Characteristic Logics for Behavioural Metrics via Fuzzy Lax Extensions

Paul Wild

Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany
paul.wild@fau.de

Lutz Schröder

Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany
lutz.schroeder@fau.de

Abstract

Behavioural distances provide a fine-grained measure of equivalence in systems involving quantitative data, such as probabilistic, fuzzy, or metric systems. Like in the classical setting of crisp bisimulation-type equivalences, the wide variation found in system types creates a need for generic methods that apply to many system types at once. Approaches of this kind are emerging within the paradigm of universal coalgebra, based either on lifting pseudometrics along set functors or on lifting general real-valued (*fuzzy*) relations along functors by means of *fuzzy lax extensions*. An immediate benefit of the latter is that they allow bounding behavioural distance by means of fuzzy bisimulations that need not themselves be (pseudo-)metrics, in analogy to classical bisimulations (which need not be equivalence relations). The known instances of generic pseudometric liftings, specifically the generic Kantorovich and Wasserstein liftings, both can be extended to yield fuzzy lax extensions, using the fact that both are effectively given by a choice of quantitative modalities. Our central result then shows that in fact all fuzzy lax extensions are Kantorovich extensions for a suitable set of quantitative modalities, the so-called *Moss modalities*. For *non-expansive* fuzzy lax extensions, this allows for the extraction of quantitative modal logics that characterize behavioural distance, i.e. satisfy a quantitative version of the Hennessy-Milner theorem; equivalently, we obtain expressiveness of a quantitative version of Moss' coalgebraic logic.

2012 ACM Subject Classification Theory of computation → Modal and temporal logics

Keywords and phrases Modal logic, behavioural distance, coalgebra, bisimulation, lax extension

Digital Object Identifier 10.4230/LIPIcs.CONCUR.2020.27

Related Version A full version of the paper is available at `https://arxiv.org/abs/2007.01033`.

Funding Work forms part of the DFG project *Probabilistic description logics as a fragment of probabilistic first-order logic* (SCHR 1118/6-2).

1 Introduction

Branching-time equivalences on reactive systems are typically governed by notions of *bisimilarity* [43, 37]. For systems involving quantitative data, such as transition probabilities, fuzzy truth values, or labellings in metric spaces, it is often appropriate to use more fine-grained, *quantitative* measures of behavioural similarity, arriving at notions of *behavioural distance*. Distance-based approaches in particular avoid the problem that small quantitative deviations in behaviour will typically render two given systems inequivalent under two-valued notions of equivalence, losing information about their similarity.

Behavioural distances serve evident purposes in system verification, allowing as they do for a reasonable notion of a specification being satisfied up to an acceptable margin of deviation (e.g. [24]). Applications have also been proposed in differential privacy [9] and conformance testing of hybrid systems [30]. Like their two-valued counterparts, behavioural distances have been introduced for quite a range of system types, such as various forms of

31st International Conference on Concurrency Theory (CONCUR 2020).
Editors: Igor Konnov and Laura Kovács; Article No. 27; pp. 27:1–27:23
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

probabilistic labelled transition systems or labelled Markov processes [25, 53, 14, 16]; systems combining nondeterministic and probabilistic branching variously known as nondeterministic probabilistic transition systems [8], probabilistic automata [13], and Markov decision processes [22]; weighted automata [3]; fuzzy transition systems [7] and fuzzy Kripke models [21]; and various forms of metric transition systems [12, 20, 19]. This range of variation creates a need for unifying concepts and methods. The present work contributes to developing such a unified view within the framework of universal coalgebra, which is based on abstracting a wide range of system types (including all the mentioned ones) as set functors.

Specifically, we fix a generic notion of *quantitative bisimulation* via the key notion of *non-expansive (fuzzy) lax extension* of a functor. While existing coalgebraic approaches to behavioural pseudometrics rely on pseudometric liftings of functors [2], fuzzy lax extensions act on unrestricted quantitative relations. Hence, quantitative bisimulations need not themselves be pseudometrics, in analogy to classical bisimulations not needing to be equivalence relations, and thus may serve as small certificates for low behavioural distance. We show that two known systematic constructions of functor liftings from chosen sets of modalities, the generic Wasserstein and Kantorovich liftings, both extend to yield non-expansive fuzzy lax extensions (it is essentially known that the Wasserstein lifting yields a fuzzy lax extension [26]). As our main result, we then establish that *every* fuzzy lax extension of a finitary functor is a Kantorovich extension induced by a suitable set of modalities, the so-called Moss modalities.

This result may be seen as a quantitative version of previous results asserting the existence of separating sets of two-valued modalities for finitary functors [47, 32, 35], which allow for generic Hennessy-Milner-type theorems stating that states in finitely branching systems (coalgebras) are behaviourally equivalent iff they satisfy the same modal formulae [44, 47]. Indeed our main result similarly allows *extracting characteristic quantitative modal logics* from given behavioural metrics, where a logic is *characteristic* or *expressive* if the induced logical distance of states coincides with behavioural distance. This result may equivalently be phrased as expressiveness of a quantitative version of Moss' coalgebraic logic [42], which provides a coalgebraic generalization of the classical relational ∇-modality (which e.g. underlies the $a \to \Psi$ notation used in Walukiewicz's μ-calculus completeness proof [55]). We relax the standard requirement of finite branching, i.e. use of finitary functors, to an approximability condition called *finitary separability*, and hence in particular cover countable probabilistic branching.

Organization. We recall basic concepts on pseudometrics, coalgebraic bisimilarity, and coalgebraic logic in Section 2. The central notion of (nonexpansive) fuzzy lax extension is introduced in Section 3, and the arising principle of quantitative bisimulation in Section 4. The generic Kantorovich and Wasserstein liftings are discussed in Sections 5 and 6, respectively. Our central result showing that every lax extension is a Kantorovich lifting is established in Section 7. In Section 8, we show how our results amount to extracting characteristic modal logics from given non-expansive lax extensions. Proofs are sometimes omitted or only sketched; some additional proofs are in Appendix A, a full version with all proofs is available [56].

Related Work. Probabilistic quantitative characteristic modal logics go back to Desharnais et al. [16]; they relate to fragments of quantitative μ-calculi [29, 38, 40]. A further well-known class of quantitative modal logics are fuzzy modal and description logics (e.g. [41, 23, 49, 34]). Van Breugel and Worrell [53] prove a Hennessy-Milner theorem for quantitative probabilistic modal logic. Quantitative Hennessy-Milner-type theorems have since been established for fuzzy modal logic with Gödel semantics [21], for systems combining probability and non-determinism [17], and for Heyting-valued modal logics [18] as introduced by Fitting [23]. König and Mika-Michalski [31] provide a quantitative Hennessy-Milner theorem in coalgebraic generality for the case where behavioural distance is induced by the pseudometric Kantorovich lifting defined by the same set of modalities as the logic, a result that we complement by showing that in fact all fuzzy lax extensions are Kantorovich.

Fuzzy lax extensions are a quantitative version of lax extensions [35, 50, 33], which in turn belong to an extended strand of research on relation liftings [28, 50, 33]. They appear to go back to work on monoidal topology [27], and have been used in work on applicative bisimulation [24]; as indicated above, Hofmann [26] effectively already introduces the generic Wasserstein lax extension (without using the term but proving the relevant properties, except non-expansiveness). Our notion of *non-expansive* lax extension, which is central to the connection with characteristic logics, appears to be new. Our method of extracting quantitative modalities from fuzzy lax extensions generalizes the construction of two-valued Moss liftings for (two-valued) lax extensions [32, 35].

2 Preliminaries

We recall basic notions on pseudometrics, universal coalgebra [46], and the generic treatment of two-valued bisimilarity. Basic knowledge of category theory (e.g. [1]) will be helpful.

Pseudometric Spaces. A *(1-bounded) pseudometric* on a set X is a function $d\colon X\times X \to [0,1]$ satisfying $d(x,x)=0$ (reflexivity), $d(x,y)=d(y,x)$ (symmetry), and $d(x,z)\le d(x,y)+d(y,z)$ (triangle inequality) for $x,y,z\in X$. If moreover $d(x,y)=0$ implies $x=y$, then d is a *metric*. The pair (X,d) is a *(pseudo-)metric space*. The unit interval $[0,1]$ is a metric space under Euclidean distance $d_E(x,y)=|x-y|$. The *supremum distance* of functions $f,g\colon X\to[0,1]$ is $\|f-g\|_\infty=\sup_{x\in X}|f(x)-g(x)|$. A map $f\colon X\to Y$ of pseudometric spaces (X,d_1), (Y,d_2), is *nonexpansive* (notation: $f\colon (X,d_1)\to_1 (Y,d_2)$) if $d_2(f(x),f(y))\le d_1(x,y)$ for all $x,y\in X$.

Universal Coalgebra is a uniform framework for a broad range of state-based system types. It is based on encapsulating the transition type of a system as an (endo-)functor, for the present purposes on the category of sets: A *functor* T assigns to each set X a set TX, and to each map $f\colon X\to Y$ a map $Tf\colon TX\to TY$, preserving identities and composition. We may think of TX as a parametrized datatype; e.g. the *(covariant) powerset functor* $\mathcal{P}$ assigns to each set X its powerset $\mathcal{P}X$, and to $f\colon X\to Y$ the direct image map $\mathcal{P}f\colon\mathcal{P}X\to\mathcal{P}Y, A\mapsto f[A]$; and the *distribution functor* $\mathcal{D}$ maps each set X to the set of discrete probability distributions on X. Recall that a discrete probability distribution on X is given by a *probability mass function* $\mu\colon X\to[0,1]$ such that $\sum_{x\in X}\mu(x)=1$ (implying that the *support* $\{x\in X\mid \mu(x)>0\}$ of μ is at most countable); we abuse μ to denote also the induced probability measure, writing $\mu(A)=\sum_{x\in A}\mu(x)$ for $A\subseteq X$. Moreover, $\mathcal{D}$ maps $f\colon X\to Y$ to $\mathcal{D}f\colon\mathcal{D}X\to\mathcal{D}Y$, $\mu\mapsto\mu f^{-1}$ where the *image measure* μf^{-1} is given by $\mu f^{-1}(B)=\mu(f^{-1}[B])$ for $B\subseteq Y$.

Systems of a transition type T are then cast as T*-coalgebras* (A,α), consisting of a set A of *states* and a *transition function* $\alpha\colon A\to TA$, thought of as assigning to each state a structured collection of successors. E.g. a $\mathcal{P}$-coalgebra $\alpha\colon A\to\mathcal{P}A$ assigns to each state a a set $\alpha(a)$ of successors, so is just a (non-deterministic) transition system. Similarly, a $\mathcal{D}$-coalgebra assigns to each state a distribution over successor states, and thus is a probabilistic transition system or a Markov chain. A *morphism* $f\colon(A,\alpha)\to(B,\beta)$ of T-coalgebras (A,α) and (B,β) is a map $f\colon A\to B$ such that $\beta\circ f=Tf\circ\alpha$, where $\circ$ denotes the usual (applicative) composition of functions; e.g. morphisms of $\mathcal{P}$-coalgebras are functional bisimulations.

A functor T is *finitary* if for each set X and each $t \in TX$, there exists a finite subset $Y \subseteq X$ such that $t = Ti(t')$ for some $t' \in TY$, where $i\colon Y \to X$ is the inclusion map. Intuitively, T is finitary if every element of TX mentions only finitely many elements of X. Every set functor T has a *finitary part* T_ω given by $T_\omega X = \bigcup\{Ti[TY] \mid Y \subseteq X \text{ finite}, i\colon Y \to X \text{ inclusion}\}$. E.g. $\mathcal{P}_\omega$, the *finite powerset functor*, maps a set to the set of its finite subsets, and $\mathcal{D}_\omega$, the *finite distribution functor*, maps a set X to the set of discrete probability distributions on X with finite support. Coalgebras for finitary functors generalize finitely branching systems, and hence feature in Hennessy-Milner type theorems, which typically fail under infinite branching.

Bisimilarity and Lax Extensions. Coalgebras come with a canonical notion of observable equivalence: States $a \in A$, $b \in B$ in T-coalgebras (A,α), (B,β) are *behaviourally equivalent* if there exist a coalgebra (C,γ) and morphisms $f\colon(A,\alpha) \to (C,\gamma)$, $g\colon(B,\beta) \to (C,\gamma)$ such that $f(a) = g(b)$. Behavioural equivalence can often be characterized in terms of bisimulation relations, which may provide small witnesses for behavioural equivalence of states and in particular need not form equivalence relations. The most general known way of treating bisimulation coalgebraically is via *lax extensions* L of the functor T, which map relations $R \subseteq X \times Y$ to $LR \subseteq TX \times TY$ subject to a number of axioms (monotonicity, preservation of relational converse, lax preservation of composition, extension of function graphs) [35]; L *preserves diagonals* if $L\Delta_X = \Delta_{TX}$ for each set X, where for any set Y, Δ_Y denotes the *diagonal* $\{(y,y) \mid y \in Y\}$. The *Barr extension* $\overline{T}$ of T [4, 51] is defined by

$$\overline{T}R = \{(T\pi_1(r), T\pi_2(r)) \mid r \in TR\}$$

for $R \subseteq X \times Y$, where $\pi_1\colon R \to X$ and $\pi_2\colon R \to Y$ are the projections; $\overline{T}$ preserves diagonals, and is a lax extension if T preserves weak pullbacks. E.g., the Barr extension $\overline{\mathcal{P}}$ of the powerset functor $\mathcal{P}$ is the well-known Egli-Milner extension, given by

$$(V,W) \in \overline{\mathcal{P}}(R) \iff (\forall x \in V.\, \exists y \in W.\, (x,y) \in R) \wedge (\forall y \in W.\, \exists x \in V.\, (x,y) \in R)$$

for $R \subseteq X \times Y$, $V \in \mathcal{P}(X)$, $W \in \mathcal{P}(Y)$. An L-*bisimulation* between T-coalgebras (A,α), (B,β) is a relation $R \subseteq A \times B$ such that $(\alpha(a),\beta(b)) \in LR$ for all $(a,b) \in R$; e.g. for $L = \overline{\mathcal{P}}$, we obtain exactly Park/Milner bisimulation on transition systems. If L preserves diagonals, then two states are behaviourally equivalent iff they are related by some L-bisimulation [35].

Coalgebraic Logic serves as a generic framework for the specification of state-based systems [11]. It is based on interpreting custom *modalities* of given finite arity over coalgebras for a functor T as n-ary *predicate liftings*, which are families of maps

$$\lambda_X\colon (2^X)^n \to 2^{TX}$$

(subject to a naturality condition) where $2 = \{\bot,\top\}$ and for any set Y, 2^Y is the set of 2-valued predicates on Y. We do not distinguish notationally between modalities and the associated predicate liftings. Satisfaction of a formula of the form $\lambda(\phi_1,\dots,\phi_n)$ (in some ambient logic) in a state $a \in A$ of a T-coalgebra (A,α) is then defined inductively by

$$a \models \lambda(\phi_1,\dots,\phi_n) \text{ iff } \alpha(a) \in \lambda_A([\![\phi_1]\!],\dots,[\![\phi_n]\!]) \tag{1}$$

where for any formula ψ, $[\![\psi]\!] = \{c \in A \mid c \models \psi\}$. E.g. the standard diamond modality $\diamond$ is interpreted over the powerset functor $\mathcal{P}$ by the predicate lifting $\diamond_X(Y) = \{Z \in \mathcal{P}(X) \mid \exists x \in Z.\, Y(x) = \top\}$, which according to (1) induces precisely the usual semantics of $\diamond$ over transition systems ($\mathcal{P}$-coalgebras). The standard Hennessy-Milner theorem is generalized coalgebraically [44, 47] as saying that two states in T-coalgebras are behaviourally equivalent iff they satisfy the same Λ-formulae, provided that T is finitary (which corresponds to the usual assumption of finite branching) and Λ is *separating*, i.e. for any set X, every $t \in TX$ is uniquely determined (within TX) by the set

$$\{(\lambda, Y_1, \dots, Y_n) \mid \lambda \in \Lambda \ n\text{-ary}, Y_1, \dots, Y_n \in 2^X, t \in \lambda(Y_1, \dots, Y_n)\}.$$

For finitary T, a separating set of modalities always exists [47].

3 Fuzzy Relations and Lax Extensions

We next introduce the central notion of the paper, concerning extensions of *fuzzy* (or *real-valued*) relations along a *set functor T, which we fix for the remainder of the paper.* We begin by fixing basic concepts and notation on fuzzy relations. Pseudometrics can be viewed as particular fuzzy relations, forming a quantitative analogue of equivalence relations.

▶ **Definition 3.1.** Let A and B be sets. A *fuzzy relation* between A and B is a map $R: A \times B \to [0,1]$, also written $R: A \nrightarrow B$. We say that R is *crisp* if $R(a,b) \in \{0,1\}$ for all $a \in A, b \in B$ (and generally apply the term crisp to concepts that live in the standard two-valued setting). The *converse* relation $R^\circ: B \nrightarrow A$ is given by $R^\circ(b,a) = R(a,b)$. For $R, S: A \nrightarrow B$, we write $R \le S$ if $R(a,b) \le S(a,b)$ for all $a \in A, b \in B$.

▶ **Convention 3.2.** Crisp relations are just ordinary relations. However, since we are working in a pseudometric setting, it will be more natural to use the convention that elements $a \in A, b \in B$ are related by a crisp relation R if $R(a,b) = 0$, in which case we write aRb.

▶ **Convention 3.3** (Composition)**.** We write composition of fuzzy relations diagrammatically, using ';'. Explicitly, the composite $R_1; R_2 : A \nrightarrow C$ of $R_1: A \nrightarrow B$ and $R_2: B \nrightarrow C$ is defined by

$$(R_1; R_2)(a,c) = \inf_{b \in B} R_1(a,b) \oplus R_2(b,c),$$

where $\oplus$ denotes Łukasiewicz disjunction: $x \oplus y = \min(x+y, 1)$. We reserve the applicative composition operator $\circ$ for composition of functions. In particular, $R: A \nrightarrow B$ is viewed as a function $A \times B \to [0,1]$ whenever $\circ$ is applied to R.

▶ **Definition 3.4** (Functions as relations)**.** The *ϵ-graph* of a function $f: A \to B$ is the fuzzy relation $\mathrm{Gr}_{\epsilon,f}: A \nrightarrow B$ given by $\mathrm{Gr}_{\epsilon,f}(a,b) = \epsilon$ if $f(a) = b$, and $\mathrm{Gr}_{\epsilon,f}(a,b) = 1$ otherwise. The ϵ-graph of the identity function id_A is also called the *ϵ-diagonal* of A, and denoted by $\Delta_{\epsilon,A}$. We refer to $\mathrm{Gr}_{0,f}$ simply as the *graph* of f, also denoted Gr_f, and to $\Delta_{0,A}$ as the *diagonal* of A, which we continue to denote as Δ_A.

Using the notation assembled, we can rephrase the definition of pseudometric as follows.

▶ **Lemma 3.5.** *A fuzzy relation $d: X \nrightarrow X$ is a pseudometric iff*

$$\begin{aligned} d &\le \Delta_X && \textit{(reflexivity)} \\ d^\circ &= d && \textit{(symmetry)} \\ d &\le d; d && \textit{(triangle inequality).} \end{aligned}$$

We now introduce our central notion of non-expansive lax extension:

▶ **Definition 3.6** (Fuzzy lax extensions)**.** A *(fuzzy) relation lifting* L of T maps each fuzzy relation $R\colon A \nrightarrow B$ to a fuzzy relation $LR\colon TA \nrightarrow TB$ such that

(L0) $L(R^\circ) = (LR)^\circ$

for all R. We say that L is a *fuzzy lax extension* if it additionally satisfies

(L1) $R_1 \le R_2 \Rightarrow LR_1 \le LR_2$
(L2) $L(R;S) \le LR; LS$
(L3) $L\mathrm{Gr}_f \le \mathrm{Gr}_{Tf}$

for all sets A, B, and $R, R_1, R_2\colon A \nrightarrow B$, $S\colon B \nrightarrow C$, $f\colon A \to B$. A fuzzy lax extension L is *non-expansive*, and then briefly called a *non-expansive lax extension*, if

(L4) $L\Delta_{\epsilon,A} \le \Delta_{\epsilon,TA}$

for all sets A and $\epsilon > 0$.

Axioms (L0)–(L3) are straightforward quantitative generalizations of the axiomatization of two-valued lax extensions [35]; fuzzy lax extensions in this sense have also been called $[0,1]$-*relators* [24, 27] (in the more general setting of quantale-valued relations). Axiom (L4) has no two-valued analogue; its role and the terminology are explained by the following characterization:

▶ **Lemma 3.7.** *Let L be a fuzzy lax extension of T. Then the following are equivalent.*
1. *L satisfies Axiom (L4) (i.e. is non-expansive).*
2. *For all functions $f\colon A \to B$ and all $\epsilon > 0$, $L Gr_{\epsilon,f} \le Gr_{\epsilon,Tf}$.*
3. *For all sets A, B, the map $R \mapsto LR$ is non-expansive w.r.t. the supremum metric on $A \nrightarrow B$.*

This characterization is an important prerequisite for the Hennessy-Milner theorem. Its proof relies on the following basic property [27, Corollary III.1.4.4]:

▶ **Lemma 3.8** (Naturality)**.** *Let L be a fuzzy lax extension of T, let $R\colon A' \nrightarrow B'$ be a fuzzy relation, and let $f\colon A \to A', g\colon B \to B'$. Then $L(R \circ (f \times g)) = LR \circ (Tf \times Tg)$.*

As indicated previously, existing approaches to behavioural metrics (e.g. [53, 2]) are based on lifting functors to pseudometric spaces. Every lax extension induces such a functor lifting:

▶ **Lemma 3.9.** *Let L be a fuzzy lax extension, and let $d\colon X \nrightarrow X$ be a pseudometric. Then Ld is a pseudometric on TX. Moreover, for every non-expansive map $f\colon (X, d_1) \to (Y, d_2)$ of pseudometric spaces, the map $Tf\colon (TX, Ld_1) \to (TY, Ld_2)$ is non-expansive.*

That is, every fuzzy lax extension of $T\colon \mathsf{Set} \to \mathsf{Set}$ gives rise to a functor $\overline{T}\colon \mathsf{PMet} \to \mathsf{PMet}$ on the category PMet of pseudometric spaces and nonexpansive maps that *lifts* T in the sense that $U \circ \overline{T} = T \circ U$, where $U\colon \mathsf{PMet} \to \mathsf{Set}$ is the functor that forgets the pseudometric.

Much of the development will be based on finitary functors; for instance, we need a finitary functor so we can give an explicit syntax for the characterizing logic of a lax extension. The following notion captures a broader class of functors than just the finitary ones.

▶ **Definition 3.10.** A fuzzy lax extension L for the functor T is *finitarily separable* if for every set X, $T_\omega X$ is a dense subset of TX wrt. the pseudometric $L\Delta_X$.

Clearly, any lax extension of a finitary functor is finitarily separable. The prototypical example of a finitarily separable lax extension of a non-finitary functor is the Kantorovich lifting of the discrete distribution functor $\mathcal{D}$ (Example 5.8.1). We conclude the section with a basic example of a non-expansive lax extension, deferring further examples to the sections on systematic constructions of such extensions (Sections 5 and 6):

▶ **Example 3.11** (Hausdorff lifting). The *Hausdorff lifting* is the relation lifting H for the powerset functor $\mathcal{P}$, defined for fuzzy relations $R\colon A \nrightarrow B$ by

$$HR(U,V) = \max(\sup_{a\in U}\inf_{b\in V} R(a,b),\ \sup_{b\in V}\inf_{a\in U} R(a,b)).$$

for $U \subseteq A, V \subseteq B$. The Hausdorff lifting can be viewed as a quantitative analogue of the Egli-Milner extension (Section 2), where sup replaces universal quantification and inf replaces existential quantification. It is shown already in [27] that H is a fuzzy lax extension. Indeed, it is easy to see that H is also non-expansive. These properties will also follow from the results of Section 6, where we show that H is in fact an instance of the Wasserstein lifting. H is not finitarily separable, because for every set X we have $H\Delta_X = \Delta_{\mathcal{P}X}$.

4 Quantitative Bisimulations

We next identify a notion of bisimulation based on a lax extension L of the functor T; similar concepts appear in work on quantitative applicative bisimilarity [24]. We define behavioural distance based on this notion, and show coincidence with the distance defined via the pseudometric lifting induced by L according to Lemma 3.9.

▶ **Definition 4.1.** Let L be a lax extension of T, and let $\alpha\colon A \to TA$ and $\beta\colon B \to TB$ be coalgebras.

1. A fuzzy relation $R\colon A \nrightarrow B$ is an L-*bisimulation* if $LR \circ (\alpha \times \beta) \leq R$.
2. We define L-*behavioural distance* $d^L_{\alpha,\beta}\colon A \nrightarrow B$ to be the infimum of all L-bisimulations:

   $$d^L_{\alpha,\beta} = \inf\{R\colon A \nrightarrow B \mid R \text{ is an } L\text{-bisimulation}\}.$$

   If $\alpha = \beta$, we write $d^L_\alpha = d^L_{\alpha,\beta}$ instead.

▶ Remark 4.2. Putting Definition 4.1 in other words, an L-bisimulation is precisely a prefix point for the map $F(R) = LR \circ (\alpha \times \beta)$. Note that F is monotone by (L1). This means that, according to the Knaster-Tarski fixpoint theorem, $d^L_{\alpha,\beta}$ is itself a prefix point (i.e. an L-bisimulation), and also the least fixpoint of F, i.e. $d^L_{\alpha,\beta} = Ld^L_{\alpha,\beta} \circ (\alpha \times \beta)$.

As L-behavioural distance is the least L-bisimulation, we get:

▶ **Lemma 4.3.** *For every coalgebra $\alpha\colon A \to TA$, d^L_α is a pseudometric.*

▶ Remark 4.4. As announced above, existing generic notions of behavioural distance defined via functor liftings [2] agree with the one given above (when both apply). Specifically, when applied to the functor lifting induced by a lax extension L of T according to Lemma 3.9, the definition of behavioural distance via functor liftings amounts to taking the same least fixpoint as in Definition 4.1 but only over pseudometrics instead of over fuzzy relations.

▶ Remark 4.5. Every fuzzy lax extension L induces a crisp lax extension L_c, where for any crisp relation R, $L_cR = (LR)^{-1}[\{0\}] \subseteq TA \times TB$ (recall Convention 3.2). It is easily checked that L_c preserves diagonals (Section 2) iff

$$L\Delta_A \text{ is a metric for each set } A. \tag{2}$$

By results on lax extensions cited in Section 2, L_c-bisimilarity coincides with behavioural equivalence in this case, i.e. if L satisfies (2), then L characterizes behavioural equivalence: Two states $a \in A$, $b \in B$ in coalgebras (A,α), (B,β) are behaviourally equivalent iff $d^L_{\alpha,\beta}(a,b) = 0$.

▶ **Example 4.6** (Small bisimulations). We give an example for the functor $TX = [0,1] \times \mathcal{P}X$. Coalgebras for T are Kripke frames where each state is labelled with a number from the unit interval. This T has a non-expansive lax extension L, defined for fuzzy relations $R\colon A \twoheadrightarrow B$ by

$$LR((p,U),(q,V)) = \tfrac{1}{2}(|p-q| + HR(U,V)),$$

where $p,q \in [0,1], U \subseteq A, V \subseteq B$, and H is the Hausdorff lifting (Example 3.11). The motivating idea behind this definition is that the L-behavioural distance of two states is the supremum of the accumulated branching-time differences between state labels over all runs of a process starting at these states. The factor $\frac{1}{2}$ ensures that the total distance is at most 1 by discounting the differences at later stages with exponentially decreasing factors.

Now consider the T-coalgebras (A,α) and (B,β) below:

We put $R(a_1,b_1) = 0.2, R(a_2,b_3) = 0.1, R(a_3,b_2) = 0.05$ and $R(a_i,b_j) = 1$ in all other cases. Then R is an L-bisimulation witnessing that $d^L_{\alpha,\beta}(a_1,b_1) \le 0.2$, but is neither reflexive nor symmetric, nor transitive on the disjoint union of the systems.

As indicated previously, quantitative Hennessy-Milner theorems can only be expected to hold for non-expansive lax extensions. The key observation is the following. By standard fixpoint theory, L-behavioural distance can be approximated from below by an ordinal-indexed increasing chain. Crucially, if L is non-expansive and finitarily separable, then this chain stabilizes after ω steps. Formally:

▶ **Theorem 4.7.** *Let L be a non-expansive finitarily separable lax extension of T. Given T-coalgebras (A,α), (B,β), define a sequence $(d_n\colon A \twoheadrightarrow B)_{n<\omega}$ and $d_\omega\colon A \twoheadrightarrow B$ by*

$$d_0 = 0, \qquad d_{n+1} = Ld_n \circ (\alpha \times \beta), \qquad d_\omega = \sup\nolimits_{n<\omega} d_n.$$

Then
- **(i)** $Ld_\omega \circ (\alpha \times \beta) = d_\omega$, *and*
- **(ii)** *L-behavioural distance $d^L_{\alpha,\beta}$ equals d_ω.*

Proof (sketch). In case T is finitary, we can exploit the fact that under restriction to a finite subset of $A \times B$ the pointwise convergence of $(d_n)_{n<\omega}$ becomes uniform, so (i) follows from nonexpansivity of L using Lemma 3.7.3. To generalize to the non-finitary case, one can use an *unravelling* construction and approximate the unravelled T-coalgebra by a T_ω-coalgebra such that the series of accumulated errors converges to a fixed ϵ. Claim (ii) is immediate from (i) by the fixpoint definition of $d^L_{\alpha,\beta}$. ◀

5 The Kantorovich Lifting

As a pseudometric lifting, the Kantorovich lifting is standard in the probabilistic setting: Given a metric d on a set X, the Kantorovich distance $Kd(\mu_1,\mu_2)$ between discrete distributions μ_1,μ_2 on X is defined by

$$Kd(\mu_1,\mu_2) = \sup\{\mathbb{E}_{\mu_1}(f) - \mathbb{E}_{\mu_2}(f) \mid f\colon (X,d) \to ([0,1],d_E) \text{ nonexpansive}\}$$

where $\mathbb{E}$ takes expected values. The coalgebraic generalization of the Kantorovich lifting, both in the pseudometric setting [31] and in the present setting of fuzzy relations, is based on fuzzy predicate liftings, a quantitative analogue of two-valued predicate liftings (Section 2) that goes back to work on coalgebraic fuzzy description logics [48]. Fuzzy predicate liftings will feature in the generic quantitative modal logics that we extract from fuzzy lax extensions (Section 8).

Recall that the *contravariant fuzzy powerset functor* $\mathcal{Q}\colon \mathsf{Set} \to \mathsf{Set}$ is defined on sets X as $\mathcal{Q}X = (X \to [0,1])$ and on functions $f\colon X \to Y$ as $\mathcal{Q}f(h) = h \circ f$.

▶ **Definition 5.1** (Fuzzy predicate liftings). Let $n \in \mathbb{N}$.

1. An *n-ary (fuzzy) predicate lifting* is a natural transformation

   $$\lambda\colon \mathcal{Q}^n \Rightarrow \mathcal{Q} \circ T,$$

   where the exponent n denotes n-fold cartesian product.
2. The *dual* of λ is the n-ary predicate lifting $\bar{\lambda}$ given by $\bar{\lambda}(f_1,\dots,f_n) = 1 - \lambda(1-f_1,\dots,1-f_n)$.
3. We call λ *monotone* if for all sets X and all functions $f_1,\dots,f_n,g_1,\dots,g_n \in \mathcal{Q}X$ such that $f_i \le g_i$ for all i, $\lambda_X(f_1,\dots,f_n) \le \lambda_X(g_1,\dots,g_n)$.
4. We call λ *nonexpansive* if for all sets X and all functions $f_1,\dots,f_n,g_1,\dots,g_n \in \mathcal{Q}X$,

   $$\|\lambda_X(f_1,\dots,f_n) - \lambda_X(g_1,\dots,g_n)\|_\infty \le \max(\|f_1 - g_1\|_\infty,\dots,\|f_n - g_n\|_\infty).$$

▶ Remark 5.2. By the Yoneda lemma, unary predicate liftings are equivalent to the *evaluation functions* $e\colon T[0,1] \to [0,1]$ used in work on pseudometric functor liftings [2, 47] and on the generic Wasserstein lifting [26]; more generally, an n-ary predicate lifting is equivalent to a generalized form of evaluation function, of type $T[0,1]^n \to [0,1]$ [47].

Before we can prove that the Kantorovich lifting is a lax extension, we first need to generalize it so that it lifts arbitrary fuzzy relations instead of just pseudometrics. To this end, we introduce the notion of nonexpansive pairs (a similar idea appears already in [54, Section 5]):

▶ **Definition 5.3.** Let $R\colon A \nrightarrow B$. A pair (f,g) of functions $f\colon A \to [0,1]$ and $g\colon B \to [0,1]$ is *R-nonexpansive* if $f(a) - g(b) \le R(a,b)$ for all $a \in A, b \in B$.

Given a function and a fuzzy relation, we can construct a *nonexpansive companion*:

▶ **Definition 5.4.** Let $R\colon A \nrightarrow B$ and $f\colon A \to [0,1]$. Then we define $R[f]\colon B \to [0,1]$ by

$$R[f](b) = \sup\nolimits_{a \in A} f(a) \ominus R(a,b),$$

where for $x,y \in [0,1]$, $x \ominus y = \max(x-y,0)$.

▶ **Definition 5.5.** Let Λ be a set of monotone predicate liftings that is closed under duals. The *Kantorovich lifting* K_Λ is defined as follows: for $R\colon A \nrightarrow B$, $K_\Lambda R\colon TA \nrightarrow TB$ is given by

$$K_\Lambda R(t_1,t_2) = \sup\{\lambda_A(f_1,\dots,f_n)(t_1) - \lambda_B(g_1,\dots,g_n)(t_2) \mid \lambda \in \Lambda \text{ } n\text{-ary}, (f_1,g_1),\dots(f_n,g_n) \text{ } R\text{-nonexpansive}\}.$$

We show in the appendix that closure under duals guarantees that $K_\Lambda R(t_1,t_2) \ge 0$ always.

▶ **Theorem 5.6.** *Let Λ be a set of monotone predicate liftings that is closed under duals. The Kantorovich lifting K_Λ is a lax extension. If all $\lambda \in \Lambda$ are nonexpansive, then K_Λ is nonexpansive as well.*

Proof (sketch). We sketch the proofs for (L2) and (L4). For (L2), one observes that given a nonexpansive pair (f, h) for $R; S$, one can obtain nonexpansive pairs (f, g) for R and (g, h) for S using the nonexpansive companion $g = R[f]$. For (L4), we note that $\Delta_{\epsilon,A}$-nonexpansivity of (f, g) implies that $f(a) - g(a) \leq \epsilon$ for all $a \in A$. By monotonicity of predicate liftings, we can assume that w.l.o.g. $g(a) = f(a) \ominus \epsilon$. In this case, for every $\lambda \in \Lambda$ (for simplicity, unary) and $t \in TA$,

$$\lambda(f)(t) - \lambda(g)(t) \leq \|\lambda(f) - \lambda(g)\|_\infty \leq \|f - g\|_\infty \leq \epsilon.$$ ◀

▶ Remark 5.7 (Kantorovich for pseudometrics). On pseudometrics, the Kantorovich lifting K_Λ as given by Definition 5.5 agrees with the usual Kantorovich distance $-^{\uparrow T}$ [2, Definition 5.4] defined for pseudometrics. If $d: A \nrightarrow A$ is a pseudometric, then $d^{\uparrow T}(t_1, t_2)$ equals

$$\sup\{|\lambda_A(f_1, \dots, f_n)(t_1) - \lambda_A(f_1, \dots, f_n)(t_2)| \mid \lambda \in \Lambda, f_1, \dots, f_n: (A, d) \to_1 ([0,1], d_E)\}$$

▶ **Example 5.8** (Kantorovich liftings).

1. The standard Kantorovich lifting K of the discrete distribution functor $\mathcal{D}$ is an instance of the generic one, for the single predicate lifting $\Diamond(f)(\mu) = \mathbb{E}_\mu(f)$. Crucially, K is finitarily separable, by the observation that for every discrete distribution $\mu \in \mathcal{D}X$ and every $\epsilon > 0$, there are only finitely many points x with $\mu(x) > \epsilon$.
2. The *fuzzy neighbourhood functor* is the (covariant) functor $\mathcal{N} = \mathcal{Q} \circ \mathcal{Q}$; the elements of $\mathcal{N}X$ are called *fuzzy neighbourhood systems*, and their coalgebras *fuzzy neighbourhood frames* [45, 10]. The *monotone (nonexpansive) fuzzy neighbourhood functor* $\mathcal{M}$ is the subfunctor $\mathcal{M}$ of $\mathcal{N}$ given by $\mathcal{M}X$ consisting of the fuzzy neighbourhood systems that are monotone and nonexpansive as maps $A: \mathcal{Q}X \to [0,1]$. We put

$$LR(A, B) = \max(\sup_{f \in \mathcal{Q}X} A(f) \ominus B(R[f]), \sup_{g \in \mathcal{Q}X} B(g) \ominus A(R^\circ[g]))$$

for $R: A \nrightarrow B$, $A \in \mathcal{M}X$, $B \in \mathcal{M}Y$ (recall Definition 5.4). Then L is a nonexpansive lax extension of $\mathcal{M}$; specifically, $L = K_{\{\lambda\}}$ where λ is the predicate lifting given by $\lambda_X(f)(A) = A(f)$.

6 The Wasserstein Lifting

The other generic construction for lax extensions arises in a similar way, by generalizing the generic Wasserstein lifting for pseudometrics [2] to lift arbitrary fuzzy relations instead of just pseudometrics; our construction slightly generalizes one given by Hofmann [26]. Compared to the case of the Kantorovich lifting, where we needed to work with nonexpansive pairs, the generalization from pseudometric lifting to relation lifting is much more direct. In the same way as for the original construction of pseudometric Wasserstein liftings, additional constraints, both on the functor and the set of predicate liftings involved, are needed for the Wasserstein lifting to be a lax extension. Indeed, the Wasserstein lifting may be seen as a quantitative analogue of the two-valued Barr extension (Section 2), and like the latter works only for functors that preserve weak pullbacks. In particular, Wasserstein liftings are based on the central notion of coupling:

▶ **Definition 6.1.** Let $t_1 \in TA, t_2 \in TB$ for sets A, B. The set of *couplings* of t_1 and t_2 is $\mathsf{Cpl}(t_1, t_2) = \{t \in T(A \times B) \mid T\pi_1(t) = t_1, T\pi_2(t) = t_2\}$.

Like the Kantorovich lifting, the Wasserstein lifting is based on a choice of predicate liftings. It is, however, built in a quite different manner, and in particular appears to make sense only for unary predicate liftings, so unlike in some other places in the paper, the restriction to unary liftings in the next definition is not just for readability.

▶ **Definition 6.2** (Wasserstein lifting)**.** Let Λ be a set of unary predicate liftings. The *generic Wasserstein lifting* is the relation lifting W_Λ of T defined for $R\colon A \nrightarrow B$ by

$$W_\Lambda R(t_1, t_2) = \sup_{\lambda \in \Lambda} \inf\{\lambda_{A\times B}(R)(t) \mid t \in \mathsf{Cpl}(t_1, t_2)\}.$$

This construction is similar to [26, Definition 3.4] except that we admit more than one modality. On pseudometrics, the Wasserstein lifting coincides with the pseudometric lifting $-^{\downarrow T}$ as defined in [2, Definition 5.12]. We will see that the following conditions ensure that the Wasserstein lifting is a fuzzy lax extension:

▶ **Definition 6.3.** Let λ be a unary predicate lifting.

1. λ is *subadditive* if for all sets X and all $f, g \in \mathcal{Q}X$, $\lambda_X(f \oplus g) \le \lambda_X(f) \oplus \lambda_X(g)$.
2. λ *preserves the zero function* if for all sets X, $\lambda_X(0_X) = 0_{TX}$, where $0_X\colon x \mapsto 0$.
3. λ is *standard* if it is monotone, subadditive, and preserves the zero function.

▶ Remark 6.4. Baldan et al. give conditions under which the Wasserstein lifting arising from some set of evaluation functions (Remark 5.2) preserves pseudometrics. For this purpose they consider the notion of a *well-behaved evaluation function* [2, Definition 5.14]. We show in Appendix A that this amounts to a slightly stronger condition than standardness of the corresponding predicate lifting. Similar conditions also feature in Hofmann's *topological theories* [26, Definition 3.1], which consist of a monad acting on a quantale via an evaluation function and on which his generic Wasserstein extension is based. We show in the appendix that, ignoring some monad-specific axioms, the conditions imposed on the functor and evaluation function are equivalent to standardness of the associated predicate lifting.
Now indeed we have

▶ **Theorem 6.5.** *If T preserves weak pullbacks and Λ is a set of standard predicate liftings, then the Wasserstein lifting W_Λ is a lax extension. If additionally all $\lambda \in \Lambda$ are nonexpansive, then W_Λ is nonexpansive as well.*

Proof (sketch). The proofs of (L0)–(L3) are similar to [26, Theorem 3.5]. In particular, (L3) follows by preservation of the zero function, and(L2) is based on subadditivity of predicate liftings and weak pullback preservation of T. The latter is a prerequisite for the so-called gluing lemma (e.g. [2, Lemma 5.18]), which gives a canonical way of producing couplings $t_{13} \in \mathsf{Cpl}(t_1, t_3)$ from couplings $t_{12} \in \mathsf{Cpl}(t_1, t_2)$ and $t_{23} \in \mathsf{Cpl}(t_2, t_3)$. The proof of (L4) is by nonexpansivity of predicate liftings. ◀

▶ **Example 6.6** (Wasserstein liftings)**.**

1. The Hausdorff lifting H (Example 3.11) is the Wasserstein lifting $W_{\{\lambda\}}$ for $\mathcal{P}$, where $\lambda_X(f)(A) = \sup f[A]$ for $A \subseteq X$.
2. The convex powerset functor $\mathcal{C}$, whose coalgebras combine probabilistic branching and nondeterminism [6], maps a set X to the set of nonempty convex subsets of $\mathcal{D}X$. The Wasserstein lifting $W_{\{\lambda\}}$, where $\lambda_X(f)(A) = \sup_{\mu \in A} \mathbb{E}_\mu(f)$ for $A \in \mathcal{C}X$, is a non-expansive lax extension of $\mathcal{C}$. Of course, λ is just the composite of the predicate liftings respectively defining the standard Kantorovich and Hausdorff liftings. As we show in the appendix, $W_{\{\lambda\}}$ indeed coincides with the composite of these liftings (for which a quantitative equational axiomatization has recently been given by Mio and Vignudelli [39]).

7 Lax Extensions as Kantorovich Liftings

We proceed to establish the central result that every fuzzy lax extension is a Kantorovich lifting for some suitable set Λ of predicate liftings, and moreover we characterize the Kantorovich liftings induced by non-expansive predicate liftings as precisely the non-expansive lax extensions. For a given fuzzy lax extension L, the equality $K_\Lambda R = LR$ splits into two inequalities, one of which is characterized straightforwardly:

▶ **Definition 7.1.** An n-ary predicate lifting λ *preserves nonexpansivity* if for all fuzzy relations R and all R-nonexpansive pairs $(f_1, g_1), \dots, (f_n, g_n)$, the pair $(\lambda_A(f_1, \dots, f_n), \lambda_B(g_1, \dots, g_n))$ is LR-nonexpansive. A set Λ of predicate liftings *preserves nonexpansivity* if all $\lambda \in \Lambda$ preserve nonexpansivity.

▶ **Lemma 7.2.** *We have $K_\Lambda R \le LR$ if and only if Λ preserves nonexpansivity.*

▶ **Definition 7.3** (Separation)**.** A set Λ of predicate liftings is *separating* for L if $K_\Lambda R \ge LR$ for all fuzzy relations R.

To motivate Definition 7.3, recall from Section 2 that in the two-valued setting a set Λ of predicate liftings (for simplicity, assumed to be unary) is separating if

$$t_1 \neq t_2 \implies \exists \lambda \in \Lambda, A' \subseteq A \text{ such that } t_1 \in \lambda_A(A') \not\Leftrightarrow t_2 \in \lambda_A(A')$$

for $t_1, t_2 \in TA$. Analogously, unfolding definitions in the inequality $K_\Lambda R \ge LR$ (and again assuming unary liftings), we arrive at the condition that for all $t_1 \in TA, t_2 \in TB$, $\epsilon > 0$,

$$LR(t_1, t_2) > \epsilon \implies \exists \lambda \in \Lambda, (f, g) \ R\text{-nonexpansive such that } \lambda_A(f)(t_1) - \lambda_B(g)(t_2) > \epsilon.$$

We are now ready to state our main result, which says that all lax extensions are Kantorovich:

▶ **Theorem 7.4.** *If L is a finitarily separable lax extension of T, then there exists a set Λ of monotone predicate liftings that preserves nonexpansivity and is separating for L, i.e. $L = K_\Lambda$. Moreover, L is nonexpansive iff Λ can be chosen in such a way that all $\lambda \in \Lambda$ are nonexpansive.*

This result can be seen as a fuzzy version of the statements that every finitary functor has a separating set of two-valued modalities (and hence an expressive two-valued coalgebraic modal logic) [47, Corollary 45], and that more specifically, every finitary functor equipped with a diagonal-preserving lax extension has a separating set of two-valued *monotone* predicate liftings [35, Theorem 14]. We will detail in Section 8 how Theorem 7.4 implies the existence of characteristic modal logics. The proof of Theorem 7.4 uses a quantitative version of the so-called Moss modalities [32, 35]. The construction of these modalities relies on the fact that T_ω can be presented by algebraic operations of finite arity:

▶ **Definition 7.5.** A *finitary presentation* of T_ω consists of a *signature* Σ of operations with given finite arities, and for each $\sigma \in \Sigma$ of arity n a natural transformation $\sigma\colon (-)^n \Rightarrow T_\omega$ such that every element of $T_\omega X$ has the form $\sigma_X(x_1, \dots, x_n)$ for some $\sigma \in \Sigma$.

For the remainder of this section, we fix a finitary presentation of T_ω (such a presentation always exists) and assume a finitarily separable fuzzy lax extension L of T.

▶ **Definition 7.6.** Let $\sigma \in \Sigma$ be n-ary. The *Moss lifting* $\mu^\sigma\colon \mathcal{Q}^n \Rightarrow \mathcal{Q} \circ T$ is defined by

$$\mu^\sigma_X(f_1, \dots, f_n)(t) = L\mathsf{ev}_X(\sigma_{\mathcal{Q}X}(f_1, \dots, f_n), t),$$

where $\mathsf{ev}_X\colon \mathcal{Q}X \nrightarrow X$ is given by $\mathsf{ev}_X(f, x) = f(x)$.

We take Λ to be the set of all Moss liftings and their duals:

$$\Lambda = \{\mu^\sigma \mid \sigma \in \Sigma\} \cup \{\overline{\mu^\sigma} \mid \sigma \in \Sigma\}$$

Now Theorem 7.4 is immediate from

▶ **Lemma 7.7.** *Λ is a set of monotone predicate liftings that preserves nonexpansivity and is separating for L. If L is nonexpansive, then so are all predicate liftings in Λ.*

Proof (sketch).

- *Λ is separating*: Let $s\colon B \to \mathcal{Q}A, s(b)(a) = R(a,b)$ and $\epsilon > 0$. Because the set of Σ-terms over $\mathcal{Q}A$ generates $T_\omega \mathcal{Q}A$ and L is finitarily separable, there exists some $\sigma \in \Sigma$ (for simplicity unary) and some $f \in \mathcal{Q}A$ such that $L\Delta_{\mathcal{Q}A}(\sigma(f), Ts(t_2)) \le \epsilon$. If we put $g = R[f]$, then $|\mu^\sigma(f)(t_1) - LR(t_1,t_2)| \le \epsilon$ and $\mu^\sigma(g)(t_2) \le \epsilon$. Then let $\epsilon \to 0$.
- *Nonexpansivity of Moss liftings*: This is based on the observation that for any set A and any $f, g \in \mathcal{Q}A$, $\|f - g\|_\infty \le \epsilon$ iff both (f,g) and (g,f) are $\Delta_{\epsilon,A}$-nonexpansive pairs.

For the remaining properties we refer to the full version of this paper [56]. ◀

8 Real-valued Coalgebraic Modal Logic

We next recall the generic framework of *real-valued coalgebraic modal logic*, which lifts two-valued coalgebraic modal logic (Section 2) to the quantitative setting, and will yield characteristic quantitative modal logics for all non-expansive lax extensions. The framework goes back to work on fuzzy description logics [48]. The present version, characterized by a specific choice of propositional operators, appears in work on the coalgebraic quantitative Hennessy-Milner theorem [31], and generalizes quantitative probabilistic modal logic [53].

Given a set Λ of (fuzzy) predicate liftings, the set $\mathcal{L}_\Lambda$ of modal (Λ)-formulae is given by

$$\phi, \psi ::= c \mid \phi \ominus c \mid \neg\phi \mid \phi \wedge \psi \mid \lambda(\phi_1, \dots, \phi_n) \tag{3}$$

where $c \in \mathbb{Q} \cap [0,1]$ and $\lambda \in \Lambda$ has arity n. The semantics assigns to each formula ϕ and each coalgebra (A, α) a real-valued map $[\![\phi]\!]_{A,\alpha}\colon A \to [0,1]$, or just $[\![\phi]\!]$, defined by

$$\begin{aligned} [\![c]\!](a) &= c & [\![\phi \wedge \psi]\!](a) &= \min([\![\phi]\!](a), [\![\psi]\!](a)) \\ [\![\phi \ominus c]\!](a) &= \max([\![\phi]\!](a) - c, 0) & [\![\lambda(\phi_1,\dots,\phi_n)]\!](a) &= \lambda_A([\![\phi_1]\!], \dots, [\![\phi_n]\!])(\alpha(a)) \\ [\![\neg\phi]\!](a) &= 1 - [\![\phi]\!](a) \end{aligned}$$

▶ Remark 8.1. We thus adopt what is often called *Zadeh semantics* for the propositional operators. This choice is pervasive in characteristic logics for behavioural distances (including [53, 31, 57]) – in particular, the more general *Łukasiewiecz semantics* fails to be nonexpansive w.r.t. behavioural distance, and indeed induces a discrete logical distance [57].

In the two-valued setting, one can sometimes restrict the propositional base in characteristic logics; notably, two-valued probabilistic modal logic characterizes (event) bisimilarity of probabilistic transition systems even with conjunction as the only propositional connective [15]. No similar results appear to be known in the quantitative case; e.g. van Breugel and Worrell's characteristic logic for behavioural distance of probabilistic transition systems [53] does feature essentially the same propositional operators as our grammar (3).

▶ **Example 8.2.**

1. *Fuzzy modal logic* may be seen as a basic fuzzy description logic [34]. Eliding propositional atoms for brevity (they may be added as nullary modalities), we take $\Lambda = \{\Diamond\}$. Models are fuzzy relational structures, i.e. coalgebras for the *covariant* fuzzy powerset functor $\mathcal{F}$ given

by $\mathcal{F}X = [0,1]^X$ and $\mathcal{F}f(g)(y) = \sup_{f(x)=y} g(x)$, and $\Diamond$ is interpreted as the predicate lifting $\Diamond_A(f)(g) = \sup_{a\in A}\min(g(a), f(a))$. Hennessy-Milner-type results necessarily apply only to finitely branching models, i.e. coalgebras for $\mathcal{F}_\omega$.

2. *Probabilistic modal logic:* Take models to be probabilistic transition systems with possible deadlocks, i.e. coalgebras for the functor $1 + \mathcal{D}$, where $\mathcal{D}A$ is the set of discrete probability distributions on A (Section 2); and $\Lambda = \{\Diamond\}$, with

$$\Diamond_A(f)(*) = 0 \quad \text{for } * \in 1 \text{, and} \qquad \Diamond_A\,(f)(\mu) = \mathbb{E}_\mu(f) = \textstyle\sum_{a\in A}\mu(a)\cdot f(a).$$

When extended with propositional atoms, this induces (up to restricting to discrete probabilities) van Breugel et al.'s contraction-free quantitative probabilistic modal logic [52].

In the two-valued setting, modal logic is typically invariant under bisimulation, i.e. bisimilar states satisfy the same modal formulae. In the quantitative setting, this corresponds to non-expansiveness of formula evaluation, which may be phrased as saying that logical distance is below behavioural distance:

▶ **Definition 8.3.** The Λ-*logical distance* between states $a \in A$, $b \in B$ in T-coalgebras (A, α), (B, β) is $d^\Lambda(a, b) = \sup\{|[\![\phi]\!](a) - [\![\phi]\!](b)| \mid \phi \in \mathcal{L}_\Lambda\}$.

▶ **Lemma 8.4** (Non-expansiveness of quantitative modal logic)**.** *If Λ preserves non-expansiveness w.r.t. a lax extension L, then $d^\Lambda \le d^L$.*

Finally, we show how the characterization of lax extensions as Kantorovich extensions can be used to define characteristic logics for nonexpansive lax extensions. We use a Hennessy-Milner result by König and Mika-Michalski [31], for the (pseudometric) Kantorovich lifting:

▶ **Theorem 8.5.** *Let Λ be a set of predicate liftings such that iterative approximation of the fixpoint d^{K_Λ} as in Theorem 4.7 stabilizes in ω steps. Then $d^\Lambda \ge d^{K_\Lambda}$.*

We combine this result with our Theorems 4.7 and 5.6 to obtain, complementing Lemma 8.4, a criterion phrased directly in terms of conditions on the lax extension and the modalities:

▶ **Corollary 8.6** (Coalgebraic quantitative Hennessy-Milner theorem)**.** *Let L be a finitarily separable fuzzy lax extension, and let Λ be a separating set of monotone non-expansive predicate liftings for L. Then $d^\Lambda \ge d^L$.*

▶ **Example 8.7.** Since we only require L to be finitarily separable (rather than T finitary), Example 5.8.1 implies that we recover expressiveness [52, 53] of quantitative probabilistic modal logic over countably branching discrete probabilistic transition systems (Example 8.2.2) as an instance of Corollary 8.6.

▶ Remark 8.8. In [31], Theorem 8.5 is in fact only shown for the case of distances d^L_α defined on a single coalgebra. The general case of distances $d^L_{\alpha,\beta}$ between two possibly distinct coalgebras can be recovered by working on their coproduct (disjoint union), using that both L-behavioural distance and formula evaluation are preserved under morphisms.

Applying Lemma 8.4 and Corollary 8.6 to $L = K_\Lambda$ and using our result that all lax extensions are Kantorovich extensions for their Moss liftings (Theorem 7.4), which moreover are monotone and nonexpansive in case L is nonexpansive (Lemma 7.7), we obtain expressive logics for finitarily separable nonexpansive lax extensions:

▶ **Corollary 8.9.** *If L is a finitarily separable nonexpansive lax extension of a functor T, then $d^L = d^\Lambda$ for the set Λ of Moss liftings.*

We can see the coalgebraic modal logic of Moss liftings as concrete syntax for a more abstract logic where we incorporate functor elements into the syntax directly, as in Moss' coalgebraic logic [42]. The set $\mathcal{L}_L$ of formulae in the arising *quantitative Moss logic* is generated by the same propositional operators as above, and additionally by a modality Δ that applies to $\Phi \in T\mathcal{L}_0$ for finite $\mathcal{L}_0 \subseteq \mathcal{L}_L$, with semantics

$$[\![\Delta\Phi]\!](a) = L\mathsf{ev}_A(\Phi, \alpha(a)).$$

The dual of Δ is denoted ∇, and behaves like a quantitative analogue of Moss' two-valued ∇. From Corollary 8.9, it is immediate that this logic is expressive:

▶ **Corollary 8.10** (Expressiveness of quantitative Moss logic)**.** *Let L be a finitarily separable nonexpansive lax extension of a functor T. Then L-behavioural distance d^L coincides with logical distance in quantitative Moss logic, i.e. for all states $a \in A$, $b \in B$ in $(A, \alpha), (B, \beta)$ of coalgebras, and all $a \in A, b \in B$, $d^L_{\alpha,\beta}(a, b) = \sup\{|[\![\phi]\!](a) - [\![\phi]\!](b)| \mid \phi \in \mathcal{L}_L\}$.*

▶ **Example 8.11.**

1. We equip the finite fuzzy powerset functor $\mathcal{F}_\omega$ with the Wasserstein lifting $W_\Diamond$ for $\Diamond$ as in Example 8.2.1, in analogy to the Hausdorff lifting (Example 6.6.1). Then ∇ applies to finite fuzzy sets Φ of formulae, and

   $$[\![\nabla\Phi]\!](a) = \sup_{t \in \mathsf{Cpl}(\Phi, \alpha(a))} \inf_{(\phi, a') \in \mathcal{L}_L \times A} \max(1 - t(\phi, a'), \phi(a'))$$

   for a state a in an $\mathcal{F}$-coalgebra (A, α), i.e. in a finitely branching fuzzy relational structure.
2. Let C_{fg} be the subfunctor of the convex powerset functor $\mathcal{C}$ given by the finitely generated convex sets of (not necessarily finite) discrete distributions, equipped with the Wasserstein lifting described in Example 6.6.2. Then ∇ applies to finite sets of finite distributions on formulae, understood as spanning a convex polytope. By Corollary 8.10, the arising instance of quantitative Moss logic is expressive for all C_{fg}-coalgebras.

9 Conclusions

We have developed a systematic theory of behavioural distances based on fuzzy lax extensions, identifying the key notion of *non-expansive lax extension*, which we believe has good claims to being the right notion of quantitative relation lifting in this context. We give two general constructions of non-expansive lax extensions, respectively generalizing the classical Kantorovich and Wasserstein distances and strengthening previous generalizations where only pseudometrics are lifted [2]. Our construction of the Kantorovich lifting is based in particular on the key notion of non-expansive pair (implicit in recent work on optimal transportation [54]). Our main result shows that every non-expansive lax extension is a Kantorovich lifting for a suitable choice of modalities, the so-called Moss modalities. Moreover, one can extract from a given non-expansive lax extension a characteristic modal logic satisfying a strong form of quantitative Hennessy-Milner property. Future work will concern the extension of the systematic study of behavioural distances beyond branching-time distances as exemplified in previous work on the concrete case of metric transition systems [20], possibly using a quantitative variant of graded monads [36]; as well as a further generalization to quantale-valued metrics.

References

1. Jiří Adámek, Horst Herrlich, and George Strecker. *Abstract and Concrete Categories*. Wiley Interscience, 1990. Available as *Reprints Theory Appl. Cat.* 17 (2006), pp. 1-507.
2. Paolo Baldan, Filippo Bonchi, Henning Kerstan, and Barbara König. Coalgebraic Behavioral Metrics. *Logical Methods in Computer Science*, Volume 14, Issue 3, September 2018. doi:10.23638/LMCS-14(3:20)2018.
3. Borja Balle, Pascale Gourdeau, and Prakash Panangaden. Bisimulation metrics for weighted automata. In *International Colloquium on Automata, Languages, and Programming, ICALP 2017*, volume 80 of *LIPIcs*, pages 103:1–103:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.ICALP.2017.103.
4. Michael Barr. Relational algebras. In *Proc. Midwest Category Seminar*, volume 137 of *LNM*. Springer, 1970.
5. Michael Barr. Terminal coalgebras in well-founded set theory. *Theoret. Comput. Sci.*, 114:299–315, 1993.
6. Filippo Bonchi, Alexandra Silva, and Ana Sokolova. The power of convex algebras. In *Concurrency Theory, CONCUR 2017*, volume 85 of *LIPIcs*, pages 23:1–23:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.CONCUR.2017.23.
7. Yongzhi Cao, Sherry Sun, Huaiqing Wang, and Guoqing Chen. A behavioral distance for fuzzy-transition systems. *IEEE Trans. Fuzzy Systems*, 21(4):735–747, 2013. doi:10.1109/TFUZZ.2012.2230177.
8. Valentina Castiglioni, Daniel Gebler, and Simone Tini. Logical characterization of bisimulation metrics. In *Quantitative Aspects of Programming Languages and Systems, QAPL 2016*, volume 227 of *EPTCS*, pages 44–62, 2016. doi:10.4204/EPTCS.227.
9. Konstantinos Chatzikokolakis, Daniel Gebler, Catuscia Palamidessi, and Lili Xu. Generalized bisimulation metrics. In *Concurrency Theory, CONCUR 2014*, volume 8704 of *LNCS*, pages 32–46. Springer, 2014. doi:10.1007/978-3-662-44584-6.
10. Petr Cintula, Carles Noguera, and Jonas Rogger. From Kripke to neighborhood semantics for modal fuzzy logics. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU 2016*, volume 611 of *CCIS*, pages 95–107. Springer, 2016. doi:10.1007/978-3-319-40581-0_9.
11. Corina Cîrstea, Alexander Kurz, Dirk Pattinson, Lutz Schröder, and Yde Venema. Modal logics are coalgebraic. *Comput. J.*, 54:31–41, 2011. doi:10.1093/comjnl/bxp004.
12. Luca de Alfaro, Marco Faella, and Mariëlle Stoelinga. Linear and branching system metrics. *IEEE Trans. Software Eng.*, 35(2):258–273, 2009. doi:10.1109/TSE.2008.106.
13. Yuxin Deng, Tom Chothia, Catuscia Palamidessi, and Jun Pang. Metrics for action-labelled quantitative transition systems. In *Quantitative Aspects of Programming Languages, QAPL 2005*, volume 153 of *ENTCS*, pages 79–96. Elsevier, 2006. doi:10.1016/j.entcs.2005.10.033.
14. Josée Desharnais. *Labelled Markov processes*. PhD thesis, McGill University, Montreal, November 1999.
15. Josee Desharnais, Abbas Edalat, and Prakash Panangaden. A logical characterization of bisimulation for labeled Markov processes. In *Logic in Computer Science, LICS 1998*, pages 478–487. IEEE Computer Society, 1998.
16. Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled Markov processes. *Theor. Comput. Sci.*, 318:323–354, 2004.
17. Wenjie Du, Yuxin Deng, and Daniel Gebler. Behavioural pseudometrics for nondeterministic probabilistic systems. In *Dependable Software Engineering: Theories, Tools, and Applications, SETTA 2016*, volume 9984 of *LNCS*, pages 67–84. Springer, 2016. doi:10.1007/978-3-319-47677-3.
18. Pantelis Eleftheriou, Costas Koutras, and Christos Nomikos. Notions of bisimulation for Heyting-valued modal languages. *J. Log. Comput.*, 22(2):213–235, 2012.
19. Uli Fahrenberg and Axel Legay. The quantitative linear-time-branching-time spectrum. *Theor. Comput. Sci.*, 538:54–69, 2014. doi:10.1016/j.tcs.2013.07.030.

20 Uli Fahrenberg, Axel Legay, and Claus Thrane. The quantitative linear-time–branching-time spectrum. In *Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2011*, volume 13 of *LIPIcs*, pages 103–114. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2011. doi:10.4230/LIPIcs.FSTTCS.2011.103.

21 T. Fan. Fuzzy bisimulation for Gödel modal logic. *IEEE Trans. Fuzzy Sys.*, 23:2387–2396, December 2015. doi:10.1109/TFUZZ.2015.2426724.

22 Norm Ferns, Prakash Panangaden, and Doina Precup. Metrics for finite Markov decision processes. In *Uncertainty in Artificial Intelligence, UAI 2004*, pages 162–169. AUAI Press, 2004.

23 Melvin Fitting. Many-valued modal logics. *Fundam. Inform.*, 15:235–254, 1991.

24 Francesco Gavazzo. Quantitative behavioural reasoning for higher-order effectful programs: Applicative distances. In *Logic in Computer Science, LICS 2018*, pages 452–461. ACM, 2018. doi:10.1145/3209108.

25 Alessandro Giacalone, Chi-Chang Jou, and Scott Smolka. Algebraic reasoning for probabilistic concurrent systems. In *Programming concepts and methods, PCM 1990*, pages 443–458. North-Holland, 1990.

26 Dirk Hofmann. Topological theories and closed objects. *Adv. Math.*, 215(2):789–824, 2007. doi:10.1016/j.aim.2007.04.013.

27 Dirk Hofmann, Gavin Seal, and Walter Tholen, editors. *Monoidal Topology: A Categorical Approach to Order, Metric, and Topology*. Cambridge University Press, 2014.

28 Jesse Hughes and Bart Jacobs. Simulations in coalgebra. *Theor. Comput. Sci.*, 327(1-2):71–108, 2004. doi:10.1016/j.tcs.2004.07.022.

29 Michael Huth and Marta Kwiatkowska. Quantitative analysis and model checking. In *Logic in Computer Science, LICS 1997*, pages 111–122. IEEE, 1997.

30 Narges Khakpour and Mohammad Mousavi. Notions of conformance testing for cyber-physical systems: Overview and roadmap (invited paper). In *Concurrency Theory, CONCUR 2015*, volume 42 of *LIPIcs*, pages 18–40. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015. doi:10.4230/LIPIcs.CONCUR.2015.18.

31 Barbara König and Christina Mika-Michalski. (Metric) bisimulation games and real-valued modal logics for coalgebras. In Sven Schewe and Lijun Zhang, editors, *Concurrency Theory, CONCUR 2018*, volume 118 of *LIPIcs*, pages 37:1–37:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.CONCUR.2018.37.

32 Alexander Kurz and Raul Leal. Equational coalgebraic logic. In *Mathematical Foundations of Programming Semantics, MFPS 2009*, volume 249 of *ENTCS*, pages 333–356. Elsevier, 2009.

33 Paul Levy. Similarity quotients as final coalgebras. In *Foundations of Software Science and Computational Structures, FOSSACS 2011*, volume 6604 of *LNCS*, pages 27–41. Springer, 2011. doi:10.1007/978-3-642-19805-2.

34 Thomas Lukasiewicz and Umberto Straccia. Managing uncertainty and vagueness in description logics for the semantic web. *J. Web Sem.*, 6(4):291–308, 2008. doi:10.1016/j.websem.2008.04.001.

35 Johannes Marti and Yde Venema. Lax extensions of coalgebra functors and their logic. *J. Comput. Syst. Sci.*, 81(5):880–900, 2015.

36 Stefan Milius, Dirk Pattinson, and Lutz Schröder. Generic trace semantics and graded monads. In *Algebra and Coalgebra in Computer Science, CALCO 2015*, Leibniz International Proceedings in Informatics, 2015.

37 Robin Milner. *Communication and Concurrency*. Prentice Hall, 1989.

38 Matteo Mio and Alex Simpson. Łukasiewicz μ-calculus. *Fund. Inf.*, 150(3–4):317–346, 2017.

39 Matteo Mio and Valeria Vignudelli. Monads and quantitative equational theories for nondeterminism and probability, 2020. This volume; full version available as arXiv e-print arXiv:2005.07509. doi:10.4230/LIPIcs.CONCUR.2020.28.

40 Carroll Morgan and Annabelle McIver. A probabilistic temporal calculus based on expectations. In Lindsay Groves and Steve Reeves, editors, *Formal Methods Pacific, FMP 1997*. Springer, 1997.

41 Charles Morgan. Local and global operators and many-valued modal logics. *Notre Dame J. Formal Log.*, 20:401–411, 1979.

42 Lawrence Moss. Coalgebraic logic. *Ann. Pure Appl. Logic*, 96:277–317, 1999.

43 David Park. Concurrency and automata on infinite sequences. In *Theoretical Computer Science, 5th GI-Conference*, volume 104 of *LNCS*, pages 167–183. Springer, 1981. doi:10.1007/BFb0017288.

44 Dirk Pattinson. Expressive logics for coalgebras via terminal sequence induction. *Notre Dame J. Formal Log.*, 45:19–33, 2004.

45 Ricardo Rodriguez and Lluis Godo. Modal uncertainty logics with fuzzy neighborhood semantics. In *Weighted Logics for Artiticial Intelligence, WL4AI 2013 (Workshop at IJCAI 2013)*, pages 79–86, 2013.

46 Jan Rutten. Universal coalgebra: A theory of systems. *Theoret. Comput. Sci.*, 249:3–80, 2000.

47 Lutz Schröder. Expressivity of coalgebraic modal logic: The limits and beyond. *Theoret. Comput. Sci.*, 390:230–247, 2008.

48 Lutz Schröder and Dirk Pattinson. Description logics and fuzzy probability. In Toby Walsh, editor, *Int. Joint Conf. Artificial Intelligence, IJCAI 2011*, pages 1075–1081. AAAI, 2011.

49 Umberto Straccia. A fuzzy description logic. In *Artificial Intelligence, AAAI 1998*, pages 594–599. AAAI Press / MIT Press, 1998.

50 Albert Thijs. *Simulation and fixpoint semantics*. PhD thesis, University of Groningen, 1996.

51 Věra Trnková. General theory of relational automata. *Fund. Inform.*, 3:189–234, 1980.

52 Franck van Breugel, Claudio Hermida, Michael Makkai, and James Worrell. Recursively defined metric spaces without contraction. *Theor. Comput. Sci.*, 380(1-2):143–163, 2007. doi:10.1016/j.tcs.2007.02.059.

53 Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *Theor. Comput. Sci.*, 331:115–142, 2005.

54 Cédric Villani. *Optimal Transport: Old and New*. Springer, 2008.

55 Igor Walukiewicz. Completeness of Kozen's axiomatisation of the propositional μ-calculus. In *Logic in Computer Science, LICS 1995*, pages 14–24. IEEE Computer Society, 1995. doi:10.1109/LICS.1995.523240.

56 Paul Wild and Lutz Schröder. Characteristic logics for behavioural metrics via fuzzy lax extensions, 2020. arXiv e-print. arXiv:2007.01033.

57 Paul Wild, Lutz Schröder, Dirk Pattinson, and Barbara König. A van Benthem theorem for fuzzy modal logic. In *Logic in Computer Science, LICS 2018*, pages 909–918. ACM, 2018.

A Appendix

We assume w.l.o.g. that all functors preserve injective maps [5].

Proof of Lemma 3.7

- 1. $\iff$ 2.: The implication "$\Leftarrow$" is trivial; we prove "$\Rightarrow$". We have

$$\begin{aligned} L\mathrm{Gr}_{\epsilon,f} &= L(\Delta_{\epsilon,B} \circ (\mathsf{id}_B \times f)) \\ &= L\Delta_{\epsilon,B} \circ (\mathsf{id}_{TB} \times Tf) && \text{(Lemma 3.8)} \\ &\le \Delta_{\epsilon,TB} \circ (\mathsf{id}_{TB} \times Tf) \\ &= \mathrm{Gr}_{\epsilon,Tf} \end{aligned}$$

- 1. $\Longrightarrow$ 3.: Let $R_1, R_2 : A \nrightarrow B$ and $\epsilon > 0$ such that $\|R_1 - R_2\|_\infty \leq \epsilon$; we need to show that $\|LR_1 - LR_2\|_\infty \leq \epsilon$. The assumption implies $R_1 \leq R_2; \Delta_{\epsilon,B}$, hence $LR_1 \leq L(R_2; \Delta_{\epsilon,B}) \leq LR_2; L\Delta_{\epsilon,B} \leq LR_2; \Delta_{\epsilon,TB}$ using (L1), (L2), and (L4). Symmetrically, we show $LR_2 \leq LR_1; \Delta_{\epsilon,TB}$, so that $\|LR_1 - LR_2\|_\infty \leq \epsilon$.
- 3. $\Longrightarrow$ 1.: We have $\|\Delta_{\epsilon,A} - \Delta_A\|_\infty = \epsilon$, and hence by assumption $\|L\Delta_{\epsilon,A} - L\Delta_A\|_\infty \leq \epsilon$. In particular, $L\Delta_{\epsilon,A} - L\Delta_A \leq \Delta_{\epsilon,TA}$, so

  $$L\Delta_{\epsilon,A} \leq L\Delta_A; \Delta_{\epsilon,TA} \leq \Delta_{TA}; \Delta_{\epsilon,TA} = \Delta_{\epsilon,TA}$$

  using (L3).

Proof of Theorem 4.7

By the fixpoint definition of $d^L_{\alpha,\beta}$, (ii) is immediate from (i). We prove (i), i.e. that $Ld_\omega(\alpha(a), \beta(b)) = d_\omega(a, b)$ for all $a \in A$, $b \in B$. We begin by assuming that T is finitary, and generalise to the non-finitary case later.

Since T is finitary, there exist finite subsets $A_0 \subseteq A$, $B_0 \subseteq B$ and $s \in TA_0$, $t \in TB_0$ such that $\alpha(a) = Ti(s)$ and $\beta(b) = Tj(t)$, where $i : A_0 \to A$ and $j : B_0 \to B$ are the inclusion maps. We then have $Ld_\omega(\alpha(a), \beta(b)) = L(d_\omega \circ (i \times j))(s, t)$ by naturality (Lemma 3.8). Now the $d_n \circ (i \times j)$ converge to $d_\omega \circ (i \times j)$ pointwise, and therefore also under the supremum metric (i.e. uniformly), since $A_0 \times B_0$ is finite. Since the assumptions imply that L is continuous w.r.t. the supremum metric, it follows that

$$\begin{aligned} &L(d_\omega \circ (i \times j))(s, t) \\ &= \sup\nolimits_{n<\omega} L(d_n \circ (i \times j))(s, t) \\ &= \sup\nolimits_{n<\omega} Ld_n(\alpha(a), \beta(b)) && \text{(naturality)} \\ &= \sup\nolimits_{n<\omega} d_{n+1}(a, b) = d_\omega(a, b). \end{aligned}$$

For non-finitary T, we refer to the full version.

Details for Remark 5.2

An evaluation function $e : T[0,1] \to [0,1]$ gives rise to a unary predicate lifting λ_e by putting $\lambda_e(f) = e \circ Tf$. Conversely, an evaluation function for $\lambda : \mathcal{Q} \Rightarrow \mathcal{Q} \circ T$ can be defined via $e_\lambda = \lambda_{[0,1]}(\mathsf{id})$.

Generalizing to higher arities, an n-ary evaluation function is a map $e : T([0,1]^n) \to [0,1]$, and gives rise to a predicate lifting $\lambda_e(f_1, \dots, f_n) = e \circ T\langle f_1, \dots f_n \rangle$, while for each n-ary predicate lifting λ the corresponding evaluation function is $e_\lambda = \lambda_{[0,1]^n}(\pi_1, \dots, \pi_n)$.

Details for Definition 5.5

From the definition it is clear that $K_\Lambda R(t_1, t_2) \in [-1, 1]$ for all t_1 and t_2. To see that $K_\Lambda R(t_1, t_2) \geq 0$, consider the maps $h_X : X \to [0,1], x \mapsto \frac{1}{2}$ for any set X. The pair (h_A, h_B) is clearly R-nonexpansive and so, for some arbitrary unary $\lambda \in \Lambda$

$$\begin{aligned} K_\Lambda R(t_1, t_2) &\geq \max(\lambda_A(h_A)(t_1) - \lambda_B(h_B)(t_2), \bar{\lambda}_A(h_A)(t_1) - \bar{\lambda}_B(h_B)(t_2)) \\ &= |\lambda_A(h_A)(t_1) - \lambda_B(h_B)(t_2)| \geq 0. \end{aligned}$$

(If λ has higher arity, just supply more copies of h_A and h_B.)

Proof of Theorem 5.6

The following lemma will be used in the proof of (L2):

▶ **Lemma A.1.** *Let $R\colon A \nrightarrow B, S\colon B \nrightarrow C$. Then for every $(R;S)$-nonexpansive pair (f,h) there exists some function $g\colon B \to [0,1]$ such that (f,g) is R-nonexpansive and (g,h) is S-nonexpansive.*

Proof. For each $b \in B$ the value $g(b)$ can be chosen arbitrarily in the interval

$$[\sup_{a\in A} f(a) \ominus R(a,b), \inf_{c\in C} h(c) \oplus S(b,c)],$$

so for instance we can use the nonexpansive companion $g := R[f]$ (Definition 5.4). This interval is non-empty because by assumption

$$\begin{aligned} f(a) - h(c) &\le (R;S)(a,c) \\ &\le \inf_{b'\in B} R(a,b') + S(b',c) \\ &\le R(a,b) + S(b,c) \end{aligned}$$

for all $a \in A, c \in C$, so $f(a) - R(a,b) \le h(c) + S(b,c)$ by rearranging. Similar rearranging also shows that choosing $g(b)$ in this way ensures that (f,g) is R-nonexpansive and (g,h) is S-nonexpansive. ◀

Now we are ready for the main proof.

Proof. For readability, we pretend that all $\lambda \in \Lambda$ are unary although the proof works just as well for unrestricted arities, whose treatment requires no more than adding indices. We show the five properties one by one:

- (L0): Let $R\colon A \nrightarrow B$ and $t_1 \in TA, t_2 \in TB$. Note that a pair (g,f) is R°-nonexpansive iff $(1-f, 1-g)$ is R-nonexpansive. Now, using that Λ is closed under duals,

  $$\begin{aligned} K_\Lambda(R^\circ)(t_2,t_1) &= \sup\{\lambda_B(g)(t_2) - \lambda_A(f)(t_1) \mid \lambda \in \Lambda, (g,f)\ R^\circ\text{-nonexp.}\} \\ &= \sup\{\bar\lambda_A(f)(t_1) - \bar\lambda_B(g)(t_2) \mid \lambda \in \Lambda, (f,g)\ R\text{-nonexp.}\} = K_\Lambda R(t_1,t_2) \end{aligned}$$

- (L1): Let $R_1 \le R_2$. Then every R_1-nonexpansive pair is also R_2-nonexpansive. Thus $K_\Lambda R_1 \le K_\Lambda R_2$, because the supremum on the left side is taken over a subset of that on the right side.
- (L2): Let $R\colon A \nrightarrow B, S\colon B \nrightarrow C$ and $t_1 \in TA, t_2 \in TB, t_3 \in TC$. Let $\lambda \in \Lambda$ and let (f,h) be $(R;S)$-nonexpansive. Let g be given by Lemma A.1. Then it is enough to observe that:

  $$\begin{aligned} \lambda_A(f)(t_1) - \lambda_C(h)(t_3) &= (\lambda_A(f)(t_1) - \lambda_B(g)(t_2)) + (\lambda_B(g)(t_2) - \lambda_C(h)(t_3)) \\ &\le K_\Lambda R(t_1,t_2) + K_\Lambda S(t_2,t_3). \end{aligned}$$

- (L3): Let $h\colon A \to B$ and $t \in TA$. We need to show that $K_\Lambda \mathrm{Gr}_h(t, Th(t)) = 0$. Let $\lambda \in \Lambda$ and let (f,g) be Gr_h-nonexpansive, implying $f \le g \circ h$. Then

  $$\lambda_A(f)(t) \le \lambda_A(g\circ h)(t) = \lambda_B(g)(Th(t)),$$

  by monotonicity and naturality of λ.
- (L4): Let A be a set, $t \in TA$ and $\epsilon > 0$. We need to show that $K_\Lambda \Delta_{\epsilon,A}(t,t) \le \epsilon$. Let $\lambda \in \Lambda$ and let (f,g) be $\Delta_{\epsilon,A}$-nonexpansive, implying $f(a) - g(a) \le \epsilon$ for all $a \in A$. By monotonicity of λ, we can restrict our attention to the case $g(a) = f(a) \ominus \epsilon$. In this case,

  $$\lambda(f)(t) - \lambda(g)(t) \le \|\lambda(f) - \lambda(g)\|_\infty \le \|f - g\|_\infty \le \epsilon.$$ ◀

Details for Remark 5.7

First, note that if (f, g) with $f, g\colon A \to [0,1]$ is d-nonexpansive, then $f(a) - g(a) \le d(a,a) = 0$ for all $a \in A$, so $f \le g$. By monotonicity of the $\lambda \in \Lambda$ the value of the supremum in Definition 5.5 does not change if we restrict the choice of (f, g) to the case $f = g$. Finally, in the case $f = g$, d-nonexpansivity implies that $f(a) - f(b) \le d(a,b)$ and $f(b) - f(a) \le d(b,a) = d(a,b)$ for every $a, b \in A$, which means that f is in fact a nonexpansive map $f\colon (A,d) \to_1 ([0,1], d_E)$. Also the supremum does not change when taking the absolute value, because f is nonexpansive iff $1 - f$ is and Λ is closed under duals.

Details for Example 5.8.1

We show that K is finitarily separable. Let $\mu \in \mathcal{D}X$ and $\epsilon > 0$. We need to find $\mu_\epsilon \in \mathcal{D}X$ with finite support such that $K\Delta_X(\mu, \mu_\epsilon) \le \epsilon$. Note that a pair (f, g) is Δ_X-nonexpansive iff $f \le g$, so by monotonicity

$$K\Delta_X(\mu, \mu_\epsilon) = \sup\{\textstyle\sum_{x\in X} f(x)(\mu(x) - \mu_\epsilon(x)) \mid f\colon X \to [0,1]\} \le \textstyle\sum_{x \in X} |\mu(x) - \mu_\epsilon(x)|.$$

Because μ is discrete, there exists a finite set $Y \subseteq X$ with $\sum_{x\in Y} \mu(x) \ge 1 - \frac{\epsilon}{2}$. If $Y = X$, then we can just put $\mu_\epsilon = \mu$. Otherwise, let $x_0 \in X \setminus Y$. Then we define μ_ϵ as follows: $\mu_\epsilon(x_0) = \mu(Y)$, $\mu_\epsilon(x) = \mu(x)$ for $x \in Y$, and $\mu_\epsilon(x) = 0$ otherwise. In this case,

$$\textstyle\sum_{x\in X} |\mu(x) - \mu_\epsilon(x)| \le 2\mu(Y) \le \epsilon.$$

Details for Remark 6.4

We recall the definition of well-behaved evaluation functions from [2]:

▶ **Definition A.2.** An evaluation function $e\colon T[0,1] \to [0,1]$ is *well-behaved* if it satisfies the following:

1. The predicate lifting λ_e is monotone.
2. For all $t \in T([0,1]^2)$, we have $d_E(e(t_1), e(t_2)) \le \lambda_e(d_E)(t)$, where $t_1 = T\pi_1(t)$ and $t_2 = T\pi_2(t)$.
3. $e^{-1}[\{0\}] = Ti[T\{0\}]$, where $i\colon \{0\} \to [0,1]$ is the inclusion map.

This notion is almost equivalent to that of a standard predicate lifting in the following sense:

▶ **Lemma A.3.** *e is a well-behaved evaluation function iff the predicate lifting λ_e is standard and $e^{-1}[\{0\}] \subseteq Ti[T\{0\}]$.*

Proof. First, note that monotonicity of λ_e features in both notions and λ_e preserves zero iff $e^{-1}[\{0\}] \supseteq Ti[T\{0\}]$. It remains to relate Item 2 of Definition A.2 with subadditivity of λ. Reformulating in terms of λ_e gives

$$|\lambda_e(\pi_1)(t) - \lambda_e(\pi_2)(t)| \le \lambda_e(d_E)(t) \quad \text{for } t \in T([0,1]^2). \tag{4}$$

We show that (4) is equivalent to subadditivity of λ_e, given that λ_e is monotone:

- "⇒": Let $f, g \in \mathcal{Q}X, t \in TX$. Put $t' := T\langle f \oplus g, f\rangle(t) \in T([0,1]^2)$. Then, by naturality, we have $\lambda_e(\pi_1)(t') = \lambda_e(f \oplus g)(t)$ and $\lambda_e(\pi_2)(t') = \lambda_e(f)(t)$ and

  $$\lambda_e(d_E)(t') = \lambda_e(d_E \circ \langle f \oplus g, f\rangle)(t) \le \lambda_e(g)(t),$$

  where we used monotonicity of λ_e in the last step. Therefore, $\lambda(f \oplus g)(t) - \lambda(f)(t) \le \lambda(g)(t)$ by (4).

- "⇐": Put $f = d_E, g = \pi_1 : [0,1]^2 \to [0,1]$. Then it is easily checked that $f \oplus g \geq \pi_2$ and therefore

$$\lambda_e(\pi_2) \leq \lambda_e(f \oplus g) \leq \lambda_e(f) + \lambda_e(g) = \lambda_e(d_E) + \lambda_e(\pi_1)$$

by monotonicity and subadditivity of λ_e, so $\lambda_e(\pi_1) - \lambda_e(\pi_2) \leq \lambda_e(d_E)$. Similarly, we can show that $\lambda_e(\pi_2) - \lambda_e(\pi_1) \leq \lambda_e(d_E)$ by swapping the roles of π_1 and π_2. ◀

In [26, Definition 3.1], topological theories are defined as triples consisting of a monad T, a quantale V, and a map $\xi : TV \to V$ satisfying a number of axioms. We only consider the case of the quantale $[0,1]^{\mathrm{op}}$, with the order given by $\geq$ and the monoid structure by $\oplus$. The first two axioms state that ξ is a T-algebra and can be ignored for our purposes. The remaining axioms instantiate as follows, where as usual $\lambda_\xi(f) = \xi \circ Tf$ is the predicate lifting associated with ξ:

($Q_\otimes$) $\quad \oplus \circ \langle \lambda_\xi(\pi_1), \lambda_\xi(\pi_2) \rangle \geq \lambda_\xi(\oplus)$

(Q_k) $\quad 0 \geq \lambda_\xi(0_1)(t)$ for every $t \in T1$, where 1 is a singleton set

($Q'_\vee$) $\quad \lambda_\xi$ is a monotone natural transformation

Using a similar idea as in Lemma A.3, we see that ($Q_\otimes$) is equivalent to subadditivity of λ_ξ and (Q_k) is equivalent to preservation of the zero function. Finally note that [26, Theorem 3.5 (d)] (which states that the Wasserstein lifting satisfies (L2)) requires that the functor satisfies the *Beck-Chevalley condition*, i.e. preserves weak pullbacks.

Proof of Example 6.6.1

Let $R : A \nrightarrow B$, and let $U \subseteq A$ and $V \subseteq B$. We show $HR(U,V) = W_{\{\lambda\}}R(U,V)$. There are two inequalities:

- "≤": Let $Z \in \mathsf{Cpl}(U,V)$. Then for every $a \in U$ there exists $b \in V$ such that $(a,b) \in Z$, so $\inf_{b \in V} R(a,b) \leq \sup R[Z]$. Thus, we have $\sup_{a \in U} \inf_{b \in V} R(a,b) \leq \sup R[Z]$, and, by a symmetrical argument, $\sup_{b \in V} \inf_{a \in U} R(a,b) \leq \sup R[Z]$.
- "≥": It is enough to find for each $\epsilon > 0$ a coupling $Z \in \mathsf{Cpl}(U,V)$ such that $\sup R[Z] \leq HR(U,V) + \epsilon$. So let $\epsilon > 0$. We construct functions $f : U \to V$ and $g : V \to U$ as follows: For each $a \in U$ choose $f(a) \in V$ such that $R(a, f(a)) \leq \inf_{b \in V} R(a,b) + \epsilon$. Similarly, for each $b \in V$ choose $g(b) \in U$ such that $R(g(b), b) \leq \inf_{a \in U} R(a,b) + \epsilon$. Now put $Z = \{(a, f(a)) \mid a \in U\} \cup \{(g(b), b) \mid b \in V\}$. Clearly, $Z \in \mathsf{Cpl}(U,V)$ and by construction,

$$\sup R[Z] = \max(\sup_{a \in U} R(a, f(a)), \sup_{b \in V} R(g(b), b)) \leq HR(U,V) + \epsilon.$$

Details for Example 6.6.2

We denote the Wasserstein lifting of the distribution functor $\mathcal{D}$ by W. Let $R : A \nrightarrow B$, and let $U \in \mathcal{C}A$ and $V \in \mathcal{C}B$. We show $W_{\{\lambda\}}(R)(U,V) = HW(R)(U,V)$. There are two inequalities:

- "≥": Let $Z \in \mathsf{Cpl}_{\mathcal{C}}(U,V)$. We put $Y = \mathcal{P}\langle \mathcal{D}\pi_1, \mathcal{D}\pi_2 \rangle(Z)$. Then $\mathcal{P}\pi_1(Y) = \mathcal{P}\mathcal{D}\pi_1(Z) = \mathcal{C}\pi_1(Z) = U$ and similarly $\mathcal{P}\pi_2(Y) = V$, so that $Y \in \mathsf{Cpl}_{\mathcal{P}}(U,V)$. Now, note that for every $\mu \in \mathcal{D}(A \times B)$ we have that $\mathbb{E}_\mu(R) \geq WR(\mathcal{D}\pi_1(\mu), \mathcal{D}\pi_2(\mu))$ and therefore

$$\sup_{\mu \in Z} \mathbb{E}_\mu(R) \geq \sup_{(\mu_1,\mu_2) \in Y} WR(\mu_1,\mu_2) \geq HW(R)(U,V).$$

- "≤": Let $Y \in \mathsf{Cpl}_{\mathcal{P}}(U,V)$. It is enough to find for each $\epsilon > 0$ some $Z \in \mathsf{Cpl}_{\mathcal{C}}(\mu_1,\mu_2)$ such that

$$\sup_{\mu \in Z} \mathbb{E}_\mu(R) \leq \sup_{(\mu_1,\mu_2) \in Y} WR(\mu_1,\mu_2) + \epsilon.$$

For every $(\mu_1, \mu_2) \in \mathcal{D}A \times \mathcal{D}B$ there exists some $\mu \in \mathsf{Cpl}_{\mathcal{D}}(U, V)$ such that $\mathbb{E}_\mu(R) \leq WR(\mu_1, \mu_2) + \epsilon$. Let Z' be a set consisting of one such μ for every pair $(\mu_1, \mu_2) \in Y$ and put $Z = \mathrm{conv}(Z')$, where conv is convex hull. Then we have

$$\mathcal{C}\pi_1(Z) = \mathcal{P}\mathcal{D}\pi_1(\mathrm{conv}(Z')) = \mathrm{conv}(\mathcal{P}\mathcal{D}\pi_1(Z')) = \mathrm{conv}(U) = U.$$

Here we made use of the fact that $\mathcal{D}\pi_1$ is linear when considered as a map $\mathbb{R}^{A\times B} \to \mathbb{R}^A$, and linear maps preserve convex sets. We similarly get $\mathcal{C}\pi_2(Z) = V$, so that $Z \in \mathsf{Cpl}_{\mathcal{C}}(U, V)$. Finally, we note that taking expected values is a linear operation, so if $\mu = \sum_{i=1}^n p_i \mu_i$ is a convex combination of probability measures, then $\mathbb{E}_\mu = \sum_{i=1}^n p_i \mathbb{E}_{\mu_i} \leq \max_{i=1}^n \mathbb{E}_{\mu_i}$. Therefore we have, as desired,

$$\sup_{\mu \in Z} \mathbb{E}_\mu(R) = \sup_{\mu \in Z'} \mathbb{E}_\mu(R) \leq \sup_{(\mu_1,\mu_2)\in Y} WR(\mu_1, \mu_2) + \epsilon.$$

Proof of Lemma 8.4

Immediate from the following lemma:

▶ **Lemma A.4.** *Let ϕ be a modal Λ-formula, and let $a \in A$, $b \in B$ be states in T-coalgebras (A, α), (B, β). Then $[\![\phi]\!]_{A,\alpha}(a) - [\![\phi]\!]_{B,\beta}(b)| \leq d^{K_\Lambda}_{\alpha,\beta}(a, b)$.*

Proof. Induction on ϕ, with trivial Boolean cases (in Zadeh semantics, all propositional operators on $[0, 1]$ are non-expansive). For the modal case, we have (for readability, restricting to unary $\lambda \in \Lambda$ and omitting subscripts)

$$\begin{aligned}
|[\![\lambda(\phi)]\!](a) - [\![\lambda(\phi)]\!](b)| &= |\lambda_A([\![\phi]\!])(\alpha(a)) - \lambda_B([\![\phi]\!](\beta(b))| \\
&\leq K_\Lambda d^{K_\Lambda}(\alpha(a), \beta(b)) && \text{(definition, IH)} \\
&= d^{K_\Lambda}(a, b) && \text{(definitionup)}
\end{aligned}$$

◀

Proof of Corollary 8.10

Immediate from Corollary 8.9 once one notes that the closure under duals incorporated in the Definition of the Kantorovich distance is ensured by the presence of negation in quantitative Moss logic.