Commonsense Reasoning by Distance Semantics

Ofer Arieli

Department of Computer Science,
The Academic College of Tel-Aviv,
4 Antokolski street, Tel-Aviv 61161, Israel.
oarieli@mta.ac.il

Abstract

We introduce a uniform distance semantics for different paradigms that require nonmonotonic and paraconsistent reasoning, among which are mediators of independent data-sources, integrators of prioritized knowledge-bases, operators for iterated belief revision, and analytic tools of decision support systems. We show that the consequence relations that are induced by our framework share some desirable properties and demonstrate this by relevant applications.

1 INTRODUCTION

Many realistic decision aid problems are fraught with facets of ambiguity, anomalies and conflicts. This commands the incorporation of declarative means for expressing how to reason in various contexts involving decision or choice. Classical logic, the most advocated formalism for reasoning with mathematical theories, is not useful for this task as, for instance, *any* conclusion classically follows from an inconsistent set of assumptions. Additionally, by its definition, classical logic is monotonic, while human thinking is non-monotonic in nature, that is, the set of conclusions is not necessarily non-decreasing in the size of the premises.

Distance semantics is a subtle way of handling dynamically evolving and possibly contradictory information, as it provides quantitative means for drawing 'rational' conclusions form a given set of

assumptions. Indeed, reasoning with distance semantics is a cornerstone behind many formalisms for maintaining uncertainty, such as operators for modelling belief revision (e.g., [7, 16, 18, 24, 29]), methodologies for merging independent data-sources [20, 21, 25], consistent query answering in database systems [1, 3, 4, 9, 10, 26], and several areas in the context of social choice theory, such as group decision making [22], preference representation [23], and judgment aggregation [17, 27].

The goal of this paper is to introduce a uniform framework for the above paradigms and many other formalisms that are based on distance semantics. For this, we define a possible-world semantics, derived by distance considerations, and consider different metrics on the universe (i.e., all the possible worlds). Each metric induces a different consequence relation for reasoning with the underlying information.

The rest of this paper is organized as follows: in the next section we introduce the framework and the family of distance-based entailments that it induces. In that section we also show some basic properties of the distance entailments. In Section 3 we consider situations in which the available data is prioritized, and extend the definition of the distance-based entailments accordingly. In Section 4 we demonstrate the applicative potential of our framework and in Section 5 we conclude.

2 DISTANCE-BASED ENTAILMENTS

The intuition behind our approach is simple. Given a distance function d on a space of inter-

pretations, reasoning with a given set of premises Γ is based on those interpretations that are '*d-closest*' to Γ (called the *most plausible* interpretations of Γ). This set of interpretations determine the Γ-conclusions, which means that, unlike classical logic, our formalisms are non-monotonic and do not become trivial in the presence of contradictions. For instance, it is intuitively clear that interpretations in which q is true should be closer to $\Gamma = \{p, \neg p, q\}$ than interpretations in which q is false, and so q should follow from Γ while $\neg q$ should *not* follow from Γ, although Γ is not consistent. In what follows we formalize this idea.

2.1 PRELIMINARIES

In the sequel, we consider *finite* sets of premises (theories) in a propositional language $\mathcal{L}$ with a finite set Atoms of atomic formulas. We denote by Λ the space of the two-valued interpretations on Atoms. The set of atomic formulas that occur in the formulas of a theory Γ is denoted Atoms(Γ), and the set of models of Γ (that is, the interpretations $\nu \in \Lambda$ such that $\nu(\psi) = \mathsf{t}$ for every $\psi \in \Gamma$) is denoted $mod(\Gamma)$.

Definition 1 A *pseudo distance* on Λ is a total function $d : \Lambda \times \Lambda \to \mathbb{R}^+$ that is symmetric ($\forall u, v \in \Lambda \ d(u, v) = d(v, u)$) and preserves identity ($\forall u, v \in \Lambda \ d(u, v) = 0$ iff $u = v$). A *distance* on Λ is a pseudo distance on Λ that satisfies the triangular inequality ($\forall u, v, w \in \Lambda \ d(u, v) \leq d(u, w) + d(w, v)$).

Example 2 It is easy to verify that the following two functions are distances on Λ.

- *The drastic distance:*
  $d^U(\nu, \mu) = 0$ if $\nu = \mu$, otherwise $d^U(\nu, \mu) = 1$.

- *The Hamming distance:*
  $d^H(\nu, \mu) = |\{p \in \mathsf{Atoms} \mid \nu(p) \neq \mu(p)\}|$. [1]

Definition 3 A *numeric aggregation function* f is a total function that accepts a multiset of real numbers and returns a real number. In addition, f is non-decreasing in the values of its argument,[2]

---

[1] I.e., $d^H(\nu, \mu)$ is the number of atoms p such that $\nu(p) \neq \mu(p)$. This function is also known as the Dalal distance [12].

[2] That is, the function value is non-decreasing when an element in the multiset is replaced by a larger element.

$f(\{x_1, \ldots, x_n\}) = 0$ iff $x_1 = \ldots = x_n = 0$, and $\forall x \in \mathbb{R} \ f(\{x\}) = x$.

The aggregation function in Definition 3 may be, e.g., a summation or the average of the distances, the maximum value among those distances (which yields a worst case analysis), a median value (for mean case analysis), and so forth. Such functions are common in data integration systems (see, e.g., Section 4).

Definition 4 Given a theory $\Gamma = \{\psi_1, \ldots, \psi_n\}$, an interpretation $\nu \in \Lambda$, a pseudo-distance d on Λ, and an aggregation function f, define:

- $d(\nu, \psi_i) = \min\{d(\nu, \mu) \mid \mu \in mod(\psi_i)\}$

- $\delta_{d,f}(\nu, \Gamma) = f(\{d(\nu, \psi_1), \ldots, d(\nu, \psi_n)\})$.

Definition 5 A (pseudo) distance d is *unbiased*, if for every formula ψ and every two-valued interpretations ν_1, ν_2, if $\nu_1(p) = \nu_2(p)$ for every $p \in \mathsf{Atoms}(\psi)$, then $d(\nu_1, \psi) = d(\nu_2, \psi)$.

The last property assures that a distance between an interpretation and a formula depends only on the atoms that appear in the formula, and so it is not 'biased' by irrelevant atoms. Note, e.g., that the distances in Example 2 are unbiased.

The next definition captures the intuition behind distance semantics that the relevant interpretations of a theory Γ are those ones that are $\delta_{d,f}$-closest to Γ.

Definition 6 The *most plausible* valuations of Γ (with respect to a pseudo distance d and an aggregation function f) are the interpretations $\nu \in \Lambda$ that belong to the following set:

$$\Delta_{d,f}(\Gamma) = \{\nu \in \Lambda \mid \forall \mu \in \Lambda \ \delta_{d,f}(\nu, \Gamma) \leq \delta_{d,f}(\mu, \Gamma)\}.$$

Corresponding consequence relations are now defined as follows.

Definition 7 For a pseudo distance d and an aggregation function f, define $\Gamma \models_{d,f} \psi$ if $\Delta_{d,f}(\Gamma) \subseteq mod(\psi)$. That is, conclusions should follow from *all* the most plausible valuations of the premises.

Example 8 Let $\Gamma = \{p, q, r, \neg p \vee \neg q, r \wedge s\}$. This theory is not consistent, and so everything classically follows from it, including, e.g., $\neg r$, which

seems to be a very strange conclusion in this case.[3] Using distance-based semantics, this anomaly can be lifted. The following table lists the distances between the relevant valuations and Γ according to several common metrics:

	p	q	r	s	$\delta_{d^U,\Sigma}$	$\delta_{d^H,\Sigma}$	$\delta_{d^H,\max}$
ν_1	t	t	t	t	**1**	**1**	**1**
ν_2	t	t	t	f	2	2	1
ν_3	t	t	f	t	3	3	1
ν_4	t	t	f	f	3	4	2
ν_5	t	f	t	t	**1**	**1**	**1**
ν_6	t	f	t	f	2	2	1
ν_7	t	f	f	t	3	3	1
ν_8	t	f	f	f	3	4	2
ν_9	f	t	t	t	**1**	**1**	**1**
ν_{10}	f	t	t	f	2	2	1
ν_{11}	f	t	f	t	3	3	1
ν_{12}	f	t	f	f	3	4	2
ν_{13}	f	f	t	t	2	2	1
ν_{14}	f	f	t	f	3	3	1
ν_{15}	f	f	f	t	4	4	1
ν_{16}	f	f	f	f	4	5	2

Here, $\Delta_{d^U,\Sigma}(\Gamma) = \Delta_{d^H,\Sigma}(\Gamma) = \{\nu_1, \nu_5, \nu_9\}$, thus $\Gamma \models_{d^U,\Sigma} r$ and $\Gamma \models_{d^H,\Sigma} r$, while $\Gamma \not\models_{d^U,\Sigma} \neg r$ and $\Gamma \not\models_{d^H,\Sigma} \neg r$. As intuitively expected, s behaves similarly. Note also that the atoms p, q that are involved in the inconsistency are not deducible from Γ, nor their complements. The entailment $\models_{d^H,\max}$ is more cautious; it does not allow to infer neither $\neg r$ (as expected) nor r, but the weaker conclusion $r \vee s$ is deducible.

2.2 BASIC PROPERTIES OF $\models_{d,f}$

2.2.1 Paraconsistency

Paraconsistent logics [11, 28] are formalisms that tolerate inconsistency and do not become trivial in the presence of contradictions. As human knowledge and thinking necessarily requires inconsistency, conflicting data is unavoidable in practice, but (unlike the case of classical logic) this should not block the ability to draw 'rational' conclusions from contradictory theories. For instance, in Example 8, although Γ is not consistent, there is no reason to infer $\neg r$ from it.

---

[3]Indeed, r is not part of the inconsistent fragment of Γ, therefore it is not sensible in this case to conclude its complement.

Proposition 9 *For every pseudo distance d and aggregation function f, $\models_{d,f}$ is paraconsistent.*

Proof. By the fact that for every Γ, $\Delta_{d,f}(\Gamma) \neq \emptyset$ (as the minimal $\delta_{d,f}$-distance from Γ over a finite space of interpretations is always obtained). Thus, for every formula ψ such that there exists a valuation $\nu \in \Delta_{d,f}(\Gamma)$ for which $\nu(\psi) = \mathsf{f}$, it holds that $\Gamma \not\models_{d,f} \psi$. □

Unbiasedness of the distance function allows us to strengthen Proposition 9, as in this case a formula never follows from a theory unless they share propositional atoms.

Proposition 10 *Let d be an unbiased pseudo distance and f an aggregation function. For every theory Γ and a non-tautological formula ψ in $\mathcal{L}$ such that $\mathsf{Atoms}(\Gamma) \cap \mathsf{Atoms}(\{\psi\}) = \emptyset$, it holds that $\Gamma \not\models_{d,f} \psi$.*

Proof. Let $\Gamma = \{\psi_1, \ldots, \psi_n\}$ and $\nu \in \Delta_{d,f}(\Gamma)$. If $\nu(\psi) = \mathsf{f}$ we are done. Otherwise, consider a valuation μ that is the same as ν on $\mathsf{Atoms}(\Gamma)$ and $\mu(\psi) = \mathsf{f}$. Such a valuation exists since ψ is not a tautology, the value of $\mu(\psi)$ depends only on the assignments of μ on $\mathsf{Atoms}(\{\psi\})$, and $\mathsf{Atoms}(\Gamma) \cap \mathsf{Atoms}(\{\psi\}) = \emptyset$. Now, as d is unbiased, $d(\mu, \psi_i) = d(\nu, \psi_i)$ for every $\psi_i \in \Gamma$. Thus, $\delta_{d,f}(\mu, \Gamma) = \delta_{d,f}(\nu, \Gamma)$, and so $\mu \in \Delta_{d,f}(\Gamma)$ as well. Hence, $\Gamma \not\models_{d,f} \psi$. □

On the other hand, when the set of premises is consistent, there is no reason to draw conclusions that are different from those of standard classical logic. The following propositions show that this is exactly the situations in our framework:

Proposition 11 *Let d be a pseudo distance and f an aggregation function. For every consistent Γ, $\Delta_{d,f}(\Gamma) = mod(\Gamma)$.*

Proof. Let $\Gamma = \{\psi_1, \ldots, \psi_n\}$. If ν is a model of Γ, then $d(\nu, \psi_i) = 0$ for every $1 \leq i \leq n$, and so $\delta_{d,f}(\nu, \Gamma) = 0$ as well. Now, since for every valuation μ, $\delta_{d,f}(\mu, \Gamma) \geq 0$, necessarily $\nu \in \Delta_{d,f}(\Gamma)$.
For the converse, suppose that $\nu \notin mod(\Gamma)$. Then ν does not satisfy ψ_j for some $1 \leq j \leq n$, and so $d(\nu, \psi_j) > 0$. By Definition 3, f is strictly positive whenever it has at least one strictly positive argument and the other arguments are non-negative. We have, then, that $\delta_{d,f}(\nu, \Gamma) > 0$. On the other

hand, as $mod(\Gamma) \neq \emptyset$, there is a model μ of Γ, for which $\delta_{d,f}(\mu, \Gamma) = 0$. It follows that $\nu \notin \Delta_{d,f}(\Gamma)$. $\square$

Corollary 12 *Denote by $\models$ the standard entailment of classical logic. For every classically consistent set of formulas Γ and formula ψ, $\Gamma \models \psi$ iff $\Gamma \models_{d,f} \psi$.*

By Proposition 9 and Corollary 12 we have, therefore, the following desirable property:

Corollary 13 *Let d be an unbiased pseudo distance and f an aggregation function. Then $\models_{d,f}$ is the same as the classical entailment for consistent premises and is non-trivial otherwise.*

2.2.2 Nonmonotonicity

Another characteristic property of $\models_{d,f}$ is its non-monotonic nature, which implies that conclusions may be retracted in light of new information. In Example 8, for instance, we have that for $d = d^U$ or $d = d^H$, $\Gamma \models_{d,\Sigma} r$, but it is easy to check that $\Gamma, \neg r \not\models_{d,\Sigma} r$. As the following proposition shows, this is not a coincidence.

Proposition 14 *For every unbiased pseudo distance d and aggregation function f, $\models_{d,f}$ is non-monotonic.*

2.2.3 Adaptivity

Ataptivity, due to Batens [5, 6], is the ability to handle contradictory theories in a nontrivial way and at the same time to presuppose the consistency of all the formulas 'unless and until proven otherwise'. Consequence relations with this property *adapt* to the *specific* inconsistencies that occur in the theories.

Example 15 The Disjunctive Syllogism should *not* be applied for concluding q from $\{p, \neg p, \neg p \vee q\}$. On the other hand, in the case of $\{p, \neg p, r, \neg r \vee q\}$, applying the Disjunctive Syllogism to r and $\neg r \vee q$ may be justified by the fact that the subset of formulas to which the Disjunctive Syllogism is applied is not affected by the inconsistency of the whole theory, therefore inference rules that are classically valid can be applied to it.

The next proposition shows that for hereditary aggregation functions (see Definition 16 below) $\models_{d,f}$

is adaptive: if a given theory can be split up to a consistent and an inconsistent parts, then every assertion that is not related to the inconsistent part, and which classically follows from the consistent part, is entailed by the whole theory.

Definition 16 An aggregation function f is called *hereditary*, if whenever $f(\{x_1, \ldots, x_n\}) < f(\{y_1, \ldots, y_n\})$, also $f(\{x_1, \ldots, x_n, z_1, \ldots, z_m\}) < f(\{y_1, \ldots, y_n, z_1, \ldots, z_m\})$. [4]

Proposition 17 *Let d be an unbiased pseudo distance and f a hereditary aggregation function. Suppose that Γ is a theory that can be represented as $\Gamma' \cup \Gamma''$, where Γ' is a classically consistent theory and $\mathsf{Atoms}(\Gamma') \cap \mathsf{Atoms}(\Gamma'') = \emptyset$. Then for every formula ψ such that $\mathsf{Atoms}(\psi) \cap \mathsf{Atoms}(\Gamma'') = \emptyset$, it holds that if $\Gamma' \models \psi$ then $\Gamma \models_{d,f} \psi$.*

Proof . We first show the following lemma:

Lemma 17–A. *Let d be a non-biased pseudo distance and f a hereditary aggregation function. If $\Gamma \models_{d,f} \psi$ then $\Gamma, \phi \models_{d,f} \psi$ for every ϕ such that $\mathsf{Atoms}(\Gamma \cup \{\psi\}) \cap \mathsf{Atoms}(\phi) = \emptyset$.*

Indeed, let $\Gamma = \{\psi_1, \ldots, \psi_n\}$. If ψ is a tautology the lemma trivially holds. Otherwise, let μ be a valuation such that $\mu(\psi) = \mathsf{f}$. As $\Gamma \models_{d,f} \psi$, necessarily $\mu \notin \Delta_{d,f}(\Gamma)$, and so there is a $\nu \in \Delta_{d,f}(\Gamma)$ for which $d_f(\nu, \Gamma) < d_f(\mu, \Gamma)$. In oder words,

$$f(\{d(\nu, \psi_1), \ldots, d(\nu, \psi_n)\}) < \\ f(\{d(\mu, \psi_1), \ldots, d(\mu, \psi_n)\}).$$

Again, since $\Gamma \models_{d,f} \psi$, $\nu(\psi) = \mathsf{t}$. Now, consider a valuation σ, defined for every atom p as follows:

$$\sigma(p) = \begin{cases} \nu(p) & \text{if } p \in \mathsf{Atoms}(\Gamma \cup \psi), \\ \mu(p) & \text{otherwise.} \end{cases}$$

Note that $\sigma(p) = \nu(p)$ for all $p \in \mathsf{Atoms}(\psi)$, thus $\sigma(\psi) = \mathsf{t}$. Now, as $\mathsf{Atoms}(\Gamma \cup \{\psi\}) \cap \mathsf{Atoms}(\phi) = \emptyset$, d is non-biased, and f is hereditary, we have that

$$\begin{aligned} d_f(\sigma&, \Gamma \cup \{\phi\}) \\ &= f(\{d(\sigma, \psi_1), \ldots, d(\sigma, \psi_n), d(\sigma, \phi)\}) \\ &= f(\{d(\nu, \psi_1), \ldots, d(\nu, \psi_n), d(\mu, \phi)\}) \\ &< f(\{d(\mu, \psi_1), \ldots, d(\mu, \psi_n), d(\mu, \phi)\}) \\ &= d_f(\mu, \Gamma \cup \{\phi\}). \end{aligned}$$

---

[4]Note that hereditary, unlike monotonicity, is defined by strict inequalities. Thus, for instance, summation of distances is hereditary (as distances are non-negative), while the maximum function is not.

Thus, for every valuation μ such that $\mu(\psi) = \mathsf{f}$ there is a valuation σ such that $\sigma(\psi) = \mathsf{t}$ and $d_f(\sigma, \Gamma \cup \{\phi\}) < d_f(\mu, \Gamma \cup \{\phi\})$. It follows that the elements of $\Delta_{d,f}(\Gamma \cup \{\phi\})$ must satisfy ψ, and so $\Gamma, \phi \models_{d,f} \psi$.

Now, to complete the proof of Proposition 17, suppose that $\Gamma' \models \psi$. Then by Corollary 12, $\Gamma' \models_{d,f} \psi$. Thus, as $\mathsf{Atoms}(\Gamma' \cup \{\psi\}) \cap \mathsf{Atoms}(\Gamma'') = \emptyset$, we have, by the Lemma 17–A, that $\Gamma \models_{d,f} \psi$. $\qquad \square$

Note 18 The condition on the aggregation function in Proposition 17 is indeed necessary. To see this, consider the theory Γ in Example 8. This theory can be split up to a consistent subtheory $\Gamma' = \{r, r \wedge s\}$ and an inconsistent subtheory $\Gamma'' = \{p, q, \neg p \vee \neg q\}$. Also, $\mathsf{Atoms}(\Gamma') \cap \mathsf{Atoms}(\Gamma'') = \emptyset$. Yet, although $\Gamma' \models r$, we have that $\Gamma \not\models_{d^H, \max} r$. Indeed, max is not a hereditary function, so adaptivity is not assured here. It follows, then, that unlike $\models_{d^H, \Sigma}$, which is adaptive by Proposition 17, $\models_{d^H, \max}$ is not adaptive.

3   PRIORITIZED THEORIES

We now extend the distance-based semantics considered above to prioritized theories. Reasoning with prioritized information is very frequent in everyday life, and it is a cornerstone of many paradigms of knowledge representation and reasoning, such as annotated logic [30] and possibilistic logic [15]. Moreover, in many cases the nature of the underlying data dictates preferences. This is the case, for instance, in iterated belief revision [13, 14, 19], where more recent data is considered more accurate and so it is preferred over older one, in database systems, where integrity constraints have higher precedence over database facts, and in consistency restoration of ranked knowledge-bases [2, 8], where pieces of information may come from different sources with different reliability. In what follows we introduce a unifying distance-based approach for defining the underlying semantics of such systems.

Definition 19 An *n-prioritized theory* is a set $\Gamma^{\langle n \rangle}$ of formulas in $\mathcal{L}$, partitioned into $n \geq 1$ pairwise disjoint sub-theories Γ_i $(1 \leq i \leq n)$. We denote this by $\Gamma^{\langle n \rangle} = \Gamma_1 \oplus \Gamma_2 \oplus \ldots \oplus \Gamma_n$.

In what follows we shall usually write Γ instead of $\Gamma^{\langle n \rangle}$. Intuitively, formulas in higher levels of Γ are of higher priority than those in lower levels. That is, if $1 \leq i < j \leq n$, then every formula in Γ_j is preferred over the formulas in Γ_i.[5] Next we formalize this intuition by a distance semantics.

Definition 20 Let d be a pseudo distance and f an aggregation function. For an n-prioritized theory $\Gamma = \Gamma_1 \oplus \Gamma_2 \oplus \ldots \oplus \Gamma_n$ consider the following n sets of interpretations:

- $\Delta_{d,f}^n(\Gamma) = \big\{ \nu \in \Lambda \mid \forall \mu \in \Lambda$
$$\delta_{d,f}(\nu, \Gamma_n) \leq \delta_{d,f}(\mu, \Gamma_n) \big\}$$

- for every $1 \leq i < n$,
$\Delta_{d,f}^{n-i}(\Gamma) = \big\{ \nu \in \Delta_{d,f}^{n-i+1}(\Gamma) \mid \forall \mu \in \Delta_{d,f}^{n-i+1}(\Gamma)$
$$\delta_{d,f}(\nu, \Gamma_{n-i}) \leq \delta_{d,f}(\mu, \Gamma_{n-i}) \big\}$$

The sequence $\Delta_{d,f}^n(\Gamma), \ldots, \Delta_{d,f}^1(\Gamma)$ is clearly non-increasing, as sets with smaller indices are subsets of those with higher indices. This reflects the intuitive idea that higher levelled formulas are preferred over lower levelled formulas, thus the interpretations of the latter are determined by the interpretations of the former.

Let $\Delta_{d,f}(\Gamma)$ be the last set in the sequence (that is, $\Delta_{d,f}(\Gamma) = \Delta_{d,f}^1(\Gamma)$). The elements of $\Delta_{d,f}(\Gamma)$ are the *most plausible interpretations* of Γ. Again, these interpretations determine the Γ-conclusions:

Definition 21 Let d be a pseudo distance and f an aggregation function. A formula ψ *follows* from an an n-prioritized theory Γ (notation: $\Gamma \models_{d,f} \psi$) if $\Delta_{d,f}(\Gamma) \subseteq mod(\psi)$.

Example 22 Consider the following three-leveled theory:

$$\Gamma^{\langle 3 \rangle} = \Gamma_1 \oplus \Gamma_2 \oplus \Gamma_3 = \{\neg q, r\} \oplus \{\neg p \vee q\} \oplus \{p\}.$$

Then:

$$\Gamma^{\langle 3 \rangle} \models_{d^H, \Sigma} p, \ \Gamma^{\langle 3 \rangle} \models_{d^H, \Sigma} q, \text{ and } \Gamma^{\langle 3 \rangle} \models_{d^H, \Sigma} r,$$

while

$$\Gamma^{\langle 3 \rangle} \not\models_{d^H, \Sigma} \neg p, \ \Gamma^{\langle 3 \rangle} \not\models_{d^H, \Sigma} \neg q, \text{ and } \Gamma^{\langle 3 \rangle} \not\models_{d^H, \Sigma} \neg r.$$

Clearly, Definition 21 is a conservative extension of Definition 7:

---

[5] Note that in this writing the precedence is *right-hand increasing*.

Proposition 23 *For $n = 1$ we have that $\Gamma^{\langle n \rangle} \models_{d,f} \psi$ iff $\Gamma \models_{d,f} \psi$.*

Definition 24 Let Γ be an n-prioritized theory. For any $1 \leq i \leq n$, denote by $\Gamma_{\geq i}$ the $n - i + 1$ highest levels of Γ, that is: $\Gamma_{\geq i} = \Gamma_i \oplus \ldots \oplus \Gamma_n$. The *consistency level* con of Γ is the minimal value $i \leq n$ such that $\Gamma_i \cup \ldots \cup \Gamma_n$ is consistent. If there is no such value, let con $= n+1$ (and then $\Gamma_{\geq \mathsf{con}} = \emptyset$).

The following proposition is the analogue, for the prioritized case, of Proposition 11.

Proposition 25 *Let d be a pseudo distance and f an aggregation function. For every n-prioritized theory Γ with a consistency level con, $\Delta_{d,f}(\Gamma_{\geq \mathsf{con}}) = mod(\cup_{\mathsf{con} \leq i \leq n} \Gamma_i)$.*

Proof. The proof is obtained by a simple accommodation to the prioritized case of the proof of Proposition 11. $\square$

Some relations to the entailment $\models$ of classical logic are listed below:

Proposition 26 *Let d be a pseudo distance and f an aggregation function. For every n-prioritized theory Γ with a consistency level con, we have that*

1. *$\Gamma_{\geq \mathsf{con}} \models \psi$ iff $\Gamma_{\geq \mathsf{con}} \models_{d,f} \psi$.* [6]

2. *$\Gamma_{\geq \mathsf{con}} \models \psi$ implies that $\Gamma \models_{d,f} \psi$.*

3. *$\Gamma \models_{d,f} \psi$ implies that $\Gamma \models \psi$.*

Proof. The first part of the proposition is a clear consequence of Proposition 25.
For the second part, note that the definition of $\Delta_{d,f}$ implies that $\Delta_{d,f}(\Gamma) \subseteq \Delta_{d,f}(\Gamma_{\geq \mathsf{con}})$. Again, by Proposition 25, $\Delta_{d,f}(\Gamma_{\geq \mathsf{con}}) = mod(\Gamma_{\geq \mathsf{con}})$, and so $\Delta_{d,f}(\Gamma) \subseteq mod(\Gamma_{\geq \mathsf{con}})$. Now, if $\Gamma_{\geq \mathsf{con}} \models \psi$, then ψ is true in every element of $mod(\Gamma_{\geq \mathsf{con}})$, which implies that ψ must be true in every element of $\Delta_{d,f}(\Gamma)$, thus $\Gamma \models_{d,f} \psi$.
The last part holds since if Γ is consistent, i.e., con $= 1$, then by Proposition 25 $\Delta_{d,f}(\Gamma) = mod(\Gamma)$, thus $\Gamma \models_{d,f} \psi$ iff $\Gamma \models \psi$. If Γ is not consistent, then the claim trivially holds, since for *every* formula ψ, $\Gamma \models \psi$. $\square$

---

[6] Here and in what follows, $\Gamma_{\geq \mathsf{con}} \models \psi$ means that every model of the formulas of $\Gamma_{\geq \mathsf{con}}$ is a model of ψ. That is, for the standard satisfaction relation $\models$ one ignores the priorities among the premises.

Note also, that although $\models_{d,f}$ is non-monotonic, we have, by items 1 and 2 of Proposition 26, that $\Gamma_{\geq \mathsf{con}} \models_{d,f} \psi$ implies that $\Gamma \models_{d,f} \psi$. [7]

4 APPLICATIONS

As noted above, distance semantics is in the heart of different disciplines involving knowledge representation and reasoning. Below, we demonstrate the application potential of our framework in the context of merging and analyzing constraint data sources. In this case we have *several* theories (each one corresponds to a different source) that may have different priorities. That is, the underlying n-prioritized theory is now of the form

$$\Gamma = \{T_1^1, \ldots, T_{k_1}^1\} \oplus \ldots \oplus \{T_1^n, \ldots, T_{k_n}^n\},$$

where each T_j^i is a different theory, theories with the same superscript have the same precedence, and T^i is preferred over T^j iff $i > j$. To handle this situation we further generalize Definition 20 by incorporating another aggregation function for merging the sources:

Definition 27 Let d be a pseudo distance and f, g two aggregation functions. For an n-prioritized theory $\Gamma = \{T_1^1, \ldots, T_{k_1}^1\} \oplus \ldots \oplus \{T_1^n, \ldots, T_{k_n}^n\}$ and every $1 \leq i \leq n$, let

$$\delta_{d,f,g}(\nu, \overline{T^i}) = g\big(\{\delta_{d,f}(\nu, T_1^i), \ldots, \delta_{d,f}(\nu, T_{k_i}^i)\}\big).$$

Now, consider the following n sets of interpretations:

- $\Delta_{d,f,g}^n(\Gamma) = \big\{\nu \in \Lambda \mid \forall \mu \in \Lambda$
$$\delta_{d,f,g}(\nu, \overline{T^n}) \leq \delta_{d,f,g}(\mu, \overline{T^n})\big\}$$

- for every $1 \leq i < n$,
$\Delta_{d,f,g}^{n-i}(\Gamma) = \big\{\nu \in \Delta_{d,f,g}^{n-i+1}(\Gamma) \mid \forall \mu \in \Delta_{d,f,g}^{n-i+1}(\Gamma)$
$$\delta_{d,f,g}(\nu, \overline{T^{n-i}}) \leq \delta_{d,f,g}(\mu, \overline{T^{n-i}})\big\}$$

The *most plausible valuations* (with respect to d, f, g) of Γ are the valuations in $\Delta_{d,f,g}^1(\Gamma)$.

Example 28 Consider the following scenario (borrowed from [21]) regarding speculations on the

---

[7] This also follows from Definition 20. In fact, for every $1 \leq i \leq j \leq n$, if $\Gamma_{\geq j} \models_{d,f} \psi$ then $\Gamma_{\geq i} \models_{d,f} \psi$.

stock exchange: Suppose that an investor (represented by the mediator system) consults four financial experts about their opinion regarding four different shares, denoted s_1, s_2, s_3 and s_4. The opinion of expert i is represented by the theory (data-source) T_i:

$$T_1 = T_2 = \{s_1, s_2, s_3\},$$
$$T_3 = \{\neg s_1, \neg s_2, \neg s_3, \neg s_4\},$$
$$T_4 = \{s_1, s_2, \neg s_4\}.$$

For instance, in our case expert 4 suggests to buy shares s_1 and s_2, doesn't recommend to buy share s_4, and doesn't have a particular opinion about s_3.

In addition, the investor may have further restrictions about his or her investments. For instance, if some share, say s_4, is considered risky, buying it may be balanced by purchasing at least two out of the three other shares and vice-versa. This policy may be represented by the following integrity constraint:

$$\mathcal{IC} = \{s_4 \longleftrightarrow \big((s_1 \wedge s_2) \vee (s_2 \wedge s_3) \vee (s_1 \wedge s_3)\big)\}.$$

Assuming that the experts are equally faithful, their suggestions may be represented by the 2-prioritized theory $\Gamma = \{T_1, T_2, T_3, T_4\} \oplus \{\mathcal{IC}\}$, in which the investor's constraint about the purchasing policy is of higher precedence than the experts' opinions. Now, taking the drastic distance d^U and the summation function for the aggregation functions f and g, we get, by Proposition 25, that the most plausible valuations of Γ belong to the following set:

$$\Delta_{d^U, \Sigma, \Sigma}(\Gamma) = \{\nu \in mod(\mathcal{IC}) \mid \forall \mu \in mod(\mathcal{IC})$$
$$\delta_{d, \Sigma, \Sigma}(\nu, \{T_i \mid 1 \le i \le 4\}) \le$$
$$\delta_{d, \Sigma, \Sigma}(\mu, \{T_i \mid 1 \le i \le 4\})\}.$$

The models of $\mathcal{IC}$ and their distances to $\overline{T} = \{T_1, \ldots, T_4\}$ are given below.

	s_1	s_2	s_3	s_4	$\delta_{d^U, \Sigma, \Sigma}(\nu_i, \overline{T})$
ν_1	t	t	t	t	**5**
ν_2	t	t	f	t	7
ν_3	t	f	t	t	7
ν_4	t	f	f	f	7
ν_5	f	t	t	t	7
ν_6	f	t	f	f	6
ν_7	f	f	t	f	6
ν_8	f	f	f	f	8

Thus $\Delta_{d^U, \Sigma, \Sigma}(\Gamma) = \{\nu_1\}$, and so the investor will purchase all the four shares.

Clearly, the experts could have different reputations, and this may affect the investor's decision. For instance, assuming that expert 4 has a better reputation in the eye of the investor, his opinion may get a higher precedence, yielding the following 3-prioritized theory: $\Gamma' = \{T_1, T_2, T_3\} \oplus \{T_4\} \oplus \{\mathcal{IC}\}$. It is interesting to note that in this case the recommendation of the most significant expert (number 4) does not comply with the investor's restriction. Using the same setting as before (i.e., $d = d^U$ and $f = g = \Sigma$), the investor ends up with a different investment policy, according to the following tables:

	s_1	s_2	s_3	s_4	$\delta_{d^U, \Sigma, \Sigma}(\nu_i, T_4)$
ν_1	t	t	t	t	**1**
ν_2	t	t	f	t	**1**
ν_3	t	f	t	t	2
ν_4	t	f	f	f	**1**
ν_5	f	t	t	t	2
ν_6	f	t	f	f	**1**
ν_7	f	f	t	f	2
ν_8	f	f	f	f	2

	s_1	s_2	s_3	s_4	$\delta_{d^U, \Sigma, \Sigma}(\nu_i, \{T_1, T_2, T_3\})$
ν_1	t	t	t	t	0+0+4 = 4
ν_2	t	t	f	t	1+1+3 = 5
ν_3	t	f	t	t	N.A.
ν_4	t	f	f	f	**1+1+1 = 3**
ν_5	f	t	t	t	N.A.
ν_6	f	t	f	f	**1+1+1 = 3**
ν_7	f	f	t	f	N.A.
ν_8	f	f	f	f	N.A.

Thus $\Delta_{d^U, \Sigma, \Sigma}(\Gamma') = \{\nu_4, \nu_6\}$, and the decision would be to purchase either s_1 or s_2 *but not both*, which seems as a 'fair balance' between the investor's restriction and the recommendation of the most significant expert (taking into account also the other recommendations).

5 CONCLUSION

The principle of minimal change is a primary motif in commonsense reasoning, and it is often implicitly derived by distance considerations. In this paper, we introduced a simple and natural framework for representing this principle in an explicit way,

and explored the main logical properties of the corresponding consequence relations. It is shown that the entailments induced by distance semantics sustain different aspects of human thinking, such as non-monotonicity, paraconsistency, and adaptivity, and are useful for handling incomplete, inconsistent, and prioritized information.

References

[1] M. Arenas, L. Bertossi, and J. Chomicki. Consistent query answers in inconsistent databases. In *Proc. PODS'99*, pages 68–79, 1999.

[2] O. Arieli. Four-valued logics for reasoning with uncertainty in prioritized data. In B. Bouchon-Meunier, R. Yager, and L. Zadeh, editors, *Information, Uncertainty, Fusion*, pages 293–304. 1999.

[3] O. Arieli, M. Denecker, and M. Bruynooghe. Distance-based repairs of databases. In *Proc. JELIA'06*, LNAI 4160, pages 43–55. Springer, 2006.

[4] O. Arieli, M. Denecker, B. Van Nuffelen, and M. Bruynooghe. Computational methods for database repair by signed formulae. *Annals of Mathematics and Artificial Intelligence*, 46(1–2):4–37, 2006.

[5] D. Batens. Dynamic dialectical logics. In G. Priest, R. Routely, and J. Norman, editors, *Paraconsistent logic. Essay on the Inconsistent*, pages 187–217. Philosophia Verlag, 1989.

[6] D. Batens. Inconsistency-adaptive logics. In E. Orlowska, editor, *Logic at Work*, pages 445–472. Physica Verlag, 1998.

[7] J. Ben Naim. Lack of finite characterizations for the distance-based revision. In *Proc. KR'06*, pages 239–248. AAAI Press, 2006.

[8] S. Benferhat, D. Dubois, and H. Prade. How to infer from inconsistent beliefs without revising? In *Proc. IJCAI'95*, pages 1449–1455, 1995.

[9] J Chomicki. Consistent query answering: Five easy pieces. In *Proc. ICDT'07*, LNCS 4353, pages 1–17. Springer, 2007.

[10] J. Chomicki and J. Marchinkowski. Minimal-change integrity maintenance using tuple deletion. *Information and Computation*, 197(1–2):90–121, 2005.

[11] N. C. A. da Costa. On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic*, 15:497–510, 1974.

[12] M. Dalal. Investigations into a theory of knowledge base revision. In *Proc. AAAI'88*, pages 475–479. AAAI Press, 1988.

[13] A. Darwiche and J. Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89:1–29, 1997.

[14] J. Delgrande, D. Dubois, and J. Lang. Iterated revision and prioritized merging. In *Proc. KR'06*, pages 210–220. AAAI Press, 2006.

[15] D. Dubois, J. Lang, and H. Prade. Possibilistic logic. In D. Gabbay, C. Hogger, and J. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3, pages 439–513. 1994.

[16] D. Dubois and H. Prade. Belief change and possibility theory. In P. Gärdenfors, editor, *Belief Revision*, pages 142–182. Cambridge Press, 1992.

[17] D. Eckert and G. Pigozzi. Belief merging, judgment aggregation, and some links with social choice theory. In J. Delgrande, J. Lang, H. Rott, and J. Tallon, editors, *Proc. Dagstuhl semniar No.05321*. 2005.

[18] A. Grove. Two modellings for theory change. *Journal of Philosophical Logic*, 17:157–180, 1988.

[19] Y. Jin and M. Thielscher. Iterated revision, revised. In *Proc. IJCAI'05*, pages 478–483, 2005.

[20] S. Konieczny, J. Lang, and P. Marquis. Distance-based merging: A general framework and some complexity results. In *Proc. KR'02*, pages 97–108. Morgan Kaufmann Publishers, 2002.

[21] S. Konieczny and R. Pino Pérez. Merging information under constraints: A logical framework. *Logic and Computation*, 12(5):773–808, 2002.

[22] C. Lafage and J. Lang. Logical representation of preference for group decision making. In *Proc. KR'2000*, pages 457–468. Morgan Kaufmann Publishers, 2000.

[23] C. Lafage and J. Lang. Propositional distances and preference representation. In *Proc. ECSQARU'01*, LNAI 2143, pages 48–59. Springer, 2001.

[24] D. Lehmann, M. Magidor, and K. Schlechta. Distance semantics for belief revision. *Journal of Symbolic Logic*, 66(1):295–317, 2001.

[25] J. Lin and A. O. Mendelzon. Knowledge base merging by majority. In *Dynamic Worlds: From the Frame Problem to Knowledge Management*. Kluwer, 1999.

[26] A. Lopatenko and L. Bertossi. Complexity of consistent query answering in databases under cardinality-based and incremental repair semantics. In *Proc. ICDT'07*, LNCS 4353, pages 179–193. Springer, 2007.

[27] G. Pigozzi. Two aggregation paradoxes in social decision making: the ostrogorski paradox and the discursive dilemma. *Episteme: A Journal of Social Epistemology*, 2(2):33–42, 2005.

[28] G. Priest. Paraconsistent logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 6, pages 287–393. Kluwer, 2002.

[29] W. Spohn. Ordinal conditional functions: a dynamic theory of epistemic states. In W. L. Harper and B. Skyrms, editors, *Belief Change and Statistics*, volume II, pages 105–134. Kluwer, 1988.

[30] V. S. Subrahmanian. Mechanical proof procedures for many valued lattice-based logic programming. *Journal of Non-Classical Logic*, 7:7–41, 1990.