Markov Processes as Function Transformers

Part I: Overview and Categories of Cones

Prakash Panangaden[1]

[1]School of Computer Science
McGill University

CIRM Marseille 20 - 24 February 2012

Outline

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
- 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.
- 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*.**

Motivation

Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

- hybrid control systems; e.g. flight management systems.
- telecommunication systems with spatial variation; e.g. cell phones
- performance modelling,
- continuous time systems,
- probabilistic process algebra with recursion.

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

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

Approximation

- Our main result is a systematic approximation scheme for labelled Markov processes. The set of **LMP**s is a Polish space.
- Furthermore, our approximation results allow us to approximate integrals of continuous functions by computing them on finite approximants.
- For any **LMP** we explicitly provide a (countable) sequence of approximants to it such that: for every $\mathcal{L}_0$ property satisfied by a process, there is an element of the chain that also satisfies the property.
- The sequence of approximants converges – in a certain metric – to the process that is being approximated.

Domain-theoretic Results

- We establish the following equivalence of categories:

$$\mathbf{LMP} \simeq \mathbf{Proc}$$

where $\mathbf{LMP}$ is the category with objects $\mathbf{LMP}$s and with morphisms simulations; and $\mathbf{Proc}$ is the solution to the recursive domain equation

$$\mathbf{Proc} \simeq \prod_{\mathtt{Labels}} \mathcal{P}_{JP}(\mathbf{Proc}).$$

- We show that there is a perfect match between: bisimulation and equality in $\mathbf{Proc}$,
- simulation and the partial order of $\mathbf{Proc}$ and strict simulation and way below in $\mathbf{Proc}$.

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.
- 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.
- 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.
- 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.
- The theory works much more smoothly as I hope to show.

Banach spaces

- A *norm* on a vector space V is a function $\|\cdot\| : V \longrightarrow \mathbb{R}^{\geq 0}$ such that:
  1. $\|v\| = 0$ iff $v = 0$
  2. $\|r \cdot v\| = | r | \|v\|$ and
  3. $\|x + y\| \leq \|x\| + \|y\|$.
- The norm induces a metric: $d(u, v) = \|u - v\|$ and, hence, a topology. This topology is called the *norm topology*.
- If V is complete in this metric it is called a Banach space.
- In quantum mechanics the state spaces are Hilbert spaces – hence automatically Banach spaces – but the spaces of operators are not Hilbert spaces, they are Banach spaces.

Maps between Banach spaces

- A linear map $T : U \longrightarrow V$ is *bounded* if there exists a positive real number α such that $\forall u \in U, \|Tu\| \leq \alpha \|u\|$.
- A lineap map between normed spaces is *continuous* iff it is *bounded*.
- Given a bounded linear map between normed spaces $T : U \longrightarrow V$ we define $\|T\| = \sup\{\|Tu\| \mid u \in U, \|u\| \leq 1\}$.
- This is a norm on the space of bounded linear maps and is called the *operator norm*.
- With this norm the space of bounded linear maps between Banach spaces forms a Banach space.

Duality for Banach spaces

- The space of bounded ($=$ continuous) linear maps from V, a Banach space, to $\mathbb{R}$ is itself a Banach space, called the *dual space*, V^*.
- For any two vector spaces U, V we say that they are in *algebraic duality* if there is a bilinear form $\langle \cdot, \ \cdot \rangle : U \times V \longrightarrow \mathbb{R}$ such that spaces of functionals $\langle \cdot, \ V \rangle$ and $\langle U, \ \cdot \rangle$ separates points of U and V.
- We say two Banach spaces are in *duality* if $\langle \cdot, \ V \rangle \subseteq U^*$ and $\langle U, \ \cdot \rangle \subseteq V^*$.
- For V a Banach space, the spaces V and V^* are in duality.
- The bilinear form is $\langle v, \ \phi \rangle = \phi(v)$.
- There is a canonical injection $\iota : V \longrightarrow V^{**}$; if this is an isometry we say that the Banach space V is reflexive.
- Infinite dimensional Banach spaces are **not necessarily** reflexive.
- Finite dimensional Banach spaces are always reflexive.

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!
- 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.
- 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.

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.

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**.

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.

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_∞.
- 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
- 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 $\mathcal{M}^{\ll p}(X)$.

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 $\mathcal{M}^p_{\mathsf{UB}}(X)$.

Corollary

$\mathcal{M}^p_{\mathsf{UB}}(X)$ 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:

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

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,
- 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,
- 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

$$\begin{tikzcd} \mathcal{M}^{\ll p}(X) \arrow[r, leftrightarrow, "\sim"] \arrow[d, leftrightarrow, dotted] & L_1^+(X,p) \arrow[r, leftrightarrow, "\sim"] \arrow[d, leftrightarrow, dotted] & L_\infty^{+,*}(X,p) \arrow[d, leftrightarrow, dotted] \\ \mathcal{M}^p_{\mathrm{UB}} \arrow[r, leftrightarrow, "\sim"] & L_\infty^+(X,p) \arrow[r, leftrightarrow, "\sim"] & L_1^{+,*}(X,p) \end{tikzcd} \tag{1}$$

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