Quantitative equational algebra

Radu Mardare Prakash Panangaden Gordon Plotkín LICS 2016

Approximate equations: $a =_{\epsilon} b$, a is within ϵ of b.

Approximate equations: $a =_{\epsilon} b$, a is within ϵ of b.

Definitely not an equivalence relation; it defines a *uniformity*.

Approximate equations: $a =_{\epsilon} b$, a is within ϵ of b.

Definitely not an equivalence relation; it defines a *uniformity*.

Quantitative analogue of equational reasoning.

Approximate equations: $a =_{\epsilon} b$, a is within ϵ of b.

Definitely not an equivalence relation; it defines a *uniformity*.

Quantitative analogue of equational reasoning.

Birkhoff-like completeness theorem, uniformity results,

Approximate equations: $a =_{\epsilon} b$, a is within ϵ of b.

Definitely not an equivalence relation; it defines a *uniformity*.

Quantitative analogue of equational reasoning.

Birkhoff-like completeness theorem, uniformity results, monads ...

Approximate equations: $a =_{\epsilon} b$, a is within ϵ of b.

Definitely not an equivalence relation; it defines a *uniformity*.

Quantitative analogue of equational reasoning.

Birkhoff-like completeness theorem, uniformity results, monads ...

Many examples: mostly probabilistic

Moggi 1988: How to incorporate "effects" into denotational semantics? Answer: monads!

Moggi 1988: How to incorporate "effects" into denotational semantics? Answer: monads!

Plotkin, Power (and then many others): computational effects algebraically. Monads are given by operations and equations.

Moggi 1988: How to incorporate "effects" into denotational semantics? Answer: monads!

Plotkin, Power (and then many others): computational effects algebraically. Monads are given by operations and equations.

Categorically: equational presentations are Lawvere theories.

Moggi 1988: How to incorporate "effects" into denotational semantics? Answer: monads!

Plotkin, Power (and then many others): computational effects algebraically. Monads are given by operations and equations.

Categorically: equational presentations are Lawvere theories.

A monad of great interest: Lawvere (1964) The category of probabilistic mappings (unpublished).

Moggi 1988: How to incorporate "effects" into denotational semantics? Answer: monads!

Plotkin, Power (and then many others): computational effects algebraically. Monads are given by operations and equations.

Categorically: equational presentations are Lawvere theories.

A monad of great interest: Lawvere (1964) The category of probabilistic mappings (unpublished).

Later (1981) Giry: monad on measure spaces and also on Polish spaces.

Probabilistic reasoning requires measure theory but,

Probabilistic reasoning requires measure theory but, measure theory works best on metric (Polish) spaces.

Probabilistic reasoning requires measure theory but, measure theory works best on metric (Polish) spaces.

Metric ideas present in semantics from the start: Jaco de Bakker's school.

Probabilistic reasoning requires measure theory but, measure theory works best on metric (Polish) spaces.

Metric ideas present in semantics from the start: Jaco de Bakker's school.

Our goal (motivated by probability): develop the theory of effects in a metric setting.

Probabilistic reasoning requires measure theory but, measure theory works best on metric (Polish) spaces.

Metric ideas present in semantics from the start: Jaco de Bakker's school.

Our goal (motivated by probability): develop the theory of effects in a metric setting.

Algebras will come with metric structure and

Probabilistic reasoning requires measure theory but, measure theory works best on metric (Polish) spaces.

Metric ideas present in semantics from the start: Jaco de Bakker's school.

Our goal (motivated by probability): develop the theory of effects in a metric setting.

Algebras will come with metric structure and

quantitative equational theories will define monads on \mathbf{Met} .

(M. H. Stone 1949)

(M. H. Stone 1949)

Signature: $\{+_{\epsilon} | \epsilon \in [0, 1]\}$

(M. H. Stone 1949)

Signature: $\{+_{\epsilon} | \epsilon \in [0, 1]\}$ (a *set* of operations)

(M. H. Stone 1949)

Signature: $\{+_{\epsilon} | \epsilon \in [0, 1]\}$ (a set of operations)

Axioms:

$$(B_1) \vdash t + t' = t$$

$$(B_2) \vdash t + t' = t$$

$$(SC) \vdash t + t' = t' + t' = t' + t$$

$$(SA) \vdash (t + t') + t'' = t + t' + t'' = t + t''$$

(M. H. Stone 1949)

Signature: $\{+_{\epsilon} | \epsilon \in [0, 1]\}$ (a set of operations)

Axioms:

$$(B_1) \vdash t + t' = t$$

$$(B_2) \vdash t + t' = t$$

$$(SC) \vdash t + t' = t' + t' = t' + t = t' + t' = t' + t' = t + t' = t + t' = t + t'' = t + t'' = t + t'' = t''' = t'' = t''' = t'' = t''' = t'' = t'' = t''' = t''$$

Equations define an equivalence relation on terms. Our quantitative equations will define a metric on terms.

Deducibility relations

(**Refl**) $\emptyset \vdash t =_0 t$ (Symm) $\{t =_{\epsilon} s\} \vdash s =_{\epsilon} t.$ (Triang) $\{t =_e s, s =_{e'} u\} \vdash t =_{e+e'} u$. (Max) For e' > 0, $\{t =_e s\} \vdash t =_{e+e'} s$. (Arch) For $e \ge 0$, $\{t =_{e'} s \mid e' > e\} \vdash t =_{e} s$. (NExp) For $f : n \in \Omega$, $\{t_1 =_e s_1, \ldots, t_n =_e s_n\} \vdash f(t_1, \ldots, t_i, \ldots, t_n) =_e f(s_1, \ldots, s_i, \ldots, s_n)$ **(Subst)** If $\sigma \in \Sigma(X)$, $\Gamma \vdash t =_e s$ implies $\sigma(\Gamma) \vdash \sigma(t) =_e \sigma(s)$. (Cut) If $\Gamma \vdash \phi$ for all $\phi \in \Gamma'$ and $\Gamma' \vdash \psi$, then $\Gamma \vdash \psi$. (Assumpt) If $\phi \in \Gamma$, then $\Gamma \vdash \phi$.

Signature Ω , variables X we get terms $\mathbb{T}X$

Signature Ω , variables X we get terms $\mathbb{T}X$ Quantitative equations: $\mathcal{V}(\mathbb{T}X)$

 $s =_e t, s, t \in \mathbb{T}X, e \in \mathbb{Q} \cap [0, 1].$

Signature Ω , variables X we get terms $\mathbb{T}X$ Quantitative equations: $\mathcal{V}(\mathbb{T}X)$

 $s =_e t, s, t \in \mathbb{T}X, e \in \mathbb{Q} \cap [0, 1].$

Quantitative inferences: $\mathcal{E}(\mathbb{T}X) = \mathcal{P}_{\mathsf{fin}}(\mathcal{V}(\mathbb{T}X)) \times \mathcal{V}(\mathbb{T}X)$

$$\{s_1 =_{e_1} t_1, \dots, s_n =_{e_n} t_n\} \vdash s =_e t$$

Signature Ω , variables X we get terms $\mathbb{T}X$ Quantitative equations: $\mathcal{V}(\mathbb{T}X)$

 $s =_e t, s, t \in \mathbb{T}X, e \in \mathbb{Q} \cap [0, 1].$

Quantitative inferences: $\mathcal{E}(\mathbb{T}X) = \mathcal{P}_{\mathsf{fin}}(\mathcal{V}(\mathbb{T}X)) \times \mathcal{V}(\mathbb{T}X)$

$$\{s_1 =_{e_1} t_1, \dots, s_n =_{e_n} t_n\} \vdash s =_e t$$

Given $S \subset \mathcal{E}(\mathbb{T}X), \vdash_S$: smallest deducibility relation containing S.

Signature Ω , variables X we get terms $\mathbb{T}X$ Quantitative equations: $\mathcal{V}(\mathbb{T}X)$

 $s =_e t, s, t \in \mathbb{T}X, e \in \mathbb{Q} \cap [0, 1].$

Quantitative inferences: $\mathcal{E}(\mathbb{T}X) = \mathcal{P}_{\mathsf{fin}}(\mathcal{V}(\mathbb{T}X)) \times \mathcal{V}(\mathbb{T}X)$

$$\{s_1 =_{e_1} t_1, \dots, s_n =_{e_n} t_n\} \vdash s =_e t$$

Given $S \subset \mathcal{E}(\mathbb{T}X), \vdash_S$: smallest deducibility relation containing S.

Equational theory: $\mathcal{U} = \vdash_S \bigcap \mathcal{E}(\mathbb{T}X)$

Quantitative Algebra

Ω: signature; $\mathcal{A} = (A, d)$: A an Ω-algebra and (A, d) a metric space

Quantitative Algebra

 Ω : signature; $\mathcal{A} = (A, d)$: A an Ω -algebra and (A, d) a metric space all functions in Ω are nonexpansive.

Quantitative Algebra

 Ω : signature; $\mathcal{A} = (A, d)$: A an Ω -algebra and (A, d) a metric space all functions in Ω are nonexpansive.

Morphisms are Ω -algebra homomorphisms that are nonexpansive.

 Ω : signature; $\mathcal{A} = (A, d)$: A an Ω -algebra and (A, d) a metric space all functions in Ω are nonexpansive.

Morphisms are Ω -algebra homomorphisms that are nonexpansive.

 $\mathbb{T}X$ is an Ω -algebra. $\sigma: \mathbb{T}X \to A, \Omega$ -homomorphism.

 Ω : signature; $\mathcal{A} = (A, d)$: A an Ω -algebra and (A, d) a metric space all functions in Ω are nonexpansive.

Morphisms are Ω -algebra homomorphisms that are nonexpansive.

 $\mathbb{T}X$ is an Ω -algebra. $\sigma: \mathbb{T}X \to A, \Omega$ -homomorphism.

(A,d) satisfies $\{s_i =_{e_i} t_i / i = 1, \ldots, n\} \vdash s =_e t$ if

 Ω : signature; $\mathcal{A} = (A, d)$: A an Ω -algebra and (A, d) a metric space all functions in Ω are nonexpansive.

Morphisms are Ω -algebra homomorphisms that are nonexpansive.

 $\mathbb{T}X$ is an Ω -algebra. $\sigma: \mathbb{T}X \to A, \Omega$ -homomorphism.

(A,d) satisfies $\{s_i =_{e_i} t_i / i = 1, \ldots, n\} \vdash s =_e t$ if

 $\forall \sigma, d(\sigma(s_i), \sigma(t_i)) \leq e_i, i = 1, \dots, n$ implies $d(\sigma(s), \sigma(t)) \leq e_i$.

 Ω : signature; $\mathcal{A} = (A, d)$: A an Ω -algebra and (A, d) a metric space all functions in Ω are nonexpansive.

Morphisms are Ω -algebra homomorphisms that are nonexpansive.

 $\mathbb{T}X$ is an Ω -algebra. $\sigma: \mathbb{T}X \to A, \Omega$ -homomorphism.

(A,d) satisfies $\{s_i =_{e_i} t_i / i = 1, \ldots, n\} \vdash s =_e t$ if

$$\forall \sigma, d(\sigma(s_i), \sigma(t_i)) \leq e_i, i = 1, \dots, n$$

implies $d(\sigma(s), \sigma(t)) \leq e$.

We write $\mathbb{K}(\mathcal{U}, \Omega)$ for the algebras satisfying \mathcal{U} .

$\forall \mathcal{A} \in \mathbb{K}(\mathcal{U}, \Omega), \mathcal{A} \text{ satisfies } \Gamma \vdash \phi \text{ iff } \Gamma \vdash \phi \in \mathcal{U}.$

 $\forall \mathcal{A} \in \mathbb{K}(\mathcal{U}, \Omega), \mathcal{A} \text{ satisfies } \Gamma \vdash \phi \text{ iff } \Gamma \vdash \phi \in \mathcal{U}.$

Analogue of the usual completeness theorem for equational logic,

 $\forall \mathcal{A} \in \mathbb{K}(\mathcal{U}, \Omega), \mathcal{A} \text{ satisfies } \Gamma \vdash \phi \text{ iff } \Gamma \vdash \phi \in \mathcal{U}.$

Analogue of the usual completeness theorem for equational logic, but the proof needs to deal

 $\forall \mathcal{A} \in \mathbb{K}(\mathcal{U}, \Omega), \mathcal{A} \text{ satisfies } \Gamma \vdash \phi \text{ iff } \Gamma \vdash \phi \in \mathcal{U}.$

Analogue of the usual completeness theorem for equational logic, but the proof needs to deal with quantitative issues and inequalities.

Quotient by the kernel to get a metric space $\mathbb{T}[X]$.

Quotient by the kernel to get a metric space $\mathbb{T}[X]$.

Thm: $(\mathbb{T}[X], d)$ is in $\mathbb{K}(\mathcal{U}, \Omega)$.

Quotient by the kernel to get a metric space $\mathbb{T}[X]$.

Thm: $(\mathbb{T}[X], d)$ is in $\mathbb{K}(\mathcal{U}, \Omega)$.

Starting from a **metric space** (M, d) we can define $\mathbb{T}M$ by adding $m \in M$ as constants and $\emptyset \vdash m =_e n$ as axioms for every rational $d(m, n) \leq e$.

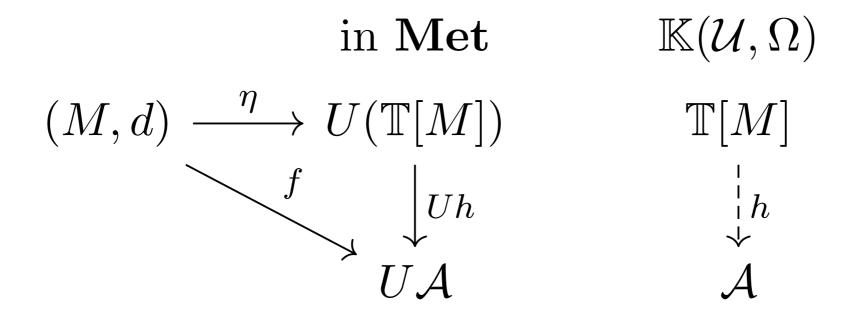
Quotient by the kernel to get a metric space $\mathbb{T}[X]$.

Thm: $(\mathbb{T}[X], d)$ is in $\mathbb{K}(\mathcal{U}, \Omega)$.

Starting from a **metric space** (M, d) we can define $\mathbb{T}M$ by adding $m \in M$ as constants and $\emptyset \vdash m =_e n$ as axioms for every rational $d(m, n) \leq e$.

From this we can construct $(\mathbb{T}[M], d_M) \in \mathbb{K}(\mathcal{U}, \Omega)$.

Universal property



 $U: \mathbb{K}(\mathcal{U}, \Omega) \to \mathbf{Met}$: forgetful functor

 ${\mathbb T}$ defines a monad on ${\bf Met}$

Example 1: Semiadditive barycentric algebras Signature: $\mathcal{B} = \{+_{\epsilon} | \epsilon \in [0, 1]\}$ (a set of operations)

Equations:

$$(B_1) \vdash t + t' =_0 t$$

$$(B_2) \vdash t + t =_0 t$$

$$(SC) \vdash t + t' =_0 t' + t =_0 t' + t =_0 t' + t =_0 t' + t' =_0 t' =_0$$

Example 1: Semiadditive barycentric algebras Signature: $\mathcal{B} = \{+_{\epsilon} | \epsilon \in [0, 1]\}$ (a set of operations)

Equations:

$$(B_1) \vdash t + t' =_0 t$$

$$(B_2) \vdash t + t =_0 t$$

$$(SC) \vdash t + t' =_0 t' + t =_0 t' + t =_0 t' + t =_0 t' + t' =_0 t' =_0 t' + t' =_0 t' =$$

The freely generated algebra is the space of probability distributions with the *total variation* metric.

Same signature. Equations:

 $(B_1), (B_2), (SC), (SA)$ and

Same signature. Equations:

> $(B_1), (B_2), (SC), (SA) \text{ and}$ $(W)\{x_1 =_{e_1} y_1, x_2 =_{e_2} y_2\} \vdash x_1 +_{\epsilon} x_2 =_{e'} y_1 +_{\epsilon} y_2$ where $(\epsilon \cdot e_1^p + (1 - \epsilon)e_2^p)^{\frac{1}{p}} \leq e'$

Same signature. Equations:

> $(B_1), (B_2), (SC), (SA) \text{ and}$ $(W)\{x_1 =_{e_1} y_1, x_2 =_{e_2} y_2\} \vdash x_1 +_{\epsilon} x_2 =_{e'} y_1 +_{\epsilon} y_2$ where $(\epsilon \cdot e_1^p + (1 - \epsilon)e_2^p)^{\frac{1}{p}} \leq e'$

Special case p = 1:

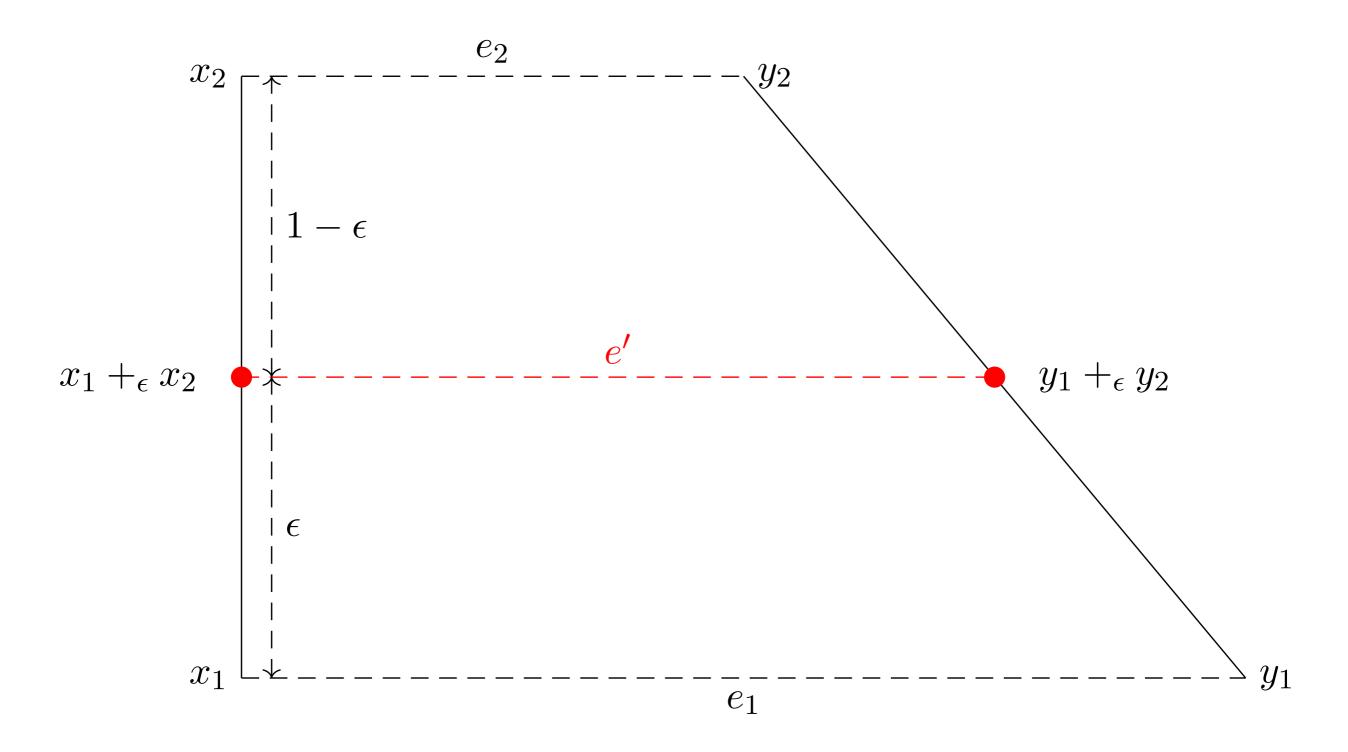
Same signature. Equations:

> $(B_1), (B_2), (SC), (SA) \text{ and}$ $(W)\{x_1 =_{e_1} y_1, x_2 =_{e_2} y_2\} \vdash x_1 +_{\epsilon} x_2 =_{e'} y_1 +_{\epsilon} y_2$ where $(\epsilon \cdot e_1^p + (1 - \epsilon)e_2^p)^{\frac{1}{p}} \leq e'$

Special case p = 1:

 $(K) \{ x_1 =_{e_1} y_1, x_2 =_{e_2} y_2 \} \vdash x_1 +_{\epsilon} x_2 =_{e'} y_1 +_{\epsilon} y_2$ where $(\epsilon \cdot e_1 + (1 - \epsilon)e_2) \le e'$

A picture of equation (K)



(M, d): metric space, $\Pi[M]$: finitely supported probability measures on M, $\mu +_{\epsilon} \nu = \epsilon \mu + (1 - \epsilon) \nu$, convex sum.

(M, d): metric space, $\Pi[M]$: finitely supported probability measures on M, $\mu +_{\epsilon} \nu = \epsilon \mu + (1 - \epsilon) \nu$, convex sum.

$$W^{p}(\mu,\nu) = \inf_{\omega \in \mathcal{C}(\mu,\nu)} \int_{M \times M} d(x,y)^{p} d\omega$$

(M, d): metric space, $\Pi[M]$: finitely supported probability measures on M, $\mu +_{\epsilon} \nu = \epsilon \mu + (1 - \epsilon) \nu$, convex sum.

$$W^{p}(\mu,\nu) = \inf_{\omega \in \mathcal{C}(\mu,\nu)} \int_{M \times M} d(x,y)^{p} d\omega$$

where $C(\mu, \nu)$ is the space of couplings: probability distributions on $M \times M$ with marginals matching μ, ν .

(M, d): metric space, $\Pi[M]$: finitely supported probability measures on M, $\mu +_{\epsilon} \nu = \epsilon \mu + (1 - \epsilon)\nu$, convex sum.

$$W^{p}(\mu,\nu) = \inf_{\omega \in \mathcal{C}(\mu,\nu)} \int_{M \times M} d(x,y)^{p} d\omega$$

where $C(\mu, \nu)$ is the space of couplings: probability distributions on $M \times M$ with marginals matching μ, ν .

Prop. $(\Pi[M], W^p)$ is in \mathbb{K} .

Thm. $(\Pi[M], W^p)$ is the *M*-generated barycentric algebra *i.e.* it has the universal property.

Thm. $(\Pi[M], W^p)$ is the *M*-generated barycentric algebra *i.e.* it has the universal property.

Proof ideas:
Couplings form a convex set.
Convexity properties of spaces of measures.
Linearity of integration.
Splitting lemma.
Non-expansiveness – induction on size of the support.

Need topology on $\Delta[M]$ = Borel measures on M.

Need topology on $\Delta[M]$ = Borel measures on M.

 $\lim_{i \to \infty} \mu_i = \mu$

Need topology on $\Delta[M]$ = Borel measures on M.

 $\lim_{i \to \infty} \mu_i = \mu \quad \text{if for all bounded continuous functions}$

Need topology on $\Delta[M]$ = Borel measures on M.

 $\lim_{i \to \infty} \mu_i = \mu \quad \text{if for all bounded continuous functions}$

 $f: M \to \mathbf{R}$ we have $\int f d\mu_i \to \int f d\mu$.

Need topology on $\Delta[M]$ = Borel measures on M.

 $\lim_{i \to \infty} \mu_i = \mu \quad \text{if for all bounded continuous functions}$

 $f: M \to \mathbf{R}$ we have $\int f d\mu_i \to \int f d\mu$.

Thm. If M is complete and separable then $\Pi[M]$ is dense in $\Delta[M]$ in the weak topology. Moreover, W^p metrizes the weak topology.

Need topology on $\Delta[M]$ = Borel measures on M.

 $\lim_{i \to \infty} \mu_i = \mu \quad \text{if for all bounded continuous functions}$

 $f: M \to \mathbf{R}$ we have $\int f d\mu_i \to \int f d\mu$.

Thm. If M is complete and separable then $\Pi[M]$ is dense in $\Delta[M]$ in the weak topology. Moreover, W^p metrizes the weak topology.

Using this we can extend the barycentric algebra results to the continuous case.

Other examples

Other examples

1. Pointed barycentric algebras (subprobability distributions)

Other examples

1. Pointed barycentric algebras (subprobability distributions)

2. Hausdorff metric from (quantitative) semilattices

Other examples

1. Pointed barycentric algebras (subprobability distributions)

2. Hausdorff metric from (quantitative) semilattices

3. Markov chains (not in the paper) with trace-based metric

Other examples

1. Pointed barycentric algebras (subprobability distributions)

2. Hausdorff metric from (quantitative) semilattices

- 3. Markov chains (not in the paper) with trace-based metric
- 4. Exceptions, state, IO quantitative analogues (not in the paper)

van Breugel et al. (2007) Gives equational (in the ordinary sense) presentation of Hausdorff and Kantorovich and exhibits monads on the category of complete metric spaces.

van Breugel et al. (2007) Gives equational (in the ordinary sense) presentation of Hausdorff and Kantorovich and exhibits monads on the category of complete metric spaces.

Adamek et al. (2012) shows the finitely nature of these monads.

van Breugel et al. (2007) Gives equational (in the ordinary sense) presentation of Hausdorff and Kantorovich and exhibits monads on the category of complete metric spaces.

Adamek et al. (2012) shows the finitely nature of these monads.

We use barycentric axioms rather than mean-value axioms.

van Breugel et al. (2007) Gives equational (in the ordinary sense) presentation of Hausdorff and Kantorovich and exhibits monads on the category of complete metric spaces.

Adamek et al. (2012) shows the finitely nature of these monads.

We use barycentric axioms rather than mean-value axioms.

We develop a general theory of quantitative equations.

van Breugel et al. (2007) Gives equational (in the ordinary sense) presentation of Hausdorff and Kantorovich and exhibits monads on the category of complete metric spaces.

Adamek et al. (2012) shows the finitely nature of these monads.

We use barycentric axioms rather than mean-value axioms.

We develop a general theory of quantitative equations.

We can easily adapt Kantorovich to Wasserstein.

Most of our results work for metric taking values in $[0,\infty]$ except for Wasserstein/Kantorovich.

Most of our results work for metric taking values in $[0,\infty]$ except for Wasserstein/Kantorovich.

We need to extend Kantorovich-Rubinstein duality results.

Most of our results work for metric taking values in $[0,\infty]$ except for Wasserstein/Kantorovich.

We need to extend Kantorovich-Rubinstein duality results.

Markov processes and bisimulation metrics. Perhaps need many-sorted version of the theory (?)

Most of our results work for metric taking values in $[0,\infty]$ except for Wasserstein/Kantorovich.

We need to extend Kantorovich-Rubinstein duality results.

Markov processes and bisimulation metrics. Perhaps need many-sorted version of the theory (?)

Many tempting examples lurking in the shadows.

Most of our results work for metric taking values in $[0,\infty]$ except for Wasserstein/Kantorovich.

We need to extend Kantorovich-Rubinstein duality results.

Markov processes and bisimulation metrics. Perhaps need many-sorted version of the theory (?)

Many tempting examples lurking in the shadows.

Quantitative theory of effects: contribute to probabilistic programming languages.

Most of our results work for metric taking values in $[0,\infty]$ except for Wasserstein/Kantorovich.

We need to extend Kantorovich-Rubinstein duality results.

Markov processes and bisimulation metrics. Perhaps need many-sorted version of the theory (?)

Many tempting examples lurking in the shadows.

Quantitative theory of effects: contribute to probabilistic programming languages.

Birkhoff variety theorem?