# Behavioural Metrics – A Coalgebraic Approach

Barbara König

Universität Duisburg-Essen, Germany

Joint work with Paolo Baldan, Filippo Bonchi, Henning Kerstan

# Overview

# Behavioural Equivalences

Behavioural equivalences (bisimilarity, trace equivalence, ...) relate states with the same behaviour

## Applications

- Comparing a system with its specification
- Minimizing the state space
- Analysis of model transformations
- Verification of cryptographic protocols (are two protocols equivalent from the point of view of an external observer, a.k.a. the attacker?)

# Behavioural Metrics

Finding a quantitative notion of behavioural equivalence ...

- Do not insist on the exact same behaviour.
- Measure the behavioural distance between two states.
- Make statements such as "the behaviour of two states differs only by $\varepsilon$".

$\rightsquigarrow$ behavioural metrics

# Behavioural Metrics

## Pseudo-metric space

Let $X$ be a set, $\mathbb{R}_0^\infty = \mathbb{R}_0 \cup \{\infty\}$. A pseudo-metric is a function $d\colon X \times X \to \mathbb{R}_0^\infty$ where for all $x, y, z \in X$:

1. $d(x,x) = 0$ (identity) (metric if $(d(x,y) = 0 \Rightarrow x = y)$)
2. $d(x,y) = d(y,x)$ (symmetry)
3. $d(x,z) \le d(x,y) + d(y,z)$ (triangle inequality)

A (pseudo-)metric space is a pair $(X,d)$ where $X$ is a set and $d$ is a (pseudo-)metric on $X$.

## Non-expansive function

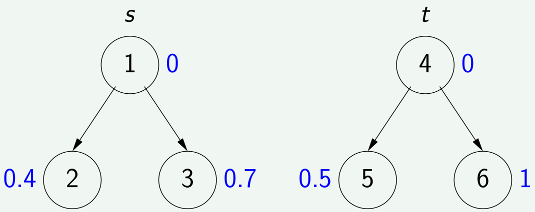
A non-expansive function $f\colon X \to Y$ between two (pseudo-)metric spaces $(X, d_X), (Y, d_Y)$ satisfies for $x, y \in X$

$$d_X(x,y) \ge d_Y(f(x), f(y))$$

# Metric Transition Systems

**Metric transition system [de Alfaro et al., 2009] (slightly simplified)**

Let $(X, d_r)$ be a metric space. A metric transition system is a tuple $M = (S, \tau, [\cdot])$, where $S$ is a set of states, $\tau \subseteq S \times S$ is a transition relation and every state $s$ is assigned an element $[s] \in X$.




Metric space $X = [0, 1]$ with Euclidean metric.

# Metric Transition Systems

**Hausdorff metric (metric on finite sets)**

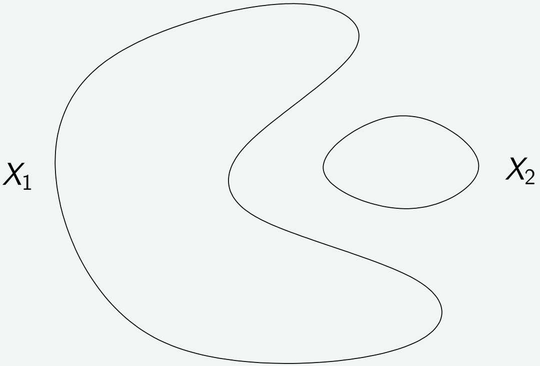
Lifting a metric space $(X, d)$ to $(\mathcal{P}_{fin}(X), d')$: for $X_1, X_2 \subseteq X$:

$$d^H(X_1, X_2) = \max\{ \max_{x \in X_1} \min_{y \in X_2} d(x, y), \max_{y \in X_2} \min_{x \in X_1} d(x, y) \}$$

- For each element $x$ (in $X_1, X_2$) take the closest element $y$ in the other set and measure the distance $d(x, y)$
- Take the maximum of all such distances.

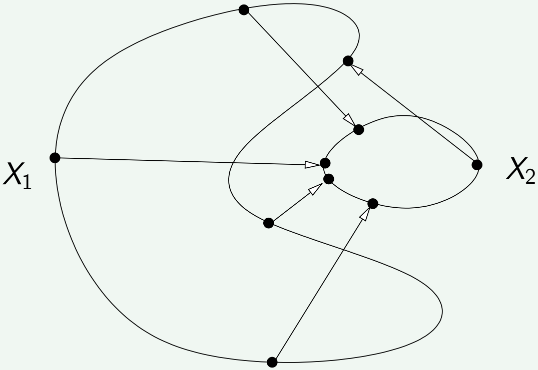
# Metric Transition Systems

## Example: Hausdorff metric



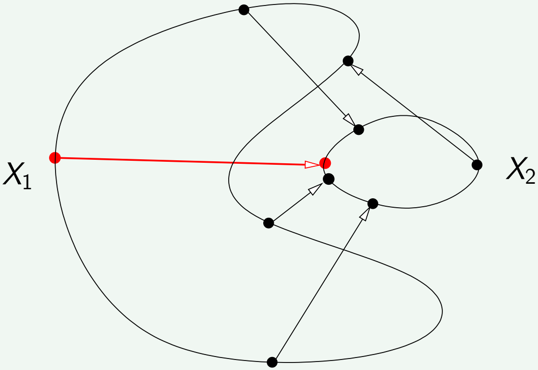
# Metric Transition Systems

## Example: Hausdorff metric



# Metric Transition Systems

## Example: Hausdorff metric

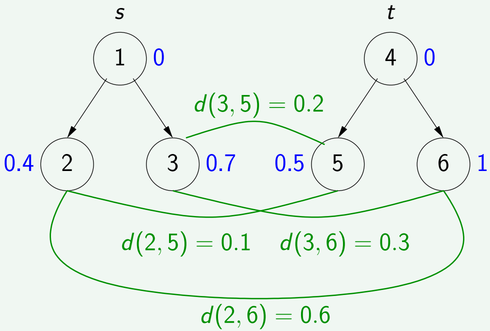


# Metric Transition Systems

**Distance of states in a metric transition system**

Compute the smallest fixed-point of

$$d(s,t) = \max\{\ d_r([s],[t]),\ d^H(\tau(s),\tau(t))\}$$

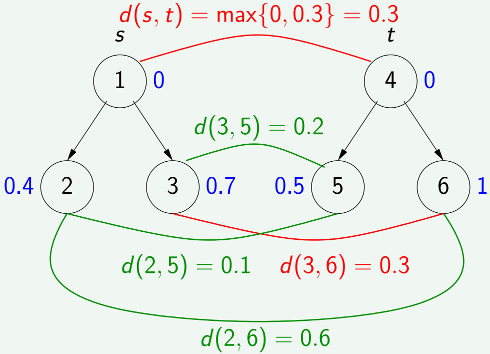


# Metric Transition Systems

## Distance of states in a metric transition system

Compute the smallest fixed-point of

$$d(s,t) = \max\{\ d_r([s],[t]),\ d^H(\tau(s),\tau(t))\}$$



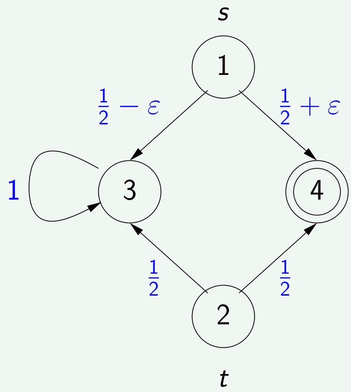
# Probabilistic Transition Systems

**Probabilistic transition system**

A probabilistic transition system is a tuple $P = (S, T, p.)$, where $S$ is a set of states, $T \subseteq S$ is the set of terminal states and every state $s \notin T$ is assigned a probability distribution $p_s \colon S \to [0, 1]$.

Studied by Larsen/Skou [Larsen and Skou, 1989], van Breugel/Worrell [van Breugel and Worrell, 2005] (again simplified)

# Probabilistic Transition Systems




Terminal state: 4

What is the distance between states 1 and 2? $\rightsquigarrow$ distance $\varepsilon$

# Probabilistic Transition Systems

**Distance of states in a probabilistic transition system**

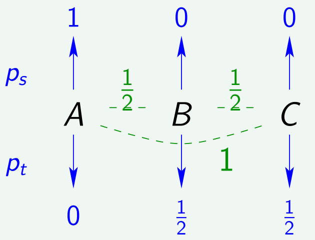
Compute the smallest fixed-point of

$$d(s,t) = \begin{cases} 1 & \text{if } s \in T, t \notin T \text{ or } s \notin T, t \in T \\ 0 & \text{if } s, t \in T \\ d^P(p_s, p_t) & \text{otherwise} \end{cases}$$

What does it mean to compute the distance between two probability distributions $p_s, p_t$ on a metric space?

# Transportation Problem & Duality [Villani, 2009]

**Lift metric to prob. distr.**




distances between states
probabilities of states

Interpret $p_s$ as supply and $p_t$ as demand. Transporting a unit along a distance $d$ costs $d$.

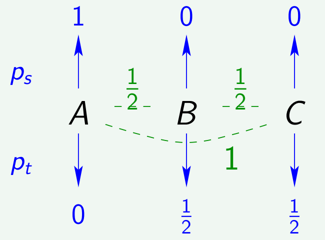
What is the minimal possible cost?

- transport $\frac{1}{2}$ from $A$ to $B$: cost $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
- transport $\frac{1}{2}$ from $A$ to $C$: cost $1 \cdot \frac{1}{2} = \frac{1}{2}$

Overall cost: $\frac{3}{4}$ (= distance $d^P(p_s, p_t)$)

# Transportation Problem & Duality [Villani, 2009]

Alternative: you have a logistics firm and handle transportation. You do this by setting a price (per unit) for locations $A, B, C$ $(pr_A, pr_B, pr_B)$. You buy and sell for this price at every location. Your prices have to satisfy: $pr_B - pr_A \leq d(A, B)$ (otherwise you do not get the contract).




distances between states
probabilities of states

You want to maximize your profit. Which prices do you set?
$\rightsquigarrow pr_A = 0,\ pr_B = \frac{1}{2},\ pr_C = 1$

- you get: $\frac{1}{2} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{3}{4}$
- you pay: $0 \cdot 1 = 0$

Profit: $\frac{3}{4}$

# Transportation Problem & Duality [Villani, 2009]

Duality in transportation theory (Kantorovich-Rubinstein duality)

The following values coincide for a metric $d\colon X \times X \to [0,1]$ and two probability distributions $p, q\colon X \to [0,1]$:

The minimum of $\sum_{x,y} P(x,y) \cdot d(x,y)$

for all probability distributions $P\colon X \times X \to [0,1]$ (couplings, indicating transport from $x$ to $y$), such that $\sum_{y \in X} P(x,y) = p(x)$, $\sum_{x \in X} P(x,y) = q(y)$ (marginal distributions are $p, q$)

The maximum of $|\sum_{x \in X} f(x) \cdot p(x) - \sum_{x \in X} f(x) \cdot q(x)|$

for all nonexpansive functions $f\colon X \to [0,1]$

# Generalization of Metric Transition Systems

This leads to the following questions:

- Are these metrics canonical/natural in some way?
- How can we define other metric transition systems (with different branching types)?
- Are there generic methods to compute metrics?

$\rightsquigarrow$ use coalgebra, a general theory of behavioural equivalences, to answer these questions.

Coalgebra offers a toolbox from which transition systems with different branching types can be constructed and analyzed.

# Functors

## Typical examples of functors

- (finite) powerset functor $\mathcal{P}_{fin}(X) = \{Y \mid Y \subseteq X, Y \text{ finite}\}$
- probability distribution functor $\mathcal{D}(X) = \{p\colon X \to [0,1] \mid \sum_{x \in X} p(x) = 1\}$
- product functor $F(X) = A \times X$ (for a fixed set $X$)
- coproduct functor (disjoint union) $F(X) = X + B$ (for a fixed set $B$)
- combinations of these functors

The functor defines the branching type of the transition system:

- powerset functor $\rightsquigarrow$ non-determinism
- probability distribution functor $\rightsquigarrow$ probabilistic branching
- product functor $\rightsquigarrow$ labelling
- coproduct functor $\rightsquigarrow$ termination, exceptions, failure

# Coalgebras & Coalgebra Homomorphisms

Transition systems are now called coalgebras:

## Coalgebra & Coalgebra Homomorphism

Let $F$ be a given functor. A coalgebra is a function $\alpha \colon S \to F(S)$ (where $S$ is the state set).

A coalgebra homomorphism between two coalgebras $\alpha \colon S \to F(S)$, $\beta \colon S' \to F(S')$ is a function $f \colon S \to S'$ satisfying $F(f) \circ \alpha = \beta \circ f$.

$$\begin{array}{ccc} S & \xrightarrow{\alpha} & F(S) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle F(f)} \\ S' & \xrightarrow{\beta} & F(S') \end{array}$$

Coalgebra homomorphisms are functions between transition systems that preserve branching. They correspond to functional bisimulations.

# Coalgebras & Coalgebra Homomorphisms

Our examples can be represented as coalgebras in the following way:

**Metric transition systems**

$$\alpha\colon S \to X \times \mathcal{P}(S)$$

where $X$ is a fixed metric space.

**Probabilistic transition systems**

$$\beta\colon S \to \mathcal{D}(S) + 1$$

where 1 is a singleton set ($1 = \{\surd\}$), representing termination.

# Coalgebras & Coalgebra Homomorphisms

**Final Coalgebra**

The final colgebra $\omega\colon \Omega \to F(\Omega)$ is a coalgebra such that there is a unique coalgebra homomorphism from any other coalgebra into $\omega$.

The final coalgebra can be considered as the universe of all possible behaviours. The mapping into the final coalgebra maps a state to its behaviour.
Final coalgebras do not necessarily exist, but they exist for our example functors. E.g., for the finite powerset functor: take all possible transition systems and quotient by bisimilarity.

Final coalgebras are useless for algorithmic purposes. But they induce a canonical notion of behavioural equivalence (two states are equivalent if they are mapped to the same state in the final coalgebra).

# Coalgebras in (Pseudo-)Metric Spaces

Idea:

- Define metric transition systems as coalgebras in **PMet** (the category of pseudo-metric spaces and non-expansive functions)
- Lift existing functors on **Set** to functors on **PMet** (transform metric on $S$ to metric on $F(S)$)
- Pseudo-metric on the final coalgebra should be a metric (since all states in the final coalgebra have different behaviour)

Existing results:

- Final coalgebra result by Rutten for contractive functors [Rutten, 1998]
- Theory of probabilistic distances [van Breugel and Worrell, 2005]

# Coalgebras in (Pseudo-)Metric Spaces

Our idea: general methods for lifting a functor $F$ to metric spaces

$\rightsquigarrow$ Wasserstein lifting, Kantorovich lifting

**Evaluation function**

We need one parameter: an evaluation function (algebra)

$$ev\colon F(\mathbb{R}_0^\infty) \to \mathbb{R}_0^\infty$$

# Coalgebras in (Pseudo-)Metric Spaces

## Wasserstein lifting

Let $d\colon X \times X \to \mathbb{R}_0^\infty$ be a pseudo-metric and $t_1, t_2 \in F(S)$:

$$d^{\downarrow F}(t_1, t_2) = \inf\{ev(F(d)(t)) \mid t \in F(S \times S), F(\pi_i)(t) = t_i\}$$

## Kantorovich lifting

Let $d\colon X \times X \to \mathbb{R}_0^\infty$ be a pseudo-metric and $t_1, t_2 \in F(S)$:

$$d^{\uparrow F}(t_1, t_2) = \sup\{d_e(ev(F(f)(t_1)), ev(F(f)(t_2))) \mid f\colon (X, d) \to (\mathbb{R}_0^\infty, d_e) \text{ non-expansive}\}$$

where $d_e(x, y) = |x - y|$ for $x, y \in \mathbb{R}_0^\infty$.

# Coalgebras in (Pseudo-)Metric Spaces

## Results

- $d^{\uparrow F}$, $d^{\downarrow F}$ are both pseudo-metrics (for the Wasserstein lifting we need some constraints on the evaluation function and weak pullback preservation)
- $d^{\uparrow F} \leq d^{\downarrow F}$
  There are cases where $d^{\uparrow F} < d^{\downarrow F}$, i.e., the Kantorovich-Rubinstein duality does not necessarily hold.
- Non-expansive functions and isometries (distance-preserving functions) are preserved by lifting.
- The Wasserstein lifting preserves metrics (if the infimum is always a minimum).

# Coalgebras in (Pseudo-)Metric Spaces

Several standard metrics can be recovered by lifting. In each of these cases the Kantorovich-Rubinstein duality holds.

| functor | evaluation fct. | resulting metric |
|---|---|---|
| $\mathcal{P}_{fin}$ | $ev(R \subseteq \mathbb{R}_0^\infty) = \max R$ | Hausdorff |
| $\mathcal{D}$ | $ev(p\colon \mathbb{R}_0^\infty \to [0,1])$ | |
| | $= \sum_{x \in \mathbb{R}_0^\infty} x \cdot p(x)$ | Kantorovich |
| $X + Y$ | $ev(x \in \mathbb{R}_0^\infty) = x$ | distance on disjoint union |
| $X \times Y$ | $ev(x, y) = \max\{x, y\}$ | maximum of distances |
| $X \times Y$ | $ev(x, y) = x + y$ | sum of distances |

Last three cases: bifunctor lifting

# Computing Distances in Coalgebras

**Compute metrics in a coalgebraic setting**

Given a coalgebra in $\alpha\colon S \to F(S)$ compute its associated metric $d\colon S \times S \to \mathbb{R}_0^\infty$ as the smallest fixed-point of:

$$d(s,t) = d^F(\alpha(s), \alpha(t))$$

where $d^F$ is an appropriate lifting (preserving isometries and metrics).

If we compute the metric $d_\omega$ for the final coalgebra $\omega$, we obtain a final coalgebra in the category of (pseudo-)metric spaces.

If we compute the pseudo-metric $d_\alpha$ for any other coalgebra $\alpha$, we obtain the pseudo-metric induced by the coalgebra homomorphism $f$ from $\alpha$ into the final coalgebra $\omega$, i.e.,

$$d_\alpha(s,t) = d_\omega(f(s), f(t))$$

# Trace Metrics

Ideas:

- Work with coalgebras that model both implicit and explicit branching
  Coalgebras of the form $\alpha\colon S \to F(T(S))$
  ($F$: explicit branching, $T$ – monad: implicit branching)
  Example: $F(S) = 2 \times S^{\Sigma}$, $T(S) = \mathcal{P}_{fin}(S)$
  (non-deterministic automata)
- How to obtain the "right" notion of behavioural equivalence (here: trace equivalence)?
  First determinize the coalgebra, obtaining a coalgebra

$$\alpha^{\#} = F(\mu_S) \circ \lambda_{T(S)} \circ T(\alpha)\colon T(S) \to F(T(S))$$

  where $\lambda\colon TF \Rightarrow FT$ is a distributive law and $\mu$ is the multiplication of the monad.
  Then determine behavioural equivalences, behavioural metrics, etc. on the determinized coalgebra.

# Trace Metrics

Formally: embed **Set** into an Eilenberg-Moore category

**Eilenberg-Moore category $\mathcal{EM}(T)$ of a monad $T$**

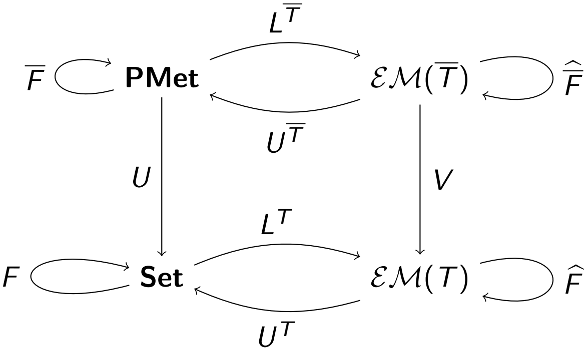
- Objects: algebras $a\colon T(S) \to S$
  with $a \circ \mu_S = id_S$, $a \circ Ta = a \circ \mu_S$.
- Arrows: Algebra homomorphisms

Embedding from **Set** to $\mathcal{EM}(T)$: $S \mapsto \mu_S\colon T(T(S)) \to S$

- Lift the monad $T$ to a monad $\overline{T}$ on **PMet** (under certain conditions monad lift to monads).
- Lift the distributive law (i.e., natural transformation) to **PMet**.
- Lift the functor $F$ to **PMet** and then to $\mathcal{EM}(\overline{T})$ (using the lifted distributive law).
- Determinize the coalgebra and compute behavioural distances in $\mathcal{EM}(T)$.

# Trace Metrics

Summary:



# Trace Metrics

Examples for trace metrics, obtained by defining suitable evaluation functions:

- Non-deterministic automata:
  We obtain the usual ultrametric on words, lifted to languages:

  $$d(L_1, L_2) = c^{|w|}$$

  where $L_1, L_2 \subseteq \Sigma^*$, $0 < c < 1$ and $w$ is the shortest word such that $w \in L_1$, $w \notin L_2$ (or vice versa).
- Probabilistic automata:
  We obtain the total variation distance:

  $$d(p_1, p_2) = \frac{1}{2} \cdot \sum_{w \in \Sigma^*} |p_1(w) - p_2(w)|$$

  where $p_1, p_2 \colon \Sigma^* \to [0, 1]$ are weighted languages.

# Conclusion

## Other issues

- Logical characterization of distances
- A fibrational view on behavioural metrics
- Quantitative linear-time/branching-time spectrum [Fahrenberg et al., 2011]
- Distances different from real numbers (monoids, quantales, ...) [Fahrenberg and Legay, 2013]
- Directed metrics (simulation distances) [de Alfaro et al., 2009]
- Algorithms (polynomial-time [Chen et al., 2012], on-the-fly [Bacci et al., 2013], ...)

Bacci, G., Bacci, G., Larsen, K. G., and Mardare, R. (2013).
On-the-fly exact computation of bisimilarity distances.
In *Proc. of TACAS '13*, pages 1–15. Springer.
LNCS/ARCoSS 7795.

Chen, D., van Breugel, F., and Worrell, J. (2012).
On the complexity of computing probabilistic bisimilarity.
In *Proc. of FOSSACS '12*, pages 437–451. Springer.
LNCS/ARCoSS 7213.

de Alfaro, L., Faella, M., and Stoelinga, M. (2009).
Linear and branching system metrics.
*IEEE Transactions on Software Engineering*, 25(2).

Fahrenberg, U. and Legay, A. (2013).
Generalized quantitative analysis of metric transition systems.
In *Proc. of APLAS '13*, pages 192–208. Springer.
LNCS 8301.

Fahrenberg, U., Legay, A., and Thrane, C. (2011).

The quantitative linear-time–branching-time spectrum.
In *Proc. of FSTTCS '11*, volume 13 of *LIPIcs*, pages 103–114. Schloss Dagstuhl – Leibniz Center for Informatics.

Larsen, K. G. and Skou, A. (1989).
Bisimulation through probabilistic testing (preliminary report).
In *Proc. of POPL '89*, pages 344–352. ACM.

Rutten, J. (1998).
Relators and metric bisimulations.
In *Proc. of CMCS '98 (Workshop on Coalgebraic Methods in Computer Science)*, number 11 in ENTCS, pages 252–258.

van Breugel, F. and Worrell, J. (2005).
Approximating and computing behavioural distances in probabilistic transition systems.
*Theoretical Computer Science*, 360:373–385.

Villani, C. (2009).

*Optimal Transport – Old and New*, volume 338 of *A Series of Comprehensive Studies in Mathematics*.
Springer.