Logical Methods in Computer Science
Volume 18, Issue 2, 2022, pp. 19:1–19:35
https://lmcs.episciences.org/

Submitted Apr. 12, 2021
Published Jun. 15, 2022

CHARACTERISTIC LOGICS FOR BEHAVIOURAL HEMIMETRICS VIA FUZZY LAX EXTENSIONS

PAUL WILD AND LUTZ SCHRÖDER

Friedrich-Alexander-Universität Erlangen-Nürnberg
e-mail address: paul.wild@fau.de, lutz.schroeder@fau.de

Abstract. In systems involving quantitative data, such as probabilistic, fuzzy, or metric systems, behavioural distances provide a more fine-grained comparison of states than two-valued notions of behavioural equivalence or behaviour inclusion. Like in the two-valued case, 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 (bi-)simulations that need not themselves be hemi- or pseudometrics; this is analogous to classical simulations and bisimulations, which need not be preorders or equivalence relations, respectively. The known 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 *nonexpansive* 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. All our results explicitly hold also for asymmetric distances (*hemimetrics*), i.e. notions of quantitative simulation.

1. Introduction

Branching-time equivalences on reactive systems are typically governed by notions of *bisimilarity* [Par81, Mil89]. 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. We note in passing that

Key words and phrases: modal logic, behavioural distance, coalgebra, bisimulation, lax extension.

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

DOI:10.46298/LMCS-18(2:19)2022
© P. Wild and L. Schröder
Creative Commons

behavioural distances are typically *pseudo*metrics, i.e. distinct states can have distance 0 if their behaviours are exactly equivalent.

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. [Gav18]). Applications have also been proposed in differential privacy [CGPX14] and conformance testing of hybrid systems [KM15]. Like their two-valued counterparts, behavioural distances have been introduced for quite a range of system types, such as various forms of probabilistic labelled transition systems or labelled Markov processes [GJS90, vBW05, Des99, DGJP04]; systems combining nondeterministic and probabilistic branching variously known as nondeterministic probabilistic transition systems [CGT16], probabilistic automata [DCPP06], and Markov decision processes [FPP04]; weighted automata [BGP17]; fuzzy transition systems [CSWC13] and fuzzy Kripke models [Fan15]; and various forms of metric transition systems [dAFS09, FLT11, FL14], which are nondeterministic transition systems with additional quantitative information, e.g. a metric on the labels and/or the states. Besides symmetric notions of behavioural distance, there are asymmetric variants, which correspond to quantitative notions of simulation, e.g. for rational- [CHR12], real- [TFL10], and lattice-weighted transition systems [PLC15].

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 work with a generic notion of *quantitative simulation* via the key notion of *nonexpansive (fuzzy) lax extension* of a functor. Fuzzy and quantale-valued generalizations of lax extensions have been studied in the past [Gav18, HST14]; we identify a new criterion for such lax extensions to be *nonexpansive* (equivalently *strong* in the sense of Gavazzo [Gav18]) that allows us to relate lax extensions to fuzzy logics featuring nonexpansive modalities via a Hennessy-Milner thoerem. Given a fuzzy lax extension, behavioural distance is defined as the greatest quantitative simulation; in general, behavioural distance is a *hemimetric*, i.e. obeys the usual axioms of a pseudometric except symmetry, or equivalently a generalized metric space in the sense of Lawvere [Law73].

For instance, on weighted transition systems with labels in a finite metric space (M, d_M) Larsen et al. [LFT11] consider a *simulation distance* defined as the least fixed point of the equation

$$d(s,t) = \sup_{s \xrightarrow{m} s'} \inf_{t \xrightarrow{n} t'} d_M(m,n) + \lambda d(s',t'),$$

where $0 \leq \lambda < 1$ is a discount factor. We shall later see that this simulation distance arises via a nonexpansive fuzzy lax extension and thus forms an instance of this framework.

For lax extensions obeying a suitable symmetry axiom, quantitative simulations are in fact quantitative bisimulations in the sense that their relational converse is also a simulation, and the induced behavioural distance is symmetric, i.e. forms a pseudometric. Existing coalgebraic approaches to behavioural pseudometrics rely on pseudometric liftings of functors [BBKK18], and in particular lift only pseudometrics; contrastingly, fuzzy lax extensions act on unrestricted quantitative relations. Hence, quantitative (bi-)simulations need not themselves be hemi- or 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 nonexpansive fuzzy lax extensions (it is essentially known that the Wasserstein lifting yields a fuzzy lax extension [Hof07]). 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. Notably, the definition of the Moss modalities involves application of the given lax extension to the quantitative elementhood relation, and hence centrally relies on lifting quantitative relations that fail to be hemi- or pseudometrics.

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 [Sch08, KL09, MV15], 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 [Pat04, Sch08]. Indeed, for nonexpansive lax extensions our main result similarly allows *extracting characteristic quantitative modal logics* from given behavioural hemi- or pseudometrics, 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 [Mos99], 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 [Wal95]). 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. Moreover, we emphasize that we obtain characteristic logics also for asymmetric distances, i.e. notions of quantitative simulation.

Organization. We recall basic concepts on hemi- and 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 (bi-)simulation 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 nonexpansive lax extensions.

Related Work. Probabilistic quantitative characteristic modal logics go back to Desharnais et al. [DGJP04]; they relate to fragments of quantitative μ-calculi [HK97, MS17, MM97]. A further well-known class of quantitative modal logics are fuzzy modal and description logics (e.g. [Mor79, Fit91, Str98, LS08]). Van Breugel and Worrell [vBW05] 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 [Fan15], for systems combining probability and non-determinism [DDG16], and for Heyting-valued modal logics [EKN12] as introduced by Fitting [Fit91]. König and Mika-Michalski [KMM18] 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. The assumptions of König and Mika-Michalski's theorem require that behavioural distance be approximable in ω steps. We give a sufficient criterion for this property: The predicate liftings need to be nonexpansive, and the given lax extension needs to be finitarily separable (as mentioned above). Again, we remove any assumption of symmetry, obtaining an expressiveness criterion for characteristic logics of quantitative simulations; in this sense, our work relates also to (coalgebraic and specific) results on characteristic logics for two-valued notions of similarity, e.g. [vG90, Bal00, Cîr06, KKV12, FMS21].

Fuzzy lax extensions are a quantitative version of lax extensions [MV15, Thi96, Lev11, BdBH+91], which in turn belong to an extended strand of research on relation liftings [HJ04, Thi96, Lev11]. They appear to go back to work on monoidal topology [HST14], and have been used in work on applicative bisimulation [Gav18]; as indicated above, Hofmann [Hof07] effectively already introduces the generic Wasserstein lax extension (without using the term but proving the relevant properties, except nonexpansiveness). Our notion of *nonexpansive* lax extension, which is central to the connection with characteristic logics, appears to be new, but as indicated above it can be seen to relate to a condition involving the strength of the underlying functor as considered by Gavazzo [Gav18]. Our method of extracting quantitative modalities from fuzzy lax extensions generalizes the construction of two-valued Moss liftings for (two-valued) lax extensions [KL09, MV15].

This paper is an extended and revised version of a previous conference publication [WS20]. Besides containing additional discussion and full proofs, the present version generalizes the overall technical treatment including the main results to the asymmetric setting, thus covering not only quantitative notions of bisimulation but also quantitative notions of simulation.

2. Preliminaries

We recall basic notions on metrics, pseudometrics (where distinct points may have distance 0), and hemimetrics (where additionally distance is not required to be symmetric). Moreover, we give a brief introduction to universal coalgebra [Rut00] and the generic treatment of two-valued bisimilarity. Basic knowledge of category theory (e.g. [AHS90]) will be helpful.

Hemimetric Spaces. For the present purposes, we are interested only in bounded distance functions, and then normalize distances to lie in the unit interval. Thus, a *(1-bounded) hemimetric* on a set X is a function $d: X \times X \to [0,1]$ satisfying $d(x,x) = 0$ (reflexivity), and $d(x,z) \leq d(x,y) + d(y,z)$ (triangle inequality) for $x,y,z \in X$. If additionally $d(x,y) = d(y,x)$ for all $x,y \in X$ (symmetry), then d is a *pseudometric*. If moreover for all $x,y \in X$, $d(x,y) = 0$ implies $x = y$, then d is a *metric*. The pair (X,d) is a *hemimetric space*, or respectively a *(pseudo-)metric space* if d is a (hemi-/pseudo-)metric. We write $\ominus$ for truncated subtraction on the unit interval, i.e. $x \ominus y = \max(x-y, 0)$ for $x,y \in [0,1]$. Then $d_\ominus(x,y) = x \ominus y$ defines a hemimetric $d_\ominus$ on $[0,1]$; moreover, $[0,1]$ is a metric space under Euclidean distance $d_E(x,y) = |x-y|$. The *supremum distance* of functions $f,g: X \to [0,1]$ is $\|f-g\|_\infty = \sup_{x\in X} |f(x) - g(x)|$. A map $f: X \to Y$ of hemimetric spaces (X,d_1), (Y,d_2), is *nonexpansive* (notation: $f: (X,d_1) \to_1 (Y,d_2)$) if $d_2(f(x),f(y)) \leq d_1(x,y)$ 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 $f: X \to Y$ a map $Tf: TX \to TY$, 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 $f: X \to Y$ the direct image map $\mathcal{P}f: \mathcal{P}X \to \mathcal{P}Y, A \mapsto f[A]$; 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* $\mu: X \to [0,1]$ 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 $f: X \to Y$ to $\mathcal{D}f: \mathcal{D}X \to \mathcal{D}Y$,

$\mu \mapsto \mu f^{-1}$ where the *image measure* μf^{-1} is given by $\mu f^{-1}(B) = \mu(f^{-1}[B])$ for $B \subseteq Y$. We will introduce further examples later.

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, also known as *p-morphisms* or *bounded morphisms*.

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 (this is equivalent to the more categorically phrased condition that T preserves directed colimits). 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) [MV15]; 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 [Bar70, Trn80] 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 $L = \overline{\mathcal{P}}$, L-bisimulations are precisely Park/Milner bisimulations on transition systems. If a lax extension L preserves diagonals, then two states are behaviourally equivalent iff they are related by some L-bisimulation [MV15].

Coalgebraic Logic serves as a generic framework for the specification of state-based systems [CKP+11]. For our present purposes, we are primarily interested in its simplest incarnation as a modal next-step logic, dubbed *coalgebraic modal logic*, and its role as a characteristic logic for behavioural equivalence in generalization of Hennessy-Milner logic [HM85]. We briefly recall the syntax and semantics of coalgebraic modal logic, as well as basic results. The framework 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

$$a \vDash \lambda(\phi_1, \dots, \phi_n) \text{ iff } \alpha(a) \in \lambda_A([\![\phi_1]\!], \dots, [\![\phi_n]\!]) \tag{2.1}$$

where for any formula ψ, $[\![\psi]\!] = \{c \in A \mid c \vDash \psi\}$. 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 (2.1) induces precisely the usual semantics of $\Diamond$ over transition systems ($\mathcal{P}$-coalgebras). The standard Hennessy-Milner theorem is generalized coalgebraically [Pat04, Sch08] 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

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

For finitary T, a separating set of modalities always exists [Sch08].

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. Hemimetrics can be viewed as particular fuzzy relations, forming a quantitative analogue of preorders; correspondingly, pseudometrics may be seen as 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 $R \colon A \times B \to [0,1]$, also written $R \colon A \nrightarrow B$. 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 $R^\circ \colon B \nrightarrow A$ is given by $R^\circ(b,a) = R(a,b)$. For $R, S \colon A \nrightarrow B$, 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 metric 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 $R_1; R_2 \colon A \nrightarrow C$ of $R_1 \colon A \nrightarrow B$ and $R_2 \colon B \nrightarrow C$ 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: $x \oplus y = \min(x+y, 1)$. Note that given our previous convention on crisp relations, the restriction of this composition operator to crisp relations is precisely the standard relational composition. 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. Throughout the paper, we use the fact that composition is monotone, that is, for $R_1 \leq R_1'$ and $R_2 \leq R_2'$ we have $R_1; R_2 \leq R_1'; R_2'$.

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.

The following is now straightforward.

Lemma 3.5.

(1) *For every function $f: A \to B$, we have $\Delta_B \leq Gr_f^\circ; Gr_f$ and $Gr_f; Gr_f^\circ \leq \Delta_A$.*
(2) *For every $R: A' \nrightarrow B'$, $f: A \to A'$ and $g: B \to B'$, we have $R \circ (f \times g) = Gr_f; R; Gr_g^\circ$.*

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

Lemma 3.6. *Let $d: X \nrightarrow X$ be a fuzzy relation.*

(1) *d is a hemimetric iff $d \leq \Delta_X$ (reflexivity) and $d \leq d; d$ (triangle inequality).*
(2) *d is a pseudometric iff it is a hemimetric and additionally $d^\circ = d$ (symmetry).*

We now introduce our central notion of nonexpansive lax extension:

Definition 3.7 (Fuzzy relation liftings and lax extensions)**.** A *(fuzzy) relation lifting* L of T maps each fuzzy relation $R: A \nrightarrow B$ to a fuzzy relation $LR: TA \nrightarrow TB$.

(1) We say that L *preserves converse* if for all R we have

$$\text{(L0)} \quad L(R^\circ) = (LR)^\circ.$$

(2) We say that L is a *(fuzzy) lax extension* if it satisfies

$$\text{(L1)} \quad R_1 \leq R_2 \Rightarrow LR_1 \leq LR_2$$

$$\text{(L2)} \quad L(R; S) \leq LR; LS$$

$$\text{(L3)} \quad L\mathrm{Gr}_f \leq \mathrm{Gr}_{Tf} \text{ and } L(\mathrm{Gr}_f^\circ) \leq \mathrm{Gr}_{Tf}^\circ$$

for all sets A, B, and $R, R_1, R_2: A \nrightarrow B$, $S: B \nrightarrow C$, $f: A \to B$.

(3) A fuzzy lax extension L is *nonexpansive*, and then briefly called a *nonexpansive lax extension*, if

$$\text{(L4)} \quad L\Delta_{\epsilon,A} \leq \Delta_{\epsilon,TA}$$

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

Axioms (L0)–(L3) are straightforward quantitative generalizations of the axiomatization of two-valued lax extensions [MV15]; fuzzy lax extensions in this sense have also been called $[0,1]$-*relators* [Gav18, HST14] (in the more general setting of quantale-valued relations). Compared to [WS20], we do not require fuzzy lax extensions to satisfy Axiom (L0) in general; examples of this will be shown in Example 3.14. This necessitates the addition of the second clause in Axiom (L3) (which of course is implied by the first clause in presence of (L0)).

Axiom (L4) has no two-valued analogue; its role and the terminology are explained by Lemma 3.9 below.

The axioms (L1)–(L3) imply the following basic property [HST14, 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*

$$L(R \circ (f \times g)) = LR \circ (Tf \times Tg).$$

Proof. We need to show two inequalities. For '$\le$', we have

$$\begin{aligned} L(R \circ (f \times g)) &= L(\mathrm{Gr}_f; R; \mathrm{Gr}_g^\circ) && \text{(Lemma 3.5.2)} \\ &\le L\mathrm{Gr}_f; LR; L(\mathrm{Gr}_g^\circ) && \text{(L2)} \\ &\le \mathrm{Gr}_{Tf}; LR; \mathrm{Gr}_{Tg}^\circ && \text{(L3)} \\ &= LR \circ (Tf \times Tg). && \text{(Lemma 3.5.2)} \end{aligned}$$

For '$\ge$', we have

$$\begin{aligned} LR \circ (Tf \times Tg) &= \mathrm{Gr}_{Tf}; LR; \mathrm{Gr}_{Tg}^\circ && \text{(Lemma 3.5.2)} \\ &= \mathrm{Gr}_{Tf}; L(\Delta_{A'}; R; \Delta_{B'}); \mathrm{Gr}_{Tg}^\circ && (\Delta \text{ neutral for } ;) \\ &\le \mathrm{Gr}_{Tf}; L(\mathrm{Gr}_f^\circ; \mathrm{Gr}_f; R; \mathrm{Gr}_g^\circ; \mathrm{Gr}_g); \mathrm{Gr}_{Tg}^\circ && \text{(Lemma 3.5.1) and (L1)} \\ &\le \mathrm{Gr}_{Tf}; \mathrm{Gr}_{Tf}^\circ; L(\mathrm{Gr}_f; R; \mathrm{Gr}_g^\circ); \mathrm{Gr}_{Tg}; \mathrm{Gr}_{Tg}^\circ && \text{(L2) and (L3)} \\ &\le L(\mathrm{Gr}_f; R; \mathrm{Gr}_g^\circ) && \text{(Lemma 3.5.2)} \\ &\le L(R \circ (f \times g)). && \text{(Lemma 3.5.1)} \end{aligned}$$

∎

Using Lemma 3.8, we can prove the following characterization of Axiom (L4), which is an important prerequisite for the Hennessy-Milner theorem.

Lemma 3.9. *Let L be a fuzzy lax extension of T. Then the following are equivalent:*

(1) *L satisfies Axiom (L4) (i.e. is nonexpansive).*
(2) *For all functions $f\colon A \to B$ and all $\epsilon > 0$, $LGr_{\epsilon,f} \le Gr_{\epsilon,Tf}$.*
(3) *For all sets A, B, the map $R \mapsto LR$ is nonexpansive w.r.t. the supremum metric on $A \nrightarrow B$.*

Proof.

- *(1) $\iff$ (2):* The implication'$\Leftarrow$' is trivial; we prove '$\Rightarrow$'. We have

$$\begin{aligned} L\mathrm{Gr}_{\epsilon,f} &= L(\Delta_{\epsilon,B} \circ (f \times \mathsf{id}_B)) && \text{(Definition 3.4)} \\ &= L\Delta_{\epsilon,B} \circ (Tf \times \mathsf{id}_{TB}) && \text{(Lemma 3.8)} \\ &\le \Delta_{\epsilon,TB} \circ (Tf \times \mathsf{id}_{TB}) && \text{(1)} \\ &= \mathrm{Gr}_{\epsilon,Tf}. && \text{(Definition 3.4)} \end{aligned}$$

- *(1) $\Longrightarrow$ (3):* Let $R_1, R_2\colon A \nrightarrow B$ and $\epsilon > 0$ such that $\|R_1 - R_2\|_\infty \le \epsilon$; we need to show that $\|LR_1 - LR_2\|_\infty \le \epsilon$. The assumption implies $R_1 \le R_2; \Delta_{\epsilon,B}$, hence, using (L1), (L2), and (1),

$$LR_1 \le L(R_2; \Delta_{\epsilon,B}) \le LR_2; L\Delta_{\epsilon,B} \le LR_2; \Delta_{\epsilon,TB}.$$

Symmetrically, we show $LR_2 \le LR_1; \Delta_{\epsilon,TB}$, so that $\|LR_1 - LR_2\|_\infty \le \epsilon$.

- *(3) ⟹ (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, $L\Delta_{\epsilon,A} \leq L\Delta_A; \Delta_{\epsilon,TA}$, so

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

using (L3). □

Remark 3.10. As stated in the introduction, nonexpansiveness of lax extensions relates to conditions on certain lax extensions for strong monads introduced by Gavazzo [Gav18], called *L-continuous V-relators* (for V a quantale). Specifically, as T is a set functor, it has a *tensorial strength* $\mathsf{str}_{A,B}: A \times TB \to T(A \times B)$ given by $\mathsf{str}_{A,B}(a,t) = T(b \mapsto (a,b))(t)$. Instantiating to the unit interval and using our notation, the axioms of an L-continuous $[0,1]$-relator L require that strength is nonexpansive, i.e. that for all sets A, B, X and Y and all fuzzy relations $R: A \nrightarrow B$ and $S: X \nrightarrow Y$ we have

$$L(R \oplus S) \circ (\mathsf{str}_{A,X} \times \mathsf{str}_{B,Y}) \leq R \oplus LS, \tag{3.1}$$

where $\oplus$ is taken pointwise. We say that L is *strong* if it satisfies (3.1) and show that L is strong iff it is nonexpansive:

Let $a \in A$, $b \in B$ and let $p: X \to A \times X$ and $q: Y \to B \times Y$ be the maps $x \mapsto (a,x)$ and $y \mapsto (b,y)$, respectively. Then, unfolding the definition of str and applying Lemma 3.8, we have, for $t_1 \in TX$ and $t_2 \in TY$,

$$L(R \oplus S)(\mathsf{str}_{A,X}(a,t_1), \mathsf{str}_{B,Y}(b,t_2)) = L(R \oplus S)(Tp(t_1), Tq(t_2)) = L((R \oplus S) \circ (p \times q))(t_1,t_2).$$

Put $\epsilon := R(a,b)$. Then we have, for all $x \in X$ and $y \in Y$,

$$((R \oplus S) \circ (p \times q))(x,y) = R(a,b) \oplus S(x,y) = (\Delta_{\epsilon,X}; S)(x,y),$$

so that $(R \oplus S) \circ (p \times q) = \Delta_{\epsilon,X}; LS$. Similarly, for a and b fixed like this we have that $(R \oplus LS)((a,t_1),(b,t_2)) = (\Delta_{\epsilon,TX}; LS)(t_1,t_2)$. Thus, (3.1) is equivalent to the requirement that

$$L(\Delta_{\epsilon,X}; S) \leq \Delta_{\epsilon,TX}; LS \tag{3.2}$$

for all sets X, Y, all $S: X \nrightarrow Y$ and all $\epsilon \geq 0$. Finally, we show that (3.2) is equivalent to (L4):

- '⇒': We have $L(\Delta_{\epsilon,X}) = L(\Delta_{\epsilon,X}; \Delta_X) \leq \Delta_{\epsilon,TX}; L\Delta_X \leq \Delta_{\epsilon,TX}; \Delta_{TX} = \Delta_{\epsilon,TX}$, using (3.2) and (L3) in the inequalities and neutrality of Δ in the equalities.
- '⇐': By (L2) and (L4), we have $L(\Delta_{\epsilon,X}; S) \leq L\Delta_{\epsilon,X}; LS \leq \Delta_{\epsilon,TX}; LS$.

As indicated previously, many existing approaches to behavioural metrics (e.g. [vBW05, BBKK18]) are based on lifting functors to pseudometric spaces. In the present framework, every lax extension induces a functor lifting to hemimetric spaces; or to pseudometric spaces if the lax extension preserves converse:

Lemma 3.11. *Let L be a fuzzy lax extension.*

1. *Let $d: X \nrightarrow X$ be a hemimetric. Then Ld is a hemimetric on TX. If d is a pseudometric and L preserves converse, then Ld is a pseudometric as well.*
2. *For every nonexpansive map $f: (X, d_1) \to (Y, d_2)$ of hemimetric spaces, the map $Tf: (TX, Ld_1) \to (TY, Ld_2)$ is nonexpansive.*

Proof.

1. Using Lemma 3.6 and the laws of lax extensions, we have $Ld \leq L\Delta_X \leq \Delta_{TX}$ and $Ld \leq L(d;d) \leq Ld; Ld$, so Ld is a hemimetric. If L preserves converse and d is a pseudometric, then $(Ld)^\circ = L(d^\circ) = Ld$, so that Ld is a pseudometric.

(2) Let $f\colon (X, d_1) \to (Y, d_2)$ be a nonexpansive map, that is $d_2 \circ (f \times f) \leq d_1$. Then Tf is nonexpansive as well, by naturality (Lemma 3.8) and monotonicity:

$$Ld_2 \circ (Tf \times Tf) = L(d_2 \circ (f \times f)) \leq Ld_1. \qquad \square$$

As a consequence of Lemma 3.11, every fuzzy lax extension of $T\colon \mathsf{Set} \to \mathsf{Set}$ gives rise to a functor $\overline{T}\colon \mathsf{HMet} \to \mathsf{HMet}$ on the category HMet of hemimetric spaces and nonexpansive maps that *lifts* T in the sense that $U \circ \overline{T} = T \circ U$, where $U\colon \mathsf{HMet} \to \mathsf{Set}$ is the functor that forgets the hemimetric. Similarly, every converse-preserving fuzzy lax extension induces a lifting of T to the category of pseudometric spaces.

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. We can capture a broader class of functors, specifically those functors that are suitably approximated by their finitary parts in the sense that the finitary part forms a dense subset of the unrestricted functor.

Definition 3.12. Let (X, d) be a hemimetric space. A set $A \subseteq X$ is *dense* if for all $x \in X$ and all $\epsilon > 0$ there exists some $a \in A$ such that both $d(x, a) \leq \epsilon$ and $d(a, x) \leq \epsilon$.

This notion of density for hemimetrics coincides with an existing one for quantale-valued distances [FK97]. In particular, it is essential to require both inequalities in Definition 3.12, as otherwise certain pathological cases of dense subsets may occur. For instance, if we left out the second inequality from the above definition, then the singleton set $\{1\}$ would be a dense subset of the unit interval $[0, 1]$ under the hemimetric $d_\ominus$ (Section 2).

Equipped with this definition of density, we proceed to introduce the following condition which allows for the treatment of lax extensions of certain non-finitary functors.

Definition 3.13. A fuzzy lax extension L of the functor T is *finitarily separable* if for every set X, $T_\omega X$ is a dense subset of TX w.r.t. the hemimetric $L\Delta_X$.

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 5.11.1); that is, every discrete distribution can be approximated, under the usual Kantorovich metric, by finitely supported distributions.

We conclude the section with a basic example of a nonexpansive lax extension, deferring further examples to the sections on systematic constructions of such extensions (Sections 5 and 6):

Example 3.14 (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 [HST14] that H is a fuzzy lax extension. Indeed, it is easy to see that H is also converse-preserving and nonexpansive. 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 $H\Delta_X = \Delta_{\mathcal{P}X}$.

We may also consider asymmetric versions of the Hausdorff lifting by simply omitting one of the two terms in the definition, putting

$$H^{\leftarrow}R(U,V) = \sup_{a\in U}\inf_{b\in V} R(a,b) \qquad \text{and} \qquad H^{\rightarrow}R(U,V) = \sup_{b\in V}\inf_{a\in U} R(a,b)$$

for $U \subseteq A, V \subseteq B$. Both $H^{\leftarrow}$ and $H^{\rightarrow}$ are nonexpansive fuzzy lax extensions, but neither of them preserves converse.

4. Quantitative Simulations

We next identify a notion of simulation based on a lax extension L of the functor T; similar concepts appear in work on quantitative applicative bisimilarity [Gav18]. 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 3.11.

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-*simulation* if $LR \circ (\alpha \times \beta) \le R$.
(2) R is an L-*bisimulation* if both R and R° are L-simulations.
(3) We define L-*behavioural distance* $d^L_{\alpha,\beta}\colon A \nrightarrow B$ to be the infimum of all L-simulations:

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

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-simulation 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-simulation), and also the least fixpoint of F, i.e. $d^L_{\alpha,\beta} = Ld^L_{\alpha,\beta} \circ (\alpha \times \beta)$. In particular, the infimum in Definition 4.1.3 is always a minimum.

Example 4.3. The weighted transition systems discussed in the introduction can be modelled as coalgebras for the functor $\mathcal{P}_\omega(M \times -)$, and the simulation distance given there is then L-behavioural distance for the fuzzy lax extension L defined for $R\colon A \nrightarrow B$ by

$$LR(U,V) = \sup_{(m,a)\in U}\ \inf_{(n,b)\in V} d_M(m,n) + \lambda R(a,b),$$

where $U \subseteq M \times A, V \subseteq M \times B$. To ensure that all values of LR lie in the unit interval $[0,1]$, we require that $d_M(m,n) \le 1 - \lambda$ for all $m,n \in M$. If M is finite (as is the case in [LFT11]) this can always be achieved by rescaling.

We note the following facts about L-simulations:

Lemma 4.4. *Let L be a fuzzy lax extension, and let $\alpha\colon A \to TA$, $\beta\colon B \to TB$ and $\gamma\colon C \to TC$ be coalgebras. Then*

(1) *Δ_A is an L-simulation.*
(2) *For any L-simulations $R\colon A \nrightarrow B$ and $S\colon B \nrightarrow C$, $R;S$ is an L-simulation.*

Proof. For Item (1), we have

$$L\Delta_A \circ (\alpha \times \alpha) \leq \Delta_{TA} \circ (\alpha \times \alpha) = \mathrm{Gr}_\alpha^\circ; \Delta_{TA}; \mathrm{Gr}_\alpha = \mathrm{Gr}_\alpha^\circ; \mathrm{Gr}_\alpha \leq \Delta_A$$

by (L3) and both parts of Lemma 3.5. For Item (2), we compute

$$\begin{aligned}
&L(R;S) \circ (\alpha \times \gamma) && \\
&\quad \leq (LR;LS) \circ (\alpha \times \gamma) && \text{(L2)} \\
&\quad = \mathrm{Gr}_\alpha; LR; LS; \mathrm{Gr}_\gamma^\circ && \text{(Lemma 3.5.2)} \\
&\quad \leq \mathrm{Gr}_\alpha; LR; \mathrm{Gr}_\beta^\circ; \mathrm{Gr}_\beta; LS; \mathrm{Gr}_\gamma^\circ && \text{(Lemma 3.5.1)} \\
&\quad = LR \circ (\alpha \times \beta); LS \circ (\beta \times \gamma) && \text{(Lemma 3.5.2)} \\
&\quad \leq R; S. && \text{(assumption)}
\end{aligned}$$

□

For converse-preserving lax extensions, this notion of simulation is actually one of bisimulation, more precisely:

Lemma 4.5. *If L preserves converse, then every L-simulation is an L-bisimulation.*

Proof. Let $\alpha\colon A \to TA$ and $\beta\colon B \to TB$ be coalgebras and let R be an L-simulation. Then by (L0) we have

$$L(R^\circ) \circ (\beta \times \alpha) = (LR)^\circ \circ (\beta \times \alpha) = (LR \circ (\alpha \times \beta))^\circ \leq R^\circ.$$

□

As L-behavioural distance is the least L-simulation, we have

Lemma 4.6. *For every coalgebra $\alpha\colon A \to TA$, d_α^L is a hemimetric. If L preserves converse, then d_α^L is a pseudometric.*

Proof. Since d_α^L is an L-simulation, both Δ_A and $d_\alpha^L; d_\alpha^L$ are L-simulations by Lemma 4.4. As d_α^L is the *least* L-simulation, this implies $d_\alpha^L \leq \Delta_A$ and $d_\alpha^L \leq d_\alpha^L; d_\alpha^L$, so that d_α^L is a hemimetric by Lemma 3.6.

In the converse-preserving case, we additionally have that $(d_\alpha^L)^\circ$ is an L-simulation by Lemma 4.5, and therefore $d_\alpha^L \leq (d_\alpha^L)^\circ$, making d_α^L a pseudometric by Lemma 3.6. □

Remark 4.7. As announced above and as we show next, existing generic notions of behavioural distance defined via functor liftings [BBKK18] agree with the one given above (when both apply). Specifically, when applied to the functor lifting induced by a converse-preserving lax extension L of T according to Lemma 3.11, 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 [BBKK18, Lemma 6.1]. Now let (A, α) be a coalgebra and denote the behavioural distance on A according to the definition in [BBKK18] by $\bar{d}_\alpha$. We claim that $\bar{d}_\alpha = d_\alpha^L$, with d_α^L defined according to Definition 4.1. Indeed, '≥' is trivial since $\bar{d}_\alpha$ is, by definition, an L-bisimulation, and '≤' is immediate from d_α^L being a pseudometric (Lemma 4.6).

Remark 4.8. Every converse-preserving fuzzy lax extension L induces a crisp lax extension L_c, where for any crisp relation R, $L_c R = (LR)^{-1}[\{0\}] \subseteq TA \times TB$ (recall Convention 3.2). It is easily checked that L_c preserves diagonals (Section 2) iff

$$L\Delta_A \text{ is a metric for each set } A. \tag{4.1}$$

By results on lax extensions cited in Section 2, L_c-bisimilarity coincides with behavioural equivalence in this case, i.e. if L satisfies (4.1), then L characterizes behavioural equivalence:

Two states $a \in A$ and $b \in B$ in coalgebras (A, α) and (B, β) are behaviourally equivalent iff $d^L_{\alpha,\beta}(a, b) = 0$.

Example 4.9 (Small bisimulations)**.** We give an example of a bisimulation for a lax extension of 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. They are similar to the weighted transition systems from [LFT11], except that here the labels are on the states rather than on the transitions. This T has a converse-preserving nonexpansive lax extension L, defined for fuzzy relations $R\colon A \nrightarrow B$ 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 3.14). The 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:

$$\begin{tikzcd}
 & a_1\ 0.7 \arrow[dl] \arrow[dr] & \\
a_2\ 0.2 & & a_3\ 0.8
\end{tikzcd}
\qquad
\begin{tikzcd}
 & b_1\ 0.4 \arrow[dl] \arrow[dr] & \\
b_2\ 0.7 & & b_3\ 0
\end{tikzcd}$$

We put $R(a_1, b_1) = 0.2, R(a_2, b_3) = R(a_3, b_2) = 0.1$, and $R(a_i, b_j) = 1$ in all other cases. We show that R is an L-bisimulation witnessing that $d^L_{\alpha,\beta}(a_1, b_1) \leq 0.2$, even though it is clearly neither reflexive nor symmetric on the disjoint union of the systems (it is easy to come up with similar but slightly larger examples where R also fails to be transitive, i.e. to satisfy the triangle inequality).

Specifically, we need to show for each a_i and b_j that $LR(\alpha(a_i), \beta(b_j)) \leq R(a_i, b_j)$. The cases with $R(a_i, b_j) = 1$ are trivial; in the other cases we have:

$$\begin{aligned}
HR(\{a_2,a_3\},\{b_2,b_3\}) &= \max(\max_{a \in \{a_2,a_3\}} \min_{b \in \{b_2,b_3\}} R(a,b), \max_{b \in \{b_2,b_3\}} \min_{a \in \{a_2,a_3\}} R(a,b)) \\
&= \max(0.1, 0.1) = 0.1 \\
LR(\alpha(a_1),\beta(b_1)) &= LR((0.7,\{a_2,a_3\}),(0.4,\{b_2,b_3\})) \\
&= \tfrac{1}{2}(|0.7-0.4| + HR(\{a_2,a_3\},\{b_2,b_3\})) \\
&= \tfrac{1}{2}(0.3+0.1) = 0.2 = R(a_1,b_1) \\
LR(\alpha(a_2),\beta(b_3)) &= LR((0.2,\varnothing),(0,\varnothing)) = \tfrac{1}{2}0.2 = 0.1 = R(a_2,b_3) \\
LR(\alpha(a_3),\beta(b_2)) &= LR((0.8,\varnothing),(0.7,\varnothing)) = \tfrac{1}{2}0.1 \leq 0.1 = R(a_3,b_2)
\end{aligned}$$

As indicated previously, quantitative Hennessy-Milner theorems can only be expected to hold for nonexpansive 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:

Definition 4.10. Let L be a lax extension of T, and let (A, α), (B, β) be T-coalgebras. The sequence of *approximants* of (L-behavioural distance) $d^L_{\alpha,\beta}$ are the fuzzy relations $d_\kappa\colon A \nrightarrow B$, indexed over ordinal numbers κ, inductively defined by

$$d_0 = 0, \qquad d_{\kappa+1} = Ld_\kappa \circ (\alpha \times \beta), \qquad d_\lambda = \sup\nolimits_{\kappa<\lambda} d_\kappa \quad (\lambda \text{ limit ordinal}).$$

We show some basic properties of these sequences:

Lemma 4.11. *Let L be a lax extension of T, let (A,α), (B,β) be T-coalgebras, and let $(d_\kappa\colon A \nrightarrow B)_\kappa$ be the sequence of approximants of $d^L_{\alpha,\beta}$. Then:*

(1) *The sequence $(d_\kappa)_\kappa$ is increasing.*
(2) *We have $d_\kappa \le d^L_{\alpha,\beta}$ for each ordinal κ.*
(3) *Let (A',α'), (B',β') be T-coalgebras, let $f\colon A \to A', g\colon B \to B'$ be coalgebra morphisms, and let $(d'_\kappa\colon A' \nrightarrow B')_\kappa$ be the sequence of approximants of $d^L_{\alpha',\beta'}$. Then $d_\kappa = d'_\kappa \circ (f \times g)$ for each ordinal κ.*

Proof.

(1) Because $d_\lambda = \sup_{\kappa<\lambda} d_\kappa$ for limit ordinals λ, it is enough to show that $d_\kappa \le d_{\kappa+1}$ for each ordinal κ. We proceed by induction; there are three cases. $d_0 \le d_1$ holds trivially. For successor ordinals $\kappa+1$, we have $d_{\kappa+1} = Ld_\kappa \circ (\alpha \times \beta) \le Ld_{\kappa+1} \circ (\alpha\times\beta) = d_{\kappa+2}$, where we used (L1) and the inductive hypothesis in the inequality. Finally, for limit ordinals λ, we have $Ld_\kappa \le Ld_\lambda$ for each $\kappa < \lambda$ by (L1), so that $\sup_{\kappa<\lambda} Ld_\kappa \le Ld_\lambda$. Therefore, $d_\lambda = \sup_{\kappa<\lambda} d_\kappa \le \sup_{\kappa<\lambda} d_{\kappa+1} = \sup_{\kappa<\lambda} Ld_\kappa \circ (\alpha\times\beta) \le Ld_\lambda \circ (\alpha\times\beta) = d_{\lambda+1}$.
(2) We proceed by induction; the cases for 0 and limit ordinals are trivial. For successor ordinals, we have $d_{\kappa+1} = Ld_\kappa \circ (\alpha\times\beta) \le Ld^L_{\alpha,\beta} \circ (\alpha\times\beta) = d^L_{\alpha,\beta}$ by the inductive hypothesis, by (L1), and by definition of $d^L_{\alpha,\beta}$.
(3) Again, we proceed by induction; the cases for 0 and limit ordinals are immediate from the definition. For $\kappa+1$ a successor ordinal, we compute

$$
\begin{aligned}
d_{\kappa+1} &= Ld_\kappa \circ (\alpha\times\beta) && \text{(definition of } d_{\kappa+1})\\
&= L(d'_\kappa \circ (f\times g)) \circ (\alpha\times\beta) && \text{(IH)}\\
&= Ld'_\kappa \circ (Tf \times Tg) \circ (\alpha\times\beta) && \text{(Lemma 3.8)}\\
&= Ld'_\kappa \circ (\alpha\times\beta) \circ (f\times g) && (f, g \text{ morphisms})\\
&= d'_{\kappa+1} \circ (f\times g). && \text{(definition of } d'_{\kappa+1})
\end{aligned}
$$

□

Crucially, if L is nonexpansive and finitarily separable, then the chain of approximants stabilizes after ω steps. Formally:

Theorem 4.12. *Let L be a nonexpansive finitarily separable lax extension of T. Given T-coalgebras (A,α), (B,β), let $(d_\kappa\colon A \nrightarrow B)_\kappa$, be the approximants of $d^L_{\alpha,\beta}$. Then*

(i) *$d_{\omega+1} = d_\omega$, and*
(ii) *L-behavioural distance $d^L_{\alpha,\beta}$ equals d_ω.*

To prove Theorem 4.12 in the case of non-finitary T, we make use of unravellings of coalgebras:

Definition 4.13 (Unravelling)**.** Let (C,γ) be a T-coalgebra and put $C^+ = \bigcup_{m\ge1} C^m$.

(1) For $\bar c = (c_1,\dots,c_m) \in C^m$ and $c \in C$ we put $\mathsf{last}(\bar c) = c_m$ and $\mathsf{app}_{\bar c}(c) = (c_1,\dots,c_m,c)$, defining maps $\mathsf{last}\colon C^+ \to C$ for each $m \ge 1$ and $\mathsf{app}_{\bar c}\colon C \to C^+$ for each $m \ge 1$ and $\bar c \in C^m$.
(2) The *unravelling* of (C,γ) is the T-coalgebra (C^+,γ^+), where $\gamma^+\colon C^+ \to TC^+$ is given by

$$\gamma^+(\bar c) = T\mathsf{app}_{\bar c}(\gamma(\mathsf{last}(\bar c))).$$

Every coalgebra is behaviourally equivalent to its unravelling:

Lemma 4.14. *For every T-coalgebra (C,γ),*

(i) *the map* $\mathsf{last}\colon (C^+,\gamma^+) \to (C,\gamma)$ *is a coalgebra morphism; and*
(ii) *every state* $c \in C$ *is behaviourally equivalent to the state* $(c) \in C^+$*.*

This fact is essentially standard; we give a proof for the sake of completeness:

Proof.

(i) Let $\bar{c} \in C^+$. Then clearly $\mathsf{last} \circ \mathsf{app}_{\bar{c}} = \mathsf{id}$ by definition and therefore

$$(T\mathsf{last} \circ \gamma^+)(\bar{c}) = T\mathsf{last}(T\mathsf{app}_{\bar{c}}(\gamma(\mathsf{last}(\bar{c})))) = T\mathsf{id}(\gamma(\mathsf{last}(\bar{c}))) = (\gamma \circ \mathsf{last})(\bar{c}).$$

(ii) This is immediate from (i), as behavioural equivalence is witnessed by the coalgebra morphisms $\mathsf{id}\colon C \to C$ and $\mathsf{last}\colon C^+ \to C$. □

Proof of Theorem 4.12. By the fixpoint definition of $d^L_{\alpha,\beta}$ and Lemma 4.11.2, (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 generalize 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 $i\colon A_0 \to A$ and $j\colon B_0 \to B$ 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). By Lemma 4.11.1, the maps $d_n \circ (i \times j)$ 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 L is nonexpansive, it is also continuous w.r.t. the supremum metric by Lemma 3.9, so it follows that

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

This covers the finitary case. In the general case, we make use of the unravellings (A^+,α^+) and (B^+,β^+), as well as the sequence $(d^+_\kappa\colon A^+ \nrightarrow B^+)_\kappa$ of approximants of $d^L_{\alpha^+,\beta^+}$. We can assume w.l.o.g. that $A \neq \emptyset \neq B$; then the inclusions $A^m \hookrightarrow A^+$, $B^m \hookrightarrow B^+$ (for $m \geq 1$) are preserved by T, and for readability we assume in the following that TA^m is in fact a subset of TA^+; similarly for B^m and T_ω, with naturality of L guaranteeing that the identification does not affect lifted distance. Now let $\epsilon > 0$. As L is finitarily separable, we can construct T_ω-coalgebras $\alpha^\epsilon\colon A^+ \to T_\omega A^+$ and $\beta^\epsilon\colon B^+ \to T_\omega B^+$ approximating α^+ and β^+ respectively. Specifically, for every $\bar{a} \in A^m$ we have $\alpha^+(\bar{a}) \in TA^{m+1}$ by definition, and as $T_\omega A^{m+1}$ is dense in TA^{m+1}, we can choose an element $\alpha^\epsilon(\bar{a}) \in T_\omega A^{m+1}$ such that

$$L\Delta_{A^+}(\alpha^+(\bar{a}),\alpha^\epsilon(\bar{a})) \leq \epsilon \cdot 3^{-m} \quad \text{and} \quad L\Delta_{A^+}(\alpha^\epsilon(\bar{a}),\alpha^+(\bar{a})) \leq \epsilon \cdot 3^{-m}. \tag{4.2}$$

Similarly, for each $\bar{b} \in B^m$ we choose $\beta^\epsilon(\bar{b}) \in T_\omega B^{m+1}$ such that

$$L\Delta_{B^+}(\beta^+(\bar{b}),\beta^\epsilon(\bar{b})) \leq \epsilon \cdot 3^{-m} \quad \text{and} \quad L\Delta_{B^+}(\beta^\epsilon(\bar{b}),\beta^+(\bar{b})) \leq \epsilon \cdot 3^{-m}. \tag{4.3}$$

We denote the sequence of approximants of $d^L_{\alpha^\epsilon,\beta^\epsilon}$ by $(d^\epsilon_\kappa\colon A^+ \nrightarrow B^+)_\kappa$ and show by induction that the d^ϵ_κ approximate the d^+_κ in the following sense: for all $m \geq 1$ and all $\bar{a} \in A^m, \bar{b} \in B^m$,

$$|d^\epsilon_\kappa(\bar{a},\bar{b}) - d^+_\kappa(\bar{a},\bar{b})| \leq \epsilon \cdot 3^{1-m} \tag{4.4}$$

for all ordinals κ.

For $\kappa = 0$ this clearly holds. For the inductive step from κ to $\kappa+1$, we note again that for $\bar{a} \in A^m$ and $\bar{b} \in B^m$ we have $\alpha^+(\bar{a}) \in TA^{m+1}$ and $\beta^+(\bar{b}) \in TB^{m+1}$ by definition . Therefore, by Lemma 3.9.3 and the inductive hypothesis, we have

$$|Ld^\epsilon_\kappa(\alpha^+(\bar{a}), \beta^+(\bar{b})) - Ld^+_\kappa(\alpha^+(\bar{a}), \beta^+(\bar{b}))| \leq \epsilon \cdot 3^{1-(m+1)} = \epsilon \cdot 3^{-m}, \tag{4.5}$$

so that we compute:

$$\begin{aligned}
&d^\epsilon_{\kappa+1}(\bar{a}, \bar{b}) \\
&= Ld^\epsilon_\kappa(\alpha^\epsilon(\bar{a}), \beta^\epsilon(\bar{b})) && \text{(definition of } d^\epsilon_{\kappa+1}) \\
&= L(\Delta_{A^+}; d^\epsilon_\kappa; \Delta_{B^+})(\alpha^\epsilon(\bar{a}), \beta^\epsilon(\bar{b})) && (\Delta \text{ neutral for } ;) \\
&\leq L\Delta_{A^+}(\alpha^\epsilon(\bar{a}), \alpha^+(\bar{a})) + Ld^\epsilon_\kappa(\alpha^+(\bar{a}), \beta^+(\bar{b})) + L\Delta_{B^+}(\beta^+(\bar{b}), \beta^\epsilon(\bar{b})) && \text{(L2)} \\
&\leq Ld^\epsilon_\kappa(\alpha^+(\bar{a}), \beta^+(\bar{b})) + 2\epsilon \cdot 3^{-m} && \text{(4.2) and (4.3)} \\
&\leq Ld^+_\kappa(\alpha^+(\bar{a}), \beta^+(\bar{b})) + \epsilon \cdot 3^{-m} + 2\epsilon \cdot 3^{-m} && \text{(4.5)} \\
&= d^+_{\kappa+1}(\bar{a}, \bar{b}) + \epsilon \cdot 3^{1-m}. && \text{(definition of } d^+_{\kappa+1})
\end{aligned}$$

We can symmetrically derive $d^+_{\kappa+1}(\bar{a}, \bar{b}) \leq d^\epsilon_{\kappa+1}(\bar{a}, \bar{b}) + \epsilon \cdot 3^{1-m}$, this time using the other inequalities in (4.2) and (4.3), so (4.4) holds for $\kappa+1$ as claimed. Finally, if κ is a limit ordinal, then (4.4) also follows inductively, as taking suprema is a nonexpansive operation.

Since the functor T_ω is finitary, we know from the finitary case that $d^\epsilon_\omega = d^\epsilon_{\omega+1}$. Therefore we have, for all $\bar{a} \in A^k, \bar{b} \in B^k$,

$$|d^+_\omega(\bar{a}, \bar{b}) - d^+_{\omega+1}(\bar{a}, \bar{b})| \leq |d^+_\omega(\bar{a}, \bar{b}) - d^\epsilon_\omega(\bar{a}, \bar{b})| + |d^\epsilon_{\omega+1}(\bar{a}, \bar{b}) - d^+_{\omega+1}(\bar{a}, \bar{b})| \leq 2\epsilon \cdot 3^{1-k} \leq 2\epsilon.$$

Because this holds for all $\epsilon > 0$, we have $d^+_\omega = d^+_{\omega+1}$. Thus, using Lemma 4.11.3 twice,

$$d_{\omega+1} \circ (\mathsf{last} \times \mathsf{last}) = d^+_{\omega+1} = d^+_\omega = d_\omega \circ (\mathsf{last} \times \mathsf{last}).$$

As last is surjective, this implies $d_{\omega+1} = d_\omega$. □

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 and $d_E(x, y) = |x - y|$ is Euclidean distance. The coalgebraic generalization of the Kantorovich lifting, both in the pseudometric setting [KMM18] 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 [SP11]. 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* $\mathcal{Q}\colon \mathsf{Set}^{\mathrm{op}} \to \mathsf{Set}$ is defined on sets X as $\mathcal{Q}X = (X \to [0,1])$ and on functions $f\colon X \to Y$ as $\mathcal{Q}f\colon \mathcal{Q}Y \to \mathcal{Q}X$, $\mathcal{Q}f(h) = h \circ f$.

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 $f_i \leq g_i$ for all i,
$$\lambda_X(f_1, \dots, f_n) \leq \lambda_X(g_1, \dots, g_n).$$

(4) We call λ *nonexpansive* if for all sets X and all functions $f_1, \dots, f_n, g_1, \dots, g_n \in \mathcal{Q}X$,
$$\|\lambda_X(f_1, \dots, f_n) - \lambda_X(g_1, \dots, g_n)\|_\infty \leq \max(\|f_1 - g_1\|_\infty, \dots, \|f_n - g_n\|_\infty).$$

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 [BBKK18, Sch08] and on the generic Wasserstein lifting [Hof07]; more generally, an n-ary predicate lifting is equivalent to a generalized form of evaluation function, of type $T([0,1]^n) \to [0,1]$ [Sch08].

More precisely, an evaluation function $e\colon T[0,1] \to [0,1]$ gives rise to a unary predicate lifting λ_e given by $\lambda_e(f) = e \circ Tf$. Conversely, the evaluation function corresponding to $\lambda\colon \mathcal{Q} \Rightarrow \mathcal{Q} \circ T$ is $e_\lambda = \lambda_{[0,1]}(\mathsf{id})$.

In the more general setting with higher arities, an n-ary evaluation function is a map $e\colon T([0,1]^n) \to [0,1]$, giving 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 $e_\lambda = \lambda_{[0,1]^n}(\pi_1, \dots, \pi_n)$.

Before we can show 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 [Vil08, 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 $f(a) - g(b) \leq R(a, b)$ for all $a \in A, b \in B$.

This notion is compatible with our previous use of the term: When $A = B$ and $d\colon A \nrightarrow A$ is a hemimetric, then $f\colon (A, d) \to ([0,1], d_\ominus)$ is nonexpansive in the sense used so far (cf. Section 2) iff the pair (f, f) is d-nonexpansive in the sense defined above. If d is a pseudometric, then this is moreover equivalent to f being nonexpansive as a map $(A, d) \to ([0,1], d_E)$.

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
$$R[f](b) = \sup\nolimits_{a \in A} f(a) \ominus R(a, b)$$
(recall from Section 2 that $\ominus$ denotes truncated subtraction).

We note some basic properties of nonexpansive pairs and nonexpansive companions. In particular, the nonexpansive companion of some function f is the least function (in pointwise order) forming a nonexpansive pair with f.

Lemma 5.5. *Let $R\colon A \nrightarrow B$. Then the following hold:*

(1) *If $f' \leq f$ and $g \leq g'$ and (f, g) is R-nonexpansive, then (f', g') is R-nonexpansive.*

(2) *(f, g) is R-nonexpansive if and only if $R[f] \leq g$.*

Definition 5.6. Let Λ be a set of monotone predicate liftings. 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

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

To show that the Kantorovich lifting is a lax extension, we need the following fact about nonexpansive pairs that will be used in the proof of the triangle inequality (L2).

Lemma 5.7. *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

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

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. □

We are now ready to prove the central result of the section, stating that the Kantorovich lifting is always a fuzzy lax extension. In general, it does not preserve converse, but does if the set of predicate liftings contains all duals of predicate liftings.

Theorem 5.8. *Let Λ be a set of monotone predicate liftings. The Kantorovich lifting K_Λ is a lax extension. If Λ is closed under duals, then K_Λ preserves converse. If all $\lambda \in \Lambda$ are nonexpansive, then K_Λ is nonexpansive as well.*

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:

- (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 5.7. Then it is enough to observe that:
$$\begin{aligned}\lambda_A(f)(t_1) \ominus \lambda_C(h)(t_3) &\le (\lambda_A(f)(t_1) \ominus \lambda_B(g)(t_2)) + (\lambda_B(g)(t_2) \ominus \lambda_C(h)(t_3))\\ &\le K_\Lambda R(t_1,t_2) + K_\Lambda S(t_2,t_3).\end{aligned}$$
- (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. The proof for Gr_h° is analogous, noting that a pair (f,g) is Gr_h°-nonexpansive iff $f \circ h \le g$.

- (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 $g(a) = f(a) \ominus \epsilon$, so that we have $\|f-g\|_\infty \le \epsilon$. In this case,
$$\lambda_A(f)(t) \ominus \lambda_A(g)(t) \le \|\lambda_A(f) - \lambda_A(g)\|_\infty \le \|f-g\|_\infty \le \epsilon.$$
- (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,

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

Remark 5.9 (Kantorovich for pseudometrics)**.** On pseudometrics, the Kantorovich lifting K_Λ as given by Definition 5.6 agrees with the usual Kantorovich distance $-^{\uparrow T}$ defined for pseudometrics [BBKK18, Definition 5.4]. If $d\colon A \nrightarrow A$ is a pseudometric, then

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

Lemma 5.10. *If Λ is closed under duals, then $K_\Lambda(d) = d^{\uparrow T}$ for every pseudometric d.*

Proof. 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 5.6 thus does not change if we restrict the choice of (f,g) to the case $f = g$. Finally, in case $f = g$, d-nonexpansiveness 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. $\square$

Example 5.11 (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 $\Diamond(f)(\mu) = \mathbb{E}_\mu(f)$. We claim that K is finitarily separable. To see this, 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

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

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

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

Following Remark 4.8, we can also see that the Kantorovich lifting characterizes behavioural equivalence for probabilistic transition systems, i.e. *probabilistic bisimilarity* [LS91]: To see that K satisfies (4.1), by Lemma 3.11 it suffices to show that $K\Delta_X(\mu_1,\mu_2) > 0$ for any $\mu_1 \ne \mu_2 \in \mathcal{D}X$. W.l.o.g. assume $\mu_1(x) > \mu_2(x)$ for some $x \in X$ and let $f \in \mathcal{Q}X$ be such that $f(x) = 1$ and $f(x') = 0$ otherwise. Then, as (f,f) is Δ_X-nonexpansive, we have $K\Delta_X(\mu_1,\mu_2) \ge f(x)(\mu_1(x) - \mu_2(x)) > 0$.

(2) The asymmetric Hausdorff lifting $H^{\leftarrow}$ (Example 3.14) is equal to the Kantorovich lifting for the single predicate lifting $\Diamond_X(f)(A) = \sup f[A]$. Let $R\colon A \nrightarrow B$ and let $U \subseteq A$, $V \subseteq B$. We show $H^{\leftarrow}R(U,V) = K_{\{\Diamond\}}R(U,V)$.
   - '$\leq$': Let (f,g) be an R-nonexpansive pair. Then
   $$\sup_{a\in U} f(a) \ominus \sup_{b\in V} g(b) \leq \sup_{a\in U} \inf_{b\in V} f(a) \ominus g(b) \leq \sup_{a\in U} \inf_{b\in V} R(a,b) = H^{\leftarrow}R(a,b).$$
   - '$\geq$': Let $a \in U$ and let $f \in \mathcal{Q}A$ be the indicator function of $\{a\}$, that is $f(a') = 1$ if $a' = a$ and $f(a') = 0$ otherwise. Put $g = R[f]$, so that $g(b) = 1 \ominus R(a,b)$ for each $b \in B$. Then, as (f,g) is R-nonexpansive (Lemma 5.5),
   $$K_{\{\Diamond\}}R(U,V) \geq \sup_{a\in U} f(a) \ominus \sup_{b\in V} g(b) = 1 \ominus \sup_{b\in V}(1 \ominus R(a,b)) = \inf_{b\in V} R(a,b).$$

   Dually, the other asymmetric form $H^{\rightarrow}$ of the Hausdorff lifting is thus the Kantorovich lifting for the single predicate lifting $\Box_X(f)(A) = \inf f[A]$. It follows immediately that the symmetric Hausdorff lifting H is the Kantorovich lifting K_Λ for $\Lambda = \{\Box, \Diamond\}$.

(3) 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* [RG13, CNR16]. 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 $A\colon \mathcal{Q}X \to [0,1]$. We put
   $$LR(A,B) = \sup\nolimits_{f\in\mathcal{Q}X} A(f) \ominus B(R[f])$$
   for $R\colon X \nrightarrow Y$, $A \in \mathcal{M}X$, $B \in \mathcal{M}Y$ (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 [BBKK18] to lift arbitrary fuzzy relations instead of just pseudometrics; our construction slightly generalizes one given by Hofmann [Hof07]. Like the Kantorovich lifting, the Wasserstein lifting is based on a choice of predicate liftings. Compared to the case of the Kantorovich lifting, where we needed to work with nonexpansive pairs, the generalization from lifting pseudometrics to lifting relations is much more direct for the Wasserstein lifting. 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 $\mathsf{Cpl}(t_1,t_2) = \{t \in T(A\times B) \mid T\pi_1(t) = t_1, T\pi_2(t) = t_2\}$.

The Wasserstein lifting uses predicate liftings in a quite different manner from the Kantorovich lifting, and in particular appears to make sense only for unary predicate liftings, so unlike elsewhere 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 [Hof07, 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 [BBKK18, Definition 5.12] (again up to the fact that we admit more than one modality). 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 $f, g \in \mathcal{Q}X$, $\lambda_X(f \oplus g) \leq \lambda_X(f) \oplus \lambda_X(g)$.
(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.

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* [BBKK18, Definition 5.14].

Definition 6.4. An evaluation function $e\colon T[0,1] \to [0,1]$ is *well-behaved* if it satisfies the following conditions.

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

This amounts to a slightly stronger condition than standardness of the corresponding predicate lifting:

Lemma 6.5. *An evaluation function $e\colon T[0,1] \to [0,1]$ is well-behaved 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 6.4 with subadditivity of λ_e. Reformulating in terms of λ_e gives

$$|\lambda_e(\pi_1)(t) - \lambda_e(\pi_2)(t)| \leq \lambda_e(d_E)(t) \quad \text{for } t \in T([0,1]^2). \tag{6.1}$$

We show that (6.1) is equivalent to subadditivity of λ_e, given that λ_e is monotone:

- '$\Rightarrow$': Let $f, g \in \mathcal{Q}X, t \in TX$. Put $t' \coloneqq T\langle f \oplus g, f\rangle(t) \in T([0,1]^2)$. 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
$$\lambda_e(d_E)(t') = \lambda_e(d_E \circ \langle f \oplus g, f\rangle)(t) \leq \lambda_e(g)(t),$$
where the last step is by monotonicity of λ_e. Therefore, $\lambda(f \oplus g)(t) - \lambda(f)(t) \leq \lambda(g)(t)$ by (6.1).
- '$\Leftarrow$': Put $f = d_E, g = \pi_1\colon [0,1]^2 \to [0,1]$. 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. ∎

Similar conditions also feature in Hofmann's *topological theories* [Hof07, 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. Explicitly, a topological theory is defined as a triple consisting of a monad T, a quantale V, and a map $\xi\colon TV \to V$ 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:

$$\begin{aligned}
&(Q_\otimes) && \oplus \circ \langle \lambda_\xi(\pi_1), \lambda_\xi(\pi_2)\rangle \geq \lambda_\xi(\oplus)\\
&(Q_k) && 0 \geq \lambda_\xi(0_1)(t) \text{ for every } t \in T1, \text{ where } 1 \text{ is a singleton set}\\
&(Q'_\vee) && \lambda_\xi \text{ is a monotone natural transformation}
\end{aligned}$$

Using a similar idea as in Lemma 6.5, 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 [Hof07, 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.

If T preserves weak pullbacks, the following so-called gluing lemma holds [BBKK18, Lemma 5.18]:

Lemma 6.6 (Gluing)**.** *Let A, B and C be sets, and let $t_1 \in TA, t_2 \in TB, t_3 \in TC$. Let $t_{12} \in \mathsf{Cpl}(t_1,t_2)$ and $t_{23} \in \mathsf{Cpl}(t_2,t_3)$. Then there exists $t_{123} \in \mathsf{Cpl}(t_1,t_2,t_3)$ such that*

$$T\langle \pi_1,\pi_2\rangle(t_{123}) = t_{12} \quad \textit{and} \quad T\langle \pi_2,\pi_3\rangle(t_{123}) = t_{23},$$

where the π_j are the projections of the product $A \times B \times C$. Moreover, $t_{13} \coloneqq T\langle \pi_1,\pi_3\rangle(t_{123}) \in \mathsf{Cpl}(t_1,t_3)$.

Using Lemma 6.6, we can now show that the Wasserstein lifting is a fuzzy lax extension. Unlike the Kantorovich lifting, the Wasserstein lifting always preserves converse, without any further restrictions on the set of predicate liftings.

Theorem 6.7. *If T preserves weak pullbacks and Λ is a set of standard predicate liftings, then the Wasserstein lifting W_Λ is a converse-preserving lax extension. If additionally all $\lambda \in \Lambda$ are nonexpansive, then W_Λ is nonexpansive as well.*

Proof. We show the five properties one by one:

- (L0): Let $\mathsf{swap} = \langle \pi_2,\pi_1\rangle\colon A \times B \to B \times A$. Then $T\mathsf{swap}$ is an isomorphism between $\mathsf{Cpl}(t_1,t_2)$ and $\mathsf{Cpl}(t_2,t_1)$ and it suffices to observe that for every $\lambda \in \Lambda$ and $t \in T(A \times B)$, $\lambda_{B\times A}(R^\circ)(T\mathsf{swap}(t)) = \lambda_{A\times B}(R)(t)$ by naturality of λ.
- (L1): Immediate from the definition of W_Λ and monotonicity of the predicate liftings.
- (L2): Let $R\colon A \nrightarrow B, S\colon B \nrightarrow C$ and let $t_1 \in TA, t_2 \in TB, t_3 \in TC$. We need to show that $W_\Lambda(R;S)(t_1,t_3) \leq W_\Lambda R(t_1,t_2) + W_\Lambda S(t_2,t_3)$. Let $\lambda \in \Lambda$, $t_{12} \in \mathsf{Cpl}(t_1,t_2)$ and $t_{23} \in \mathsf{Cpl}(t_2,t_3)$, and let t_{123} and t_{13} be as in Lemma 6.6. We need to show
$$\lambda_{A\times C}(R;S)(t_{13}) \leq \lambda_{A\times B}(R)(t_{12}) + \lambda_{B\times C}(S)(t_{23}). \tag{6.2}$$
We define three functions $f_{12}, f_{13}, f_{23}\colon A\times B\times C \to [0,1]$ by $f_{12}(a,b,c) = R(a,b)$, $f_{23}(a,b,c) = S(b,c)$, and $f_{13}(a,b,c) = (R;S)(a,c)$. Then, as $f_{13} \leq f_{12} \oplus f_{23}$, we obtain
$$\lambda_{A\times B\times C}(f_{13})(t_{123}) \leq \lambda_{A\times B\times C}(f_{12})(t_{123}) + \lambda_{A\times B\times C}(f_{23})(t_{123})$$
by monotonicity and subadditivity of λ, which is equivalent to (6.2) by naturality of λ.

- (L3): Let $f\colon A \to B$, $t_1 \in TA$ and $\lambda \in \Lambda$. We need to find $t \in \mathsf{Cpl}(t_1, Tf(t_1))$ such that $\lambda_{A\times B}(\mathrm{Gr}_f)(t) = 0$. Indeed, take $t = T\langle \mathsf{id}_A, f\rangle(t_1)$. Then $T\pi_1(t) = T\mathsf{id}_A(t_1) = t_1$ and $T\pi_2(t) = Tf(t_1)$, and, as λ is natural and preserves zero,
$$\lambda_{A\times B}(\mathrm{Gr}_f)(t) = \lambda_A(\mathrm{Gr}_f \circ \langle \mathsf{id}_A, f\rangle)(t_1) = \lambda_A(0_A)(t_1) = 0.$$
This proves $W_\Lambda(\mathrm{Gr}_f) \le \mathrm{Gr}_{Tf}$. The second clause $W_\Lambda(\mathrm{Gr}^\circ_f) \le \mathrm{Gr}^\circ_{Tf}$ now follows using (L0).
- (L4): Let A be a set, $\epsilon > 0$, $t_1 \in TA$ and $\lambda \in \Lambda$. It is enough to find $t \in \mathsf{Cpl}(t_1, t_1)$ such that $\lambda_{A\times A}(\Delta_{\epsilon,A})(t) \le \epsilon$. Indeed, take $t = T\langle \mathsf{id}_A, \mathsf{id}_A\rangle(t_1)$. Then $T\pi_1(t) = T\pi_2(t) = t_1$, and with $\epsilon_A\colon A \to [0,1]$ being the constant map $a \mapsto \epsilon$ we derive
$$\lambda_{A\times A}(\Delta_{\epsilon,A})(t) = \lambda_A(\epsilon_A)(t_1) \le \|\lambda_A(\epsilon_A) - \lambda_A(0_A)\|_\infty \le \epsilon,$$
using that λ is natural, nonexpansive and preserves zero. □

Example 6.8 (Wasserstein liftings)**.**

(1) Similar to the case of the standard Kantorovich lifting K (Example 5.11.1), the standard Wasserstein lifting W of the discrete distribution functor $\mathcal{D}$ arises as an instance of the generic Wasserstein lifting, for the same predicate lifting $\Diamond(f)(\mu) = \mathbb{E}_\mu(f)$. In fact, it is well known [Vil08, Theorem 5.10] that $K = W$, a fact known as *Kantorovich-Rubinstein duality*.

(2) The Hausdorff lifting H (Example 3.14) is the Wasserstein lifting $W_{\{\lambda\}}$ for $\mathcal{P}$, where $\lambda_X(f)(A) = \sup f[A]$ for $A \subseteq X$. To see this, let $R\colon A \nrightarrow B$, and let $U \subseteq A$ and $V \subseteq B$. Then we show that $HR(U,V) = W_{\{\lambda\}}R(U,V)$ by proving the two inequalities separately:
   - '$\le$': 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 $\inf_{b\in V} R(a,b) \le \sup R[Z]$. Thus, we have $\sup_{a\in U}\inf_{b\in V} R(a,b) \le \sup R[Z]$, and, by a symmetrical argument, $\sup_{b\in V}\inf_{a\in U} R(a,b) \le \sup R[Z]$.
   - '$\ge$': Let $\epsilon > 0$. It suffices to find a coupling $Z \in \mathsf{Cpl}(U,V)$ such that $\sup R[Z] \le HR(U,V) + \epsilon$. So let $\epsilon > 0$. We construct functions $f\colon U \to V$ and $g\colon V \to U$ as follows: For each $a \in U$ choose $f(a) \in V$ such that $R(a, f(a)) \le \inf_{b\in V} R(a,b) + \epsilon$. Similarly, for each $b \in V$ choose $g(b) \in U$ such that $R(g(b), b) \le \inf_{a\in U} R(a,b) + \epsilon$. 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,
$$\sup R[Z] = \max(\sup_{a\in U} R(a, f(a)), \sup_{b\in V} R(g(b), b)) \le HR(U,V) + \epsilon.$$

(3) The convex powerset functor $\mathcal{C}$, whose coalgebras combine probabilistic branching and nondeterminism [BSS17], 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 nonexpansive lax extension of $\mathcal{C}$. Of course, λ is just the composite of the predicate liftings respectively defining the standard Kantorovich/Wasserstein and Hausdorff liftings. $W_{\{\lambda\}}$ indeed coincides with the composite HW of these liftings (for which a quantitative equational axiomatization has recently been given by Mio and Vignudelli [MV20]):

   Let $R\colon A \nrightarrow B$, 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:
   - '$\ge$': 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 $\mathbb{E}_\mu(R) \ge WR(\mathcal{D}\pi_1(\mu), \mathcal{D}\pi_2(\mu))$ and therefore
$$\sup_{\mu\in Z}\mathbb{E}_\mu(R) \ge \sup_{(\mu_1,\mu_2)\in Y} WR(\mu_1,\mu_2) \ge HW(R)(U,V).$$

- '≤': Let $Y \in \mathsf{Cpl}_{\mathcal{P}}(U,V)$ and $\epsilon > 0$. It suffices to find $Z \in \mathsf{Cpl}_{\mathcal{C}}(\mu_1,\mu_2)$ such that
$$\sup_{\mu \in Z} \mathbb{E}_\mu(R) \le \sup_{(\mu_1,\mu_2)\in Y} WR(\mu_1,\mu_2) + \epsilon.$$
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 $\mathbb{E}_\mu(R) \le WR(\mu_1,\mu_2) + \epsilon$. 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 have $\mathcal{C}\pi_2(Z) = V$, 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 $\mathbb{E}_\mu = \sum_{i=1}^n p_i \mathbb{E}_{\mu_i} \le \max_{i=1}^n \mathbb{E}_{\mu_i}$. Therefore we have, as required,
$$\sup_{\mu\in Z} \mathbb{E}_\mu(R) = \sup_{\mu \in Z'} \mathbb{E}_\mu(R) \le \sup_{(\mu_1,\mu_2)\in Y} WR(\mu_1,\mu_2) + \epsilon.$$

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 nonexpansive predicate liftings as precisely the nonexpansive 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 nonexpansiveness* 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 nonexpansiveness* if all $\lambda \in \Lambda$ preserve nonexpansiveness.

Lemma 7.2. *We have $K_\Lambda R \le LR$ for all fuzzy relations R if and only if Λ preserves nonexpansiveness.*

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
$$t_1 \ne t_2 \implies \exists \lambda \in \Lambda, A' \subseteq A \text{ such that } t_1 \in \lambda_A(A') \not\leftrightarrow t_2 \in \lambda_A(A')$$
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 nonexpansiveness 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) [Sch08, 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 [MV15, 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 [KL09, MV15]. 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 with signature Σ (such a presentation always exists [MV15, Example 21]) and assume a finitarily separable fuzzy lax extension L of T. To derive predicate liftings from the operations in Σ, we make use of the fuzzy elementhood relation ϵ_X (indexed over arbitrary sets X), where $\epsilon_X\colon X \nrightarrow \mathcal{Q}X$ is given by $\epsilon_X(x, f) = f(x)$.

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

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

It follows from naturality of σ and L (Lemma 3.8) that μ^σ is indeed natural and therefore a predicate lifting, as shown next. Indeed, for any $g\colon A \to B$, $f_1, \dots, f_n \in \mathcal{Q}B$ and $t \in TB$ we note that $\epsilon_A \circ (\mathsf{id} \times \mathcal{Q}g) = \epsilon_B \circ (g \times \mathsf{id})$ by definition of ϵ_A and ϵ_B and thus

$$\begin{aligned}
&\mu^\sigma_A(f_1 \circ g, \dots, f_n \circ g)(t) \\
&= L\epsilon_A(t, \sigma_{\mathcal{Q}A}(f_1 \circ g, \dots, f_n \circ g)) && \text{(definition of } \mu^\sigma) \\
&= L\epsilon_A(t, T\mathcal{Q}g(\sigma_{\mathcal{Q}B}(f_1, \dots, f_n))) && (\sigma \text{ natural}) \\
&= L(\epsilon_A \circ (\mathsf{id} \times \mathcal{Q}g))(t, \sigma_{\mathcal{Q}B}(f_1, \dots, f_n)) && \text{(Lemma 3.8)} \\
&= L(\epsilon_B \circ (g \times \mathsf{id}))(t, \sigma_{\mathcal{Q}B}(f_1, \dots, f_n)) \\
&= L\epsilon_B(Tg(t), \sigma_{\mathcal{Q}B}(f_1, \dots, f_n)) && \text{(Lemma 3.8)} \\
&= \mu^\sigma_B(f_1, \dots, f_n)(Tg(t)). && \text{(definition of } \mu^\sigma)
\end{aligned}$$

We are now in a position to present the proof of Theorem 7.4: We take Λ to be the set of Moss liftings and show the required properties of Λ one by one.

Convention 7.7. Throughout this proof, all statements and proofs will be written for the case where all $\sigma \in \Sigma$ (and therefore the induced Moss liftings) are unary. This is purely in the interest of readability; the general case requires only more indexing.

Monotonicity. The proof is based on the following auxiliary fact about pairs of elements that are mapped to 0.

Lemma 7.8. *Let $\sigma \in \Sigma$ and $R\colon A \nrightarrow B$. Then for all $a \in A, b \in B$ with $R(a,b) = 0$ we have $LR(\sigma_A(a), \sigma_B(b)) = 0$.*

Proof. Put $R_0 = \{(a,b) \in A \times B \mid R(a,b) = 0\}$ and consider the projection maps $\pi_1\colon R_0 \to A$ and $\pi_2\colon R_0 \to B$. Then it is easy to see that $R \le \mathrm{Gr}^\circ_{\pi_1}; \mathrm{Gr}_{\pi_2}$ (noting again that we read 0 as 'related' and 1 as 'unrelated'; in particular recall Convention 3.3 and Definition 3.4). Using the axioms of lax extensions, we obtain

$$LR \le L(\mathrm{Gr}^\circ_{\pi_1}; \mathrm{Gr}_{\pi_2}) \le L\mathrm{Gr}^\circ_{\pi_1}; L\mathrm{Gr}_{\pi_2} \le \mathrm{Gr}^\circ_{T\pi_1}; \mathrm{Gr}_{T\pi_2}. \tag{7.1}$$

For $(a,b) \in R_0$, put $t = \sigma_{A\times B}((a,b))$, so that $T\pi_1(t) = \sigma_A(a)$ and $T\pi_2(t) = \sigma_B(b)$ by naturality of σ. This means that $(\mathrm{Gr}^\circ_{T\pi_1}; \mathrm{Gr}_{T\pi_2})(\sigma(a), \sigma(b)) = 0$, so that by (7.1) we have $LR(\sigma_A(a), \sigma_B(b)) = 0$. ☐

Lemma 7.9. *Let $\sigma \in \Sigma$. Then the Moss lifting μ^σ is monotone.*

Proof. We make use of the fuzzy relation $R\colon \mathcal{Q}X \nrightarrow \mathcal{Q}X$ given by $R(g,f) = \sup_{x\in X} f(x) \ominus g(x)$, which we claim to satisfy the following two useful properties:

$$R(g,f) = 0 \iff f \le g \tag{7.2}$$

$$\epsilon_X \;\le\; \epsilon_X; R \tag{7.3}$$

The first property is clear; the second property amounts to showing that $f(x) \le g(x) \oplus R(g,f)$ for all $x \in X$ and all $f, g \in \mathcal{Q}X$ and is easily shown by case analysis on the definition of $\oplus$.

Let $f, g \in \mathcal{Q}X$ with $f \le g$ and let $t \in TX$. First, we note that by (7.2) we have $R(g,f) = 0$ and thus $LR(\sigma_{\mathcal{Q}X}(g), \sigma_{\mathcal{Q}X}(f)) = 0$ by Lemma 7.8. Second, by (7.3) and the axioms of lax extensions we have $L\epsilon_X \le L(\epsilon_X; R) \le L\epsilon_X; LR$. Therefore:

$$\begin{aligned}\mu^\sigma_X(f)(t) = L\epsilon_X(t, \sigma_{\mathcal{Q}X}(f)) &\le (L\epsilon_X; LR)(t, \sigma_{\mathcal{Q}X}(f)) \\ &\le L\epsilon_X(t, \sigma_{\mathcal{Q}X}(g)) \oplus LR(\sigma_{\mathcal{Q}X}(g), \sigma_{\mathcal{Q}X}(f)) = \mu^\sigma_X(g)(t).\end{aligned}$$ ☐

Preservation of nonexpansiveness.

Lemma 7.10. *Let $\sigma \in \Sigma$. Then the Moss lifting μ^σ preserves nonexpansiveness.*

Proof. Let $R\colon A \nrightarrow B$ and consider the map $R[-]\colon \mathcal{Q}A \to \mathcal{Q}B, f \mapsto R[f]$. First, we show that

$$\epsilon_A \;\le\; R; \epsilon_B; \mathrm{Gr}^\circ_{R[-]}. \tag{7.4}$$

Let $f \in \mathcal{Q}A$, $g \in \mathcal{Q}B$, and let $a \in A$, $b \in B$. We need to show that

$$\epsilon_A(a,f) \le R(a,b) \oplus \epsilon_B(b,g) \oplus \mathrm{Gr}^\circ_{R[-]}(g,f).$$

If $g \ne R[f]$, this holds trivially as $\mathrm{Gr}^\circ_{R[-]}(g,f) = 1$. Otherwise, if $g = R[f]$, then we have $f(a) \ominus R(a,b) \le g(b)$ by definition, and hence

$$\epsilon_A(a,f) = f(a) \le R(a,b) \oplus g(b) \le R(a,b) \oplus \epsilon_B(b,g) \oplus \mathrm{Gr}^\circ_{R[-]}(g,f).$$

Now let (f,g) be R-nonexpansive and let $t_1 \in TA$ and $t_2 \in TB$. We need to show that

$$\mu^\sigma_A(f)(t_1) - \mu^\sigma_B(g)(t_2) \le LR(t_1, t_2). \tag{7.5}$$

By monotonicity of μ^σ and Lemma 5.5 it is enough to show this for the case $g = R[f]$. In this case we have $TR[-](\sigma_{\mathcal{Q}A}(f)) = \sigma_{\mathcal{Q}B}(g)$ by naturality of σ. Applying the lax extension laws to (7.4), we have $L\epsilon_A \le LR; L\epsilon_B; \mathrm{Gr}^\circ_{TR[-]}$, so that

$$\begin{aligned} L\epsilon_A(t_1, \sigma_{\mathcal{Q}A}(f)) &\le LR(t_1,t_2) \oplus L\epsilon_B(t_2, \sigma_{\mathcal{Q}B}(g)) \oplus \mathrm{Gr}^\circ_{TR[-]}(\sigma_{\mathcal{Q}B}(g), \sigma_{\mathcal{Q}A}(f)) \\ &= LR(t_1,t_2) \oplus L\epsilon_B(t_2, \sigma_{\mathcal{Q}B}(g)), \end{aligned}$$

and (7.5) follows by rearranging. □

Separation. To show that Λ is separating for L, we need to make use of the fact that L is finitarily separable.

Lemma 7.11. *Λ is separating for L, that is, $L \le K_\Lambda$.*

Proof. Let $R\colon A \nrightarrow B$ and $t_1 \in TA$, $t_2 \in TB$. Let $\epsilon > 0$. Put $s\colon B \to \mathcal{Q}A, s(b)(a) = R(a,b)$. Because the set of Σ-terms over $\mathcal{Q}A$ generates $T_\omega \mathcal{Q}A$ and L is finitarily separable, there exists some $\sigma \in \Sigma$ and some $f \in \mathcal{Q}A$ such that we have $L\Delta_{\mathcal{Q}A}(\sigma_{\mathcal{Q}A}(f), Ts(t_2)) \le \epsilon$ and $L\Delta_{\mathcal{Q}A}(Ts(t_2), \sigma_{\mathcal{Q}A}(f)) \le \epsilon$. Put $g = R[f]$. Then it suffices to show that

$$\mu^\sigma_A(f)(t_1) - \mu^\sigma_B(g)(t_2) + 2\epsilon \ge LR(t_1,t_2). \tag{7.6}$$

First, by construction and naturality (Lemma 3.8),

$$L\epsilon_A(t_1, Ts(t_2)) = L(\epsilon_A \circ (\mathsf{id}_A \times s))(t_1,t_2) = LR(t_1,t_2),$$

where in the second step we used that $(\epsilon_A \circ (\mathsf{id}_A \times s))(a,b) = s(b)(a) = R(a,b)$ for all $a \in A, b \in B$. By the axioms of lax extensions we also have $L\epsilon_A \le L\epsilon_A; L\Delta_{\mathcal{Q}A}$ and therefore

$$LR(t_1,t_2) = L\epsilon_A(t_1, Ts(t_2)) \le L\epsilon_A(t_1, \sigma_{\mathcal{Q}A}(f)) \oplus L\Delta_{\mathcal{Q}A}(\sigma_{\mathcal{Q}A}(f), Ts(t_2)) \le \mu^\sigma_A(f)(t_1) + \epsilon.$$

Second, by the axioms of lax extensions and using naturality again,

$$\begin{aligned} L\epsilon_B(t_2, T(R[-]\circ s)(t_2)) &= L(\epsilon_B \circ (\mathsf{id}_B \times (R[-]\circ s)))(t_2,t_2) \\ &\le L\Delta_B(t_2,t_2) = \Delta_B(t_2,t_2) = 0, \end{aligned} \tag{7.7}$$

where in the inequality we used that for all $b_1, b_2 \in B$,

$$(\epsilon_B \circ (\mathsf{id}_B \times (R[-]\circ s)))(b_1,b_2) = R[s(b_2)](b_1) = \sup\nolimits_{a\in A} R(a,b_2) \ominus R(a,b_1) \le \Delta_B(b_1,b_2).$$

As before, we have $L\epsilon_B \le L\epsilon_B; L\Delta_{\mathcal{Q}B}$. We also have $\sigma_{\mathcal{Q}B}(g) = TR[-](\sigma_{\mathcal{Q}A}(f))$ by naturality of σ. Therefore, by naturality of L:

$$\begin{aligned} &\mu^\sigma_B(g)(t_2) \\ &= L\epsilon_B(t_2, \sigma_{\mathcal{Q}B}(g)) && \text{(definition)} \\ &\le L\epsilon_B(t_2, T(R[-]\circ s)(t_2)) \oplus L\Delta_{\mathcal{Q}B}(T(R[-]\circ s)(t_2), TR[-](\sigma_{\mathcal{Q}A}(f))) && \text{(L2)} \\ &= L\Delta_{\mathcal{Q}B}(T(R[-]\circ s)(t_2), TR[-](\sigma_{\mathcal{Q}A}(f))) && \text{(7.7)} \\ &= L(\Delta_{\mathcal{Q}B} \circ (R[-]\times R[-]))(Ts(t_2), \sigma_{\mathcal{Q}A}(f)) && \text{(naturality)} \\ &\le L\Delta_{\mathcal{Q}A}(Ts(t_2), \sigma_{\mathcal{Q}A}(f)) \le \epsilon. && \text{(Lemma 3.5)} \end{aligned}$$

Our target inequality (7.6) now follows by combining and rearranging the above inequalities. □

Nonexpansiveness. We note that ϵ-diagonals characterize the supremum norm as follows:

Lemma 7.12. *Let X be a set, let $f, g: X \to [0,1]$ and let $\epsilon > 0$. Then $\|f - g\|_\infty \le \epsilon$ if and only if both (f,g) and (g,f) are $\Delta_{\epsilon,X}$-nonexpansive pairs.*

Lemma 7.13. *Let $\sigma \in \Sigma$. If L is nonexpansive, then the Moss lifting μ^σ is nonexpansive.*

Proof. Let $f, g \in \mathcal{Q}X$ with $\|f - g\|_\infty \le \epsilon$. We need to show that $\|\mu^\sigma_X(f) - \mu^\sigma_X(g)\|_\infty \le \epsilon$. By Lemma 7.12, we know that the pairs (f,g) and (g,f) are $\Delta_{\epsilon,X}$-nonexpansive. Therefore, because the Moss liftings preserve nonexpansiveness (Lemma 7.10), the pairs $(\mu^\sigma_X(f), \mu^\sigma_X(g))$ and $(\mu^\sigma_X(g), \mu^\sigma_X(f))$ are $L\Delta_{\epsilon,X}$-nonexpansive, and thus they are also $\Delta_{\epsilon,TX}$-nonexpansive by (L4). The claim now follows by another application of Lemma 7.12. ☐

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 nonexpansive lax extensions. The framework goes back to work on fuzzy description logics [SP11]. The present version, characterized by a specific choice of propositional operators, appears in work on the coalgebraic quantitative Hennessy-Milner theorem [KMM18], and generalizes quantitative probabilistic modal logic [vBW05].

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

$$\phi, \psi ::= c \mid \phi \ominus c \mid \phi \oplus c \mid \phi \wedge \psi \mid \phi \vee \psi \mid \lambda(\phi_1, \dots, \phi_n) \tag{8.1}$$

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 $[\![\phi]\!]_{A,\alpha}: A \to [0,1]$, or just $[\![\phi]\!]$, defined by

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

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 [vBW05, KMM18, WSPK18]) – in particular, the more general *Łukasiewiecz semantics* fails to be nonexpansive w.r.t. behavioural distance, and indeed induces a discrete logical distance [WSPK18].

In the same vein, we require the modalities $\lambda \in \Lambda$ to be nonexpansive to avoid situations where non-zero logical distances (Definition 8.4) can be arbitrarily blown up by repeated application of modalities, such as in the case of the doubling modality $\lambda_X(f)(x) = 2f(x)$ of the identity functor.

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 [DEP98]. 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 [vBW05] does feature essentially the same propositional operators as our grammar (8.1), if negation is defined as in Remark 8.2 below.

Following [vBW05], we restrict truth constants in formulae to rational numbers, thus ensuring that the set of formulae is countable provided Λ is countable. This countability is not needed for any of our results, and they will still hold if the truth constants come from any dense subset of $[0, 1]$ (including $[0, 1]$ itself).

Remark 8.2. The logic as defined above does not include negation. This is to be expected, as already in the classical case the characteristic logic for similarity is negation-free modal logic with $\Diamond$ as the only modality [vG90]. However, if the set Λ of predicate liftings is closed under duals (and the corresponding Kantorovich lifting therefore preserves converse), then negation $\neg\phi$ can be defined recursively using De Morgan's laws for the propositional operators and duals for the modalities:

$$\begin{aligned} \neg c &= 1 - c & \neg(\phi \wedge \psi) &= \neg\phi \vee \neg\psi \\ \neg(\phi \ominus c) &= \neg\phi \oplus c & \neg(\phi \vee \psi) &= \neg\phi \wedge \neg\psi \\ \neg(\phi \oplus c) &= \neg\phi \ominus c & \neg\lambda(\phi_1, \dots, \phi_n) &= \bar{\lambda}(\neg\phi_1, \dots, \neg\phi_n) \end{aligned}$$

With negation defined like this, this version of real-valued coalgebraic modal logic is equivalent to the one in [WS20], which includes negation as a primitive. The latter logic does not explicitly include addition, but in the presence of subtraction and negation we can define it as $\phi \oplus c = \neg(\neg\phi \ominus c)$.

Example 8.3.

(1) *Fuzzy modal logic* may be seen as a basic fuzzy description logic [LS08]. Eliding propositional atoms for brevity (they may be added as nullary modalities), we take $\Lambda = \{\Diamond, \Box\}$. 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 $\Diamond$ and $\Box$ are interpreted as the predicate liftings

$$\Diamond_A(f)(g) = \sup_{a \in A} \min(g(a), f(a)) \qquad \text{and} \qquad \Box_A(f)(g) = \inf_{a \in A} \max(1 - g(a), f(a)).$$

We note that $\Diamond$ and $\Box$ are dual, so that negation can be defined as in Remark 8.2. 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

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

Probabilistic modal logic can be extended with negation by adding the dual $\Box$ of $\Diamond$. As taking expected values is self-dual, $\Box$ only differs from $\Diamond$ on deadlocks, where $\Box_A(f)(*) = 1$. When additionally extended with propositional atoms, this induces (up to restricting to discrete probabilities) van Breugel et al.'s contraction-free quantitative probabilistic modal logic [vBHMW07].

In the two-valued setting, modal logic is typically invariant under bisimulation, i.e. bisimilar states satisfy the same modal formulae. By contrast, under the asymmetric notion of similarity, the corresponding statement is that the fragment of modal logic that only uses the $\Diamond$ and no negations is *preserved under simulation*, i.e. if some state is simulated by another state, then all formulae of this shape that are satisfied by the first state are also satisfied by the second state.

In the quantitative setting, both of these statements correspond to nonexpansiveness of formula evaluation w.r.t. the behavioural distance arising from a Kantorovich lifting, where the distinction between the two scenarios is embedded in the choice of modalities. We may also phrase this more compactly by saying that logical distance is below behavioural distance:

Definition 8.4. The Λ-*logical distance* between states $a \in A$, $b \in B$ in T-coalgebras (A, α), (B, β) is

$$d^{\Lambda}(a, b) = \sup\{[\![\phi]\!](a) \ominus [\![\phi]\!](b) \mid \phi \in \mathcal{L}_{\Lambda}\}.$$

Lemma 8.5. *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*

$$[\![\phi]\!]_{A,\alpha}(a) \ominus [\![\phi]\!]_{B,\beta}(b) \leq d^{K_{\Lambda}}_{\alpha,\beta}(a, b).$$

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

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

Lemma 8.6 (Nonexpansiveness of quantitative modal logic)**.**
If Λ preserves nonexpansiveness w.r.t. a lax extension L, then $d^{\Lambda} \leq d^{L}$.

Proof. Immediate from Lemma 8.5 and Lemma 7.2. □

Finally, we show how the characterization of lax extensions as Kantorovich extensions can be used to define characteristic logics for nonexpansive lax extensions. Recall the sequence of approximants (Definition 4.10) we used in Theorem 4.12 to approach the L-behavioural distance $d^{L}_{\alpha,\beta}$ of coalgebras $\alpha: A \to TA$ and $\beta: B \to TB$ via fixpoint iteration:

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

If $L = K_{\Lambda}$, then each individual step in this iteration can be related to the logical distance taken over some subset of $\mathcal{L}_{\Lambda}$. More precisely, if we define the *rank* of a modal formula ϕ to be the maximal nesting depth of modalities, then

Lemma 8.7. *For each $n < \omega$ and all $a \in A$, $b \in B$ we have:*

$$d_n(a, b) = \sup\{[\![\phi]\!](a) \ominus [\![\phi]\!](b) \mid \phi \in \mathcal{L}_{\Lambda},\ \phi \textit{ has rank at most } n\}.$$

A proof for the more general case of quantale-valued logics and relations can be found in [WS21, Theorem 6.1]. In that paper, this characterization of finite-depth distances forms the basis of a Hennessy-Milner theorem for the quantale-valued Kantorovich lifting of finitary functors. In the present setting, we can drop the condition that T must be finitary by combining Lemma 8.7 with Theorem 4.12 to obtain, complementing Lemma 8.6, a criterion phrased directly in terms of conditions on the lax extension and the modalities:

Theorem 8.8 (Coalgebraic quantitative Hennessy-Milner theorem)**.** *Let L be a finitarily separable fuzzy lax extension, and let Λ be a separating set of monotone nonexpansive predicate liftings for L. Then $d^\Lambda = d^L$.*

Proof. By Lemma 8.7 we have $d_\omega = \sup_{n<\omega} d_n = d^\Lambda$ and by Theorem 4.12 we have $d_\omega = d^L$. □

Example 8.9.

(1) Since we only require L to be finitarily separable (rather than T finitary), Example 5.11.1 implies that we recover expressiveness [vBHMW07, vBW05] of quantitative probabilistic modal logic over countably branching discrete probabilistic transition systems (Example 8.3.2) as an instance of Theorem 8.8.
(2) Let L be the lax extension of $T = \mathcal{P}_\omega(M \times -)$ from Example 4.3. As T is finitary, it follows by Theorem 7.4 that $L = K_\Lambda$ for the set Λ of Moss liftings of L and the logic $\mathcal{L}_\Lambda$ is characteristic for simulation distance by Theorem 8.8.

Applying Lemma 8.6 and Theorem 8.8 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, we obtain expressive logics for finitarily separable nonexpansive lax extensions:

Corollary 8.10. *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 [Mos99] and its generalization to lax extensions [MV15]. 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

$$[\![\Delta\Phi]\!](a) = L\epsilon_A(\alpha(a), \Phi).$$

The dual of Δ is denoted ∇, and behaves like a quantitative analogue of Moss' two-valued ∇. From Corollary 8.10, it is immediate that this logic is expressive:

Corollary 8.11 (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 coalgebras $\alpha: A \to TA$, $\beta: B \to TB$,*

$$d^L_{\alpha,\beta}(a, b) = \sup\{[\![\phi]\!](a) \ominus [\![\phi]\!](b) \mid \phi \in \mathcal{L}_L\}.$$

Example 8.12.

(1) We equip the finite fuzzy powerset functor $\mathcal{F}_\omega$ with the Wasserstein lifting $W_\Diamond$ for $\Diamond$ as in Example 8.3.1, in analogy to the Hausdorff lifting (Example 6.8.2). 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 6.8.3. Then ∇ applies to finite sets of finite distributions on formulae, understood as spanning a convex polytope. By Corollary 8.11, the arising instance of quantitative Moss logic is expressive for all C_{fg}-coalgebras.

9. Conclusions

We study behavioural distances based on fuzzy lax extensions, with a particular focus on *nonexpansive lax extensions*, establishing that the latter are closely related to distances based on coalgebraic modal logic. Nonexpansiveness of a lax extension can equivalently be expressed in terms of strength of the underlying functor [Gav18] or as lax preservation of ϵ-diagonals. We examine two general constructions of nonexpansive lax extensions, respectively generalizing the classical Kantorovich and Wasserstein distances and strengthening previous generalizations where only pseudometrics are lifted [BBKK18]. Our construction of the Kantorovich lifting is based in particular on generalizing nonexpansive functions on a single space to nonexpansive pairs of functions on two different spaces (implicit in work on optimal transportation [Vil08]), while the Wasserstein lifting mostly coincides with an existing construction from work on topological theories [Hof07].

Our main result shows that every nonexpansive lax extension is a Kantorovich lifting for a suitable choice of modalities, the so-called Moss modalities. Moreover, one can extract from a given nonexpansive lax extension a characteristic modal logic satisfying a quantitative Hennessy-Milner property. Using our notion of *finitarily separable* lax extension additionally allows us to extend these constructions to certain non-finitary functors such as the discrete distribution functor. All our results apply both to symmetric behavioural distances, i.e. notions of quantitative bisimulation, and to asymmetric behavioural distances, i.e. notions of quantitative simulation.

References

[AHS90] 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.

[Bal00] Alexandru Baltag. A logic for coalgebraic simulation. In Horst Reichel, editor, *Coalgebraic Methods in Computer Science, CMCS 2000*, volume 33 of *ENTCS*, pages 42–60. Elsevier, 2000. doi:10.1016/S1571-0661(05)80343-3.

[Bar70] Micheal Barr. Relational algebras. In *Proc. Midwest Category Seminar*, volume 137 of *LNM*. Springer, 1970. doi:10.1007/bfb0060439.

[BBKK18] Paolo Baldan, Filippo Bonchi, Henning Kerstan, and Barbara König. Coalgebraic behavioral metrics. *Log. Methods Comput. Sci.*, 14(3), 2018. doi:10.23638/LMCS-14(3:20)2018.

[BdBH+91] Roland Carl Backhouse, Peter J. de Bruin, Paul F. Hoogendijk, Grant Malcolm, Ed Voermans, and Jaap van der Woude. Polynomial relators (extended abstract). In Maurice Nivat, Charles Rattray, Teodor Rus, and Giuseppe Scollo, editors, *Algebraic Methodology and Software Technology (AMAST '91), Proceedings of the Second International Conference on Methodology and Software Technology, Iowa City, USA, 22-25 May 1991*, Workshops in Computing, pages 303–326. Springer, 1991.

[BGP17] Borja Balle, Pascale Gourdeau, and Prakash Panangaden. Bisimulation metrics for weighted automata. In Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl, editors, *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.

[BSS17] Filippo Bonchi, Alexandra Silva, and Ana Sokolova. The power of convex algebras. In Roland Meyer and Uwe Nestmann, editors, *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.

[CGPX14] Konstantinos Chatzikokolakis, Daniel Gebler, Catuscia Palamidessi, and Lili Xu. Generalized bisimulation metrics. In Paolo Baldan and Daniele Gorla, editors, *Concurrency Theory, CONCUR 2014*, volume 8704 of *LNCS*, pages 32–46. Springer, 2014. doi:10.1007/978-3-662-44584-6.

[CGT16] Valentina Castiglioni, Daniel Gebler, and Simone Tini. Logical characterization of bisimulation metrics. In Mirco Tribastone and Herbert Wiklicky, editors, *Quantitative Aspects of Programming Languages and Systems, QAPL 2016*, volume 227 of *EPTCS*, pages 44–62, 2016. doi:10.4204/EPTCS.227.

[CHR12] Pavol Cerný, Thomas Henzinger, and Arjun Radhakrishna. Simulation distances. *Theor. Comput. Sci.*, 413(1):21–35, 2012. doi:10.1016/j.tcs.2011.08.002.

[Cîr06] Corina Cîrstea. A modular approach to defining and characterising notions of simulation. *Inf. Comput.*, 204(4):469–502, 2006. doi:10.1016/j.ic.2005.04.005.

[CKP+11] Corina Cîrstea, Alexander Kurz, Dirk Pattinson, Lutz Schröder, and Yde Venema. Modal logics are coalgebraic. *The Computer Journal*, 54(1):31–41, 2011. doi:10.1093/comjnl/bxp004.

[CNR16] Petr Cintula, Carles Noguera, and Jonas Rogger. From Kripke to neighborhood semantics for modal fuzzy logics. In João Paulo Carvalho, Marie-Jeanne Lesot, Uzay Kaymak, Susana M. Vieira, Bernadette Bouchon-Meunier, and Ronald R. Yager, editors, *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.

[CSWC13] 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.

[dAFS09] 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.

[DCPP06] Yuxin Deng, Tom Chothia, Catuscia Palamidessi, and Jun Pang. Metrics for action-labelled quantitative transition systems. In Antonio Cerone and Herbert Wiklicky, editors, *Quantitative Aspects of Programming Languages, QAPL 2005*, volume 153(2) of *ENTCS*, pages 79–96. Elsevier, 2006. doi:10.1016/j.entcs.2005.10.033.

[DDG16] Wenjie Du, Yuxin Deng, and Daniel Gebler. Behavioural pseudometrics for nondeterministic probabilistic systems. In Martin Fränzle, Deepak Kapur, and Naijun Zhan, editors, *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.

[DEP98] 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. doi:10.1109/LICS.1998.705681.

[Des99] Josée Desharnais. *Labelled Markov processes*. PhD thesis, McGill University, 1999.

[DGJP04] Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled Markov processes. *Theor. Comput. Sci.*, 318:323–354, 2004. doi:10.1016/j.tcs.2003.09.013.

[EKN12] Pantelis Eleftheriou, Costas Koutras, and Christos Nomikos. Notions of bisimulation for Heyting-valued modal languages. *J. Log. Comput.*, 22(2):213–235, 2012. doi:10.1093/logcom/exq005.

[Fan15] Tuan-Fang Fan. Fuzzy bisimulation for Gödel modal logic. *IEEE Trans. Fuzzy Syst.*, 23(6):2387–2396, 2015. doi:10.1109/TFUZZ.2015.2426724.

[Fit91] Melvin Fitting. Many-valued modal logics. *Fund. Inform.*, 15(3-4):235–254, 1991.

[FK97] Bob Flagg and Ralph Kopperman. Continuity spaces: Reconciling domains and metric spaces. *Theoretical Computer Science*, 177(1):111 – 138, 1997. doi:10.1016/S0304-3975(97)00236-3.

[FL14] 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.

[FLT11] Uli Fahrenberg, Axel Legay, and Claus Thrane. The quantitative linear-time–branching-time spectrum. In Supratik Chakraborty and Amit Kumar, editors, *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.

[FMS21] Chase Ford, Stefan Milius, and Lutz Schröder. Behavioural preorders via graded monads. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021, Rome, Italy, June 29 - July 2, 2021*, pages 1–13. IEEE, 2021. doi:10.1109/LICS52264.2021.9470517.

[FPP04] Norm Ferns, Prakash Panangaden, and Doina Precup. Metrics for finite Markov decision processes. In David Chickering and Joseph Halpern, editors, *Uncertainty in Artificial Intelligence, UAI 2004*, pages 162–169. AUAI Press, 2004.

[Gav18] Francesco Gavazzo. Quantitative behavioural reasoning for higher-order effectful programs: Applicative distances. In Anuj Dawar and Erich Grädel, editors, *Logic in Computer Science, LICS 2018*, pages 452–461. ACM, 2018. doi:10.1145/3209108.

[GJS90] Alessandro Giacalone, Chi-Chang Jou, and Scott Smolka. Algebraic reasoning for probabilistic concurrent systems. In Manfred Broy, editor, *Programming concepts and methods, PROCOMET 1990*, pages 443–458. North-Holland, 1990.

[HJ04] 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.

[HK97] Michael Huth and Marta Kwiatkowska. Quantitative analysis and model checking. In *Logic in Computer Science, LICS 1997*, pages 111–122. IEEE, 1997.

[HM85] M. Hennessy and R. Milner. Algebraic laws for non-determinism and concurrency. *J. ACM*, 32:137–161, 1985.

[Hof07] Dirk Hofmann. Topological theories and closed objects. *Adv. Math.*, 215(2):789 – 824, 2007. doi:10.1016/j.aim.2007.04.013.

[HST14] Dirk Hofmann, Gavin Seal, and Walter Tholen, editors. *Monoidal Topology: A Categorical Approach to Order, Metric, and Topology*. Cambridge University Press, 2014.

[KKV12] Krzysztof Kapulkin, Alexander Kurz, and Jiri Velebil. Expressiveness of positive coalgebraic logic. In Thomas Bolander, Torben Braüner, Silvio Ghilardi, and Lawrence Moss, editors, *Advances in Modal Logic, AiML 2012*, pages 368–385. College Publications, 2012.

[KL09] Alexander Kurz and Raul Leal. Equational coalgebraic logic. In Samson Abramsky, Michael Mislove, and Catuscia Palamidessi, editors, *Mathematical Foundations of Programming Semantics, MFPS 2009*, volume 249 of *ENTCS*, pages 333–356. Elsevier, 2009. doi:10.1016/j.entcs.2009.07.097.

[KM15] Narges Khakpour and Mohammad Mousavi. Notions of conformance testing for cyber-physical systems: Overview and roadmap (invited paper). In Luca Aceto and David de Frutos-Escrig, editors, *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.

[KMM18] 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.

[Law73] F William Lawvere. Metric spaces, generalized logic, and closed categories. *Rendiconti del seminario matématico e fisico di Milano*, 43(1):135–166, 1973.

[Lev11] Paul Levy. Similarity quotients as final coalgebras. In Martin Hofmann, editor, *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.

[LFT11] Kim Larsen, Uli Fahrenberg, and Claus Thrane. Metrics for weighted transition systems: Axiomatization and complexity. *Theor. Comput. Sci.*, 412(28):3358–3369, 2011. doi:10.1016/j.tcs.2011.04.003.

[LS91] K. Larsen and A. Skou. Bisimulation through probabilistic testing. *Inform. Comput.*, 94(1):1–28, 1991.

[LS08] Thomas Lukasiewicz and Umberto Straccia. Managing uncertainty and vagueness in description logics for the semantic web. *J. Web Sem.*, 6(4):291–308, 2008.

[Mil89] Robin Milner. *Communication and Concurrency*. Prentice Hall, 1989.

[MM97] 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.

[Mor79] Charles Morgan. Local and global operators and many-valued modal logics. *Notre Dame J. Formal Log.*, 20(2):401–411, 1979.

[Mos99] Lawrence Moss. Coalgebraic logic. *Ann. Pure Appl. Logic*, 96:277–317, 1999.

[MS17] Matteo Mio and Alex Simpson. Łukasiewicz μ-calculus. *Fund. Inf.*, 150(3–4):317–346, 2017.

[MV15] Johannes Marti and Yde Venema. Lax extensions of coalgebra functors and their logic. *J. Comput. Syst. Sci.*, 81(5):880–900, 2015.

[MV20] Matteo Mio and Valeria Vignudelli. Monads and quantitative equational theories for nondeterminism and probability. In Igor Konnov and Laura Kovács, editors, *Concurrency Theory,* *CONCUR 2020*, volume 171 of *LIPIcs*, pages 28:1–28:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.CONCUR.2020.28.

[Par81] David Park. Concurrency and automata on infinite sequences. In Peter Deussen, editor, *Theoretical Computer Science, 5th GI-Conference*, volume 104 of *LNCS*, pages 167–183. Springer, 1981. doi:10.1007/BFb0017288.

[Pat04] Dirk Pattinson. Expressive logics for coalgebras via terminal sequence induction. *Notre Dame J. Formal Log.*, 45:19–33, 2004.

[PLC15] Haiyu Pan, Yongming Li, and Yongzhi Cao. Lattice-valued simulations for quantitative transition systems. *Int. J. Approx. Reason.*, 56:28–42, 2015. doi:10.1016/j.ijar.2014.10.001.

[RG13] Ricardo Rodriguez and Lluis Godo. Modal uncertainty logics with fuzzy neighborhood semantics. In Lluis Godo, Henri Prade, and Guilin Qi, editors, *Weighted Logics for Artiticial Intelligence, WL4AI 2013 (Workshop at IJCAI 2013)*, pages 79–86, 2013.

[Rut00] Jan Rutten. Universal coalgebra: A theory of systems. *Theor. Comput. Sci.*, 249:3–80, 2000.

[Sch08] Lutz Schröder. Expressivity of coalgebraic modal logic: The limits and beyond. *Theor. Comput. Sci.*, 390:230–247, 2008.

[SP11] 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.

[Str98] Umberto Straccia. A fuzzy description logic. In Jack Mostow and Chuck Rich, editors, *Artificial Intelligence, AAAI 1998*, pages 594–599. AAAI Press / MIT Press, 1998. URL: http://www.aaai.org/Conferences/AAAI/aaai98.php.

[TFL10] Claus Thrane, Uli Fahrenberg, and Kim Larsen. Quantitative analysis of weighted transition systems. *J. Log. Algebraic Methods Program.*, 79(7):689–703, 2010. doi:10.1016/j.jlap.2010.07.010.

[Thi96] Albert Thijs. *Simulation and fixpoint semantics.* PhD thesis, University of Groningen, 1996.

[Trn80] Věra Trnková. General theory of relational automata. *Fund. Inform.*, 3(2):189–234, 1980.

[vBHMW07] 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.

[vBW05] Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *Theor. Comput. Sci.*, 331(1):115–142, 2005. doi:10.1016/j.tcs.2004.09.035.

[vG90] Rob van Glabbeek. The linear time-branching time spectrum (extended abstract). In Jos Baeten and Jan Willem Klop, editors, *Theories of Concurrency, CONCUR 1990*, volume 458 of *LNCS*, pages 278–297. Springer, 1990.

[Vil08] Cédric Villani. *Optimal Transport: Old and New.* Springer, 2008.

[Wal95] 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.

[WS20] Paul Wild and Lutz Schröder. Characteristic logics for behavioural metrics via fuzzy lax extensions. In Igor Konnov and Laura Kovács, editors, *Concurrency Theory, CONCUR 2020*, volume 171 of *LIPIcs*, pages 27:1–27:23. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.CONCUR.2020.27.

[WS21] Paul Wild and Lutz Schröder. A quantified coalgebraic van Benthem theorem. In Stefan Kiefer and Christine Tasson, editors, *Foundations of Software Science and Computation Structures - 24th International Conference, FOSSACS 2021, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2021, Luxembourg City, Luxembourg, March 27 - April 1, 2021, Proceedings*, volume 12650 of *Lecture Notes in Computer Science*, pages 551–571. Springer, 2021. doi:10.1007/978-3-030-71995-1_28.

[WSPK18] Paul Wild, Lutz Schröder, Dirk Pattinson, and Barbara König. A van Benthem theorem for fuzzy modal logic. In Anuj Dawar and Erich Grädel, editors, *Logic in Computer Science, LICS 2018*, pages 909–918. ACM, 2018. doi:10.1145/3209108.

This work is licensed under the Creative Commons Attribution License. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/ or send a letter to Creative Commons, 171 Second St, Suite 300, San Francisco, CA 94105, USA, or Eisenacher Strasse 2, 10777 Berlin, Germany