Markov Processes as Function Transformers

Prakash Panangaden[1]

[1]School of Computer Science
McGill University

Dagstuhl Oct 2012

Outline

1. Introduction

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

What am I trying to do?

1. Present a “new” view of Markov processes as function transformers

What am I trying to do?

1. Present a “new” view of Markov processes as function transformers
2. Show a beautiful functorial presentation of expectation values

What am I trying to do?

1. Present a “new” view of Markov processes as function transformers
2. Show a beautiful functorial presentation of expectation values
3. Make bisimulation and approximation live in the same universe

What am I trying to do?

1. Present a “new” view of Markov processes as function transformers
2. Show a beautiful functorial presentation of expectation values
3. Make bisimulation and approximation live in the same universe
4. Minimal realization theory

What am I trying to do?

1. Present a "new" view of Markov processes as function transformers
2. Show a beautiful functorial presentation of expectation values
3. Make bisimulation and approximation live in the same universe
4. Minimal realization theory
5. Approximation

What I am not trying to do

- Review all the previous work

What I am not trying to do

- Review all the previous work
- Discuss metrics

What I am not trying to do

- Review all the previous work
- Discuss metrics
- Prove everything in detail

What I am not trying to do

- Review all the previous work
- Discuss metrics
- Prove everything in detail
- Deal with continuous time

What I am not trying to do

- Review all the previous work
- Discuss metrics
- Prove everything in detail
- Deal with continuous time
- Deal with nondeterminism

What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.

What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* - no probabilities associated with environment behaviour.

What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* - no probabilities associated with environment behaviour.
- We observe the interactions - not the internal states.

What are labelled Markov processes?

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* - no probabilities associated with environment behaviour.
- We observe the interactions - not the internal states.
- **In general, the state space of a labelled Markov process may be a *continuum*.**

Formal Definition of LMPs

- An LMP is a tuple $(S, \Sigma, \mathsf{L}, \forall\alpha \in \mathsf{L}.\tau_\alpha)$ where $\tau_\alpha : S \times \Sigma \longrightarrow [0, 1]$ is a *transition probability* function such that

Formal Definition of LMPs

- An LMP is a tuple $(S, \Sigma, \mathsf{L}, \forall\alpha \in \mathsf{L}.\tau_\alpha)$ where $\tau_\alpha : S \times \Sigma \longrightarrow [0,1]$ is a *transition probability* function such that
- $\forall s : S.\lambda A : \Sigma.\tau_\alpha(s, A)$ is a subprobability measure
  and
  $\forall A : \Sigma.\lambda s : S.\tau_\alpha(s, A)$ is a measurable function.

Logical Characterization

- $$\mathcal{L}_0 ::== \mathsf{T} | \phi_1 \wedge \phi_2 | \langle a \rangle_q \phi$$

Logical Characterization

- $$\mathcal{L}_0 ::== \mathsf{T} | \phi_1 \wedge \phi_2 | \langle a \rangle_q \phi$$
- We say $s \models \langle a \rangle_q \phi$ iff
$$\exists A \in \Sigma.(\forall s' \in A.s' \models \phi) \wedge (\tau_a(s, A) > q).$$

Logical Characterization

- $$\mathcal{L}_0 ::== \mathsf{T} | \phi_1 \wedge \phi_2 | \langle a \rangle_q \phi$$
- We say $s \models \langle a \rangle_q \phi$ iff
$$\exists A \in \Sigma.(\forall s' \in A.s' \models \phi) \wedge (\tau_a(s, A) > q).$$
- Two systems are bisimilar iff they obey the same formulas of $\mathcal{L}$. [DEP 1998 LICS, I and C 2002]

State transformers and predicate transformers

- A transition system $(S, \mathcal{A}, \rightarrow)$ has a natural interpretation as a *state transformer*.

State transformers and predicate transformers

- A transition system $(S, \mathcal{A}, \rightarrow)$ has a natural interpretation as a *state transformer*.
- Given $s \in S$ and $a \in \mathcal{A}$ we have $F(s)(a) = \{s' \mid s \xrightarrow{a} s'\}$.

State transformers and predicate transformers

- A transition system $(S, \mathcal{A}, \rightarrow)$ has a natural interpretation as a *state transformer*.
- Given $s \in S$ and $a \in \mathcal{A}$ we have $F(s)(a) = \{s' \mid s \xrightarrow{a} s'\}$.
- We can extend F to $Q \subset S$ by direct image.

State transformers and predicate transformers

- A transition system $(S, \mathcal{A}, \rightarrow)$ has a natural interpretation as a *state transformer*.
- Given $s \in S$ and $a \in \mathcal{A}$ we have $F(s)(a) = \{s' \mid s \xrightarrow{a} s'\}$.
- We can extend F to $Q \subset S$ by direct image.
- We can also define *predicate transformers*: given $P \subset S$ we have $\mathsf{wp}(a)(P) = \{s' \mid s' \xrightarrow{a} s\}$.

State transformers and predicate transformers

- A transition system $(S, \mathcal{A}, \rightarrow)$ has a natural interpretation as a *state transformer*.
- Given $s \in S$ and $a \in \mathcal{A}$ we have $F(s)(a) = \{s' \mid s \xrightarrow{a} s'\}$.
- We can extend F to $Q \subset S$ by direct image.
- We can also define *predicate transformers*: given $P \subset S$ we have $\mathrm{wp}(a)(P) = \{s' \mid s' \xrightarrow{a} s\}$.
- Here the flow is backward.

State transformers and predicate transformers

- A transition system $(S, \mathcal{A}, \rightarrow)$ has a natural interpretation as a *state transformer*.
- Given $s \in S$ and $a \in \mathcal{A}$ we have $F(s)(a) = \{s' \mid s \xrightarrow{a} s'\}$.
- We can extend F to $Q \subset S$ by direct image.
- We can also define *predicate transformers*: given $P \subset S$ we have $\mathsf{wp}(a)(P) = \{s' \mid s' \xrightarrow{a} s\}$.
- Here the flow is backward.
- There is a duality between state-transformer and predicate-transformer semantics.

State transformers and predicate transformers

- A transition system $(S, \mathcal{A}, \rightarrow)$ has a natural interpretation as a *state transformer*.
- Given $s \in S$ and $a \in \mathcal{A}$ we have $F(s)(a) = \{s' \mid s \xrightarrow{a} s'\}$.
- We can extend F to $Q \subset S$ by direct image.
- We can also define *predicate transformers*: given $P \subset S$ we have $\text{wp}(a)(P) = \{s' \mid s' \xrightarrow{a} s\}$.
- Here the flow is backward.
- There is a duality between state-transformer and predicate-transformer semantics.
- Here one is thinking of a "predicate" as simply a subset of S, but such a subset can be described by a logical formula.

Logic and Probability

Classical logic	Generalization
Truth values $\{0, 1\}$	Probabilities $[0, 1]$
Predicate	Random variable
State	Distribution
The satisfaction relation $\models$	Integration $\int$

The “predicate transformer” view of Markov processes

- Recall, a Markov kernel τ from (X, Σ) to (Y, Λ) is a map $\tau : X \times \Lambda \longrightarrow [0, 1]$ which is measurable in its first argument and a (subprobability) measure in the second argument.

The "predicate transformer" view of Markov processes

- Recall, a Markov kernel τ from (X, Σ) to (Y, Λ) is a map $\tau : X \times \Lambda \to [0, 1]$ which is measurable in its first argument and a (subprobability) measure in the second argument.
- Let f be a real-valued function on Y. We define $B_\tau(f)(x) = \int_Y f(y)\tau(x, dy)$; this is playing the role analogous to a predicate transformer. It is in fact an expectation transformer.

The "predicate transformer" view of Markov processes

- Recall, a Markov kernel τ from (X, Σ) to (Y, Λ) is a map $\tau : X \times \Lambda \longrightarrow [0, 1]$ which is measurable in its first argument and a (subprobability) measure in the second argument.
- Let f be a real-valued function on Y. We define $B_\tau(f)(x) = \int_Y f(y)\tau(x, dy)$; this is playing the role analogous to a predicate transformer. It is in fact an expectation transformer.
- $B_\tau(f)(x)$ is the expectation value of f after a step given that one was at x.

The "predicate transformer" view of Markov processes

- Recall, a Markov kernel τ from (X, Σ) to (Y, Λ) is a map $\tau : X \times \Lambda \longrightarrow [0, 1]$ which is measurable in its first argument and a (subprobability) measure in the second argument.
- Let f be a real-valued function on Y. We define $B_\tau(f)(x) = \int_Y f(y)\tau(x, dy)$; this is playing the role analogous to a predicate transformer. It is in fact an expectation transformer.
- $B_\tau(f)(x)$ is the expectation value of f after a step given that one was at x.
- We can also define an analogue of the forward transformer.

The "predicate transformer" view of Markov processes

- Recall, a Markov kernel τ from (X, Σ) to (Y, Λ) is a map $\tau : X \times \Lambda \to [0,1]$ which is measurable in its first argument and a (subprobability) measure in the second argument.
- Let f be a real-valued function on Y. We define $B_\tau(f)(x) = \int_Y f(y)\tau(x, dy)$; this is playing the role analogous to a predicate transformer. It is in fact an expectation transformer.
- $B_\tau(f)(x)$ is the expectation value of f after a step given that one was at x.
- We can also define an analogue of the forward transformer.
- $F_\tau(\mu)(D \in \Lambda) = \int_X \tau(x, D)\mathrm{d}\mu(x)$.

The “predicate transformer” view of Markov processes

- Recall, a Markov kernel τ from (X, Σ) to (Y, Λ) is a map $\tau : X \times \Lambda \to [0, 1]$ which is measurable in its first argument and a (subprobability) measure in the second argument.
- Let f be a real-valued function on Y. We define $B_\tau(f)(x) = \int_Y f(y)\tau(x, dy)$; this is playing the role analogous to a predicate transformer. It is in fact an expectation transformer.
- $B_\tau(f)(x)$ is the expectation value of f after a step given that one was at x.
- We can also define an analogue of the forward transformer.
- $F_\tau(\mu)(D \in \Lambda) = \int_X \tau(x, D)\mathrm{d}\mu(x)$.
- If μ is the measure representing the “current” distribution on X then after a τ-step, $F_\tau(\mu)$ is the distribution on Y.

The general plan

- We are going to view Markov processes as function transformers rather than as state transformers.

The general plan

- We are going to view Markov processes as function transformers rather than as state transformers.
- We will take the backward view; we could, perhaps equally well, have developed a forward view but we have not spelled out the details.

The general plan

- We are going to view Markov processes as function transformers rather than as state transformers.
- We will take the backward view; we could, perhaps equally well, have developed a forward view but we have not spelled out the details.
- Measure theory works much better when one deals with measurable functions rather than “points” and measures.

The general plan

- We are going to view Markov processes as function transformers rather than as state transformers.
- We will take the backward view; we could, perhaps equally well, have developed a forward view but we have not spelled out the details.
- Measure theory works much better when one deals with measurable functions rather than “points” and measures.
- We never have to worry about “almost everywhere” and other such nonsense.

The general plan

- We are going to view Markov processes as function transformers rather than as state transformers.
- We will take the backward view; we could, perhaps equally well, have developed a forward view but we have not spelled out the details.
- Measure theory works much better when one deals with measurable functions rather than “points” and measures.
- We never have to worry about “almost everywhere” and other such nonsense.
- Because of our backward view, bisimulation becomes a cospan instead of a span. But this actually makes everything easier!

The general plan

- We are going to view Markov processes as function transformers rather than as state transformers.
- We will take the backward view; we could, perhaps equally well, have developed a forward view but we have not spelled out the details.
- Measure theory works much better when one deals with measurable functions rather than "points" and measures.
- We never have to worry about "almost everywhere" and other such nonsense.
- Because of our backward view, bisimulation becomes a cospan instead of a span. But this actually makes everything easier!
- We can develop a theory of bisimulation, logical characterization, approximation and minimal realization in this framework.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \,\mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \, \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space $L_1(X, \Sigma, \mu)$ or just $L_1(\mu)$.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space $L_1(X, \Sigma, \mu)$ or just $L_1(\mu)$.
- The space $L_p(\mu)$ is the space obtained by defining the norm $\|f\|_p = (\int \mid f \mid^p \mathrm{d}\mu)^{\frac{1}{p}}$, where $0 < p < \infty$.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space $L_1(X, \Sigma, \mu)$ or just $L_1(\mu)$.
- The space $L_p(\mu)$ is the space obtained by defining the norm $\|f\|_p = (\int \mid f \mid^p \mathrm{d}\mu)^{\frac{1}{p}}$, where $0 < p < \infty$.
- The *infinity norm* of a measurable function f is $\|f\|_\infty = \inf\{M > 0 \mid \mid f(x) \mid \leq M$ for $\mu -$ almost all $x\}$.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space $L_1(X, \Sigma, \mu)$ or just $L_1(\mu)$.
- The space $L_p(\mu)$ is the space obtained by defining the norm $\|f\|_p = (\int |f|^p \mathrm{d}\mu)^{\frac{1}{p}}$, where $0 < p < \infty$.
- The *infinity norm* of a measurable function f is $\|f\|_\infty = \inf\{M > 0 \mid |f(x)| \leq M$ for $\mu -$ almost all $x\}$.
- The collection of all equivalence classes of measurable functions f with $\|f\|_\infty < \infty$ with the norm just defined is the space $L_\infty(\mu)$.

L_p spaces

- If (X, Σ, μ) is a measure space we can define integration on X: we write $\int_X f \mathrm{d}x$. We say that f is *integrable* if this is finite.
- If two functions agree everywhere except on a set of μ-measure 0 their integrals will be equal.
- We define two functions to be equivalent if they are μ-almost everywhere the same and we actually work with equivalence classes.
- The integral defines a norm on these equivalence classes and gives the Banach space $L_1(X, \Sigma, \mu)$ or just $L_1(\mu)$.
- The space $L_p(\mu)$ is the space obtained by defining the norm $\|f\|_p = (\int \mid f \mid^p \mathrm{d}\mu)^{\frac{1}{p}}$, where $0 < p < \infty$.
- The *infinity norm* of a measurable function f is $\|f\|_\infty = \inf\{M > 0 \mid \mid f(x) \mid \leq M$ for $\mu -$ almost all $x\}$.
- The collection of all equivalence classes of measurable functions f with $\|f\|_\infty < \infty$ with the norm just defined is the space $L_\infty(\mu)$.
- These are all Banach spaces.

Duality for L_p spaces

- If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then L_p and L_q are duals of each other!

Duality for L_p spaces

- If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then L_p and L_q are duals of each other!
- However, L_1 and L_∞ are *not duals*.

Duality for L_p spaces

- If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then L_p and L_q are duals of each other!
- However, L_1 and L_∞ are *not duals*.
- The dual of L_1 is L_∞ but not the other way around!

Duality for L_p spaces

- If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then L_p and L_q are duals of each other!
- However, L_1 and L_∞ are *not duals*.
- The dual of L_1 is L_∞ but not the other way around!
- We will switch to a cone view and the situation will be much improved.

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.

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.
- We could artificially embed them in a vector space, for example, by introducing signed measures.

Abstract cones d'après Selinger

Definition of Cones

A **cone** is a commutative monoid $(V, +, 0)$ with an action of $\mathbb{R}^{\geq 0}$. Multiplication by reals distributes over addition and the following cancellation law holds:

$$\forall u, v, w \in V, v + u = w + u \Rightarrow v = w.$$

The following strictness property also holds:

$$v + w = 0 \Rightarrow v = w = 0.$$

Note that every cone comes with a natural order.

An order on a cone

If $u, v \in V$, a cone, one says $u \leq v$ if and only if there is an element $w \in V$ such that $u + w = v$.

Normed cones

Definition of a normed cone

A **normed cone** C is a cone with a function $||\cdot|| : C \longrightarrow \mathbb{R}^{\geq 0}$ satisfying the usual conditions:
$||v|| = 0$ if and only if $v = 0$
$\forall r \in \mathbb{R}^{\geq 0}, v \in C, ||r \cdot v|| = r||v||$
$||u + v|| \leq ||u|| + ||v||$
$u \leq v \Rightarrow ||u|| \leq ||v||$.

Normally one uses norms to talk about convergence of Cauchy sequences. But without negation how can we talk about Cuchy sequences?

Completeness

However, order-theoretic concepts can be used instead.

Complete normed cones

An ω-**complete normed cone** is a normed cone such that if $\{a_i \mid i \in I\}$ is an increasing sequence with $\{||a_i||\}$ bounded then the lub $\bigvee_{i \in I} a_i$ exists and $\bigvee_{i \in I} ||a_i|| = ||\bigvee_{i \in I} a_i||$.

Convergence in the sense of norm and in the order theory sense match.

Completeness

However, order-theoretic concepts can be used instead.

Complete normed cones

An ω-**complete normed cone** is a normed cone such that if $\{a_i \mid i \in I\}$ is an increasing sequence with $\{||a_i||\}$ bounded then the lub $\bigvee_{i \in I} a_i$ exists and $\bigvee_{i \in I} ||a_i|| = ||\bigvee_{i \in I} a_i||$.

Convergence in the sense of norm and in the order theory sense match.

Selinger's lemma

Suppose that u_i is an ω-chain with a l.u.b. in an ω-complete normed cone and u is an upper bound of the u_i. Suppose furthermore that $\lim_{i \longrightarrow \infty} \|u - u_i\| = 0$. Then $u = \bigvee_i u_i$.

Completeness

However, order-theoretic concepts can be used instead.

Complete normed cones

An ω-**complete normed cone** is a normed cone such that if $\{a_i \mid i \in I\}$ is an increasing sequence with $\{||a_i||\}$ bounded then the lub $\bigvee_{i \in I} a_i$ exists and $\bigvee_{i \in I} ||a_i|| = || \bigvee_{i \in I} a_i||$.

Convergence in the sense of norm and in the order theory sense match.

Selinger's lemma

Suppose that u_i is an ω-chain with a l.u.b. in an ω-complete normed cone and u is an upper bound of the u_i. Suppose furthermore that $\lim_{i \longrightarrow \infty} \|u - u_i\| = 0$. Then $u = \bigvee_i u_i$.

Here we are writing $u - u_i$ informally

We really mean w_i where $u_i + w_i = u$.

Maps between cones

Continuous maps

An ω-**continuous** linear map between two cones is one that preserves least upper bounds of countable chains.

Maps between cones

Continuous maps

An ω-**continuous** linear map between two cones is one that preserves least upper bounds of countable chains.

Bounded maps

A *bounded* linear map of normed cones $f : C \longrightarrow D$ is one such that for all u in C, $||f(u)|| \leq K||u||$ for some real number K. Any linear continuous map of complete normed cones is bounded.

Maps between cones

Continuous maps

An ω-**continuous** linear map between two cones is one that preserves least upper bounds of countable chains.

Bounded maps

A *bounded* linear map of normed cones $f : C \longrightarrow D$ is one such that for all u in C, $||f(u)|| \leq K||u||$ for some real number K. Any linear continuous map of complete normed cones is bounded.

Norm of a bounded map

The norm of a bounded linear map $f : C \longrightarrow D$ is defined as $||f|| = \sup\{||f(u)|| \, : \, u \in C, \, ||u|| \leq 1\}$.

A category of normed cones

The ambient category

The ω-complete normed cones, along with ω-continuous linear maps, form a category which we shall denote ω**CC**.

A category of normed cones

The ambient category

The ω-complete normed cones, along with ω-continuous linear maps, form a category which we shall denote ω**CC**.

The subcategory of interest

we define the subcategory $\omega\mathbf{CC}_1$: the norms of the maps are all bounded by 1. Isomorphisms in this category are always isometries.

Dual cones

Dual cone

Given an ω-complete normed cone C, its dual C^* is the set of all ω-continuous linear maps from C to $\mathbb{R}_+$. We define the norm on C^* to be the operator norm.

Dual cones

Dual cone

Given an ω-complete normed cone C, its dual C^* is the set of all ω-continuous linear maps from C to $\mathbb{R}_+$. We define the norm on C^* to be the operator norm.

Basic facts

C^* is an ω-complete normed cone as well, and the cone order corresponds to the point wise order.

The duality functor

In ω**CC**, the dual operation becomes a contravariant functor.
If $f : C \longrightarrow D$ is a map of cones, we define $f^* : D^* \longrightarrow C^*$ as follows: given a map L in D^*, we define a map f^*L in C^* as $f^*L(u) = L(f(u))$.

How does this compare with Banach spaces?

This dual is stronger than the dual in usual Banach spaces, where we only require the maps to be bounded. For instance, it turns out that the dual to $L_\infty^+(X)$ (to be defined later) is isomorphic to $L_1^+(X)$, which is not the case with the Banach space $L_\infty(X)$.

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 $\mathcal{M}^p_{\mathsf{UB}}(X)$ 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 $q \leq Kp$.
- The cones $\mathcal{M}^p_{\mathsf{UB}}(X)$ and $L_\infty^+(X, \Sigma, p)$ are isomorphic.

Duality for cones

A Reisz-like theorem

The dual of the cone $L_\infty^+(X, \Sigma, p)$ is isometrically isomorphic to $\mathcal{M}^{\ll p}(X)$.

Duality for cones

A Reisz-like theorem

The dual of the cone $L_\infty^+(X, \Sigma, p)$ is isometrically isomorphic to $\mathcal{M}^{\ll p}(X)$.

Corollary

Since $\mathcal{M}^{\ll p}(X)$ is isometrically isomorphic to $L_1^+(X)$, an immediate corollary is that $L_\infty^{+,*}(X)$ is isometrically isomorphic to $L_1^+(X)$, which is of course false in general in the context of Banach spaces.

Duality for cones II

Another Reisz-like theorem

The dual of the cone $L_1^+(X, \Sigma, p)$ is isometrically isomorphic to $\mathcal{M}^p_{\mathsf{UB}}(X)$.

Duality for cones II

Another Reisz-like theorem

The dual of the cone $L_1^+(X, \Sigma, p)$ is isometrically isomorphic to $\mathcal{M}^p_{\mathrm{UB}}(X)$.

Corollary

$\mathcal{M}^p_{\mathrm{UB}}(X)$ is isometrically isomorphic to $L_\infty^+(X)$, hence immediate corollary is that $L_1^{+,*}(X)$ is isometrically isomorphic to $L_\infty^+(X)$.

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 $\lambda g.\langle f,\ g\rangle = \lambda g. \int fg\mathrm{d}p$.

Summary of cones

We fix a probability triple (X, Σ, p) and focus on six spaces of cones that are based on them. They break into two natural groups of three isomorphic spaces. The first three spaces are:

A1 $\mathcal{M}^{\ll p}(X)$ - the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p,

Summary of cones

We fix a probability triple (X, Σ, p) and focus on six spaces of cones that are based on them. They break into two natural groups of three isomorphic spaces. The first three spaces are:

A1 $\mathcal{M}^{\ll p}(X)$ - the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p,

A2 $L_1^+(X, p)$ - the cone of integrable almost-everywhere positive functions,

Summary of cones

We fix a probability triple (X, Σ, p) and focus on six spaces of cones that are based on them. They break into two natural groups of three isomorphic spaces. The first three spaces are:

- A1 $\mathcal{M}^{\ll p}(X)$ - the cone of all measures on (X, Σ, p) that are absolutely continuous with respect to p,
- A2 $L_1^+(X, p)$ - the cone of integrable almost-everywhere positive functions,
- A3 $L_\infty^{+,*}(X, p)$ - the dual cone of the the cone of almost-everywhere positive bounded measurable functions.

Summary of cones II

The next group of three isomorphic spaces are:

B1 $\mathcal{M}^p_{\mathsf{UB}}(X)$ - the cone of all measures that are uniformly less than a multiple of the measure p,

Summary of cones II

The next group of three isomorphic spaces are:

B1 $\mathcal{M}^{p}_{\mathsf{UB}}(X)$ - the cone of all measures that are uniformly less than a multiple of the measure p,

B2 $L^{+}_{\infty}(X,p)$ - the cone of almost-everywhere positive functions in the normed vector space $L_{\infty}(X,p)$,

Summary of cones II

The next group of three isomorphic spaces are:

B1 $\mathcal{M}^p_{\mathsf{UB}}(X)$ - the cone of all measures that are uniformly less than a multiple of the measure p,

B2 $L^+_\infty(X,p)$ - the cone of almost-everywhere positive functions in the normed vector space $L_\infty(X,p)$,

B3 $L^{+,*}_1(X,p)$ - the dual of the cone of almost-everywhere positive functions in the normed vector space $L_1(X,p)$.

Summary of dualities and isos

The spaces defined in A1, A2 and A3 are dual to the spaces defined in B1, B2 and B3 respectively. The situation may be depicted in the diagram

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

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

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 d'après 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 d'après Vincent Danos.
4. Make bisimulation and approximation live in the same universe

Duality is the Key

$$\begin{array}{ccccc} \mathcal{M}^{\ll p}(X) & \xleftrightarrow{\sim} & L_1^+(X,p) & \xleftrightarrow{\sim} & L_\infty^{+,*}(X,p) \\ \updownarrow & & \updownarrow & & \updownarrow \\ \mathcal{M}_{\text{UB}}^p & \xleftrightarrow{\sim} & L_\infty^+(X,p) & \xleftrightarrow{\sim} & L_1^{+,*}(X,p) \end{array} \quad (2)$$

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:

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

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.

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_∞^+

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

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 $\tau : L^+_\infty(Y) \longrightarrow L^+_\infty(X)$ 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

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

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.

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

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

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

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

- 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

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

- 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 $\mathbf{LMP}$ (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*.
- Vincent Danos proposed the idea of *event bisimulation*, which was developed by him and Desharnais, Laviolette and P.

Event bisimulation

- In measure theory one should focus on measurable sets rather than on *points*.
- Vincent Danos proposed the idea of *event bisimulation*, which was developed by him and Desharnais, Laviolette and P.

Event bisimulation

- In measure theory one should focus on measurable sets rather than on *points*.
- Vincent Danos proposed the idea of *event bisimulation*, which was developed by him and Desharnais, Laviolette and P.

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*.
- Vincent Danos proposed the idea of *event bisimulation*, which was developed by him and Desharnais, Laviolette and P.

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

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

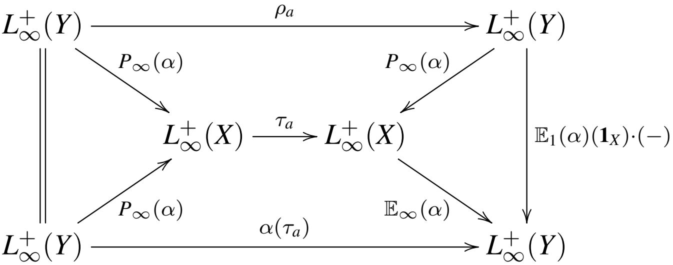
Zigzag maps

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:

$$\begin{array}{ccc} L^+_\infty(X,\Sigma,p) & \xrightarrow{\tau_a} & L^+_\infty(X,\Sigma,p) \\ {\scriptstyle P_\infty(\alpha)}\uparrow & & \uparrow{\scriptstyle P_\infty(\alpha)} \\ L^+_\infty(Y,\Lambda,q) & \xrightarrow{\rho_a} & L^+_\infty(Y,\Lambda,q) \end{array} \tag{3}$$

A key diagram

When we have a zigzag the following diagram commutes:

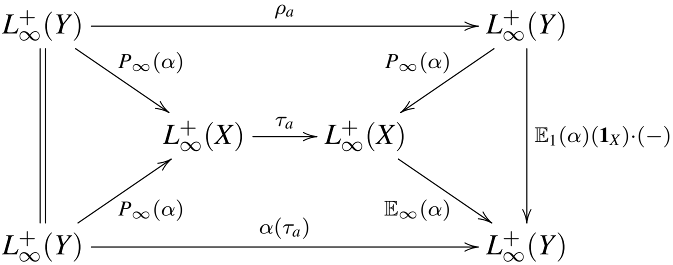


(4)

- 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:

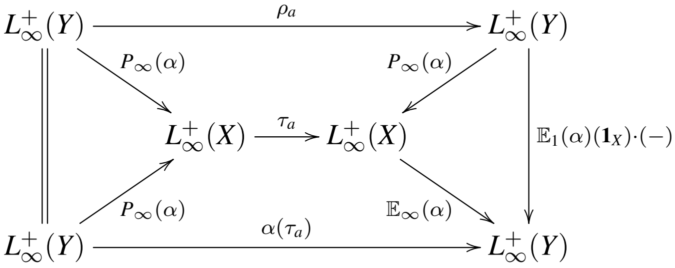


(4)

- 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:



(4)

- 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

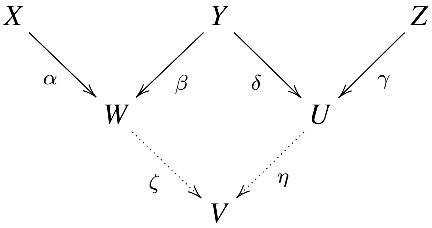
$$\alpha : (X, \Sigma, p, \tau) \longrightarrow (Z, \Gamma, r, \pi)$$
$$\beta : (Y, \Lambda, q, \rho) \longrightarrow (Z, \Gamma, r, \pi)$$

giving a cospan diagram

$$\begin{array}{ccc} (X, \Sigma, p, \tau) & & (Y, \Lambda, q, \rho) \\ & \searrow^{\alpha} \quad \swarrow_{\beta} & \\ & (Z, \Gamma, r, \pi) & \end{array} \tag{5}$$

Note that the identity function on an AMP is a zigzag, and thus that any zigzag between two AMPs X and Y implies that they are bisimilar.

Bisimulation is an equivalence



(6)

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.

Results

- Logical characterization of bisimulation holds for any measure space; not just for analytic spaces.

Results

- Logical characterization of bisimulation holds for any measure space; not just for analytic spaces.
- We can construct a unique minimal (couniversal property) version of any LAMP.

Results

- Logical characterization of bisimulation holds for any measure space; not just for analytic spaces.
- We can construct a unique minimal (couniversal property) version of any LAMP.
- We can construct approximations using expectation values to project a LAMP onto a *finite* sub-σ-algebra.

Results

- Logical characterization of bisimulation holds for any measure space; not just for analytic spaces.
- We can construct a unique minimal (couniversal property) version of any LAMP.
- We can construct approximations using expectation values to project a LAMP onto a *finite* sub-σ-algebra.
- We can show that the projective limit of the finite approximations give the minimal representation.