of the set  $M$ . Let  $\varphi \in \text{Fm}$  be any formula. The satisfiability relation  $\mathfrak{M}, \tau \models \varphi$  is defined recursively as follows:

$$\begin{aligned} \mathfrak{M}, \tau \models R(v_{i_m} : m < \text{rank}(R)) &\text{ iff } (\tau(v_{i_m}) : m < \text{rank}(R)) \in R^{\mathfrak{M}},^3 \\ \mathfrak{M}, \tau \models v_i = v_j &\text{ iff } \tau(v_i) = \tau(v_j), \\ \mathfrak{M}, \tau \models \varphi \wedge \psi &\text{ iff } \mathfrak{M}, \tau \models \varphi \text{ and } \mathfrak{M}, \tau \models \psi, \\ \mathfrak{M}, \tau \models \neg\varphi &\text{ iff } \mathfrak{M}, \tau \not\models \varphi, \\ \mathfrak{M}, \tau \models \exists v_i \varphi &\text{ iff there is } a \in M \text{ such that } \mathfrak{M}, \tau[v_i \mapsto a] \models \varphi, \end{aligned}$$

where  $\tau[v_i \mapsto a]$  is the assignment which agrees with  $\tau$  on every variable except  $\tau[v_i \mapsto a](v_i) = a$ . The cardinality of  $\mathfrak{M}$  is defined to be the cardinality of  $M$ . A formula  $\varphi$  is said to be *true* in  $\mathfrak{M}$ , in symbols  $\mathfrak{M} \models \varphi$ , iff  $\mathfrak{M}, \tau \models \varphi$ , for every assignment  $\tau$  in  $\mathfrak{M}$ . A formula  $\varphi$  is said to be a *tautology* iff it is true in every model for  $\mathcal{L}$ . The *theory* of  $\mathfrak{M}$  is defined as:

$$\text{Th}(\mathfrak{M}) \stackrel{\text{def}}{=} \{\varphi \in \text{Fm} : \mathfrak{M} \models \varphi\}.$$

We say that *two models*  $\mathfrak{M}$  and  $\mathfrak{N}$  for language  $\mathcal{L}$  are *isomorphic* iff there is a bijection  $f : M \rightarrow N$  between their underlying sets that respects the meaning of the relation symbols, i.e., for each  $R \in \mathcal{L}$ ,

$$(a_i : i < \text{rank}(R)) \in R^{\mathfrak{M}} \iff (f(a_i) : i < \text{rank}(R)) \in R^{\mathfrak{N}}.$$

### 2.3. Theories in the logic $L_\alpha^\beta$ .

**Definition 2.1.** Suppose that  $\mathcal{L}$  is a language and let  $\text{Fm}$  be its set of formulas. A *theory*  $T$  of  $\mathcal{L}$  is a set of formulas (subset of  $\text{Fm}$ ).

We use the same superscripts and subscripts for theories and their corresponding languages and formulas. For example, if we write  $T'$  is a theory, then we understand that  $T'$  is a theory of language  $\mathcal{L}'$  whose set of formulas is  $\text{Fm}'$ . A *model* for theory  $T$  is a model for  $\mathcal{L}$  in which every  $\psi \in T$  is true. We say that theory  $T$  is *consistent* iff there is at least one model for  $T$ .

<sup>2</sup>So the meaning  $P^{\mathfrak{M}}$  of a sentential constant  $P$  can be either *true* ( $T = \{\emptyset\}$ ) or *false* ( $F = \emptyset$ ).

<sup>3</sup>If  $\text{rank}(P)$  is 0, then  $(\tau(v_{i_m}) : m < \text{rank}(P))$  is the empty sequence  $\emptyset$ . Hence  $\mathfrak{M}, \tau \models P$  iff  $P^{\mathfrak{M}}$  is true.

**Definition 2.2.** Let  $T$  be a theory and let  $\kappa$  be any cardinal. The *spectrum* of  $T$ , in symbols  $I(T, \kappa)$ , is the number of its different models (up to isomorphism) of cardinality  $\kappa$ . This number is defined to be  $\infty$  if  $T$  has infinitely many non-isomorphic models of cardinality  $\kappa$ .

We say that a formula  $\varphi$  is a *theorem* of theory  $T$ , in symbols  $T \models \varphi$ , iff  $\varphi$  is true in every model for  $T$ . The set of *consequences* of theory  $T$  is defined as follows:

$$\text{Cn}(T) \stackrel{\text{def}}{=} \{\varphi \in \text{Fm} : T \models \varphi\}.$$

**Definition 2.3.** Two theories  $T_1$  and  $T_2$  are called *logically equivalent*, in symbols  $T_1 \equiv T_2$ , iff they have the same consequences, i.e.,  $\text{Cn}(T_1) = \text{Cn}(T_2)$ .

**2.4. More notions for theory-equivalence.** A *translation* of language  $\mathcal{L}_1$  into language  $\mathcal{L}_2$  is a map  $\text{tr} : \text{Fm}_1 \rightarrow \text{Fm}_2$  such that the following are true for every  $\varphi, \psi \in \text{Fm}$  and every  $v_i, v_j$ .

- $\text{tr}(v_i = v_j)$  is  $v_i = v_j$ .
- $\text{tr}$  commutes with the Boolean connectives:

$$\text{tr}(\neg\varphi) = \neg\text{tr}(\varphi) \quad \text{and} \quad \text{tr}(\varphi \wedge \psi) = \text{tr}(\varphi) \wedge \text{tr}(\psi).$$

- Finally,  $\text{tr}(\exists v_i \varphi) = \exists v_i \text{tr}(\varphi)$ .<sup>4</sup>

**Definition 2.4.** Suppose that  $T_1$  and  $T_2$  are theories in languages  $\mathcal{L}_1$  and  $\mathcal{L}_2$ , respectively, and  $\text{tr}$  is a translation of  $\mathcal{L}_1$  into  $\mathcal{L}_2$ . The translation  $\text{tr}$  is said to be an *interpretation* of  $T_1$  into  $T_2$  iff it maps theorems of  $T_1$  into theorems of  $T_2$ , i.e., for each formula  $\varphi \in \text{Fm}_1$ ,

$$T_1 \models \varphi \implies T_2 \models \text{tr}(\varphi).$$

- (a) An interpretation  $\text{tr}$  of  $T_1$  into  $T_2$  is called a *faithful interpretation* of  $T_1$  into  $T_2$  iff for each formula  $\varphi \in \text{Fm}_1$ ,

$$T_1 \models \varphi \iff T_2 \models \text{tr}(\varphi).$$

- (b) An interpretation  $\text{tr}_{12}$  of  $T_1$  into  $T_2$  is called a *definitional equivalence* between  $T_1$  and  $T_2$  iff there is an interpretation  $\text{tr}_{21}$  of  $T_2$  into  $T_1$  such that

- $T_1 \models \text{tr}_{21}(\text{tr}_{12}(\varphi)) \leftrightarrow \varphi$ ,
- $T_2 \models \text{tr}_{12}(\text{tr}_{21}(\psi)) \leftrightarrow \psi$ .

for every  $\varphi \in \text{Fm}_1$  and  $\psi \in \text{Fm}_2$ . In this case,  $\text{tr}_{21}$  is also a definitional equivalence.

**Definition 2.5.** Two theories  $T_1$  and  $T_2$  are said to be *definitionaly equivalent*, in symbols  $T_1 \rightleftharpoons T_2$ , iff there is a definitional equivalence between them.

In the literature, there are several ways to define definitional equivalence. Here, we use a variant of the definition in [Henkin et al., 1985, Definition 4.3.42 and Theorem 4.3.43]. For a discussion on the different definitions of definitional equivalence, see [Lefever and Székely, 2018b], and we refer to [Visser, 2006] for a category theory based discussion.

<sup>4</sup> In the case of ordinary first order logic (when  $\alpha = \beta = \omega$ ), to define a translation  $\text{tr} : \text{Fm}_1 \rightarrow \text{Fm}_2$ , it suffices to define  $\text{tr}$  on the basic formulas in  $\text{Fm}_1$  of the form  $v_i = v_j$  and  $R(v_0, \dots, v_{m-1})$ . Then, using Tarski's substitution observation, we can define

$$\begin{aligned} \text{tr}(R(v_{i_1}, \dots, v_{i_m})) &= \exists v_0 (v_0 = y_1 \wedge \dots \wedge \exists v_{m-1} (v_{m-1} = y_m \wedge \\ &\quad \exists y_1 (y_1 = v_{i_1} \wedge \dots \wedge \exists y_m (y_m = v_{i_m} \wedge \text{tr}(R(v_0, \dots, v_{m-1})))))), \end{aligned}$$

where  $y_i = v_{l+i}$  and  $l$  is the maximum of  $0, \dots, m-1, i_1, \dots, i_m$ . This can be extended in a unique way to a translation that covers the whole  $\text{Fm}_1$ .

**Proposition 2.6.** Let  $T_1$  and  $T_2$  be two theories and suppose that  $\text{tr}_{12} : \text{Fm}_1 \rightarrow \text{Fm}_2$  is a definitional equivalence between  $T_1$  and  $T_2$ , then  $\text{tr}_{12}$  is also a faithful interpretation.

*Proof.* Let  $T_1$  and  $T_2$  be two theories, and let  $\text{tr}_{12} : \text{Fm}_1 \rightarrow \text{Fm}_2$  be a definitional equivalence between them. Let  $\varphi \in \text{Fm}_1$ , we should show that  $T_1 \models \varphi$  iff  $T_2 \models \text{tr}_{12}(\varphi)$ . Since  $\text{tr}_{12}$  is an interpretation, we have that  $T_1 \models \varphi$  implies  $T_2 \models \text{tr}_{12}(\varphi)$ . To show the converse, let us assume that  $T_2 \models \text{tr}_{12}(\varphi)$ . By Definition 2.4, there is an interpretation  $\text{tr}_{21}$  of  $T_2$  into  $T_1$  such that  $T_1 \models \text{tr}_{21}(\text{tr}_{12}(\varphi)) \leftrightarrow \varphi$ . Since  $\text{tr}_{21}$  is an interpretation and  $T_2 \models \text{tr}_{12}(\varphi)$ , we have  $T_1 \models \text{tr}_{21}(\text{tr}_{12}(\varphi))$ . Consequently,  $T_1 \models \varphi$  since  $T_1 \models \text{tr}_{21}(\text{tr}_{12}(\varphi)) \leftrightarrow \varphi$ ; and this is what we need to show.  $\square$

**Definition 2.7.** Let  $T_1$  and  $T_2$  be two theories. We say that  $T_2$  is a *conservative extension* of  $T_1$ , in symbols  $T_1 \sqsubseteq T_2$ , iff  $\text{Fm}_1 \subseteq \text{Fm}_2$  and, for all  $\varphi \in \text{Fm}_1$ ,  $T_2 \models \varphi \iff T_1 \models \varphi$ .

We note that  $T_1 \sqsubseteq T_2$  iff the identity translation  $\text{id} : \text{Fm}_1 \rightarrow \text{Fm}_2$  is a faithful interpretation. It is also worth mentioning that  $T_1 \sqsubseteq T_2 \iff T_1 \equiv \text{Cn}(T_2) \cap \text{Fm}_1$ .

### 3. CLUSTER NETWORKS & STEP DISTANCE

Now, we introduce a general way of defining a distance on any given class  $X$ . We note that our target is to define distances on the class of all theories, thus we need to work with classes which are not necessarily sets.

**Definition 3.1.** By a **cluster**  $(X, E)$  we mean a class  $X$  equipped with an equivalence relation  $E$ .<sup>5</sup>

We are interested in distances according to which some different objects are indistinguishable. Indeed, it is natural to treat equivalent theories as if they were of distance 0 from each other. As we mentioned in the introduction, there are several notions of equivalence between theories. Such equivalence thus can be represented in the cluster of theories by the relation  $E$ .

**Definition 3.2.** A **cluster network** is a triple  $(X, E, S)$ , where  $(X, E)$  is a cluster and  $S$  is a symmetric relation on  $X$ .<sup>6</sup>

Given a cluster network  $(X, E, S)$ . A **path leading from**  $T \in X$  to  $T' \in X$  in  $(X, E, S)$  is a finite sequence  $b_1, \dots, b_m$  of 0's and 1's such that there is a sequence  $T_0, \dots, T_m$  of members of  $X$  with  $T_0 = T$ ,  $T_m = T'$  and, for each  $1 \leq i \leq m$ ,

$$b_i = 0 \iff T_{i-1} E T_i \quad \text{and} \quad b_i = 1 \iff T_{i-1} S T_i.$$

The **length of this path** is defined to be  $\sum_{i=1}^m b_i$ . Two objects  $T, T' \in X$  are **connected** in  $(X, E, S)$  iff there is a path leading from one of them to the other in  $(X, E, S)$ .

**Definition 3.3.** Let  $\mathcal{X} = (X, E, S)$  be a cluster network. The **step distance** on  $\mathcal{X}$  is the function  $d_{\mathcal{X}} : X \times X \rightarrow \mathbb{N} \cup \{\infty\}$  defined as follows. For each  $T, T' \in X$ :

- If  $T$  and  $T'$  are not connected in  $(X, E, S)$ , then  $d_{\mathcal{X}}(T, T') \stackrel{\text{def}}{=} \infty$ .

<sup>5</sup>All definitions in this section can be formulated within von Neumann–Bernays–Gödel set theory (NBG). Of course, ordered pairs of proper classes cannot be formulated even in NBG, but we do not really need ordered pairs here. Our definitions can be understood as follows: “for all classes  $X, E$ , etc., having certain properties there are classes  $d$ , etc., such that...”. We use the notations  $(-, -)$  only to make our statements easier to be read and our proofs easier to be followed.

<sup>6</sup>See footnote 5 above.

- If  $T$  and  $T'$  are connected in  $(X, E, S)$ , then

$$d_{\mathcal{X}}(T, T') \stackrel{\text{def}}{=} \min\{k \in \mathbb{N} : \exists \text{ a path leading from } T \text{ to } T' \text{ whose length is } k\}.$$

The equivalence relation  $E$  represents pairs that cannot be distinguished by the step distance, while the symmetric relation  $S$  represents the pairs of objects that are (at most) one step away from each other. The step distance then counts the minimum number of steps needed to reach an object starting from another one. We may need to say that infinitely many steps are needed, so we allow  $\infty$  in the range of the step distance.

**Example 3.4.** Let  $X$  be any class, let  $E$  be the identity relation and let  $S = X \times X$ . Then,  $\mathcal{X} = (X, E, S)$  is a cluster network and its step distance is the following discrete distance:

$$d_{\mathcal{X}}(T, T') = \begin{cases} 0 & \text{if } T = T', \\ 1 & \text{if } T \neq T'. \end{cases}$$

**Theorem 3.5.** Let  $\mathcal{X} = (X, E, S)$  be a cluster network and let  $d_{\mathcal{X}} : X \times X \rightarrow \mathbb{N} \cup \{\infty\}$  be the step distance on  $\mathcal{X}$ . The following are true for each  $T_1, T_2, T_3 \in X$ :

- (a)  $d_{\mathcal{X}}(T_1, T_2) \geq 0$ , and  $d_{\mathcal{X}}(T_1, T_2) = 0 \iff T_1 E T_2$ .
- (b)  $d_{\mathcal{X}}(T_1, T_2) = d_{\mathcal{X}}(T_2, T_1)$ .
- (c)  $d_{\mathcal{X}}(T_1, T_2) \leq d_{\mathcal{X}}(T_1, T_3) + d_{\mathcal{X}}(T_3, T_2)$ .

*Proof.* Let  $\mathcal{X} = (X, E, S)$  and  $d_{\mathcal{X}}$  be as required. Let  $T_1, T_2, T_3 \in X$ .

- (a) Clearly,  $d_{\mathcal{X}}(T_1, T_2) \geq 0$  for any two  $T_1, T_2 \in X$ , and  $d_{\mathcal{X}}(T_1, T_2) = 0$  if  $T_1 E T_2$  because then 0 is a path from  $T_1$  to  $T_2$  in  $\mathcal{X}$ . If  $d_{\mathcal{X}}(T_1, T_2) = 0$ , then there is a path  $0, \dots, 0$  from  $T_1$  to  $T_2$  in  $\mathcal{X}$ . So there is a sequence  $T'_0, \dots, T'_m \in X$  such that  $T_1 = T'_0$ ,  $T_2 = T'_m$  and  $T'_{i-1} E T'_i$  for each  $1 \leq i \leq m$ . Hence,  $T_1 E T_2$  since  $E$  is transitive.
- (b) The symmetry is satisfied because  $E$  and  $S$  are symmetric relations. Hence, if  $b_1, \dots, b_m$  is a path leading from  $T_1$  to  $T_2$  in  $\mathcal{X}$ , then  $b_m, \dots, b_1$  is a path leading from  $T_2$  to  $T_1$  in  $\mathcal{X}$ .
- (c) The triangle inequality  $d_{\mathcal{X}}(T_1, T_2) \leq d_{\mathcal{X}}(T_1, T_3) + d_{\mathcal{X}}(T_3, T_2)$  follows from the definition because, if  $b_1, \dots, b_m$  is a path leading from  $T_1$  to  $T_3$  in  $\mathcal{X}$  and  $c_1, \dots, c_k$  is a path leading from  $T_3$  to  $T_2$  in  $\mathcal{X}$ , then  $b_1, \dots, b_m, c_1, \dots, c_k$  is a path leading from  $T_1$  to  $T_2$  in  $\mathcal{X}$ .  $\square$

**Remark 3.6.** Let  $\mathcal{X} = (X, E, S)$  and  $\mathcal{X}' = (X', E', S')$  be cluster networks such that  $X \subseteq X'$ ,  $E \subseteq E' \cap (X \times X)$ , and  $S \subseteq S' \cap (X \times X)$ . Since every path in  $\mathcal{X}$  is contained in  $\mathcal{X}'$ , it is easy to see that  $d_{\mathcal{X}}(T_1, T_2) \geq d_{\mathcal{X}'}(T_1, T_2)$  for each  $T_1, T_2 \in X$ .

Now, we use the above general settings to define distances between theories. Before we start, we need the following convention: Suppose that we are given two theories  $T$  and  $T'$ . We write  $T \leftarrow T'$  iff there is  $\varphi \in \text{Fm}$  such that  $T \cup \{\varphi\} \equiv T'$ . We also write  $T - T'$  iff either  $T \leftarrow T'$  or  $T' \leftarrow T$ . Conventionally, we call the relation  $\leftarrow$  *axiom adding*, while the converse relation  $\rightarrow$  is called *axiom removal*. It is easy to see that the following are true for any theories  $T_1, T_2$  and  $T_3$ .

- (1)  $T_1 \leftarrow T_2 \ \& \ T_2 \leftarrow T_3 \implies T_1 \leftarrow T_3,$
- (2)  $T_1 \equiv T_2 \ \& \ T_2 \leftarrow T_3 \implies T_1 \leftarrow T_3,$
- (3)  $T_1 \leftarrow T_2 \ \& \ T_2 \equiv T_3 \implies T_1 \leftarrow T_3.$

**Definition 3.7.** Let  $X$  be a class of some theories (in the logic  $\mathbf{L}_\alpha^\beta$ ) and consider the cluster network  $(X, \equiv, -)$ . We call the step distance on this cluster network *axiomatic distance on  $X$* . This step distance will be denoted by  $\text{Ad}_X$ .

Let  $X$  be a class of theories. We note the following. If there is a path between  $T, T' \in X$  in the cluster network  $(X, \equiv, -)$ , then both  $T$  and  $T'$  must be formulated in the same language. In other words, if  $T, T'$  are formulated on different languages, then  $\text{Ad}_X(T, T') = \infty$ . This is because two theories can be logically equivalent only if they are formulated on the same language.

**Example 3.8.** Suppose that  $\alpha \geq 1$  or  $\beta \geq 1$ . Let  $X$  be a class of theories. Let  $T, T_\perp \in X$  be two theories formulated in the same language. Suppose that  $T$  is consistent while  $T_\perp$  is inconsistent. Then, adding a contradiction to  $T$  ensures that  $\text{Ad}_X(T, T_\perp) = 1$ .

**Example 3.9.** Let  $X$  be a class of theories. Let  $T, \emptyset \in X$  be two theories formulated in the same language such that  $\emptyset$  is an empty theory (i.e., empty set of formulas). Suppose that  $T$  is finitely axiomatizable, then we have either

$$\text{Ad}_X(T, \emptyset) = 1 \text{ or } T \equiv \emptyset.$$

Thus, in the class of all theories, the axiomatic distance between any two finitely axiomatizable theories is  $\leq 2$ .

**Example 3.10.** Suppose that  $\alpha \geq 3$  and  $\beta \geq 3$ . Let  $X$  be the set of all consistent theories of binary relations, let  $T_P$  be the theory of partial orders, and let  $T_E$  be the theory of equivalence relations. Then  $\text{Ad}_X(T_P, T_E) = 2$ . Clearly,  $\text{Ad}_X(T_P, T_E) \geq 2$  because none of  $T_P$  or  $T_E$  implies the other, and, by Example 3.9 and Theorem 3.5 (c),  $\text{Ad}_X(T_P, T_E) \leq \text{Ad}_X(T_P, \emptyset) + \text{Ad}_X(\emptyset, T_E) = 2$ .

This gives us the intuition that, in most of the cases, the axiomatic distance is either 0, 1, 2 or  $\infty$ . For example, in any class  $X$  containing only complete and consistent theories,  $\text{Ad}_X$  is either 0 or  $\infty$  because adding an axiom to a complete theory is either results in an equivalent theory or in an inconsistent one. In the remaining of this section, we assume that  $\alpha \geq 1$  or  $\beta \geq 1$ .

**Problem 1.** Let  $X$  be the class of all consistent theories in  $\mathbf{L}_\alpha^\beta$ . Is it true that, if the axiomatic distance between  $T, T' \in X$  is finite, then it must be  $\leq 2$ ?

Now, let us try to answer the above problem. We define the properties illustrated in Figure 1. Let  $X$  be a class of theories. We say that  $X$  *has the theory amalgamation property* iff for each  $T \in X$ , if there are  $T_1, T_2 \in X$  such that  $T_1 \rightarrow T \leftarrow T_2$ , then there is  $T' \in X$  such that  $T_1 \leftarrow T' \rightarrow T_2$ . Analogously, we say that  $X$  has the *theory co-amalgamation property* iff for each  $T \in X$ , if there are  $T_1, T_2 \in X$  such that  $T_1 \leftarrow T \rightarrow T_2$ , then there is  $T' \in X$  such that  $T_1 \rightarrow T' \leftarrow T_2$ .

**Theorem 3.11.** Let  $X$  be a class of theories having the theory amalgamation property or the theory co-amalgamation property. Then for all  $T, T' \in X$ , we have the following

$$(4) \quad \text{Ad}_X(T, T') = \begin{cases} 0 & \text{if } T \equiv T', \\ 1 & \text{if } T' \text{ or } T \text{ is finitely axiomatizable over the other,} \\ \infty & \text{if } T \text{ and } T' \text{ are not connected in } (X, \equiv, -), \\ 2 & \text{otherwise.} \end{cases}$$

*Proof.* Let us first assume that  $X$  has the theory amalgamation property. Suppose that  $T, T' \in X$  are connected via a path of length 3 in the cluster network  $(X, \equiv, -)$ . By (2) and (3), we can find



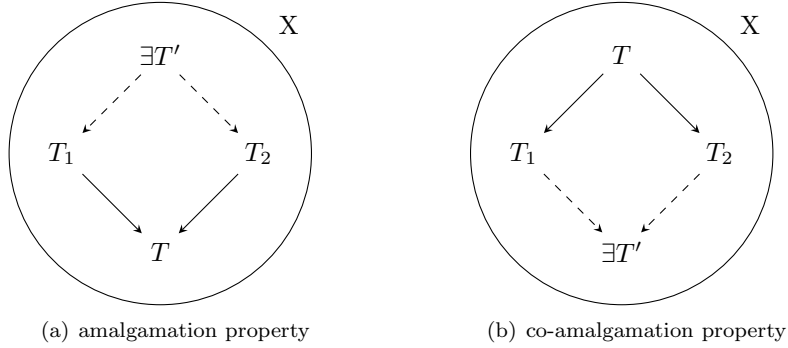


FIGURE 1. Theory amalgamation properties

$T_1, T_2 \in X$  such that  $T - T_1 - T_2 - T'$ . Note that  $T, T_1, T_2, T'$  have the same language  $\mathcal{L}$  and the same set of formulas  $\text{Fm}$ . We first show that  $T$  and  $T'$  are connected by a path of length 2. If at least two consecutive  $-$  in the path  $T - T_1 - T_2 - T'$  are in the same direction, e.g.,  $T \leftarrow T_1 \rightarrow T_2 \rightarrow T'$ , then we are done by (1). So, we may assume that we have one of the cases illustrated in Figure 2:

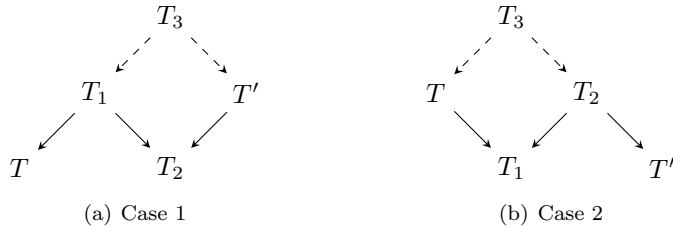


FIGURE 2. 3-paths can be replaced by 2-paths

- (a) Suppose that we are in the first case  $T \leftarrow T_1 \rightarrow T_2 \leftarrow T'$ . Then, by the theory amalgamation property, there is  $T_3 \in X$  such that  $T_1 \leftarrow T_3 \rightarrow T'$ . Hence, by (1), we have  $T \leftarrow T_3 \rightarrow T'$  which means  $T$  and  $T'$  are connected by a path of length 2.
- (b) Suppose that we are in the second case  $T \rightarrow T_1 \leftarrow T_2 \rightarrow T'$ . Then, by the theory amalgamation property, there is  $T_3 \in X$  such that  $T \leftarrow T_3 \rightarrow T_2$ . Hence, by (1), we have  $T \leftarrow T_3 \rightarrow T'$  which means  $T$  and  $T'$  are connected by a path of length 2.

Therefore, any path of length 3 can be replaced by a path of length 2. Now, we can prove the theorem. The first three cases of (4) are obvious, we need to show that otherwise the axiomatic distance is 2. For this, it is enough to show that every path can be replaced by a path of length 2. We use induction on the length of a path. If the path is of length  $l = 3$ , then we are done by the above discussion. Suppose that we have already proven that every path not longer than  $l \geq 3$  can be replaced by a path of length 2. Let path  $T - T_1 - \dots - T_l - T'$  be a path of length  $l + 1$ . By induction hypothesis, path  $T - T_1 - \dots - T_l$  can be replaced by a path of length 2. Hence, path  $T - T_1 - \dots - T_l - T'$  can be replaced by a path of length 3, which can be replaced by a path of length 2 by the induction hypothesis.

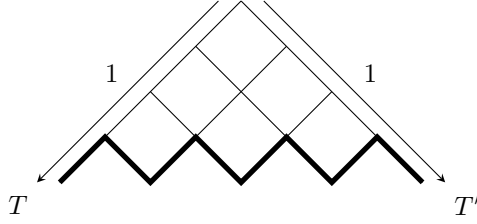


FIGURE 3. Shortening paths using the theory amalgamation property.

The above argument is illustrated in Figure 3. The proof of the case when  $X$  has the theory co-amalgamation property is completely analogous, but going downwards instead of upwards.  $\square$

Even though Theorem 3.11 shows the simplicity of axiomatic distance in plenty of cases, it leaves open the important case when  $X$  is the class of all consistent theories in  $\mathbf{L}_\alpha^\beta$  (recall Problem 1).

**Proposition 3.12.** Let  $X$  be the class of all consistent theories formulated over a fixed language  $\mathcal{L}$ . Then, the class  $X$  does not have the theory amalgamation property. Moreover, e.g., if  $\alpha = \omega$ ,  $3 \leq \beta \leq \omega$  and  $\mathcal{L}$  contains at least a binary relation symbol, then  $X$  does not have the theory co-amalgamation property either.

*Proof.* Let  $\mathcal{L}$  be a language and let  $X$  be the class of all consistent theories of language  $\mathcal{L}$ . Let  $T = \emptyset$ ,  $T_1 = \{\exists v_0 (v_0 = v_0)\}$  and  $T_2 = \{\forall v_0 (v_0 \neq v_0)\}$  be three theories in  $X$ . Then clearly,  $T_1 \rightarrow T \leftarrow T_2$ , but there is no consistent theory  $T' \in X$  for which  $T_1 \leftarrow T' \rightarrow T_2$ . Hence,  $X$  does not have the theory amalgamation property.

Now assume that  $\alpha = \omega$ ,  $3 \leq \beta \leq \omega$  and  $\mathcal{L}$  contains at least one binary relation symbol  $R$ . For every  $n \in \mathbb{N}$ , let

$$(5) \quad \Psi^{(n)} \stackrel{\text{def}}{=} \exists v_0 \exists v_1 \cdots \exists v_{n-1} \left( \left( \bigwedge_{0 \leq i \neq j \leq n-1} v_i \neq v_j \right) \wedge \forall v_n \left( \bigvee_{0 \leq i \leq n-1} v_n = v_i \right) \right).$$

The formula  $\Psi^{(n)}$  is saying that there are exactly  $n$ -many objects. Let

$$T_1 = \{\neg \Psi^{(2n)} : n \in \mathbb{N}\} \quad \text{and} \quad T_2 = \{\neg \Psi^{(2n+1)} : n \in \mathbb{N}\}$$

be two theories in  $X$ . Then a model of  $T_1$  is a model for  $\mathcal{L}$  which have odd finite or infinite cardinality. Similarly, a model of  $T_2$  is a model for  $\mathcal{L}$  which have even finite or infinite cardinality. Let  $\varphi$  be a formula in language  $\mathcal{L}$  requiring that there are infinitely many objects. Using the relation symbol  $R$  it is easy to write up such a formula by requiring that  $R$  is irreflexive, transitive and serial. Then, clearly,  $T_1 \cup \{\varphi\} \equiv T_2 \cup \{\varphi\} \equiv \{\varphi\}$  and  $T_1 \leftarrow \{\varphi\} \rightarrow T_2$ .

However, there is no theory  $T'$  for which  $T_1 \rightarrow T' \leftarrow T_2$  because (1) all the common consequences of  $T_1$  and  $T_2$  are tautologies, i.e.,  $\text{Cn}(T_1) \cap \text{Cn}(T_2) \equiv \emptyset$  since every model  $\mathfrak{M}$  for  $\mathcal{L}$  is either a model for  $T_1$  or a model for  $T_2$ , and (2) neither  $T_1$  nor  $T_2$  is finitely axiomatizable. The non-finite axiomatizability of  $T_1$  can be confirmed as follows. For each  $i \in \mathbb{N}$ , let  $\mathfrak{M}_i$  be a model for  $\mathcal{L}$  whose cardinality is  $2i$ . Clearly,  $\mathfrak{M}_i$  cannot be a model for  $T_1$ . However, the ultraproduct of the models  $\mathfrak{M}_i$ 's (over a free ultrafilter) is a model for  $T_1$  because it is of infinite cardinality (for some details about ultraproducts, we refer the reader to [van Dalen, 2013, Chapter 4.5]). Thus, by

[van Dalen, 2013, Theorem 4.5.27 (ii)], it follows that  $T_1$  is not finitely axiomatizable. Same applies to  $T_2$  by taking ultraproduct of finite models having cardinalities  $2i + 1$  for all  $i \in \mathbb{N}$ .

Therefore, the theories  $T_1$  and  $T_2$  are not finitely axiomatizable over their common consequences. Hence, there is no theory  $T'$  in  $X$  for which  $T_1 \rightarrow T' \leftarrow T_2$ . Consequently, the class  $X$  does not have the theory co-amalgamation property.  $\square$

It is worth noting that even if the range of the axiomatic distance consists only of four elements (cf., Theorem 3.11 and Problem 1), this does not mean that this distance is not interesting. However, it might be true that the distance defined in this way does not tell us much information on the nature of the axioms separating two theories, adding any axiom is considered as one step. One can overcome this problem by giving weights to the axiom adding steps, e.g., considering the addition of certain kind of axioms as two or more steps.

#### 4. CONCEPTUAL DISTANCE

In 1935, A. Tarski introduced the so-called Lindenbaum-Tarski algebras as a device to establish correspondence between logic and algebras. The *Lindenbaum-Tarski algebra* of a theory  $T$  is the quotient algebra obtained by factoring the algebra of formulas by the congruence relation of logical equivalence between formulas. These algebras are also called concept algebras, e.g., in [Andréka and Németi, 2017]. This reflects the fact that the underlying set of a Lindenbaum-Tarski algebra consists of the different concepts that can be defined within the corresponding theory.

**Definition 4.1.** A *concept* in theory  $T$  is a maximal set of logically equivalent formulas in  $T$ . In other words, a concept in  $T$  is the set  $[\varphi]_T \stackrel{\text{def}}{=} \{\psi \in \text{Fm} : T \models \varphi \leftrightarrow \psi\}$ , for some formula  $\varphi$ .

Intuitively, by a concept we mean a definition, no matter how many different ways one can write it equivalently. Now, we define conceptual size of a theory  $T$  to be the size of the underlying set of the corresponding Lindenbaum-Tarski algebra. It might be more convenient in some cases to use cardinality here instead of size.

**Definition 4.2.** Let  $T$  be a theory and let  $k \in \mathbb{N} \cup \{\infty\}$ . We say that the *conceptual size* of  $T$  is  $k$  and we write  $\text{Cz}(T) = k$  iff the set  $\{[\varphi]_T : \varphi \in \text{Fm}\}$  is of size  $k$ .

This is equivalent to saying that a theory  $T$  is of conceptual size  $k$  iff there is a maximal set  $X \subseteq \text{Fm}$  of size  $k$  such that  $T \not\models \varphi \leftrightarrow \psi$  for each  $\varphi, \psi \in X$ . It is also worthy of note that, when  $\alpha \geq \omega$ , a theory that has a model of at least two different elements cannot have a finite conceptual size.

**Proposition 4.3.** Let  $T, T'$  be two theories. If there is a faithful interpretation from  $T$  to  $T'$ , then we have  $\text{Cz}(T) \leq \text{Cz}(T')$ . Consequently,

$$T \rightleftarrows T' \implies \text{Cz}(T) = \text{Cz}(T').$$

*Proof.* Let  $T$  and  $T'$  be two theories, and suppose that  $\text{tr}$  is a faithful interpretation of  $T$  into  $T'$ . Now, there is a maximal set  $X \subseteq \text{Fm}$  of size  $\text{Cz}(T)$  such that  $T \not\models \varphi \leftrightarrow \psi$  for any  $\varphi, \psi \in X$ . Let

$$X' = \{\text{tr}(\varphi) : \varphi \in X\} \subseteq \text{Fm}'.$$

Since  $\text{tr}$  is a faithful interpretation,  $T' \not\models \text{tr}(\varphi \leftrightarrow \psi)$ . Hence  $T' \not\models \text{tr}(\varphi) \leftrightarrow \text{tr}(\psi)$  since  $\text{tr}$  is an interpretation. Consequently,  $\text{Cz}(T) \leq \text{Cz}(T')$ . If  $T \rightleftarrows T'$ , then by Proposition 2.6 there are

faithful interpretations between  $T$  and  $T'$  in both directions. Hence  $\text{Cz}(T) \leq \text{Cz}(T') \leq \text{Cz}(T)$ . Therefore,  $\text{Cz}(T) = \text{Cz}(T')$  as desired.  $\square$

As a matter of fact, two theories are definitionally equivalent iff their Lindenbaum-Tarski algebras are isomorphic [Henkin et al., 1985, Theorem 4.3.43]. So, the above Proposition is a consequence of the straightforward observation: Two algebras of different sizes cannot be isomorphic.

Now, we want to define a distance counting the minimum number of concepts that distinguish two theories  $T$  and  $T'$ . It is very natural to explain the idea within the framework of Lindenbaum-Tarski algebras, since these are concept algebras and we want to count cocnepts. For simplicity, let us assume that the Lindenbaum-Tarski algebra  $\mathfrak{A}$  of  $T$  is embeddable into the Lindenbaum-Tarski algebra  $\mathfrak{A}'$  of  $T'$ . Thus, the minimum number of concepts distinguish the two theories is equal to the minimum number of elements of  $\mathfrak{A}'$  that we can add to  $\mathfrak{A}$  (more precisely, to one of its copies inside  $\mathfrak{A}'$ ) to generate the algebra  $\mathfrak{A}'$ , see Figure 4.

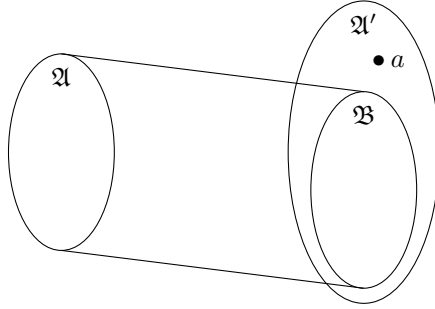


FIGURE 4. The distance between  $\mathfrak{A}$  and  $\mathfrak{A}'$  is one if  $\langle B, a \rangle \cong \mathfrak{A}'$ .

The following definitions illustrate the above idea, but in terms of logic instead of algebras. As we mentioned before, the correspondence between the following definitions and the above algebraic idea is planned to be discussed in details in a forthcoming algebra oriented paper.

**Definition 4.4.** We say that *theory  $T'$  is a **one-concept-extension** of theory  $T$*  and we write  $T \rightsquigarrow T'$  iff  $\mathcal{L}' = \mathcal{L} \cup \{R\}$ , for some relation symbol  $R$ , and  $T \sqsubseteq T'$ . We also write  $T \sim T'$  iff  $T \rightsquigarrow T'$  or  $T' \rightsquigarrow T$ , and in this case we say that  $T$  and  $T'$  are **separated by at most one concept**.

Again, we understand concept removal to be the converse of concept adding. Later, in Section 6, we are going to introduce another notion for concept removal, and a corresponding distance notion.

**Definition 4.5.** Let  $X$  be a class of theories. The step distance induced by the cluster network  $(X, \rightleftharpoons, \rightsquigarrow)$  is called **conceptual distance** on  $X$  and is denoted by  $\text{Cd}_X$ . In the case when  $X$  is the class of all theories in  $\mathbf{L}_\alpha^\beta$ , we denote the conceptual distance on  $X$  by  $\text{Cd}_\alpha^\beta$ .

By Remark 3.6, it is clear that  $\text{Cd}_\alpha^\beta(T, T') \geq \text{Cd}_\alpha^\gamma(T, T')$  for any ordinal  $\beta \leq \gamma \leq \alpha + 1$  and any theories  $T$  and  $T'$  in  $\mathbf{L}_\alpha^\beta$ . It is also apparent that an inconsistent theory is of an infinite conceptual distance from any consistent theory, because relations  $\rightleftharpoons$  and  $\rightsquigarrow$  cannot make a consistent theory inconsistent and also cannot make an inconsistent theory consistent. Now, we give more examples.

**Lemma 4.6.** Let  $k, m$  be two finite numbers. Let  $T$  and  $T'$  be two theories such that  $T \sqsubseteq T'$ . Suppose that  $\mathcal{L}' = \mathcal{L} \cup \{R\}$ , for some relation  $R$  whose rank  $\text{rank}(R) = m$ . Then,

$$I(T', k) \leq 2^{k^m} I(T, k).$$

*Proof.* This follows immediately from the following fact: If  $\mathfrak{M}$  is a model of size  $k$ , then there are at most  $2^{k^m}$  different ways of defining a relation of rank  $m$  in  $\mathfrak{M}$ .  $\square$

**Theorem 4.7.** Suppose that  $\beta \geq 1$ . For every  $n \in \mathbb{N} \cup \{\infty\}$ , there are theories  $T$  and  $T'$  in  $\mathbf{L}_\alpha^\beta$  such that  $\text{Cd}_\alpha^\beta(T, T') = n$ .

*Proof.* Let  $\mathcal{L}_\infty$  be a language for  $\mathbf{L}_\alpha^\beta$  that consists of infinitely many relation symbols (describing infinitely many different concepts), each of which is of a fixed rank  $\gamma < \beta$ .

$$(6) \quad R_1, R_2, \dots$$

For each  $n \in \mathbb{N}$ , we let  $\mathcal{L}_n \subseteq \mathcal{L}_\infty$  be the language consists of the first  $n$ -many relation symbols from the list in (6), and we let  $T_n^* = \emptyset$  be the empty theory on language  $\mathcal{L}_n$ . Let  $n \in \mathbb{N} \cup \{\infty\}$ . It is clear that  $\text{Cd}_\alpha^\beta(T_0^*, T_n^*) \leq n$ . We need to prove the other direction. We consider the following cases:

Suppose that  $\beta$  is finite. In this case, we assume that  $\gamma = \beta - 1$ . Let  $k \geq 1$  be a finite number. By Lemma 4.6, it follows that, for each pair of theories  $T_1, T_2$  in logic  $\mathbf{L}_\alpha^\beta$ ,

$$(7) \quad T_1 \rightsquigarrow T_2 \implies I(T_2, k) \leq 2^{k^{\beta-1}} I(T_1, k).$$

Moreover, it is easy to see that

$$(8) \quad I(T_n^*, k) = 2^{nk^{\beta-1}} \quad (\text{we define this entity to be } \infty \text{ if } n = \infty),$$

because there are exactly  $n$ -many relation symbols, each of which is of rank  $\beta - 1$ , in  $\mathcal{L}_n$ . Thus, at least  $n$ -many steps are needed to increase  $I(T_0^*, k) = 2^0 = 1$  to  $I(T_n^*, k) = 2^{nk^{\beta-1}}$ . Therefore, we have  $\text{Cd}_\alpha^\beta(T_0^*, T_n^*) = n$  as required.

Suppose that  $\beta$  is infinite. We do not need any extra assumption on  $\gamma$ . All what we need here is to count models of size 1. By (8), we have  $I(T_n^*, 1) = 2^n$ . Moreover, for any two theories  $T_1$  and  $T_2$ ,

$$(9) \quad T_1 \rightsquigarrow T_2 \implies I(T_2, 1) \leq 2I(T_1, 1).$$

This is true because in a model of size 1 there are at most two relations (of any fixed rank). Again, for the same reasons, we need at least  $n$ -many steps to increase  $I(T_0^*, 1) = 2^0 = 1$  to  $I(T_n^*, 1) = 2^n$ . Therefore,  $\text{Cd}_\alpha^\beta(T_0^*, T_n^*) = n$  as desired.  $\square$

Our constructions in the above proof for the case when  $\beta$  is infinite (and hence  $\alpha$  is also infinite) all have models of size 1. In Theorem 4.10 below, we investigate what happens if the theories in question do not have models of size 1. First, we need the following lemma.

**Lemma 4.8.** Suppose that  $\alpha = \beta = \omega$ . Let  $T_1, T_2$  and  $T_3$  be theories such that

$$I(T_1, 1) = I(T_2, 1) = I(T_3, 1) = 0.$$

Then, if  $T_1 \rightsquigarrow T_2 \rightsquigarrow T_3$ , then there is a theory  $T$  such that  $T_1 \rightsquigarrow T \rightleftarrows T_3$ .

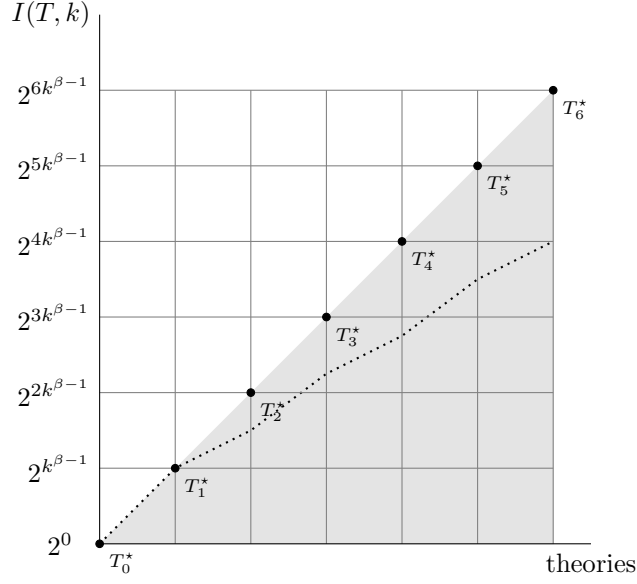


FIGURE 5. Theorem 4.7:  $\text{Cd}_\alpha^\beta(T_0^*, T_n^*) = n$

*Proof.* Suppose  $T_1$ ,  $T_2$  and  $T_3$  are as required in the statement of the lemma above, and assume that  $T_1 \rightsquigarrow T_2 \rightsquigarrow T_3$ . Then  $\mathcal{L}_3 = \mathcal{L}_1 \cup \{R, S\}$  for some relation symbols  $R$  and  $S$ . Suppose that  $\text{rank}(R) = n$  and  $\text{rank}(S) = m$ . Let  $l = \max\{n, m\} + 2$ . Let  $\mathcal{L} \stackrel{\text{def}}{=} \mathcal{L}_1 \cup \{B\}$ , for some new relation symbol  $B$  of rank  $l$ . Every model  $\mathfrak{M}$  for  $\mathcal{L}_3$  can be extended to a model  $\mathfrak{M}^+$  for  $\mathcal{L}$  by defining  $B^{\mathfrak{M}^+}$  as follows:

$$B^{\mathfrak{M}^+} \stackrel{\text{def}}{=} \{(a_0, \dots, a_{l-1}) \in M^l : \exists \text{ an assignment } \tau (\tau(v_0) = a_0, \dots, \tau(v_{l-1}) = a_{l-1} \text{ and } \mathfrak{M}, \tau \models \psi)\},$$

where  $\psi(v_0, \dots, v_{l-1}) \stackrel{\text{def}}{=} (R(v_0, \dots, v_{n-1}) \wedge v_{l-2} = v_{l-1}) \vee (S(v_0, \dots, v_{m-1}) \wedge v_{l-2} \neq v_{l-1})$ . Let

$$T \stackrel{\text{def}}{=} \{\varphi \in \text{Fm} : \mathfrak{M}^+ \models \varphi, \text{ for every model } \mathfrak{M} \text{ for } T_3\}.$$

We need to prove that  $T_1 \rightsquigarrow T$  and  $T \rightleftharpoons T_3$ . To prove that  $T_1 \rightsquigarrow T$ , it is enough to show that  $T_1 \sqsubseteq T$  (because  $\mathcal{L} = \mathcal{L}_1 \cup \{B\}$ ). Let  $\varphi \in \text{Fm}_1$ . We have

$$T_1 \models \varphi \iff T_2 \models \varphi \iff T_3 \models \varphi \iff T \models \varphi,$$

where the first two equivalences follow by the assumption  $T_1 \sqsubseteq T_2 \sqsubseteq T_3$ , and the last equivalence follows by the definition of  $T$ . To show that  $T \rightleftharpoons T_3$ , we define the translations  $\text{tr} : \text{Fm} \rightarrow \text{Fm}_3$  and  $\text{tr}' : \text{Fm}_3 \rightarrow \text{Fm}$  as follows:

$$\text{tr} : B(v_0, \dots, v_{l-2}, v_{l-1}) \mapsto \psi(v_0, \dots, v_{l-1}) \quad \text{and}$$

$$\text{tr}' : R(v_0, \dots, v_{n-1}) \mapsto \exists v_{l-1} (B(v_0, \dots, v_{l-2}, v_{l-1}) \wedge (v_{l-2} = v_{l-1}))$$

$$\text{tr}' : S(v_0, \dots, v_{m-1}) \mapsto \exists v_{l-1} (B(v_0, \dots, v_{l-2}, v_{l-1}) \wedge (v_{l-2} \neq v_{l-1}))$$

We have defined  $\text{tr}$  and  $\text{tr}'$  on specific basic formulas and these can be extended in a unique way to their domains, see Footnote 4. Let  $\mathfrak{M}$  be a model for  $T_3$ . By definition of the extension  $\mathfrak{M}^+$ ,

$$(10) \quad \mathfrak{M}^+ \models B(v_0, \dots, v_{l-2}, v_{l-1}) \leftrightarrow \psi(v_0, \dots, v_{l-2}, v_{l-1}).$$

Thus, by (10), we have

$$\begin{aligned} \mathfrak{M}^+ \models \text{tr}'(S(v_0, \dots, v_{m-1})) &\leftrightarrow \exists v_{l-1} (B(v_0, \dots, v_{l-2}, v_{l-1}) \wedge v_{l-2} \neq v_{l-1}) \\ &\leftrightarrow \exists v_{l-1} (((R(v_0, \dots, v_{n-1}) \wedge v_{l-2} = v_{l-1}) \wedge \\ &\quad (S(v_0, \dots, v_{m-1}) \wedge v_{l-2} \neq v_{l-1})) \wedge v_{l-2} \neq v_{l-1}) \\ &\leftrightarrow \exists v_{l-1} (S(v_0, \dots, v_{m-1}) \wedge v_{l-2} \neq v_{l-1}) \\ &\leftrightarrow S(v_0, \dots, v_{m-1}). \end{aligned}$$

The last  $\leftrightarrow$  follows by the assumption that the cardinality of  $\mathfrak{M}$  is at least 2 (and hence the same is true for  $\mathfrak{M}^+$ ). Similarly,  $\mathfrak{M}^+ \models \text{tr}'(R(v_0, \dots, v_{n-1})) \leftrightarrow R(v_0, \dots, v_{n-1})$ . Therefore, by the fact that  $\text{tr}$  and  $\text{tr}'$  are translations, it follows that

$$(11) \quad \mathfrak{M}^+ \models \varphi \leftrightarrow \text{tr}(\varphi) \quad \text{and} \quad \mathfrak{M}^+ \models \psi \leftrightarrow \text{tr}'(\psi)$$

for all  $\varphi \in \text{Fm}$  and  $\psi \in \text{Fm}_3$ . By (11), it is not hard to see that  $\text{tr}$  and  $\text{tr}'$  are definitional equivalences, and the desired follows.  $\square$

The above lemma is a direct consequence of the following elementary fact. In  $\mathbf{L}_\omega^\omega$  (under some conditions), for any two relations  $R$  and  $S$ , there is a relation  $M$  such that  $M$  is definable in terms of  $R$  and  $S$  and, conversely, both  $R$  and  $S$  are definable in terms of  $M$ , see [Goodman, 1943]. The idea of the above proof is distilled from [Henkin et al., 1971, Theorem 2.3.22].

**Corollary 4.9.** Suppose that  $\alpha = \beta = \omega$ . Let  $T_1, T_2, \dots, T_n$ , for some  $n \geq 2$ , be theories such that  $I(T_i, 1) = 0$ , for each  $1 \leq i \leq n$ . Then,

$$T_1 \rightsquigarrow T_2 \rightsquigarrow \dots \rightsquigarrow T_n \implies \text{there is a theory } T \text{ such that } T_1 \rightsquigarrow T \rightleftharpoons T_n.$$

*Proof.* This can be proved by a simple induction on  $n$ . If  $n = 2$ , then we are obviously done. Suppose that  $n \geq 3$  and  $T_1 \rightsquigarrow T_2 \rightsquigarrow \dots \rightsquigarrow T_{n-1} \rightsquigarrow T_n$ . If by induction hypothesis we can assume that there is  $T'$  such that  $T_2 \rightsquigarrow T' \rightleftharpoons T_n$ , then  $T_1 \rightsquigarrow T_2 \rightsquigarrow T' \rightleftharpoons T_n$ . Therefore, by the Lemma 4.8, there is theory  $T$  such that  $T_1 \rightsquigarrow T \rightleftharpoons T_n$ .  $\square$

**Theorem 4.10.** Suppose that  $\alpha = \beta = \omega$ . Let  $T$  and  $T'$  be theories on finite languages. Then,

$$I(T, 1) = I(T', 1) = 0 \implies \text{Cd}_\omega^\omega(T, T') = \infty \text{ or } \text{Cd}_\omega^\omega(T, T') \leq 2.$$

*Proof.* Let  $T$  and  $T'$  be theories on finite languages and suppose that  $I(T, 1) = I(T', 1) = 0$ . Let us assume that  $\text{Cd}_\omega^\omega(T_1, T_2) < \infty$ . Let  $\mathcal{L}$  be the empty language. For each  $\varphi \in \text{Fm}$ ,

$$(12) \quad T_1 \models \varphi \iff T_2 \models \varphi.$$

This is true because  $\text{Cd}_\omega^\omega(T, T') < \infty$  and the validity of each  $\varphi \in \text{Fm}$  is preserved under conservative extensions and definitional equivalences. Let  $T = \{\varphi \in \text{Fm} : T_1 \models \varphi \text{ and } T_2 \models \varphi\}$ . Then, by (12), it is true that  $T \sqsubseteq T_1$  and  $T \sqsubseteq T_2$ . We also claim that

$$(13) \quad \text{Cd}_\omega^\omega(T, T_1) \leq 1 \quad \text{and} \quad \text{Cd}_\omega^\omega(T, T_2) \leq 1.$$

To show (13), let us assume that  $\mathcal{L} = \{R_i : i < m\}$ , for some finite  $m$ . If  $m = 0$ , then  $T \equiv T_1$  and thus  $\text{Cd}_\omega^\omega(T, T_1) = 0$ . Assume that  $m \neq 0$ . Let  $\mathcal{L}_0^* = \{R_0\}, \dots, \mathcal{L}_{m-1}^* = \{R_0, \dots, R_{m-1}\}$  and, for

each  $0 \leq i \leq m-1$ ,  $T_i^* = \{\varphi \in \mathbf{Fm}_i^* : T_1 \models \varphi\}$ . Hence,  $T \rightsquigarrow T_0^* \rightsquigarrow \dots \rightsquigarrow T_{m-1}^* \equiv T_1$ . Clearly, for each  $0 \leq i \leq m-1$ ,  $I(T_i^*, 1) = 0$  because  $\neg\Psi^{(1)} = \neg(\exists v_0 \forall v_1 (v_0 = v_1))$  is a theorem of  $T$ , and hence is a theorem of  $T_i^*$ . Thus, by Corollary 4.9, it follows that  $\text{Cd}_\omega^\omega(T, T_1) \leq 1$ . Similarly, one can show that  $\text{Cd}_\omega^\omega(T, T_2) \leq 1$ . Therefore,  $\text{Cd}_\omega^\omega(T_1, T_2) \leq \text{Cd}_\omega^\omega(T, T_1) + \text{Cd}_\omega^\omega(T, T_2) \leq 2$ .  $\square$

We prove one more statement investigating the connection between the spectrum of theories and conceptual distance in  $\mathbf{L}_\omega^\omega$ .

**Theorem 4.11.** Suppose that  $\alpha = \beta = \omega$ . Let  $T$  and  $T'$  be two theories formulated in countable languages. Then,

$$(14) \quad \text{Cd}_\omega^\omega(T, T') < \infty \implies (\forall \text{ cardinal } \kappa) [I(T, \kappa) \neq 0 \iff I(T', \kappa) \neq 0].$$

If  $T$  and  $T'$  are formulated in finite languages, then the converse of (14) is also true.

*Proof.* In this proof, we will make use of the formulas  $\Psi^{(n)}$ 's defined in (5). Recall that with the formula  $\Psi^{(n)}$ , we require every model to be of cardinality  $n$ . Let  $T$  and  $T'$  be as required, and let  $\kappa$  be any cardinal. Suppose that  $\kappa$  is finite, then  $T$  does not have a model of size  $\kappa$  iff  $T \models \neg\Psi^{(\kappa)}$ . But,  $T' \models \neg\Psi^{(\kappa)} \iff T \models \neg\Psi^{(\kappa)}$  because of the assumption  $\text{Cd}_\omega^\omega(T, T') < \infty$  and the fact that the validity of  $\Psi^{(\kappa)}$  is preserved under conservative extensions and definitional equivalences. Hence,  $T'$  has a model of cardinality  $\kappa$  iff  $T$  has a model of cardinality  $\kappa$ . If  $\kappa$  is an infinite cardinality, then by Löwenheim–Skolem Theorem,  $T$  and  $T'$  have models of cardinality  $\kappa$  iff they have infinite models.  $T$  has an infinite model iff  $T \models \exists v_0 \exists v_1 \dots \exists v_{n-1} \left( \bigwedge_{0 \leq i \neq j \leq n-1} v_i \neq v_j \right)$  for all  $n \in \mathbb{N}$ . Again, the validity of the formulas  $\exists v_0 \exists v_1 \dots \exists v_{n-1} \left( \bigwedge_{0 \leq i \neq j \leq n-1} v_i \neq v_j \right)$  is preserved under conservative extensions and definitional equivalences. Thus, the assumption  $\text{Cd}_\omega^\omega(T, T') < \infty$  implies that

$$T' \models \exists v_0 \exists v_1 \dots \exists v_{n-1} \left( \bigwedge_{0 \leq i \neq j \leq n-1} v_i \neq v_j \right) \iff T \models \exists v_0 \exists v_1 \dots \exists v_{n-1} \left( \bigwedge_{0 \leq i \neq j \leq n-1} v_i \neq v_j \right).$$

Hence,  $T$  has an infinite model and thus a model of cardinality  $\kappa$  iff  $T'$  has such a model.

To prove the converse of (14), let us assume that  $T$  and  $T'$  are formulated in finite languages  $\mathcal{L}$  and  $\mathcal{L}'$ , respectively. Assume that, for every cardinal  $\kappa$ ,  $T$  has a model of cardinality  $\kappa$  iff  $T'$  has a model of the same cardinality. Let  $\mathcal{L}_0$  be the empty language. Then, for every  $\varphi \in \mathbf{Fm}_0$ ,

$$T \models \varphi \iff T' \models \varphi.$$

Let  $T_0 = \{\varphi \in \mathbf{Fm}_0 : T \models \varphi \text{ and } T' \models \varphi\}$ . Thus,  $T_0 \sqsubseteq T$  and  $T_0 \sqsubseteq T'$ . Now, we can add the whole  $\mathcal{L}$  to  $\mathcal{L}_0$  in finitely many steps, because there only finitely many relation symbols in  $\mathcal{L}$ , thus  $\text{Cd}_\omega^\omega(T_0, T) < \infty$ . Similarly,  $\text{Cd}_\omega^\omega(T_0, T') < \infty$ . Therefore,

$$\text{Cd}_\omega^\omega(T, T') \leq \text{Cd}_\omega^\omega(T_0, T) + \text{Cd}_\omega^\omega(T_0, T') < \infty. \quad \square$$

**Corollary 4.12.** The conceptual distance between the theories of any two finite models of different cardinalities is infinite. More precisely, if  $\mathfrak{A}$  and  $\mathfrak{B}$  are two finite models of different cardinality, then  $\text{Cd}_\omega^\omega(\text{Th}(\mathfrak{A}), \text{Th}(\mathfrak{B})) = \infty$ .

For instance, given two cyclic groups  $\langle k_1 \rangle$  and  $\langle k_2 \rangle$  of orders 5 and 7, respectively, the conceptual distance between the theories of these groups is  $\infty$ . This might seem strange; these theories are about similar structures. But if we look carefully at the statement of the above corollary, we will find



that it talks about theories of structures, not structures themselves. In other words, the conceptual distance between the theories of  $\langle k_1 \rangle$  and  $\langle k_2 \rangle$  cannot be granted as a distance between these two groups as algebraic structures. This conceptual distance can be rather considered as a distance between the Lindenbaum-Tarski algebras of the theories of these groups, which are of course of different nature than the groups themselves.

**Corollary 4.13.** There are infinitely many theories that are, in terms of conceptual distance, infinitely far from each other in  $\mathbf{L}_\omega^\omega$ .

**Problem 2.** Let  $X$  be the class of all complete and consistent theories in  $\mathbf{L}_\omega^\omega$ , and let  $T_1, T_2 \in X$ . Is it always true that

$$\text{Cd}_\omega^\omega(T_1, T_2) = \text{Cd}_X(T_1, T_2)?$$

## 5. CONCEPTUAL DISTANCE IN PHYSICS

Each physical theory is established based on some preliminary decisions. These decisions are suggested by the accumulation and the assimilation of new knowledge. The methods used to improve physical theories are intuitively conceived and applied in a fruitful way, but many obvious ambiguities have appeared. To clarify these ambiguities, it was critical to introduce the *logical foundation of physical theories*.

Even today the logic based axiomatic foundation of physical theories is intensively investigated by several research groups. For example, the Andr eka–N emeti school axiomatizes and investigates special and general relativity theories within ordinary first order logic, see, e.g., [Andr eka et al., 2002], [Andr eka et al., 2004] and [Andr eka et al., 2012]. For similar approaches related to other physical theories, see, e.g., [Baltag and Smets, 2005] and [Krause and Arenhart, 2017].

Following Andr eka–N emeti school’s traditions, two theories **ClassicalKin** and **SpecRel** are formulated in ordinary first order logic  $\mathbf{L}_\omega^\omega$  to capture the intrinsic structures of classical and relativistic kinematics. For the precise definitions of these theories, one can see [Lefever and Sz ekely, 2018a, p.67 and p. 69]. In this section, we will investigate the conceptual distance between these two theories.

In [Lefever, 2017] and [Lefever and Sz ekely, 2018a], it was shown that these two theories can be turned definitionally equivalent by the following two concept manipulating steps:

- (1) adding the concept of an observer “being stationary” to the theory of relativistic kinematics **SpecRel**, and
- (2) removing the concept of observers “not moving slower than light” from the theory of classical kinematics **ClassicalKin**.

Then, it was shown that even if observers “not moving slower than light” are removed from **ClassicalKin** the resulting theory remains definitionally equivalent to **ClassicalKin** and hence adding only the concept of “being stationary” to **SpecRel** is enough to make the two theories equivalent. Thus, it follows that the conceptual distance between relativistic and classical kinematics is 1.

**Theorem 5.1.** Classical and relativistic kinematics are distinguished from each other by only one concept, namely the existence of some distinguished observers captured by formula (15) below, i.e.,  $\text{Cd}_\omega^\omega(\text{ClassicalKin}, \text{SpecRel}) = 1$ .<sup>7</sup>

<sup>7</sup>It is worth noting that in the proof of Theorem 5.1, we add only a unary concept  $E$  to **SpecRel** to get a theory definitionally equivalent to **ClassicalKin**.

*Proof.* The key to this result is the surprising theorem stating that the only concept which needs to be added to  $\text{SpecRel}$  to make it definitional equivalent to  $\text{ClassicalKin}$  is a concept distinguishing a set of observers that are “being at absolute rest” as shown in [Lefever, 2017, p.72] and [Lefever and Székely, 2018a, p.110]. Let  $E$  be a unary relation symbol corresponding to this basic concept. Axiom  $\text{AxPrimitiveEther}$ , see [Lefever, 2017, p.46] and [Lefever and Székely, 2018a, p.87], defines  $E$  as follows:

$$(15) \quad \exists v_0 [\text{IOb}(v_0) \wedge \forall v_1 (E(v_1) \leftrightarrow [\text{IOb}(v_1) \wedge \varphi(v_0, v_1)])],$$

where  $\text{IOb}$  is a unary relation symbol represents inertial observers and  $\varphi(v_0, v_1)$  is a formula in the language of  $\text{SpecRel}$  capturing that observers  $v_0$  and  $v_1$  are stationary with respect to each other. In this proof, we only need that  $\varphi(v_0, v_1)$  is a formula with two free variables in the language of  $\text{SpecRel}$ ; its concrete definition plays no role here. Let

$$\text{SpecRel}^E = \text{SpecRel} \cup \{\text{AxPrimitiveEther}\}.$$

First, we need to prove that  $\text{SpecRel} \rightsquigarrow \text{SpecRel}^E$ . To do so, it is enough to show that  $\text{SpecRel}^E$  is a conservative extension of  $\text{SpecRel}$ , i.e.,  $\text{SpecRel} \sqsubseteq \text{SpecRel}^E$ , because the languages of these theories differ only in the unary relation symbol  $E$ . So, we need to show that for any formula  $\rho$  of the language of  $\text{SpecRel}$ ,

$$\text{SpecRel} \models \rho \iff \text{SpecRel}^E \models \rho.$$

Let  $\rho$  be an arbitrary formula of the language of  $\text{SpecRel}$ . Since  $\text{SpecRel} \subseteq \text{SpecRel}^E$ ,  $\text{SpecRel} \models \rho$  implies  $\text{SpecRel}^E \models \rho$ . We prove the other direction by proving that, if  $\text{SpecRel} \not\models \rho$ , then  $\text{SpecRel}^E \not\models \rho$ . Let  $\mathfrak{M}$  be a model of  $\text{SpecRel}$ . Since  $\text{SpecRel} \models \exists v_0 \text{IOb}(v_0)$ , there exists an  $a \in \text{IOb}^{\mathfrak{M}}$ . Let us fix such element  $a$  of  $\text{IOb}^{\mathfrak{M}}$  and let extension  $\mathfrak{M}'$  of  $\mathfrak{M}$  be defined by adding the following relation to  $\mathfrak{M}$ :

$$E^{\mathfrak{M}'} = \left\{ b \in \text{IOb}^{\mathfrak{M}} : \exists \text{ an assignment } \tau [\tau(v_0) = a, \tau(v_1) = b \text{ and } \mathfrak{M}, \tau \models \varphi] \right\},$$

where  $\varphi^{\mathfrak{M}}$  is the binary relation defined by formula  $\varphi(v_0, v_1)$  in model  $\mathfrak{M}$ . By construction,  $\mathfrak{M}'$  is a model of  $\text{SpecRel}^E$ . Therefore, if  $\mathfrak{M} \models \neg\rho$ , then  $\mathfrak{M}' \models \neg\rho$ , because  $\mathfrak{M}'$  is an extension of  $\mathfrak{M}$  means that  $\text{Th}(\mathfrak{M}) \sqsubseteq \text{Th}(\mathfrak{M}')$ . Consequently,  $\text{SpecRel} \not\models \rho$  implies  $\text{SpecRel}^E \not\models \rho$ , which is what we wanted to prove. This completes the proof of  $\text{SpecRel} \rightsquigarrow \text{SpecRel}^E$ , and hence

$$\text{Cd}_{\omega}^{\omega}(\text{SpecRel}, \text{SpecRel}^E) \leq 1.$$

By Corollary 9 in [Lefever, 2017, p.72] and [Lefever and Székely, 2018a, p.110],  $\text{SpecRel}^E$  is definitionally equivalent to  $\text{ClassicalKin}$ . Hence,

$$\text{Cd}_{\omega}^{\omega}(\text{SpecRel}^E, \text{ClassicalKin}) = 0.$$

Therefore,

$$\text{Cd}_{\omega}^{\omega}(\text{SpecRel}, \text{ClassicalKin}) \leq \text{Cd}_{\omega}^{\omega}(\text{SpecRel}, \text{SpecRel}^E) + \text{Cd}_{\omega}^{\omega}(\text{SpecRel}^E, \text{ClassicalKin}) = 1.$$

Moreover,  $\text{Cd}_{\omega}^{\omega}(\text{SpecRel}, \text{ClassicalKin})$  cannot be 0 since  $\text{SpecRel}$  and  $\text{ClassicalKin}$  are not definitionally equivalent, see Theorem 5 in [Lefever, 2017] or [Lefever and Székely, 2018a]. Consequently,

$$\text{Cd}_{\omega}^{\omega}(\text{SpecRel}, \text{ClassicalKin}) = 1$$

and thus the desired follows.  $\square$

There are several ways how one can capture the structures of relativistic and classical kinematics in first order logic. Let us now introduce another way to capture these theories. Let  $\mathbb{R}$  be the set of all real numbers. Let  $\text{Ph} \subseteq \mathbb{R}^4 \times \mathbb{R}^4$  be such that  $(\bar{x}, \bar{y}) \in \text{Ph}$  iff coordinate points  $\bar{x}$  and  $\bar{y}$  can be connected by a light signal, i.e., if  $(x_1 - y_1)^2 - (x_2 - y_2)^2 - (x_3 - y_3)^2 - (x_4 - y_4)^2 = 0$ . Let  $\text{S} \subseteq \mathbb{R}^4 \times \mathbb{R}^4$  be the simultaneity relation, i.e.,  $(\bar{x}, \bar{y}) \in \text{S}$  iff  $x_1 = y_1$ . Consider the models  $\mathfrak{R} = \langle \mathbb{R}^4, \text{Ph} \rangle$  and  $\mathfrak{R} = \langle \mathbb{R}^4, \text{S}, \text{Ph} \rangle$ , these models capture the structure of special relativity and classical kinematics, respectively.

Let  $T_n = \text{Th}(\mathfrak{R})$  and  $T_r = \text{Th}(\mathfrak{R})$ . Note that  $T_n$  is in fact a conservative extension of  $T_r$  and the conceptual distance between them is 1, i.e.,  $\text{Cd}_\omega^\omega(T_n, T_r) = 1$ .

**Problem 3** (Hajnal Andréka). Let  $X$  be the class of all theories  $T$  such that  $T_r$  is faithfully interpreted into  $T$  and  $T$  is faithfully interpreted into  $T_n$ . Is the following true: For all  $T \in X$ ,

$$\text{Cd}_\omega^\omega(T_r, T) + \text{Cd}_\omega^\omega(T, T_n) = 1?$$

If the answer to the question in the above problem is yes, then no matter which classical (i.e.,  $T_n$ -definable, but not  $T_r$ -definable) concept we add to special relativity ( $T_r$ ) we will get classical kinematics ( $T_n$ ). That would be an interesting insight for better understanding the connection between classical and relativistic concepts.

The investigation in this section opens so many questions: For any two concrete theories of physics, what is the conceptual distance between them? By Theorem 5.1, relativistic and classical kinematics are of conceptual distance one. However, the question “what is the distance between relativistic and classical dynamics?” remains open. Another natural related open problem is the following.

**Problem 4** (Jean Paul Van Bendegem). What is the conceptual distance between classical and statistical thermodynamics?

Of course, any answer to the above problems depends on the chosen axiomatizable theories capturing the physical theories in question. For an axiomatic approach of these thermodynamics theories, one can see, e.g., [Carathéodory, 1909], [Cooper, 1967] and [Lieb and Yngvason, 2000].

## 6. IDEAS FOR OTHER DISTANCES

Interpreting a theory into another one is a fundamental concept in logic. In the following definition, we define some distance that uses faithful interpretations as the minimal step between theories.

**Definition 6.1.** For any two theories  $T_1$  and  $T_2$ , we write  $T_1 \parallel T_2$  iff one of these theories can be interpreted faithfully into the other one. Let  $X$  be an arbitrary class of theories in  $\mathbf{L}_\alpha^\beta$ . The **faithful interpretation distance** on  $X$  is defined as the step distance on the cluster network  $(X, \equiv, \mathbb{I})$ .

Another natural idea one may have for defining a distance between theories is using a step that collapses two concepts into one, i.e., using the symmetric closure of following relation for the minimal single steps.

**Definition 6.2.** Let  $T, T'$  be two theories. We say that  $T'$  is the resultant of  $T$  after **collapsing two concepts** iff  $T' \equiv T \cup \{\varphi \leftrightarrow \psi\}$ , for some  $\varphi, \psi \in \text{Fm}$ .

**Example 6.3.** The theory of abelian groups (in ordinary first order logic) is the resultant of the theory of all groups after collapsing the concepts  $a \cdot b = c$  and  $b \cdot a = c$ .

One can easily see that collapsing two concepts of a theory  $T$  is a special case of adding an axiom to  $T$ . The converse is also true; adding an axiom  $\varphi$  to  $T$  is equivalent to collapsing the concept  $\varphi$  with any theorem of  $T$ . So using the symmetric closure of the above relation for generating a distance will give the axiomatic distance (Definition 3.7).

**6.1. Dropping symmetry.** In several cases, it might be natural not to assume the symmetry of distances between theories. For example, any inconsistent theory is understood to be of axiomatic distance 1 from any consistent theory; we just need to add a contradiction as an axiom. But starting from an inconsistent theory, we can never reach a consistent theory by adding axioms; so considering this distance to be  $\infty$  seems more natural.

Now, let us mimic the work of section 3 under the consideration that symmetry is not required. For instance, a *directed cluster network* is a triple  $(X, E, R)$ , where  $(X, E)$  is a cluster and  $R$  is an arbitrary relation on  $X$ . For directed cluster network  $\mathcal{X} = (X, E, R)$ , the *directed step distance*  $\vec{d}_{\mathcal{X}} : X \times X \rightarrow \mathbb{N} \cup \{\infty\}$  can be defined completely analogously to step distance. With a very similar argument to the proof of Theorem 3.5, we can see that the following are true: For each  $T_1, T_2, T_3 \in X$ ,

- (a)  $\vec{d}_{\mathcal{X}}(T_1, T_2) \geq 0$  and  $\vec{d}_{\mathcal{X}}(T_1, T_2) = 0 \iff T_1 E T_2$ .
- (b)  $\vec{d}_{\mathcal{X}}(T_1, T_2) \leq \vec{d}_{\mathcal{X}}(T_1, T_3) + \vec{d}_{\mathcal{X}}(T_3, T_2)$ .

**Definition 6.4.** Recall the axiom adding relation  $\leftarrow$  introduced on page 7 herein. Let  $X$  be a class of theories in logic  $L_{\alpha}^{\beta}$ , then  $(X, \equiv, \leftarrow)$  is a directed cluster network. Its directed step distance is called the *directed axiomatic distance* on  $X$ .

Clearly, Theorem 3.11 is no longer true if we replace the axiomatic distance by the mimicked directed one. So, this is one of the situations where dropping the symmetry might be more interesting. Not just the axiomatic distance can be directed, but also the conceptual distance, faithful interpretation distance, and so on. It might be more appropriate to call these distances “uni-directed distances”, indeed in the directed cluster network the single steps are determined by only one relation. We can also define *bi-directed distances* or *multi-directed distances* where the single steps can be determined by two or more relations.

For example, it might be useful to introduce a “new conceptual distance” that measures the minimum number of concepts needed to be added to or removed from one theory to reach the other theory up to definitional equivalence. We already have a notion for concept adding (recall  $\rightsquigarrow$  in Definition 4.4). A precise definition for concept removal is also required. So, in this case, two relations will indicate the single steps and such new conceptual distance must be a bi-directed step distance. The notion of concept-removal below is inspired by the idea how the concept of faster-than-light observers were removed from the theory capturing classical kinematics in [Lefever, 2017] and [Lefever and Székely, 2018a].

**Definition 6.5.** Let  $T, T^-$  be two theories. We say that  $T^-$  is a *concept-removal* of  $T$  and we write  $T \triangleright T^-$  iff there is  $\varphi \in \text{Fm}$  such that  $T^- = T_m \cup \{\neg\varphi\}$  for some maximal consistent subtheory of  $\text{Cn}(T)$  for which  $T_m \not\models \varphi$ .

In some cases, it is possible to remove finitely many concepts in only one step. For instance, assume that  $\alpha = 0$  and  $\beta = 1$ , and let  $\mathcal{L}_1 = \{P_1, P_2\}$  and  $\mathcal{L}_2 = \{P_1, P_2, P_3, P_4\}$  for some sentential constants  $P_1, P_2, P_3$  and  $P_4$ . For each  $i \in \{1, 2\}$ , let  $T_i$  be the empty theory on the language  $\mathcal{L}_i$ . Recall the

proof of Theorem 4.7, it was shown that we need exactly two steps of adding concepts to get  $T_2$  from  $T_1$ . However, removing a single concept, namely  $(P_2 \neq P_3) \vee (P_3 \neq P_4)$ , from  $T_2$  gives  $T_1$ .

**Definition 6.6.** Let  $X$  be a class of theories and let  $R$  be the union of the relations  $\rightsquigarrow$  and  $\triangleright$ . The *bi-directed conceptual distance* on  $X$  is the directed step distance on  $(X, \rightrightarrows, R)$ . In the case when  $X$  is the class of all theories of logic  $\mathbf{L}_\alpha^\beta$ , we denote this distance by  $\vec{C}d_\alpha^\beta$ .

The facts, in the paragraph before the above definition, that  $\vec{C}d_\alpha^\beta(T_1, T_2) = 2$  and  $\vec{C}d_\alpha^\beta(T_2, T_1) = 1$  show that the bi-directed distance is not symmetric. Hence, conceptual and bi-directed conceptual distances are different, but it is still interesting to understand how much different they are. For example, the following problem is worth investigating.

**Problem 5.** Are there two theories  $T$  and  $T'$  such that

- (1)  $Cd_\alpha^\beta(T, T')$ ,  $\vec{C}d_\alpha^\beta(T, T')$  and  $\vec{C}d_\alpha^\beta(T', T)$  are all finite,
- (2)  $Cd_\alpha^\beta(T, T') \neq \vec{C}d_\alpha^\beta(T, T')$  and  $Cd_\alpha^\beta(T, T') \neq \vec{C}d_\alpha^\beta(T', T)$ ?

It is not straightforward to find this example, each conceptual distance calculated in the present paper coincides with one of its corresponding bi-directed conceptual distances. In order to do the same with axiomatic distance, i.e., to define a *bi-directed axiomatic distance*, all we need is a precise definition for theorem removal. We propose the following theorem-removal notion.

**Definition 6.7.** Let  $T, T^-$  be two theories. We say that  $T^-$  is a *theorem-removal* of  $T$  iff there is  $\varphi \in \text{Fm}$  such that  $T \models \varphi$  and  $T^-$  is maximal consistent subtheory of  $\text{Cn}(T)$  for which  $T^- \not\models \varphi$ .

## 7. CONCLUDING PHILOSOPHICAL REMARKS

One very important topic in the philosophy of science, is how different scientific theories can be compared to each other, especially in the case of competing theories. The first criterion for theory comparison is empirical adequacy. One theory is better than another if it accounts for more of the data or phenomena than another. In the past, this has often been, or has been presented as being, a fairly straightforward matter to decide by philosophers and historians.

This is oversimplified for two reasons. One is that sometimes in the history of science, one theory accounts for some of the data or phenomena very well. Another accounts for another area of data very well. They both agree on, but have different accounts of, the same data and both have failings. The comparison of two theories in terms of empirical adequacy requires that we count the data or the phenomena. Deciding what to count, and how to assign weight to it, has some arbitrariness to it. This is very well illustrated in [Chang, 2012] where he discusses the history of the competition between the phlogiston theory of water and the compound theory of water. With Chang, we conclude that deciding that one theory is more empirically adequate than another is not at all times, and in all circumstances, simple and straightforward, and with the fragmentation of science into more and more specialized areas of research, it is increasingly rare to find empirical adequacy to be enough to decide between competing theories.

Worse: with more complicated and “cutting-edge” examples, we find that observation in science is not simple, but is informed by instruments, theory and language; making an observation is an informed and educated act. At the edges of science, we make observations using highly specialized instruments, which are constructed based on their own theories. Thus, what looked in the past to be a relatively simple judgment to make: “this scientific theory is better than this other” turns out to

be rather subtle; since it requires individuation and assigning weights, in a way that is independent with respect to the theories themselves.

Under an over simplified view of the unity of science, the subtlety threatens any pretense science has to objectivity, because what counts as a true and verifiable statement takes specialized instruments that we assume to work according to the theory we have of the instrument. If someone has an alternative account, then the explanation for the phenomenon changes. As a result, if we want to recover some semblance of objectivity in science, it is ever more pressing to receive confirmation of a theory from other directions independent of the theory.

In terms of objectivity, one reassuring feature of science is its precision. Logic is the most precise form of investigation. Under the pressure of our considerations above, when we have several logical theories that are each to some extent empirically adequate, it is not clear that we should retain one and dismiss the other. We then have pluralism in science. Pluralism in science is an obvious philosophical position when we consider that several theories are all more-or-less empirically adequate, and show merits with respect to other, incomparable, or only artificially comparable, virtues and vices. The virtues might include: simplicity (determined by language, proofs, metaphysical parsimony or concepts), meeting a particular goal of the scientists, neatness of categorization, breadth of categorisation, standardization of explanation or meeting operational opportunities and so on. Vices might include: complexity (determined by language, proofs, metaphysical elaborations that have little use within a theory or concepts), goal failure, messy or narrow categorization, non-standard explanations, being too ambitious and not going outside the constraints of the operations available. With this plethora of incommensurable theories all competing and each adequate in their own way and for their own purpose, we look elsewhere than between the theories in and of themselves and the data to make sense of the present state of science. What we then look for are other ways of comparing theories, while accepting them until such time as we come up against a good reason to give one up, such as: its being refuted by new evidence or being too remote from too many other theories to be worth pursuing (now).

For this reason, the relations between theories, independent of their relation to reality, becomes very important. Until now, this area of study has mostly been *qualitative*. In the present paper, we explore a new *quantitative* approach: the measured distance between theories. To establish this distance, we need to study the structure of the differences, i.e., the connections, between theories. By developing several such metrics based on our definitions, and noticing that some are less interesting than others, we already learn a lot. Counting axioms does not give us much information about the distance between theories. Counting concepts is much more subtle and informative.

One area of study that has a close relationship with the notion of conceptual distance is that of complexity. As we know, complexity, also can be measured in several ways: Turing complexity, in terms of the analytic hierarchy, and so on. If one theory is more complex than another in one of these measures, then it is natural to investigate the relationship between that and the distances we look at here. Some of the significance of the present work might be in its relationship to complexity theory. This is a subject of future investigation.

The idea of having a notion of distance between theories (of the same nature) seems applicable in any science. In computer science, programming languages and other systems can be seen as axiomatized theories. For more details about this, see, e.g., [Floyd, 1967], [Hoare, 1969] and [Meyer and Halpern, 1982]. Hence, it seems also natural to search for the best fit notion of equivalence between these theories. Developing this may give us insight to determine what can be one

step difference between two such theories. Having these in mind, a distance can be then defined in the same way of section 3 herein. The novelty here would be in choosing such equivalence and one step relation in a way that guarantees that the corresponding step distance is applicable.

## ACKNOWLEDGEMENT.

We shall thank Hajnal Andr eka, Jean Paul Van Bendegem, Zal an Gyenis, P eter Juh asz, Istv an N emeti and others for their fruitful comments and questions. We are especially grateful for Judit X. Madar asz because of her insightful comments leading us to significantly improve the paper.

## REFERENCES

- [Andr eka et al., 2004] Andr eka, H., Madar asz, J. X., and N emeti, I. (2004). Logical analysis of relativity theories. In *First-order Logic Revisited*, pages 7–36. Logos Verlag.
- [Andr eka et al., 2005] Andr eka, H., Madar asz, J. X., and N emeti, I. (2005). Mutual definability does not imply definitional equivalence, a simple example. *Mathematical Logic Quarterly*, 51,6:591–597.
- [Andr eka et al., 2012] Andr eka, H., Madar asz, J. X., N emeti, I., and Sz ekely, G. (2012). A logic road from special relativity to general relativity. *Synthese*, 186,3:633–649.
- [Andr eka et al., 2002] Andr eka, H., Madar asz, J. X., N emeti, I., with contributions from: Andai, A., S agi, G., Sain, I., and T oke, C. (2002). *On the logical structure of relativity theories*. Research report, Alfr ed R enyi Institute of Mathematics, Hungar. Acad. Sci., Budapest. <https://old.renyi.hu/pub/algebraic-logic/Contents.html>.
- [Andr eka and N emeti, 2017] Andr eka, H. and N emeti, I. (2017). How many varieties of cylindric algebras are there. *Trans. Amer. Math. Soc.*, 369:8903–8937.
- [Baltag and Smets, 2005] Baltag, A. and Smets, S. (2005). Complete axiomatizations for quantum actions. *International Journal of Theoretical Physics*, 44(12):2267–2282.
- [Barrett and Halvorson, 2016] Barrett, T. W. and Halvorson, H. (2016). Glymour and quine on theoretical equivalence. *Journal of Philosophical Logic*, 45(5):467–483.
- [Carath eodory, 1909] Carath eodory, C. (1909). Untersuchungen  uber die grundlagen der thermodynamik. *Math. Ann.*, 64:355–386.
- [Chang, 2012] Chang, H. (2012). *Is Water H<sub>2</sub>O? Evidence, Realism and Pluralism*. Springer, Dordrecht.
- [Cooper, 1967] Cooper, J. L. B. (1967). The foundations of thermodynamics. *Journal of Mathematical Analysis and Applications*, 17:172–193.
- [Floyd, 1967] Floyd, R. W. (1967). Assingining meanings to programs. *Proceedings of symposia in applied mathematics*, 19.
- [Goodman, 1943] Goodman, N. (1943). On the simplicity of ideas. *The Journal of Symbolic Logic*, 8(4):107–121.
- [Henkin et al., 1971] Henkin, L., Monk, J., and Tarski, A. (1971). *Cylindric Algebras Part I*. North-Holland.
- [Henkin et al., 1985] Henkin, L., Monk, J., and Tarski, A. (1985). *Cylindric Algebras Part II*. North-Holland.
- [Hoare, 1969] Hoare, C. A. R. (1969). An axiomatic basis for computer programming. *Communications of the ACM*, 12(10):576–583.
- [Japaridze and Jongh, 1998] Japaridze, G. and Jongh, D. (1998). The logic of provability. In *Handbook of proof theory, volume 137*, pages 475–550. Elsevier.
- [Krause and Arenhart, 2017] Krause, D. and Arenhart, J. R. B. (2017). *The Logical Foundations of Scientific Theories: Languages, Structures, and Models*. Routledge.
- [Lefever, 2017] Lefever, K. (2017). *Using Logical Interpretation and Definitional Equivalence to compare Classical Kinematics and Special Relativity Theory*. PhD thesis, Vrije Universiteit Brussel.
- [Lefever and Sz ekely, 2018a] Lefever, K. and Sz ekely, G. (2018a). Comparing classical and relativistic kinematics in first-order-logic. *Logique et Analyse*, 61(241):57–117.
- [Lefever and Sz ekely, 2018b] Lefever, K. and Sz ekely, G. (2018b). On generalization of definitional equivalence to languages with non-disjoint signatures. pre-print, arXiv:1802.06844.
- [Lieb and Yngvason, 2000] Lieb, E. H. and Yngvason, J. (2000). A fresh look at entropy and the second law of thermodynamics. *Physics Today*, pages 32–37.
- [Meyer and Halpern, 1982] Meyer, A. R. and Halpern, J. Y. (1982). Axiomatic def’initions of programming languages: A theoretical assessment. *Journal of the Association for Computing Machinery*, 29(2):555–576.

- [Pinter, 1978] Pinter, C. C. (1978). Properties preserved under definitional equivalence and interpretations. *Zeitschr. f. math. Logik und Grundlagen d. math.*, 24:481–488.
- [van Dalen, 2013] van Dalen, D. (2013). *Logic and Structure*. Springer-Verlag London.
- [Visser, 2006] Visser, A. (2006). Categories of theories and interpretations. In *Logic in Tehran. Proceedings of the workshop and conference on Logic, Algebra and Arithmetic, held October 18–22, 2003, volume 26 of Lecture Notes in Logic*, pages 284–341. ASL, A.K. Peters, Ltd., Wellesley, Mass.

MICHÈLE FRIEND

Columbian College of Arts & Sciences  
George Washington University  
*E-mail*: michele@gwu.edu

MOHAMED KHALED

Alfréd Rényi Institute of Mathematics  
Hungarian Academy of Sciences  
*E-mail*: khaled.mohamed@renyi.mta.hu

KOEN LEFEVER

Centre for Logic and Philosophy of Science  
Vrije Universiteit Brussel  
*E-mail*: koen.lefever@vub.be

GERGELY SZÉKELY

Alfréd Rényi Institute of Mathematics  
Hungarian Academy of Sciences  
*E-mail*: szekely.gergely@renyi.mta.hu