Approximating Probabilistic Bisimulation by Conditional Expectation

Prakash Panangaden

School of Computer Science
McGill University

CMS Meeting 5 - 8 June 2020

Outline

1. Introduction

Outline

Outline

Outline

Outline

Outline

Outline

Joint work with

Chaput, Danos and Plotkin

Philippe Chaput, Vincent Danos, Prakash Panangaden, and Gordon Plotkin. "Approximating Markov processes by averaging." Journal of the ACM (JACM) 61, no. 1 (2014): 1-45.

The idea of functorializing conditional expectation is due to Vincent.

Approximation via Averaging

1. Approximation of Markov processes should be based on “averaging”.

Approximation via Averaging

1. Approximation of Markov processes should be based on “averaging”.
2. Averages are computed by expectation values.

Approximation via Averaging

1. Approximation of Markov processes should be based on “averaging”.
2. Averages are computed by expectation values.
3. Beautiful functorial presentation of expectation values due to Vincent Danos.

Approximation via Averaging

1. Approximation of Markov processes should be based on “averaging”.
2. Averages are computed by expectation values.
3. Beautiful functorial presentation of expectation values due to Vincent Danos.
4. Make bisimulation and approximation live in the same universe

Some notation

1. Given (X, Σ, p) and (Y, Λ) and a measurable function $f : X \longrightarrow Y$ we obtain a measure q on Y by $q(B) = p(f^{-1}(B))$. This is written $M_f(p)$ and is called the *image measure* of p under f.

Some notation

1. Given (X, Σ, p) and (Y, Λ) and a measurable function $f : X \longrightarrow Y$ we obtain a measure q on Y by $q(B) = p(f^{-1}(B))$. This is written $M_f(p)$ and is called the *image measure* of p under f.
2. We say that a measure ν is **absolutely continuous** with respect to another measure μ if for any measurable set A, $\mu(A) = 0$ implies that $\nu(A) = 0$. We write $\nu \ll \mu$.

The Radon-Nikodym Theorem

The Radon-Nikodym theorem is a central result in measure theory allowing one to define a “derivative” of a measure with respect to another measure.

Radon-Nikodym

If $\nu \ll \mu$, where ν, μ are finite measures on a measurable space (X, Σ) there is a positive measurable function h on X such that for every measurable set B

$$\nu(B) = \int_B h \, \mathrm{d}\mu.$$

The function h is defined uniquely up to a set of μ-measure 0. The function h is called the Radon-Nikodym derivative of ν with respect to μ; we denote it by $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$. Since ν is finite, $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} \in L_1^+(X, \mu)$.

Notation for Radon-Nikodym

1. Given an (almost-everywhere) positive function $f \in L_1(X, p)$, we let $f \cdot p$ be the measure which has density f with respect to p.

Notation for Radon-Nikodym

1. Given an (almost-everywhere) positive function $f \in L_1(X, p)$, we let $f \cdot p$ be the measure which has density f with respect to p.
2. Two identities that we get from the Radon-Nikodym theorem are:

Notation for Radon-Nikodym

1. Given an (almost-everywhere) positive function $f \in L_1(X,p)$, we let $f \cdot p$ be the measure which has density f with respect to p.
2. Two identities that we get from the Radon-Nikodym theorem are:
   - given $q \ll p$, we have $\frac{\mathrm{d}q}{\mathrm{d}p} \cdot p = q$.

Notation for Radon-Nikodym

1. Given an (almost-everywhere) positive function $f \in L_1(X,p)$, we let $f \cdot p$ be the measure which has density f with respect to p.
2. Two identities that we get from the Radon-Nikodym theorem are:
   - given $q \ll p$, we have $\frac{\mathrm{d}q}{\mathrm{d}p} \cdot p = q$.
   - given $f \in L_1^+(X,p)$, $\frac{\mathrm{d}f \cdot p}{\mathrm{d}p} = f$

Notation for Radon-Nikodym

1. Given an (almost-everywhere) positive function $f \in L_1(X,p)$, we let $f \cdot p$ be the measure which has density f with respect to p.
2. Two identities that we get from the Radon-Nikodym theorem are:
   - given $q \ll p$, we have $\frac{\mathrm{d}q}{\mathrm{d}p} \cdot p = q$.
   - given $f \in L_1^+(X,p)$, $\frac{\mathrm{d}f \cdot p}{\mathrm{d}p} = f$
3. These two identities just say that the operations $(-) \cdot p$ and $\frac{\mathrm{d}(-)}{\mathrm{d}p}$ are inverses of each other as maps between $L_1^+(X,p)$ and $\mathcal{M}^{\ll p}(X)$ the space of finite measures on X that are absolutely continuous with respect to p.

Expectation and conditional expectation

1. The expectation $\mathbb{E}_p(f)$ of a measurable function f is the average computed by $\int f \mathrm{d}p$ and therefore it is just a number.

Expectation and conditional expectation

1. The expectation $\mathbb{E}_p(f)$ of a measurable function f is the average computed by $\int f \mathrm{d}p$ and therefore it is just a number.
2. The *conditional* expectation is not a mere number but a random variable.

Expectation and conditional expectation

1. The expectation $\mathbb{E}_p(f)$ of a measurable function f is the average computed by $\int f \mathrm{d}p$ and therefore it is just a number.
2. The *conditional* expectation is not a mere number but a random variable.
3. It is meant to measure the expected value in the presence of additional information.

Expectation and conditional expectation

1. The expectation $\mathbb{E}_p(f)$ of a measurable function f is the average computed by $\int f \mathrm{d}p$ and therefore it is just a number.
2. The *conditional* expectation is not a mere number but a random variable.
3. It is meant to measure the expected value in the presence of additional information.
4. The additional information takes the form of a sub-σ algebra, say Λ, of Σ. The experimenter knows, for every $B \in \Lambda$, whether the outcome is in B or not.

Expectation and conditional expectation

1. The expectation $\mathbb{E}_p(f)$ of a measurable function f is the average computed by $\int f \mathrm{d}p$ and therefore it is just a number.
2. The *conditional* expectation is not a mere number but a random variable.
3. It is meant to measure the expected value in the presence of additional information.
4. The additional information takes the form of a sub-σ algebra, say Λ, of Σ. The experimenter knows, for every $B \in \Lambda$, whether the outcome is in B or not.
5. Now she can recompute the expectation values given this information.

Formalizing conditional expectation

- It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.

Formalizing conditional expectation

- It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.

Kolmogorov

Let (X, Σ, p) be a measure space with p a finite measure, f be in $L_1(X, \Sigma, p)$ and Λ be a sub-σ-algebra of Σ, then there exists a $g \in L_1(X, \Lambda, p)$ such that for all $B \in \Lambda$

$$\int_B f \mathrm{d}p = \int_B g \mathrm{d}p.$$

Formalizing conditional expectation

- It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.

Kolmogorov

Let (X, Σ, p) be a measure space with p a finite measure, f be in $L_1(X, \Sigma, p)$ and Λ be a sub-σ-algebra of Σ, then there exists a $g \in L_1(X, \Lambda, p)$ such that for all $B \in \Lambda$

$$\int_B f \mathrm{d}p = \int_B g \mathrm{d}p.$$

- This function g is usually denoted by $\mathbb{E}(f|\Lambda)$.

Formalizing conditional expectation

- It is an immediate consequence of the Radon-Nikodym theorem that such conditional expectations exist.

Kolmogorov

Let (X, Σ, p) be a measure space with p a finite measure, f be in $L_1(X, \Sigma, p)$ and Λ be a sub-σ-algebra of Σ, then there exists a $g \in L_1(X, \Lambda, p)$ such that for all $B \in \Lambda$

$$\int_B f \mathrm{d}p = \int_B g \mathrm{d}p.$$

- This function g is usually denoted by $\mathbb{E}(f|\Lambda)$.
- We clearly have $f \cdot p \ll p$ so the required g is simply $\frac{\mathrm{d}f \cdot p}{\mathrm{d}p|_\Lambda}$, where $p\mid_\Lambda$ is the restriction of p to the sub-σ-algebra Λ.

Properties of conditional expectation

1. The point of requiring Λ-measurability is that it "smooths out" variations that are too rapid to show up in Λ.

Properties of conditional expectation

1. The point of requiring Λ-measurability is that it “smooths out” variations that are too rapid to show up in Λ.
2. The conditional expectation is *linear*, *increasing* with respect to the pointwise order.

Properties of conditional expectation

1. The point of requiring Λ-measurability is that it "smooths out" variations that are too rapid to show up in Λ.
2. The conditional expectation is *linear*, *increasing* with respect to the pointwise order.
3. It is defined uniquely p-almost everywhere.

What are cones?

- Want to combine linear structure with order structure.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.
- Then for any positive $v \in P$ and positive real r, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.
- Then for any positive $v \in P$ and positive real r, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$.
- We *define* a **cone** C in a vector space V to be a set with exactly these conditions.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.
- Then for any positive $v \in P$ and positive real r, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$.
- We *define* a **cone** C in a vector space V to be a set with exactly these conditions.
- Any cone defines a order by $u \leq v$ if $v - u \in C$.

What are cones?

- Want to combine linear structure with order structure.
- If we have a vector space with an order $\leq$ we have a natural notion of *positive* and *negative* vectors: $x \geq 0$ is positive.
- What properties do the positive vectors have? Say $P \subset V$ are the positive vectors, we include 0.
- Then for any positive $v \in P$ and positive real r, $rv \in P$. For $u, v \in P$ we have $u + v \in P$ and if $v \in P$ and $-v \in P$ then $v = 0$.
- We *define* a **cone** C in a vector space V to be a set with exactly these conditions.
- Any cone defines a order by $u \leq v$ if $v - u \in C$.
- Unfortunately for us, many of the structures that we want to look at are cones but are not part of any obvious vector space: *e.g.* the measures on a space.

Cones that we use I

- If μ is a measure on X, then one has the well-known Banach spaces L_1 and L_∞.

Cones that we use I

- If μ is a measure on X, then one has the well-known Banach spaces L_1 and L_∞.
- These can be restricted to cones by considering the μ-almost everywhere positive functions.

Cones that we use I

- If μ is a measure on X, then one has the well-known Banach spaces L_1 and L_∞.
- These can be restricted to cones by considering the μ-almost everywhere positive functions.
- We will denote these cones by $L_1^+(X, \Sigma, \mu)$ and $L_\infty^+(X, \Sigma)$.

Cones that we use I

- If μ is a measure on X, then one has the well-known Banach spaces L_1 and L_∞.
- These can be restricted to cones by considering the μ-almost everywhere positive functions.
- We will denote these cones by $L_1^+(X, \Sigma, \mu)$ and $L_\infty^+(X, \Sigma)$.
- These are complete normed cones.

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p
- If q is such a measure, we define its norm to be $q(X)$.

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p
- If q is such a measure, we define its norm to be $q(X)$.
- $\mathcal{M}^{\ll p}(X)$ is also an ω-complete normed cone.

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p
- If q is such a measure, we define its norm to be $q(X)$.
- $\mathcal{M}^{\ll p}(X)$ is also an ω-complete normed cone.
- The cones $\mathcal{M}^{\ll p}(X)$ and $L_1^+(X, \Sigma, p)$ are isometrically isomorphic in ω**CC**.

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p
- If q is such a measure, we define its norm to be $q(X)$.
- $\mathcal{M}^{\ll p}(X)$ is also an ω-complete normed cone.
- The cones $\mathcal{M}^{\ll p}(X)$ and $L_1^+(X, \Sigma, p)$ are isometrically isomorphic in ω**CC**.
- We write $\mathcal{M}^p_{\mathsf{UB}}(X)$ for the cone of all measures on (X, Σ) that are uniformly less than a multiple of the measure p: $q \in \mathcal{M}^p_{\mathsf{UB}}$ means that for some real constant $K > 0$ we have $q \leq Kp$.

Cones that we use II

- Let (X, Σ, p) be a measure space with finite measure p. We denote by $\mathcal{M}^{\ll p}(X)$, the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p
- If q is such a measure, we define its norm to be $q(X)$.
- $\mathcal{M}^{\ll p}(X)$ is also an ω-complete normed cone.
- The cones $\mathcal{M}^{\ll p}(X)$ and $L_1^+(X, \Sigma, p)$ are isometrically isomorphic in ω**CC**.
- We write $\mathcal{M}^p_{\text{UB}}(X)$ for the cone of all measures on (X, Σ) that are uniformly less than a multiple of the measure p: $q \in \mathcal{M}^p_{\text{UB}}$ means that for some real constant $K > 0$ we have $q \leq Kp$.
- The cones $\mathcal{M}^p_{\text{UB}}(X)$ and $L_\infty^+(X, \Sigma, p)$ are isomorphic.

The pairing

Pairing function

There is a map from the product of the cones $L_\infty^+(X,p)$ and $L_1^+(X,p)$ to $\mathbb{R}^+$ defined as follows:

$$\forall f \in L_\infty^+(X,p), g \in L_1^+(X,p) \quad \langle f,\ g\rangle = \int fg\mathrm{d}p.$$

The pairing

Pairing function

There is a map from the product of the cones $L_\infty^+(X,p)$ and $L_1^+(X,p)$ to $\mathbb{R}^+$ defined as follows:

$$\forall f \in L_\infty^+(X,p), g \in L_1^+(X,p) \quad \langle f,\ g \rangle = \int fg\mathrm{d}p.$$

This map is bilinear and is continuous and ω-continuous in both arguments; we refer to it as the pairing.

Duality expressed via pairing

This pairing allows one to express the dualities in a very convenient way. For example, the isomorphism between $L_\infty^+(X,p)$ and $(L_1^+(X,p))^*$ sends $f \in L_\infty^+(X,p)$ to $\lambda g.\langle f,\ g\rangle = \lambda g.\int fg\mathrm{d}p$.

Duality is the Key

$$\begin{array}{ccccc} \mathcal{M}^{\ll p}(X) & \xleftrightarrow{\sim} & L_1^+(X,p) & \xleftrightarrow{\sim} & L_\infty^{+,*}(X,p) \\ \updownarrow & & \updownarrow & & \updownarrow \\ \mathcal{M}^p_{\text{UB}} & \xleftrightarrow{\sim} & L_\infty^+(X,p) & \xleftrightarrow{\sim} & L_1^{+,*}(X,p) \end{array} \tag{1}$$

where the vertical arrows represent dualities and the horizontal arrows represent isomorphisms.

Pairing function

There is a map from the product of the cones $L_\infty^+(X,p)$ and $L_1^+(X,p)$ to $\mathbb{R}^+$ defined as follows:

$$\forall f \in L_\infty^+(X,p), g \in L_1^+(X,p) \quad \langle f,\ g\rangle = \int fg\mathrm{d}p.$$

Where the action happens

- We define two categories $\mathbf{Rad}_\infty$ and $\mathbf{Rad}_1$ that will be needed for the functorial definition of conditional expectation.

Where the action happens

- We define two categories $\mathbf{Rad}_\infty$ and $\mathbf{Rad}_1$ that will be needed for the functorial definition of conditional expectation.
- This will allow for L_∞ and L_1 versions of the theory.

Where the action happens

- We define two categories $\mathbf{Rad}_\infty$ and $\mathbf{Rad}_1$ that will be needed for the functorial definition of conditional expectation.
- This will allow for L_∞ and L_1 versions of the theory.
- Going between these versions by duality will be very useful.

The "infinity" category

$\mathbf{Rad}_\infty$

The category $\mathbf{Rad}_\infty$ has as objects probability spaces, and as arrows $\alpha : (X, p) \longrightarrow (Y, q)$, measurable maps such that $M_\alpha(p) \leq Kq$ for some real number K.

The reason for choosing the name $\mathbf{Rad}_\infty$ is that $\alpha \in \mathbf{Rad}_\infty$ maps to $d/dqM_\alpha(p) \in L^+_\infty(Y, q)$.

The “one” category

$\mathbf{Rad}_1$

The category $\mathbf{Rad}_1$ has as objects probability spaces and as arrows $\alpha : (X, p) \longrightarrow (Y, q)$, measurable maps such that $M_\alpha(p) \ll q$.

The "one" category

$\mathbf{Rad}_1$

The category $\mathbf{Rad}_1$ has as objects probability spaces and as arrows $\alpha : (X, p) \longrightarrow (Y, q)$, measurable maps such that $M_\alpha(p) \ll q$.

1. The reason for choosing the name $\mathbf{Rad}_1$ is that $\alpha \in \mathbf{Rad}_1$ maps to $d/dqM_\alpha(p) \in L_1^+(Y, q)$.

The "one" category

$\mathbf{Rad}_1$

The category $\mathbf{Rad}_1$ has as objects probability spaces and as arrows $\alpha : (X, p) \longrightarrow (Y, q)$, measurable maps such that $M_\alpha(p) \ll q$.

1. The reason for choosing the name $\mathbf{Rad}_1$ is that $\alpha \in \mathbf{Rad}_1$ maps to $d/dqM_\alpha(p) \in L_1^+(Y, q)$.
2. The fact that the category $\mathbf{Rad}_\infty$ embeds in $\mathbf{Rad}_1$ reflects the fact that L_∞^+ embeds in L_1^+.

Pairing function revisited

Recall the isomorphism between $L_\infty^+(X,p)$ and $L_1^{+,*}(X,p)$ mediated by the pairing function:

$$f \in L_\infty^+(X,p) \mapsto \lambda g : L_1^+(X,p).\langle f,\ g\rangle = \int fg\mathrm{d}p.$$

Precomposition

1. Now, precomposition with α in $\mathbf{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y, q)$ to $L_1^+(X, p)$.

Precomposition

1. Now, precomposition with α in $\mathbf{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y,q)$ to $L_1^+(X,p)$.
2. Dually, given $\alpha \in \mathbf{Rad}_1 : (X,p) \longrightarrow (Y,q)$ and $g \in L_\infty^+(Y,q)$ we have that $P_\infty(\alpha)(g) \in L_\infty^+(X,p)$.

Precomposition

1. Now, precomposition with α in $\mathbf{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y,q)$ to $L_1^+(X,p)$.
2. Dually, given $\alpha \in \mathbf{Rad}_1 : (X,p) \longrightarrow (Y,q)$ and $g \in L_\infty^+(Y,q)$ we have that $P_\infty(\alpha)(g) \in L_\infty^+(X,p)$.
3. Thus the subscripts on the two precomposition functors describe the *target* categories.

Precomposition

1. Now, precomposition with α in $\mathbf{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y,q)$ to $L_1^+(X,p)$.
2. Dually, given $\alpha \in \mathbf{Rad}_1 : (X,p) \longrightarrow (Y,q)$ and $g \in L_\infty^+(Y,q)$ we have that $P_\infty(\alpha)(g) \in L_\infty^+(X,p)$.
3. Thus the subscripts on the two precomposition functors describe the *target* categories.
4. Using the $*$-functor we get a map $(P_1(\alpha))^*$ from $L_1^{+,*}(X,p)$ to $L_1^{+,*}(Y,q)$ in the first case and

Precomposition

1. Now, precomposition with α in $\mathbf{Rad}_\infty$ gives a map $P_1(\alpha)$ from $L_1^+(Y,q)$ to $L_1^+(X,p)$.
2. Dually, given $\alpha \in \mathbf{Rad}_1 : (X,p) \longrightarrow (Y,q)$ and $g \in L_\infty^+(Y,q)$ we have that $P_\infty(\alpha)(g) \in L_\infty^+(X,p)$.
3. Thus the subscripts on the two precomposition functors describe the *target* categories.
4. Using the $*$-functor we get a map $(P_1(\alpha))^*$ from $L_1^{+,*}(X,p)$ to $L_1^{+,*}(Y,q)$ in the first case and
5. dually we get $(P_\infty(\alpha))^*$ from $L_\infty^{+,*}(X,p)$ to $L_\infty^{+,*}(Y,q)$.

Expectation value functor

- The **functor** $\mathbb{E}_\infty(\cdot)$ is a functor from $\mathbf{Rad}_\infty$ to ω**CC** which, on objects, maps (X, p) to $L^+_\infty(X, p)$ and on maps is given as follows:

Expectation value functor

- The **functor** $\mathbb{E}_\infty(\cdot)$ is a functor from $\mathbf{Rad}_\infty$ to ω**CC** which, on objects, maps (X, p) to $L_\infty^+(X, p)$ and on maps is given as follows:
- Given $\alpha : (X, p) \longrightarrow (Y, q)$ in $\mathbf{Rad}_\infty$ the action of the functor is to produce the map $\mathbb{E}_\infty(\alpha) : L_\infty^+(X, p) \longrightarrow L_\infty^+(Y, q)$ obtained by composing $(P_1(\alpha))^*$ with the isomorphisms between $L_1^{+,*}$ and L_∞^+

Expectation value functor

- The **functor** $\mathbb{E}_\infty(\cdot)$ is a functor from $\mathbf{Rad}_\infty$ to ω**CC** which, on objects, maps (X, p) to $L_\infty^+(X, p)$ and on maps is given as follows:
- Given $\alpha : (X, p) \longrightarrow (Y, q)$ in $\mathbf{Rad}_\infty$ the action of the functor is to produce the map $\mathbb{E}_\infty(\alpha) : L_\infty^+(X, p) \longrightarrow L_\infty^+(Y, q)$ obtained by composing $(P_1(\alpha))^*$ with the isomorphisms between $L_1^{+,*}$ and L_∞^+

$$\begin{array}{ccc} L_1^{+,*}(X,p) & \longleftarrow & L_\infty^+(X,p) \\ \downarrow{\scriptstyle (P_1(\alpha))^*} & & \downarrow{\scriptstyle \mathbb{E}_\infty(\alpha)} \\ L_1^{+,*}(Y,q) & \longrightarrow & L_\infty^+(Y,q) \end{array}$$

Consequences

1. It is an immediate consequence of the definitions that for any $f \in L_\infty^+(X,p)$ and $g \in L_1(Y,q)$

$$\langle \mathbb{E}_\infty(\alpha)(f), g\rangle_Y = \langle f, P_1(\alpha)(g)\rangle_X.$$

Consequences

1. It is an immediate consequence of the definitions that for any $f \in L_\infty^+(X,p)$ and $g \in L_1(Y,q)$

$$\langle \mathbb{E}_\infty(\alpha)(f), g\rangle_Y = \langle f, P_1(\alpha)(g)\rangle_X.$$

$$\begin{array}{ccc}
\lambda h : L_1^+(X,p).\langle f,\ h\rangle & \longleftarrow\!\dashv & f \\
\downarrow & & \vdots \\
\lambda g : L_1^+(Y,q).\langle f,\ g \circ \alpha\rangle & \longmapsto & \mathbb{E}_\infty(\alpha)(f)
\end{array}$$

Consequences

1. It is an immediate consequence of the definitions that for any $f \in L_\infty^+(X,p)$ and $g \in L_1(Y,q)$

$$\langle \mathbb{E}_\infty(\alpha)(f), g\rangle_Y = \langle f, P_1(\alpha)(g)\rangle_X.$$

$$\begin{array}{ccc} \lambda h : L_1^+(X,p).\langle f,\ h\rangle & \longleftarrow\!\!\!\dashv & f \\ \downarrow & & \vdots \\ \lambda g : L_1^+(Y,q).\langle f,\ g\circ\alpha\rangle & \longmapsto & \mathbb{E}_\infty(\alpha)(f) \end{array}$$

2. Note that since we started with α in $\mathbf{Rad}_\infty$ we get the expectation value as a map between the L_∞^+ cones.

The other expectation value functor

The **functor** $\mathbb{E}_{\mathbf{1}}(\cdot)$ is a functor from $\mathbf{Rad}_1$ to ω**CC** which maps the object (X,p) to $L_1^+(X,p)$ and on maps is given as follows:
Given $\alpha : (X,p) \longrightarrow (Y,q)$ in $\mathbf{Rad}_1$ the action of the functor is to produce the map $\mathbb{E}_1(\alpha) : L_1^+(X,p) \longrightarrow L_1^+(Y,q)$ obtained by composing $(P_\infty(\alpha))^*$ with the isomorphisms between $L_\infty^{+,*}$ and L_1^+ as shown in the diagram below

$$\begin{array}{ccc}
L_\infty^{+,*}(X,p) & \longleftarrow & L_1^+(X,p) \\
\Big\downarrow{\scriptstyle (P_\infty(\alpha))^*} & & \Big\downarrow{\scriptstyle \mathbb{E}_1(\alpha)} \\
L_\infty^{+,*}(Y,q) & \longrightarrow & L_1^+(Y,q)
\end{array}$$

Markov kernels as linear maps

1. Given τ a Markov kernel from (X, Σ) to (Y, Λ), we define $T_\tau : \mathcal{L}^+(Y) \longrightarrow \mathcal{L}^+(X)$, for $f \in \mathcal{L}^+(Y)$, $x \in X$, as $T_\tau(f)(x) = \int_Y f(z)\tau(x, dz)$.

Markov kernels as linear maps

1. Given τ a Markov kernel from (X, Σ) to (Y, Λ), we define $T_\tau : \mathcal{L}^+(Y) \longrightarrow \mathcal{L}^+(X)$, for $f \in \mathcal{L}^+(Y)$, $x \in X$, as $T_\tau(f)(x) = \int_Y f(z)\tau(x, dz)$.
2. This map is well-defined, linear and ω-continuous.

Markov kernels as linear maps

1. Given τ a Markov kernel from (X, Σ) to (Y, Λ), we define $T_\tau : \mathcal{L}^+(Y) \longrightarrow \mathcal{L}^+(X)$, for $f \in \mathcal{L}^+(Y)$, $x \in X$, as $T_\tau(f)(x) = \int_Y f(z)\tau(x, dz)$.
2. This map is well-defined, linear and ω-continuous.
3. If we write $\mathbf{1}_B$ for the indicator function of the measurable set B we have that $T_\tau(\mathbf{1}_B)(x) = \tau(x, B)$.

Markov kernels as linear maps

1. Given τ a Markov kernel from (X, Σ) to (Y, Λ), we define $T_\tau : \mathcal{L}^+(Y) \longrightarrow \mathcal{L}^+(X)$, for $f \in \mathcal{L}^+(Y)$, $x \in X$, as $T_\tau(f)(x) = \int_Y f(z)\tau(x, dz)$.
2. This map is well-defined, linear and ω-continuous.
3. If we write $\mathbf{1}_B$ for the indicator function of the measurable set B we have that $T_\tau(\mathbf{1}_B)(x) = \tau(x, B)$.
4. It encodes all the transition probability information

From linear maps to markov kernels

1. Conversely, any ω-continuous morphism L with $L(\mathbf{1}_Y) \leq \mathbf{1}_X$ can be cast as a Markov kernel by reversing the process on the last slide.

From linear maps to markov kernels

1. Conversely, any ω-continuous morphism L with $L(\mathbf{1}_Y) \leq \mathbf{1}_X$ can be cast as a Markov kernel by reversing the process on the last slide.
2. The interpretation of L is that $L(\mathbf{1}_B)$ is a measurable function on X such that $L(\mathbf{1}_B)(x)$ is the probability of jumping from x to B.

Backwards

1. We can also define an operator on $\mathcal{M}(X)$ by using τ the other way.

Backwards

1. We can also define an operator on $\mathcal{M}(X)$ by using τ the other way.
2. We define $\bar{T}_\tau : \mathcal{M}(X) \longrightarrow \mathcal{M}(Y)$, for $\mu \in \mathcal{M}(X)$ and $B \in \Lambda$, as $\bar{T}_\tau(\mu)(B) = \int_X \tau(x, B)\,\mathrm{d}\mu(x)$.

Backwards

1. We can also define an operator on $\mathcal{M}(X)$ by using τ the other way.
2. We define $\bar{T}_\tau : \mathcal{M}(X) \longrightarrow \mathcal{M}(Y)$, for $\mu \in \mathcal{M}(X)$ and $B \in \Lambda$, as $\bar{T}_\tau(\mu)(B) = \int_X \tau(x, B)\, \mathrm{d}\mu(x)$.
3. It is easy to show that this map is linear and ω-continuous.

What do they mean?

1. The operator $\bar{T}_\tau$ transforms measures “forwards in time”; if μ is a measure on X representing the current state of the system, $\bar{T}_\tau(\mu)$ is the resulting measure on Y after a transition through τ.

What do they mean?

1. The operator $\bar{T}_\tau$ transforms measures “forwards in time”; if μ is a measure on X representing the current state of the system, $\bar{T}_\tau(\mu)$ is the resulting measure on Y after a transition through τ.
2. The operator T_τ may be interpreted as a likelihood transformer which propagates information “backwards”, just as we expect from predicate transformers.

What do they mean?

1. The operator $\bar{T}_\tau$ transforms measures “forwards in time”; if μ is a measure on X representing the current state of the system, $\bar{T}_\tau(\mu)$ is the resulting measure on Y after a transition through τ.
2. The operator T_τ may be interpreted as a likelihood transformer which propagates information “backwards”, just as we expect from predicate transformers.
3. $T_\tau(f)(x)$ is just the expected value of f after one τ-step given that one is at x.

Labelled abstract Markov processes

The definition

An **abstract Markov kernel** from (X, Σ, p) to (Y, Λ, q) is an ω-continuous linear map $\tau : L_\infty^+(Y) \to L_\infty^+(X)$ with $\|\tau\| \leq 1$.

Labelled abstract Markov processes

The definition

An **abstract Markov kernel** from (X, Σ, p) to (Y, Λ, q) is an ω-continuous linear map $\tau : L^+_\infty(Y) \longrightarrow L^+_\infty(X)$ with $\|\tau\| \leq 1$.

LAMPS

A **labelled abstract Markov process** on a probability space (X, Σ, p) with a set of labels (or actions) $\mathcal{A}$ is a family of abstract Markov kernels $\tau_a : L^+_\infty(X, p) \longrightarrow L^+_\infty(X, p)$ indexed by elements a of $\mathcal{A}$.

The approximation map

The expectation value functors project a probability space onto another one with a possibly coarser σ-algebra. Given an AMP on (X, p) and a map $\alpha : (X, p) \longrightarrow (Y, q)$ in $\mathbf{Rad}_\infty$, we have the following approximation scheme:

Approximation scheme

$$
\begin{array}{ccc}
L_\infty^+(X,p) & \xrightarrow{\ \tau_a\ } & L_\infty^+(X,p) \\
{\scriptstyle P_\infty(\alpha)}\Big\uparrow & & \Big\downarrow{\scriptstyle \mathbb{E}_\infty(\alpha)} \\
L_\infty^+(Y,q) & \xrightarrow{\ \alpha(\tau_a)\ } & L_\infty^+(Y,q)
\end{array}
$$

A special case

- Take (X, Σ) and (X, Λ) with $\Lambda \subset \Sigma$ and use the measurable function $id : (X, \Sigma) \longrightarrow (X, \Lambda)$ as α.

A special case

- Take (X, Σ) and (X, Λ) with $\Lambda \subset \Sigma$ and use the measurable function $id : (X, \Sigma) \longrightarrow (X, \Lambda)$ as α.

Coarsening the σ-algebra

$$\begin{array}{ccc} L_\infty^+(X, \Sigma, p) & \xrightarrow{\tau_a} & L_\infty^+(X, \Sigma, p) \\ {\scriptstyle P_\infty(id)} \uparrow & & \downarrow {\scriptstyle \mathbb{E}_\infty(id)} \\ L_\infty^+(X, \Lambda, p) & \xrightarrow{id(\tau_a)} & L_\infty^+(X, \Lambda, p) \end{array}$$

A special case

- Take (X, Σ) and (X, Λ) with $\Lambda \subset \Sigma$ and use the measurable function $id : (X, \Sigma) \longrightarrow (X, \Lambda)$ as α.

Coarsening the σ-algebra

$$\begin{array}{ccc} L^+_\infty(X, \Sigma, p) & \xrightarrow{\tau_a} & L^+_\infty(X, \Sigma, p) \\ {\scriptstyle P_\infty(id)} \uparrow & & \downarrow {\scriptstyle \mathbb{E}_\infty(id)} \\ L^+_\infty(X, \Lambda, p) & \overset{id(\tau_a)}{\dashrightarrow} & L^+_\infty(X, \Lambda, p) \end{array}$$

- Thus $id(\tau_a)$ is the approximation of τ_a obtained by averaging over the sets of the coarser σ-algebra Λ.

A special case

- Take (X, Σ) and (X, Λ) with $\Lambda \subset \Sigma$ and use the measurable function $id : (X, \Sigma) \longrightarrow (X, \Lambda)$ as α.

Coarsening the σ-algebra

$$\begin{array}{ccc} L^+_\infty(X, \Sigma, p) & \xrightarrow{\tau_a} & L^+_\infty(X, \Sigma, p) \\ {\scriptstyle P_\infty(id)} \uparrow & & \downarrow {\scriptstyle \mathbb{E}_\infty(id)} \\ L^+_\infty(X, \Lambda, p) & \overset{id(\tau_a)}{\dashrightarrow} & L^+_\infty(X, \Lambda, p) \end{array}$$

- Thus $id(\tau_a)$ is the approximation of τ_a obtained by averaging over the sets of the coarser σ-algebra Λ.
- We now have the machinery to consider approximating along arbitrary maps α.

Bisimulation traditionally

Larsen-Skou definition

Given an **LMP** (S, Σ, τ_a) an equivalence relation R on S is called a *probabilistic bisimulation* if sRt then for every *measurable* R-closed set C we have for every a

$$\tau_a(s, C) = \tau_a(t, C).$$

This variation to the continuous case is due to Josée Desharnais and her Indian friends.

Event bisimulation

- In measure theory one should focus on measurable sets rather than on *points*.

Event bisimulation

- In measure theory one should focus on measurable sets rather than on *points*.

Event Bisimulation

Given a LMP (X, Σ, τ_a), an **event-bisimulation** is a sub-σ-algebra Λ of Σ such that (X, Λ, τ_a) is still an LMP.

Event bisimulation

- In measure theory one should focus on measurable sets rather than on *points*.

Event Bisimulation

Given a LMP (X, Σ, τ_a), an **event-bisimulation** is a sub-σ-algebra Λ of Σ such that (X, Λ, τ_a) is still an LMP.

- This means τ_a sends the subspace $L_\infty^+(X, \Lambda, p)$ to itself; where we are now viewing τ_a as a map on $L_\infty^+(X, \Lambda, p)$.

The bisimulation diagram

$$\begin{array}{ccc} L_\infty^+(X,\Sigma,p) & \xrightarrow{\tau_a} & L_\infty^+(X,\Sigma,p) \\ \uparrow & & \uparrow \\ L_\infty^+(X,\Lambda,p) & \xrightarrow{\tau_a} & L_\infty^+(X,\Lambda,p) \end{array}$$

This is a “lossless” approximation!

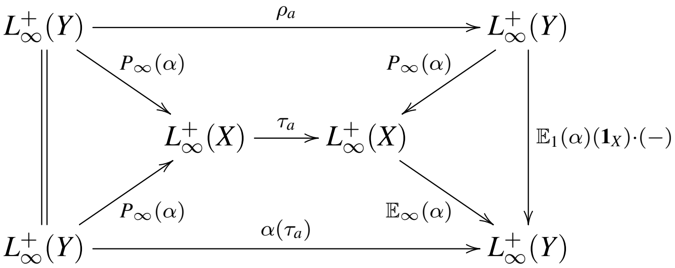
Zigzag maps

We can generalize the notion of event bisimulation by using maps other than the identity map on the underlying sets. This would be a map α from (X, Σ, p) to (Y, Λ, q), equipped with LMPs τ_a and ρ_a respectively, such that the following commutes:

$$\begin{array}{ccc} L^+_\infty(X,\Sigma,p) & \xrightarrow{\tau_a} & L^+_\infty(X,\Sigma,p) \\ \uparrow P_\infty(\alpha) & & \uparrow P_\infty(\alpha) \\ L^+_\infty(Y,\Lambda,q) & \xrightarrow{\rho_a} & L^+_\infty(Y,\Lambda,q) \end{array} \quad (2)$$

A key diagram

When we have a zigzag the following diagram commutes:

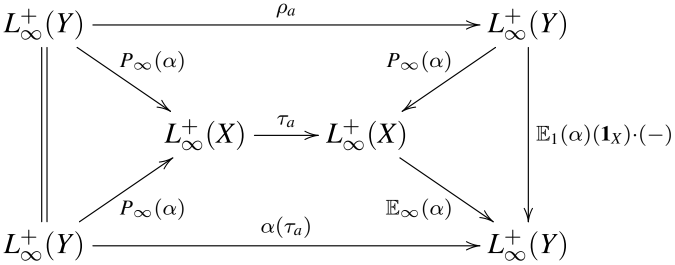


(3)

- The upper trapezium says we have a zigzag. The lower trapezium says that we have an “approximation” and the triangle on the right is an earlier lemma.

A key diagram

When we have a zigzag the following diagram commutes:



(3)

- The upper trapezium says we have a zigzag. The lower trapezium says that we have an “approximation” and the triangle on the right is an earlier lemma.
- If we “approximate” along a zigzag we actually get the exact result.

A key diagram

When we have a zigzag the following diagram commutes:

$$\begin{array}{ccccc}
L_\infty^+(Y) & & \xrightarrow{\quad\rho_a\quad} & & L_\infty^+(Y) \\
\| & \searrow^{P_\infty(\alpha)} & & \swarrow^{P_\infty(\alpha)} & \downarrow \\
\| & L_\infty^+(X) & \xrightarrow{\tau_a} & L_\infty^+(X) & \downarrow {\scriptstyle \mathbb{E}_1(\alpha)(\mathbf{1}_X)\cdot(-)} \\
\| & \nearrow_{P_\infty(\alpha)} & & \searrow_{\mathbb{E}_\infty(\alpha)} & \downarrow \\
L_\infty^+(Y) & & \xrightarrow{\quad\alpha(\tau_a)\quad} & & L_\infty^+(Y)
\end{array} \qquad (3)$$

- The upper trapezium says we have a zigzag. The lower trapezium says that we have an “approximation” and the triangle on the right is an earlier lemma.
- If we “approximate” along a zigzag we actually get the exact result.
- Approximations are approximate bisimulations.

Bisimulation as a cospan

- Zigzags give a “functional” version of bisimulation; what is the relational version.

Bisimulation as a cospan

- Zigzags give a “functional” version of bisimulation; what is the relational version.
- Use co-spans of zigzags; it is usual to use spans but co-spans give a smoother and more general theory.

Bisimulation as a cospan

- Zigzags give a “functional” version of bisimulation; what is the relational version.
- Use co-spans of zigzags; it is usual to use spans but co-spans give a smoother and more general theory.
- With spans one can prove logical characterization of bisimulation on analytic spaces.

Bisimulation as a cospan

- Zigzags give a “functional” version of bisimulation; what is the relational version.
- Use co-spans of zigzags; it is usual to use spans but co-spans give a smoother and more general theory.
- With spans one can prove logical characterization of bisimulation on analytic spaces.
- With the cospan definition we get logical characterization on *all* measurable spaces.

Bisimulation as a cospan

- Zigzags give a "functional" version of bisimulation; what is the relational version.
- Use co-spans of zigzags; it is usual to use spans but co-spans give a smoother and more general theory.
- With spans one can prove logical characterization of bisimulation on analytic spaces.
- With the cospan definition we get logical characterization on *all* measurable spaces.
- On analytic spaces the two concepts co-incide.

Bisimulation as a cospan

- Zigzags give a "functional" version of bisimulation; what is the relational version.
- Use co-spans of zigzags; it is usual to use spans but co-spans give a smoother and more general theory.
- With spans one can prove logical characterization of bisimulation on analytic spaces.
- With the cospan definition we get logical characterization on *all* measurable spaces.
- On analytic spaces the two concepts co-incide.
- Recent results show that the theory cannot be made to work with spans on general measurable spaces.

The official definition of bisimulation

Bisimulation

We say that two objects of $\mathbf{AMP}$, (X, Σ, p, τ) and (Y, Λ, q, ρ), are *bisimilar* if there is a third object (Z, Γ, r, π) with a pair of zigzags

$$\alpha : (X, \Sigma, p, \tau) \longrightarrow (Z, \Gamma, r, \pi)$$
$$\beta : (Y, \Lambda, q, \rho) \longrightarrow (Z, \Gamma, r, \pi)$$

giving a cospan diagram

$$\begin{array}{ccc} (X, \Sigma, p, \tau) & & (Y, \Lambda, q, \rho) \\ & \searrow^{\alpha} \quad {}^{\beta}\swarrow & \\ & (Z, \Gamma, r, \pi) & \end{array} \tag{4}$$

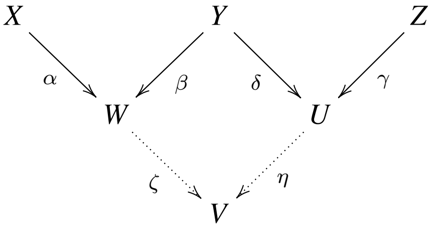
Note that the identity function on an AMP is a zigzag, so if a zigzag exists the two AMPs are bisimilar.

Fundamental categorical result

The category $\mathbf{AMP}$ has pushouts

Furthermore, if the morphisms in the span are zigzags then the morphisms in the pushout diagram are also zigzags.

Bisimulation is an equivalence



(5)

The pushouts of the zigzags β and δ yield two more zigzags ζ and η (and the pushout object V). As the composition of two zigzags is a zigzag, X and Z are bisimilar. Thus bisimulation is transitive.

What did we do with this theory?

1. We showed logical characterization of bisimulation for any measurable space.

What did we do with this theory?

1. We showed logical characterization of bisimulation for any measurable space.
2. We developed a theory of approximation by looking at finitely generated sub-σ-algebras coming form the logic: approximate bisimulations.

What did we do with this theory?

1. We showed logical characterization of bisimulation for any measurable space.
2. We developed a theory of approximation by looking at finitely generated sub-σ-algebras coming form the logic: approximate bisimulations.
3. We showed that there is a *canonical* minimal realization that arises as the projective limit of the finite approximations.