Coalgebraic Behavioral Metrics

By: Paolo Baldan, Filippo Bonchi, Henning Kerstan, and Barbara König

Johanna Maria Kirss
Matina Najafi
Wojtek Rozowski

The Adjoint School

April 17, 2023

Part I

Coalgebra

Transition systems

Transition systems are an ubiquitous object of study in theoretical computer science.

Name	Transition type	Notion of equivalence
Deterministic Finite Automata	$X \to 2 \times X^A$	Language equivalence
Markov chains	$X \to \mathcal{D}(X)$	Lumpability
Kripke Frames	$X \to \mathcal{P}_\omega(X) \times \mathcal{P}_\omega(AP)$	Zig-zag morphisms
Labelled Transition Systems	$X \to \mathcal{P}_\omega(X)^A$	(Strong) Bisimulation

All the examples above can be described as coalgebras for endofunctor F on the category Set.

Definition 1

Let $\mathcal{C}$ be a category and $\mathsf{F} : \mathcal{C} \to \mathcal{C}$ an endofunctor on $\mathcal{C}$. An F-coalgebra is a pair $(X, c : X \to \mathsf{F}X)$

Coalgebras on Set can represent state machines described above; X is a set of states, c the transition function, and F describes the type of the transitions performed.

Coalgebra homomorphisms

Definition 2

Given (X, c) and (Y, c'), two F-coalgebras on the category $\mathcal{C}$, an F-coalgebra homorphism from (X, c) to (Y, c') is an arrow $f : X \to Y$ such that $c' \circ f = \mathsf{F}f \circ c$.

Since identity map is a coalgebra homomorphism and composition of homomorphisms yields coalgebra homomorphisms then F-coalgebras along their homomorphisms form a category.

Definition 3

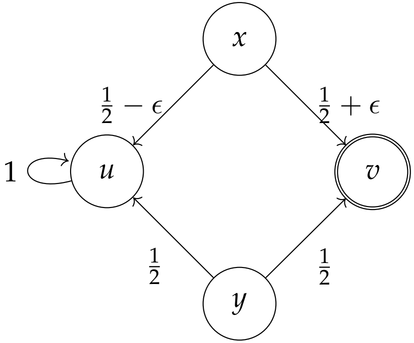
An F-coalgebra (Z, z) is final if for any other F-coalgebra (X, c) there exists a unique homomorphism $[\![\cdot]\!]_X : (X, c) \to (Z, z)$. We will omit the subscript if it is obvious from the context. We say that two elements are behaviourally equivalent if they are mapped by the final homomorphism into the same element of the carrier of the final coalgebra.

Example of the final coalgebra

Consider a Set endofunctor $\mathsf{F} = 2 \times (-)^A$. F-coalgebras are deterministic finite automata. The final coalgebra is given by 2^{A^*} the set of all languages for alphabet A. The map $\langle \epsilon?, (-)_a \rangle : 2^{A^*} \to \mathsf{F}2^{A^*}$. Given $L \in 2^{A^*}$ is called *semantic Brzozowski derivative* and is given by $\epsilon?(L) = [\epsilon \in L]$ and $L_a = \lambda w.\, L(aw)$. The final coalgebra map takes a state to the language it denotes.

More robust notion of equivalence

Consider the following probabilistic transition system (coalgebra with the transition of the type $X \to \mathcal{D}(X+1)$) where $\epsilon \in [0, \frac{1}{2}]$



States u and v are not behaviourally equivalent, x and y are behaviourally equivalent only if $\epsilon = 0$. What if ϵ is really small? Can we come up with a more robust notion of equivalence? The general problem seems to be able to go from distances on states to distances on successor distributions.

Part II

Transportation theory

Transportation Theory

For a set X, we define $d(x, y)$ for all $x, y \in X$ as distances between the elements such that $d : X \times X \to \mathbb{R}_0^+$.
If $P, Q \in \mathcal{D}(X)$ are the supply and demand functions for the elements in X, $d^{\mathcal{D}}(P, Q)$ is the distance on $\mathcal{D}(X)$ such that $d^{\mathcal{D}} : \mathcal{D}X \times \mathcal{D}X \to [0, 1]$ and $P, Q : X \to [0, 1]$.

We explore Wasserstein distance $d^{\downarrow\mathcal{D}}$ and Kantorovich distance $d^{\uparrow\mathcal{D}}$.
It can be proven that $d^{\uparrow\mathcal{D}}(P, Q) \leq d^{\downarrow\mathcal{D}}(P, Q)$.
If d is a pseudometric, then both of these distances are pseudometrics.

Wasserstein Distance

We define $t : X \times X \to [0,1]$ (transportation plan) where $t(x,y)$ is the percentage of what is transported from x to y.

- For any $x \in X$, $\sum_{y \in X} t(x,y) = P(x)$
- For any $y \in X$, $\sum_{x \in X} t(x,y) = Q(x)$

$T(P,Q)$ is the set of all possible transportation plans on X.
If $M = \sum_{x \in X} P(x) = \sum_{x \in X} Q(x)$, the total transportation cost:

$$c_d := M \cdot \sum_{x,y \in X} t(x,y) \cdot d(x,y)$$

then the distance between probability distributions would be:

$$d^{\downarrow \mathcal{D}}(P,Q) := \min\{\sum_{x,y \in X} t(x,y) \cdot d(x,y) | t \in T(P,Q)\}$$

Kantorovich Distance

We define a non-expansive price function $f : X \to \mathbb{R}_+$ such that $|f(x) - f(y)| \leq d(x, y)$. The set of all these functions is $C(d)$. Then the total profit would be:

$$g_d(f) := M \cdot \sum_{x \in X} f(x) \cdot (Q(x) - P(x))$$

The Kantorovich distance would be the maximum amount of g_d:

$$d^{\uparrow \mathcal{D}}(P, Q) := \max\{\sum_{x \in X} f(x) \cdot (Q(x) - P(x)) | f \in C(d)\}$$

Part III

Pseudometric spaces

(Pseudo)metric spaces

Definition 4

Let $\top \in]0, +\infty]$ be a fixed maximal element and X a set. We call a function $d : X \times X \to [0, \top]$ a $\top$-pseudometric on X if it satisfies:

- $d(x, x) = 0$ *(reflexivity)*
- $d(x, y) = d(y, x)$ *(symmetry)*
- $d(x, z) \leq d(x, y) + d(y, z)$ *(triangle inequality)*

for all $x, y, z \in X$. If additionally $d(x, y) = 0 \implies x = y$ then d is a $\top$-metric. A pseudometric space is a pair (X, d) where X is a set and d is a pseudometric on X.

Order structure

Pseudometrics on a given set form a poset given by pointwise order.

$$d \leq d' \iff \forall x, y \in X.\, d(x, y) \leq d'(x, y)$$

Order structure - continued

The set of pseudometrics on X, i.e., $D_x = \{d \mid d : X \times X \to [0, \top] \wedge d \text{ pseudometric}\}$ is a **complete lattice**. Joins are taken pointwise, and we can define meets in terms of joins.

Join

The join of a set $D \subseteq D_X$ is $(\sup(D))(x, y) = \sup_{d \in D}(d(x, y))$

Meet

The meet of a set $D \subseteq D_X$ is $\inf D = \sup\{d \mid d \in D_X \wedge \forall d' \in D.\ d \leq d'\}$

Pseudometrics categorically

Definition 5

A function $f : X \to Y$ between pseudometric spaces (X, d_x) and (Y, d_Y) is nonexpansive if $d_Y\big(f(x), f(y)\big) \leq d_X(x, y)$ for all $x, y \in X$ or in other words $d_Y \circ (f \times f) \leq d_X$. If equality holds, then f is an isometry.

Identity is nonexpansive and composition of nonexpansive functions yields nonexpansive functions. $\top$-pseudometrics and nonexpansive functions form a category PMet

PMet is bicomplete

PMet has all limits and colimits of small diagrams. The forgetful functor $\mathcal{U}$: PMet $\to$ Set takes them to limits and colimits in *Set*. Let $D : I \to$ PMet be a small diagram and define $(X_i, d_i) := D(i)$ for each object $i \in I$. Let $\{f_i : X \to X_i\}_{i \in I}$ be a limiting cone in Set for $\mathcal{U} \circ D$. (X, d) is a limit of D where $d = \sup_{i \in I}\{d_i \circ (f_i \times f_i)\}$.
In the colimit case, the underlying pseudometric is the supremum of the pseudometrics, which makes cocone maps nonexpansive.

Part IV

Examples

Probabilistic Transiton Systems

We are interested in coalgebras of the type $X \to \mathcal{D}(X+1)$. We set $\top = 1$.

- The *refusal functor* which takes (X, d_X) to $(X+1, d_{X+1})$ where $d_{X+1} : (X+1) \times (X+1) \to [0,1]$ is given by

$$d_{X+1}(x,y) = \begin{cases} c \cdot d_X(x,y) & x,y \in X \\ 0 & x = y = \checkmark \\ 1 & \text{otherwise} \end{cases}$$

- The following is a pseudometric on $\mathcal{D}(X+1)$:

$$d_{\mathcal{D}(X+1)}(P,Q) = d_{X+1}^{\uparrow\mathcal{D}} = \sup \left\{ \sum_{x \in X+1} |f(x) \cdot (P(x) - Q(x))| \;\middle|\; f : (X+1, d_{X+1}) \to ([0,1], d_e) \right\}.$$

- Given a coalgebra $(X, c : X \to \mathcal{D}(X+1))$ the *behavioural pseudometric for probabilistic transition systems* is the least solution to

$$d(x,y) = d_{\mathcal{D}(X+1)}(c(x), c(y)).$$

Deterministic automata

Deterministic automata are coalgebras for the functor $\mathsf{F} = 2 \times (-)^A$. Given a 1-pseudometric d on X we can lift it to pseudometric $d^{\downarrow \mathsf{F}}$ on $\mathsf{F}X$, such that for all $(o_1, s_1), (o_2, s_2) \in \mathsf{F}X$

$$d^{\downarrow \mathsf{F}}((o_1, s_1), (o_2, s_2)) = \max\{d_2(o_1, o_2), c \cdot \max_{a \in A} d(s_1(a), s_2(a))\}$$

The behavioural pseudometric for the deterministic automaton $(X, \langle o, s \rangle)$ is the least solution to the following fixpoint equation

$$d(x, y) = \max\{d_2(o(x), o(y)), c \cdot \max_{a \in A} d(s(x)(a), s(y)(a))\}$$

for states $x, y \in X$

Metric transition systems

- Let $\Sigma = \{r_1, \ldots, r_n\}$ be a finite set of *propositions* where each proposition $r \in \Sigma$ has an associated bounded metric space (M_r, d_r).
- A *valuation* is a map $u : \Sigma \to \cup_{r\in\Sigma} M_r$ such that $u(r) \in M_r$ for all $r \in \Sigma$. We can also think of valuation as a tuple $u : M_{r_1} \times \cdots \times M_{r_n}$.
- A *metric transition system* is a coalgebra of the type $X \to M_{r_1} \times \cdots \times M_{r_n} \times \mathcal{P}_\omega(X)$. We will again refer to this functor as F. We write $\pi : \mathsf{F}X \to M_{r_1} \times \cdots \times M_{r_n}$ to denote projection which gives the valuation associated with given state.
- A *directed propositional distance* between two valuations u, v is $\mathrm{pd}(u, v) = \max_{r\in R}\{d_r(u(r), v(r))\}$
- Let (X, c) be a F-coalgebra and let $s, t \in X$. The behavioural pseudometric $d_c : X \times X \to [0, \infty]$ is a solution of the following fixpoint equation

$$d_c(s,t) = \max\{\mathrm{pd}(\pi(c(s)), \pi(c(t))), \max_{s'\in c(s)} \min_{t'\in c(t)} d_c(s', t'), \max_{t'\in c(t)} \min_{s'\in c(s)} d_c(s', t')\}$$

Part V

Lifting Functors to Pseudometric Spaces

Preliminary: couplings

We will review a definition we will need later for defining the Wasserstein pseudometric.

Definition 6 (Coupling)

Let $F : \mathsf{Set} \to \mathsf{Set}$ be a functor and $n \in \mathbb{N}$. Given a set X and elements $t_i \in FX$ for $1 \leq i \leq n$, a coupling of the t_i with respect to F is an element $t \in F(X^n)$ such that $F\pi_i(t) = t_i$. We denote the set of all couplings of $t_1, \ldots, t_n$ w.r.t. F by $\Gamma_F(t_1, \ldots, t_n)$.

$$
\begin{array}{ccc}
X^n & \qquad & F(X^n) \ni t \\
\pi_i \big\downarrow & & \big\downarrow F\pi_i \\
X & & FX \ni t_i
\end{array}
$$

Lifting functors onto PMet

Definition 7

A functor $\bar{\mathsf{F}} : \mathsf{PMet} \to \mathsf{PMet}$ *is called a lifting of a functor* $F : \mathsf{Set} \to \mathsf{Set}$ *if the following diagram commutes:*

$$\begin{array}{ccc} \mathsf{PMet} & \xrightarrow{\overline{F}} & \mathsf{PMet} \\ {\scriptstyle U}\downarrow & & \downarrow{\scriptstyle U} \\ \mathsf{Set} & \xrightarrow{F} & \mathsf{Set}. \end{array}$$

Denote d^F *for the pseudometric obtained by applying the lifting of* F *to the pseudometric space* (X, d)*. In other words, we denote*

$$\overline{F}(X, d) := (FX, d^F).$$

Notes:

- The lifting is monotone, i.e. $d_1 \leq d_2$ implies $d_1^F \leq d_2^F$. No further conditions needed!
- We will use evaluation functions $ev_F : F[0, \top] \to [0, \top]$ and evaluation functors $\tilde{F} : \mathsf{Set}/[0, \top] \to \mathsf{Set}/[0, \top]$ given by:

$$\begin{array}{ccc} FX & \xrightarrow{Ff} & F[0,\top] \\ & {\scriptstyle \tilde{F}f}\searrow & \downarrow{\scriptstyle ev_F} \\ & & [0,\top] \end{array} \qquad \begin{array}{ccc} X & \xrightarrow{\varphi} & Y \\ {\scriptstyle f}\searrow & & \swarrow{\scriptstyle g} \\ & [0,\top] & \end{array} \qquad \begin{array}{ccc} FX & \xrightarrow{F\varphi} & FY \\ {\scriptstyle \tilde{F}f}\searrow & & \swarrow{\scriptstyle \tilde{F}g} \\ & [0,\top] & \end{array}$$

Kantorovich lifting

To define a *Kantorovich lifting*, we fix

- a functor $F : \mathsf{Set} \to \mathsf{Set}$,
- an evaluation function ev_F.

Then, given (X, d), we define the *Kantorovich distance* $d^{\uparrow F}$ on FX as

$$d^{\uparrow F}(t_1, t_2) := \sup\{\, d_e\left(\tilde{F}f(t_1), \tilde{F}f(t_2)\right) \mid f : (X, d) \to ([0, \top], d_e) \,\}$$

for all $t_1, t_2 \in FX$ and the *Kantorovich lifting* $\overline{F} : \mathsf{PMet} \to \mathsf{PMet}$ as the functor

- $\overline{F}(X, d) = (FX, d^{\uparrow F})$,
- $\overline{F}(f) = Ff$

for all $(X, d) \in \mathrm{Ob}(\mathsf{PMet})$ and $f : (X, d_X) \to (Y, d_Y)$ nonexpansive.

Intuition: the smallest distance d^F making all the functions $\tilde{F}f : (FX, d^F) \to ([0, \top], d_e)$ nonexpansive.

We can also check that $d^{\uparrow F}$ is a pseudometric, $\overline{F}$ is a functor, and $\overline{F}$ **preserves isometries**.

KANTOROVICH LIFTING OF THE DISTRIBUTION FUNCTOR

Consider:

- $\mathcal{D} : \mathsf{Set} \to \mathsf{Set}$, the discrete probability distribution functor

$$\mathcal{D}X = \{P : X \to [0,1] \mid \sum_{x \in X} P(x) = 1\}, \quad \mathcal{D}f(P)(y) = \sum_{x \in f^{-1}(y)} P(x),$$

- $\top = 1$,
- $ev_{\mathcal{D}} : \mathcal{D}[0,1] \to [0,1]$ as the expected value function

$$ev_{\mathcal{D}}(P) = \mathbb{E}(P) = \sum_{t \in [0,1]} t \cdot P(t).$$

Then if $g : X \to [0,1]$ is a nonexpansive function then

$$\tilde{D}g(P) = (ev_{\mathcal{D}} \circ Dg)(P) = \sum_{x \in X} g(x)P(x)$$

and so the Kantorovich distance between two discrete probability distributions P_1, P_2 is

$$d^{\uparrow \mathcal{D}}(P_1, P_2) = \sup \left\{ |\sum_{x \in X} f(x)(P_1(x) - P_2(x)) \mid f : (X, d) \to ([0,1], d_e) \right\}.$$

Note: we can also analogously define the Kantorovich distance for subdistributions, which will be helpful in the case of purely probabilistic systems.

IDEA OF THE WASSERSTEIN LIFTING

We wish to define a *Wasserstein lifting*, fixing again

- a functor $F : \mathsf{Set} \to \mathsf{Set}$,
- an evaluation function ev_F.

In particular, given (X, d), we want to define the *Wasserstein distance* $d^{\downarrow F}$ on FX as

$$d^{\downarrow F}(t_1, t_2) := \inf\{\, \tilde{F}d(t) \mid t \in \Gamma_F(t_1, t_2),\, \}$$

where $\Gamma_F(t_1, t_2)$ is the set of all *couplings* of t_1 and t_2.

Then the *Wasserstein lifting* $\overline{F} : \mathsf{PMet} \to \mathsf{PMet}$ would be defined as the functor

- $\overline{F}(X, d) = (FX, d^{\downarrow F})$,
- $\overline{F}(f) = Ff$

for all $(X, d) \in \mathrm{Ob}(\mathsf{PMet})$ and $f : (X, d_X) \to (Y, d_Y)$ nonexpansive.

Issues:

- couplings of $t_1, t_2 \in FX$ might not exist; hence it may be that

$$d^{\downarrow F}(t_1, t_1) = \inf \varnothing = \top \neq 0$$

- the triangle inequality might not be fulfilled.

We do have that the Wasserstein distance is symmetric.

Work towards the Wasserstein lifting

It turns out restricting our attention to only particular functors $F : \mathsf{Set} \to \mathsf{Set}$ and evaluation functions ev_F guarantees that $d^{\downarrow F}$ is a pseudometric.

Specifically, we wish that

- F preserves weak pullbacks,
- ev_F is well-behaved.

Definition 8

We call an evaluation function ev_F well-behaved if it satisfies the following conditions:

- *W1. $\tilde{F}$ is monotone, i.e. for $f, g : X \to [0, \top]$ with $f \leq g$, we have $\tilde{F}f \leq \tilde{F}g$,*
- *W2. for any $t \in F([0, \top]^2)$ we have $d_e(ev_F(t_1), ev_F(t_2)) \leq \tilde{F}d_e(t)$ for $t_i := F\pi_i(t)$,*
- *W3. $ev_F^{-1}[\{0\}] = Fi[F(\{0\}]$, where $i : \{0\} \hookrightarrow [0, \top]$ is the inclusion map.*

It can be shown that if ev_F fulfills condition $W3$, the Wasserstein distance is reflexive, and if F preserves weak pullbacks and ev_F fulfills conditions $W1, W2$, it also satisfies the triangle inequality:

- $W3 \implies$ reflexivity (lemma 5.15),
- $W1, W2, F$ preserves weak pullbacks $\implies$ triangle inequality (gluing lemma, lemma 5.19).

WASSERSTEIN LIFTING

With these conditions in place, the Wasserstein distance

$$d^{\downarrow F}(t_1, t_2) := \inf\{\, \tilde{F}d(t) \mid t \in \Gamma_F(t_1, t_2), \,\}$$

is indeed a pseudometric.

Theorem 1 (Wasserstein pseudometric)

Let F be an endofunctor on Set *with evaluation function ev_F. If F preserves weak pullbacks and ev_F is well-behaved then for any pseudometric space (X, d), the Wasserstein distance $d^{\downarrow F}$ is a pseudometric.*

Now we can indeed define the *Wasserstein lifting* $\overline{F}$: PMet $\to$ PMet as the functor

- $\overline{F}(X, d) = (FX, d^{\downarrow F})$,
- $\overline{F}(f) = Ff$

for all $(X, d) \in \mathrm{Ob}(\mathsf{PMet})$ and $f : (X, d_X) \to (Y, d_Y)$ nonexpansive.

It can be shown that $\overline{F}$ is a functor, and that it also **preserves isometries**. Under certain conditions w.r.t. optimal couplings, if (X, d) is a metric space, so is $(FX, d^{\downarrow F})$.

Corollary 1 (Preservation of metrics)

Let (X, d) be a metric space. If the infimum in the Wasserstein distance is always a minimum then $d^{\downarrow F}$ is a metric and thus $\overline{F}(X, d) = (FX, d^{\downarrow F})$ is a metric space.

Comparison of Kantorovich and Wasserstein

Theorem 2

Let F be an endofunctor on Set*. If ev_F satisfies Conditions W1 and W2 of the definition of a well-behaved evaluation function, then for all pseudometric spaces (X, d) we have $d^{\uparrow F} \leq d^{\downarrow F}$.*

The inequality can be

- **strict:** for S the squaring functor $X \to X \times X$, $ev_S(r_1, r_2) = r_1 + r_2$ the evaluation function, (X, d) a metric space and elements $t_1 := (x_1, x_2)$ and $t_2 := (x_1, x_2)$, $x_1 \neq x_2$ in $X \times X$, we have that

$$d^{\uparrow F}(t_1, t_2) = 0, \quad d^{\downarrow F}(t_1, t_2) = 2d(x_1, x_2) > 0.$$

- **an equality:** for $\mathcal{D}$ the distribution functor and $ev_{\mathcal{D}} : \mathcal{D}[0,1] \to [0,1]$ the expected value function and (X, d) a pseudometric space, the Wasserstein distance of two probability distributions $P_1, P_2 \in \mathcal{D}X$ is

$$d^{\downarrow \mathcal{D}}(P_1, P_2) = \inf \left\{ \sum_{x_1, x_2 \in X} d(x_1, x_2) \cdot P(x_1, x_2) \mid P \in \Gamma_{\mathcal{D}}(P_1, P_2) \right\},$$

  which coincides with the Kantorovich pseudometric by the Kantorovich-Rubinstein duality from transportation theory.

Machine functor example

- For a finite set A, define the *input functor* $-^A : \mathsf{Set} \to \mathsf{Set}$ by

$$X^A = \{f : A \to X \text{ functions}\}, \quad f^A : X^A \to Y^A,\ f^A(g) = f \circ g.$$

- Then (well-behaved) evaluation functions $[0,\top]^A \to [0,\top]$ and Wasserstein distances can be defined as follows:

$\top$	$ev_I(s)$	$d^{\downarrow I}(s_1, s_2)$
$\top \in]0,\infty]$	$\max_{a\in A} s(a)$	$\max_{a\in A} d(s_1(a), s_2(a))$
$\top = \infty$	$\sum_{a\in A} s(a)$	$\sum_{a\in A} d(s_1(a), s_2(a))$
$\top \in]0,\infty[$	$\lvert A\rvert^{-1} \sum_{a\in A} s(a)$	$\lvert A\rvert^{-1} \sum_{a\in A} d(s_1(a), s_2(a))$

- Define the *machine functor* $M_B : \mathsf{Set} \to \mathsf{Set}$ via the input functor as follows:

$$M_B(X) = B \times X^A, \quad M_B(f) = id_B \times f^A.$$

  Pick an evaluation function ev_I from above and define the evaluation function for M_B by

$$ev_{M_B} : B \times [0,\top]^A \to [0,\top], \quad ev_{M_B}(o, s) = c \cdot ev_I(s)$$

  for some $c \in (0, 1]$.

- Then the Wasserstein distance is 1 for states of different type (given that $M_B(\pi) = \mathbf{id_B} \times \pi^A$) and $c \cdot ev_I(d \circ \langle s_1, s_2\rangle)$ for the same type.

Definition 9

The product bifunctor is the bifunctor $F : Set^2 \to Set$ where:

- $F(X_1, X_2) = X_1 \times X_2$
- $F(f_1, f_2) = f_1 \times f_2, f_i : X_i \to Y_i$

$$\begin{array}{ccc} X_1 \times X_2 & \longrightarrow & X_2 \\ {\scriptstyle f_1 \times f_2}\Big\downarrow & & \Big\downarrow{\scriptstyle f_2} \\ Y_1 \times Y_2 & \longrightarrow & Y_2 \end{array}$$

This bifunctor preserves pullbacks.

Lemma 1

These evaluation functions $ev_F : [0, \top]^2 \to [0, \top]$ are well-behaved:

$\top$	*other parameters*	$ev_F(r_1, r_2)$
$\top \in]0, \infty]$	$c_1, c_2 \in]0, 1]$	$\max\{c_1 r_1, c_2 r_2\}$
$\top = \infty$	$c_1, c_2 \in]0, \infty[, p \in \mathbb{N}$	$(c_1 x_1^p + c_2 x_2^p)^{1/p}$
$\top \in]0, \infty[$	$c_1, c_2 \in]0, 1], c_1 + c_2 \leq 1, p \in \mathbb{N}$	$(c_1 x_1^p + c_2 x_2^p)^{1/p}$

Product and Coproduct Bifunctors
Product Bifunctor - continued

Lemma 2

The lifted pseudometrics $(d_1, d_2)^F : (X \times X)^2 \to [0, \top]$*:*

$ev_F(r_1, r_2)$	$(d_1, d_2)^F\big((x_1, x_2), (y_1, y_2)\big)$
$\max\{c_1 r_1, c_2 r_2\}$	$\max\{c_1 d_1(x_1, y_1), c_2 d_2(x_2, y_2)\}$
$(c_1 x_1^p + c_2 x_2^p)^{1/p}$	$\big(c_1 d_1(x_1, y_1)^p + c_2 d_2(x_2, y_2)^p\big)^{1/p}$

Lemma 3

If $c_1 = c_2 = 1$ *for the first evaluation function, given two pseudometric spaces* X_1, d_1 *and* X_2, d_2*, we obtain the lifted pseudometric* $d_\infty : (X_1 \times X_2)^2 \to [0, \top]$*, with* $d_\infty\big((x_1, x_2), (y_1, y_2)\big) = \max\{d_1(x_1, y_1), d_2(x_2, y_2)\}$*.* $(X_1 \times X_2, d_\infty) = (X_1, d_1) \times (X_2, d_2)$*.*

Product and Coproduct Bifunctors

Coproduct Bifunctor

Definition 10

The coproduct bifunctor is the bifunctor $F : Set^2 \to Set$, where:

- $F(X_1, X_2) = X_1 + X_2$
- $F(f_1, f_2) = f_1 + f_2, f_i : X_i \to Y_i$ *such that* $f_1 + f_2 : X_1 + X_2 \to Y_1 + Y_2$, $(f_1 + f_2)(x, i) = \big(f_i(x), i\big)$

This bifunctor preserves pullbacks.

Lemma 4

The evaluation function $ev_F : [0, \top] + [0, \top] \to [0, \top]$ *where* $ev_F(x, i) = x$ *is well-behaved.*

Lemma 5

The lifted pseudometric $d_+ : (X_1 + X_2)^2 \to [0, \top]$ *for the evaluation function mentioned in the previous lemma would be:*

$$d_+\big((x_1, i_1), (x_2, i_2)\big) = \begin{cases} d_i(x_1, x_2), & \textit{if } i_1 = i_2 = i \\ \top, & \textit{else} \end{cases}$$

Lemma 6

$(X_1 + X_2, d_+) = (X_1, d_1) + (X_2, d_2)$

Part VI

Bisimilarity pseudometrics

Lifting coalgebras to PMet

Let $\mathsf{F} : \mathsf{Set} \to \mathsf{Set}$ be an endofunctor on Set, which has a lifting $\bar{\mathsf{F}} : \mathsf{PMet} \to \mathsf{PMet}$. We will write $d^{\mathsf{F}} : \mathsf{F}X \times \mathsf{F}X \to [0, \top]$ for the pseudometric on $\mathsf{F}X$ obtained by applying $\bar{F}$ to some arbitrary pseudometric space (X, d).

High level idea

In general we would like to turn F-coalgebra $(X, c : X \to \mathsf{F}X)$ into an $\bar{\mathsf{F}}$-coalgebra $((X, d_c), c)$ such that $c : (X, d_c) \to (\mathsf{F}X, d_c^{\mathsf{F}})$ is nonexpansive. Observe that given any pseudometric $d : X \times X \to [0, \top]$ the assignment $d \mapsto d^{\mathsf{F}} \circ (c \times c)$ is a *monotone endomap* on the set of pseudometrics on X. By Knaster-Tarski fixpoint theorem this mapping has a least fixed point, given by

$$d_c = \inf\{d \mid d : X \times X \to [0, \top] \text{ pseudometric } \wedge d^{\mathsf{F}} \circ (c \times c) \leq d\}$$

Since $d_c = d_c^F \circ (c \times c)$, we have that c is an isometry and hence is a nonexpansive function. Now, we need to verify that homomorphisms between F-coalgebras yield homomorphisms between such defined $\bar{F}$-coalgebras.

Lifting coalgebras to PMet - continued

Lemma 7

Let (X, c) and (Y, c') be F-coalgebras. If $f : X \to Y$ is an F-coalgebra homomorphism, then $f : (X, d_c) \to (Y, d'_c)$ is nonexpansive (with d_c and $d_{c'}$ defined via fixpoint construction from the previous slide). Additionally, it is an isometry if $\bar{\mathsf{F}}$ preserves isometries.

Lemma 8

If (Z, z) is the final F-coalgebra, then $((Z, d_z), z)$ is the final $\bar{\mathsf{F}}$-coalgebra.

Definition 11

We will use $[\![.]\!]_X : X \to Z$ to denote a unique homorphism to the final coalgebra and will omit the subscript when X is obvious from the context. Given a F-coalgebra (X, c) we can define the behavioural distance to be the pseudometric $bd_c(x, y) := d_z([\![x]\!], [\![y]\!])$ for all $x, y \in X$. Observe that if $\bar{\mathsf{F}}$ preserves isometries, then $d_c = bd_c$

Behaviourally equivalent states are in zero distance

Observe that if $[\![x]\!] = [\![y]\!]$, then $bd_c([\![x]\!], [\![y]\!]) = 0$. Follows from reflexivity of d_z. The converse requires an extra condition on $\bar{\mathsf{F}}$ and we will discuss it later.

Example with deterministic automata

We are interested in coalgebras for the functor $\mathsf{F} = 2 \times (-)^A$. Final F-coalgebra is given by the set 2^{A^*} equipped with the semantic Brzozowski derivative. Consider an evaluation function:

$$\mathrm{ev}_{\mathsf{F}} : 2 \times [0,1]^A \to [0,1] \qquad \mathrm{ev}_{\mathsf{F}}(\langle o, s \rangle) = c \cdot \max_{a \in A} s(a) \qquad \text{for } 0 < c < 1$$

Given a 1-pseudometric d on 2^{A^*} the Wasserstein lifting yields the pseudometric $d^{\downarrow \mathsf{F}}$ on $\mathsf{F}2^{A^*}$, such that for all $(o_1, s_1), (o_2, s_2) \in \mathsf{F}2^{A^*}$

$$d^{\downarrow \mathsf{F}}((o_1, s_1), (o_2, s_2)) = \max\{d_2(o_1, o_2), c \cdot \max_{a \in A} d(s_1(a), s_2(a))\}$$

We write $d_2 : 2 \times 2 \to [0,1]$ for a discrete pseudometric. We want the semantic Brzozowski derivative to be an isometry, so we are looking for a least solution to the following fixpoint equation

$$d(L_1, L_2) = \max\{d_2(L_1(\epsilon), L_2(\epsilon)), c \cdot \max_{a \in A} d(\lambda w.\, L_1(aw), \lambda w.\, L_2(aw))\}$$

for languages $L_1, L_2 \in 2^{A^*}$.

Example with deterministic automata - continued

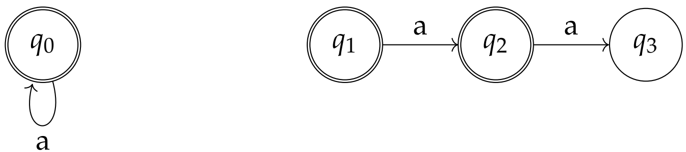
The least solution to that fixpoint equation is the pseudometric $d_{2^{A^*}} : 2^{A^*} \times 2^{A^*} \to [0,1]$ given by

$$d_{2^{A^*}}(L_1, L_2) = c^{\inf\{n \in \mathbb{N} \mid \exists w \in A^n.\ L_1(w) \neq L_2(w)\}}$$

In other words, this is c to the power of the length of the least prefix which distinguishes both languages.

Behavioural distance example

Consider following deterministic automata. Let $c = \frac{1}{2}$



$$d(q_0, q_1) = \frac{1}{4} \qquad d(q_0, q_2) = \frac{1}{2} \qquad d(q_0, q_3) = 1$$

Final chain construction

Definition 12 (Worrell 2005)

Let $\mathcal{C}$ be a category with a terminal object 1 *and limits of ordinal indexed cochains and* $\mathsf{F} : \mathcal{C} \to \mathcal{C}$ *an endofunctor on* $\mathcal{C}$*. The final sequence is an ordinal indexed sequence of objects* $\langle W_i \rangle$ *with maps* $\{p_{i,j} : W_j \to W_i\}_{i \leq j}$ *defined uniquely by the following*

- $W_{i+1} = \mathsf{F} W_i$
- $p_{i+1,j+1} = \mathsf{F} p_{i,j}$
- $\forall i \leq j \leq k.\ p_{i,k} = p_{i,j} \circ p_{j,k}$
- *If* j *is limit ordinal then the cone* $\{p_{i,j} : W_j \to W_i\}_{i \leq j}$ *is a limiting cone.*

Lemma 9 (Adamek and Koubek 1995)

If $p_{j,j+1}$ *is an isomorphism for some ordinal* j*, then* $(W_j, (p_{j,j+1})^{-1})$ *is the final* F*-coalgebra*

Lemma 10 (Worrell 2005)

We can extend F*-coalgebra* (X, c) *to a cone* $\{c_i : X \to W_i\}_i$ *over the final sequence such that* $c_{i+1} = \mathsf{F} c_i \circ c$

We prove the above by transfinite induction and show that $\forall i \leq j.\ p_{i,j} \circ c_j = c_i$. Not too hard to show that $c_j : X \to W_j$ is the final F-coalgebra homomorphism, if the final chain stabilises at j.

Final coalgebra metric

Theorem 3

Let (Z, z) be the final F-coalgebra. If $\bar{\mathsf{F}} : \mathsf{PMet} \to \mathsf{PMet}$ preserves metrics and final chain for F stabilises, then (Z, d_z) (with d_z defined as before) is a metric space.

Proof sketch.

- Endow every element of the final chain with a metric.
  - There is only one metric on the terminal object (the discrete metric).
  - If (W_j, d_j) is a metric space, then $W_{j+1} = (\mathsf{F}W_j, d_{j+1} = d_j^{\mathsf{F}})$ is also a metric space.
  - In the limit case, we need to show that $\sup_{i \leq j}(d_i \circ (p_{i,j} \times p_{i,j}))$ is a metric. If two things are in zero distance, then they agree at all $p_{i,j}$ which are jointly monic, so they must be equal.
- Show that given a F-coalgebra (X, c), the maps of the induced cone $\{c_i : X \to W_i\}_i$ are nonexpansive.
- If the final chain stabilises at j, $z_j : Z \to W_j$ is a final $\bar{\mathsf{F}}$-coalgebra homomorphism.
- If $x, y \in Z$ are in distance zero in Z, then their images under z_j are also in distance zero in W_j and hence $z_j(x) = z_j(y)$. Since, z_j is an iso, we have that $x = y$.

Bisimilarity metric for Probabilistic Transiton Systems

We are interested in coalgebras of the type $X \to \mathcal{D}(X+1)$. We set $\top = 1$.

- First, consider an identity functor with an evaluation map $\mathrm{ev}_{\mathsf{Id}} : [0,1] \to [0,1]$ given by $\mathrm{ev}_{\mathsf{Id}}(x) : c \cdot x$ for $0 < c < 1$.
- Then, take a coproduct with a discrete metric on 1. We obtain the *refusal functor* which takes (X, d_X) to $(1+X, d_{1+X})$ where $d_{X+1} : (1+X) \times (1+X) \to [0,1]$ is given by

$$d_{X+1}(x,y) = \begin{cases} c \cdot d_X(x,y) & x, y \in X \\ 0 & x = y = \checkmark \\ 1 & \text{otherwise} \end{cases}$$

- Take the *expected value evaluation function* $\mathrm{ev}_{\mathcal{D}} : \mathcal{D}[0,1] \to [0,1]$ given by $\mathrm{ev}_{\mathcal{D}}(P) = \sum_{x \in [0,1]} x \cdot P(x)$

Bisimilarity metric for Probabilistic Transition Systems - continued

- Now, we can take the Kantorovich lifting to obtain the following pseudometric on $\mathcal{D}(X+1)$

$$d_{\mathcal{D}(X+1)}(P,Q) = d_{X+1}^{\uparrow\mathcal{D}} = \sup\left\{ \sum_{x \in X+1} |f(x) \cdot (P(x) - Q(x))| \;\; | \, f : (X+1, d_{X+1}) \to ([0,1], d_e) \right\}$$

- We have Kantorovich-Rubinstein duality and preservation of metrics and isometries.
- Therefore, given a coalgebra $(X, c : X \to \mathcal{D}(X+1))$ the least solution to

$$d(x,y) = d_{\mathcal{D}(X+1)}(c(x), c(y))$$

is a *behavioural metric*. In such a way we have obtained a discrete version of the behavioural metric of Breugel and Worrell 2006

Metric transition systems

- In this example we set $\top = \infty$
- Let $\Sigma = \{r_1, \ldots, r_n\}$ be a finite set of *propositions* where each proposition $r \in \Sigma$ has an associated bounded metric space (M_r, d_r).
- A *valuation* is a map $u : \Sigma \to \cup_{r \in \Sigma} M_r$ such that $u(r) \in M_r$ for all $r \in \Sigma$. We can also think of valuation as a tuple $u : M_{r_1} \times \cdots \times M_{r_n}$.
- A *metric transition system* is a coalgebra of the type $X \to M_{r_1} \times \cdots \times M_{r_n} \times \mathcal{P}_\omega(X)$. We will again refer to this functor as F.
- A *directed propositional distance* between two valuations u, v is $\mathrm{pd}(u, v) = \max_{r \in R}\{d_r(u(r), v(r))\}$

Metric transition systems example - continued

- We can lift any pseudometric (X, d) to a metric $(\mathsf{F}X, d_F)$ combining the Hausdorff metric on $\mathcal{P}_\omega$ obtained using evaluation function $\max : \mathcal{P}_\omega[0, \infty]$ and the propositional distance. In the end we have

$$d_F(P, Q) = \max\{\mathrm{pd}(\pi(P), \pi(Q)), \max_{s' \in S} \min_{t' \in T} d(s', t'), \max_{t' \in T} \min_{s' \in S} d(s', t')\}$$

  We used $\pi : \mathsf{F}X \to M_{r_1} \times \cdots \times M_{r_n}$ to denote projection which gives the valuation associated with given state.
- Let (X, c) be a F-coalgebra and let $s, t \in X$. The behavioural pseudometric $d_c : X \times X \to [0, \infty]$ is a solution of the following fixpoint equation

$$d_c(s, t) = \max\{\mathrm{pd}(\pi(c(s)), \pi(c(t))), \max_{s' \in c(s)} \min_{t' \in c(t)} d_c(s', t'), \max_{t' \in c(t)} \min_{s' \in c(s)} d_c(s', t')\}$$

References I

Adamek, Jiri and Vaclav Koubek (1995). "On the Greatest Fixed Point of a Set Functor". In: *Theor. Comput. Sci.* 150.1, pp. 57–75. DOI: 10.1016/0304-3975(95)00011-K. URL: https://doi.org/10.1016/0304-3975(95)00011-K.

Breugel, Franck van and James Worrell (2006). "Approximating and computing behavioural distances in probabilistic transition systems". In: *Theor. Comput. Sci.* 360.1-3, pp. 373–385. DOI: 10.1016/j.tcs.2006.05.021. URL: https://doi.org/10.1016/j.tcs.2006.05.021.

Worrell, James (2005). "On the final sequence of a finitary set functor". In: *Theor. Comput. Sci.* 338.1-3, pp. 184–199. DOI: 10.1016/j.tcs.2004.12.009. URL: https://doi.org/10.1016/j.tcs.2004.12.009.