Approximating Markov Processes, Again!

Philippe Chaput[1] and Prakash Panangaden[1]
Vincent Danos[2] and Gordon Plotkin[2]

[2]Department of Computer Science
University of Edinburgh

9th Sept 2009 / CALCO 2009

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Summary

- Labelled Markov processes are probabilistic transition systems with continuous state spaces.
- We had developed a theory of bisimulation, proved a logical characterization theorem, defined metrics and developed three approximation theories.
- Proofs seemed to depend on subtle topological conditions. Why?
- Take a predicate transformer view and dualize everything.
- Everything works like magic!
- Bisimulation should never have been defined as a span!

Summary

- Labelled Markov processes are probabilistic transition systems with continuous state spaces.
- We had developed a theory of bisimulation, proved a logical characterization theorem, defined metrics and developed three approximation theories.
- Proofs seemed to depend on subtle topological conditions. Why?
- Take a predicate transformer view and dualize everything.
- Everything works like magic!
- Bisimulation should never have been defined as a span!

Summary

- Labelled Markov processes are probabilistic transition systems with continuous state spaces.
- We had developed a theory of bisimulation, proved a logical characterization theorem, defined metrics and developed three approximation theories.
- Proofs seemed to depend on subtle topological conditions. Why?
- Take a predicate transformer view and dualize everything.
- Everything works like magic!
- Bisimulation should never have been defined as a span!

Summary

- Labelled Markov processes are probabilistic transition systems with continuous state spaces.
- We had developed a theory of bisimulation, proved a logical characterization theorem, defined metrics and developed three approximation theories.
- Proofs seemed to depend on subtle topological conditions. Why?
- Take a predicate transformer view and dualize everything.
- Everything works like magic!
- Bisimulation should never have been defined as a span!

Summary

- Labelled Markov processes are probabilistic transition systems with continuous state spaces.
- We had developed a theory of bisimulation, proved a logical characterization theorem, defined metrics and developed three approximation theories.
- Proofs seemed to depend on subtle topological conditions. Why?
- Take a predicate transformer view and dualize everything.
- Everything works like magic!
- Bisimulation should never have been defined as a span!

Summary

- Labelled Markov processes are probabilistic transition systems with continuous state spaces.
- We had developed a theory of bisimulation, proved a logical characterization theorem, defined metrics and developed three approximation theories.
- Proofs seemed to depend on subtle topological conditions. Why?
- Take a predicate transformer view and dualize everything.
- Everything works like magic!
- Bisimulation should never have been defined as a span!

What is an LMP?

- Probabilistic transition system with a (possibly) *continuous* state space.
- Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

What is an LMP?

- Probabilistic transition system with a (possibly) *continuous* state space.
- Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

Some Examples

- brownian motion, gas diffusion,...
- population growth models,
- changes in stock prices,
- performance modelling,
- probabilistic process algebra with recursion,
- hybrid control systems; e.g. flight management systems.

Some Examples

- brownian motion, gas diffusion,...
- population growth models,
- changes in stock prices,
- performance modelling,
- probabilistic process algebra with recursion,
- hybrid control systems; e.g. flight management systems.

Some Examples

- brownian motion, gas diffusion,...
- population growth models,
- changes in stock prices,
- performance modelling,
- probabilistic process algebra with recursion,
- hybrid control systems; e.g. flight management systems.

Some Examples

- brownian motion, gas diffusion,...
- population growth models,
- changes in stock prices,
- performance modelling,
- probabilistic process algebra with recursion,
- hybrid control systems; e.g. flight management systems.

Some Examples

- brownian motion, gas diffusion,...
- population growth models,
- changes in stock prices,
- performance modelling,
- probabilistic process algebra with recursion,
- hybrid control systems; e.g. flight management systems.

Some Examples

- brownian motion, gas diffusion,...
- population growth models,
- changes in stock prices,
- performance modelling,
- probabilistic process algebra with recursion,
- hybrid control systems; e.g. flight management systems.

Labelled Markov Processes

- 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.
- Interaction is by synchronizing on labels. For each label there is a Markov process described by a stochastic kernel (probabilistic relation).
- We observe the interactions - not the internal states.

Labelled Markov Processes

- 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.
- Interaction is by synchronizing on labels. For each label there is a Markov process described by a stochastic kernel (probabilistic relation).
- We observe the interactions - not the internal states.

Labelled Markov Processes

- 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.
- Interaction is by synchronizing on labels. For each label there is a Markov process described by a stochastic kernel (probabilistic relation).
- We observe the interactions - not the internal states.

Labelled Markov Processes

- 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.
- Interaction is by synchronizing on labels. For each label there is a Markov process described by a stochastic kernel (probabilistic relation).
- We observe the interactions - not the internal states.

The Formal Definition

A labelled Markov process

with label set $\mathcal{A}$ is a structure

$$(S, \Sigma, i, \{\tau_a \mid a \in \mathcal{A}\}),$$

where S is the set of states, i is the initial state, and Σ is the σ-field on S, and

$$\forall a \in \mathcal{A}, \tau_a : S \times \Sigma \longrightarrow [0, 1]$$

is a transition sub-probability function.

Transition Probability Functions

$$\tau : S \times \Sigma \longrightarrow [0, 1]$$

- for fixed $s \in S$, $\tau(s, \cdot) : \Sigma \to [0, 1]$ is a subprobability measure;
- for fixed $A \in \Sigma$, $\tau(\cdot, A) : \Sigma \to [0, 1]$ is a measurable function.
- This is the stochastic analogue of a binary relation so we have the natural extension of a labelled transition system.

Transition Probability Functions

$$\tau : S \times \Sigma \longrightarrow [0,1]$$

- for fixed $s \in S$, $\tau(s,\cdot) : \Sigma \rightarrow [0,1]$ is a subprobability measure;
- for fixed $A \in \Sigma$, $\tau(\cdot, A) : \Sigma \rightarrow [0,1]$ is a measurable function.
- This is the stochastic analogue of a binary relation so we have the natural extension of a labelled transition system.

Transition Probability Functions

$$\tau : S \times \Sigma \longrightarrow [0, 1]$$

- for fixed $s \in S$, $\tau(s, \cdot) : \Sigma \to [0, 1]$ is a subprobability measure;
- for fixed $A \in \Sigma$, $\tau(\cdot, A) : \Sigma \to [0, 1]$ is a measurable function.
- This is the stochastic analogue of a binary relation so we have the natural extension of a labelled transition system.

LMPs as Coalgebras

There is a monad defined by Gìry in 1981:

$$\Gamma : \mathbf{Mes} \longrightarrow \mathbf{Mes}$$

given by

$$\Gamma((X, Sigma_X)) = \{\nu | \nu \text{ is a probability measure on } \Sigma_X\}$$

and given $f : (X, \Sigma_X) \longrightarrow (Y, \Sigma_Y)$

$$\Gamma(f)(\nu : \Gamma(X)) = \lambda B : \Sigma_Y.\nu(f^{-1}(B)).$$

LMPs are coalgebras for this monad.

LMPs as Coalgebras

There is a monad defined by Gìry in 1981:

$$\Gamma : \mathbf{Mes} \longrightarrow \mathbf{Mes}$$

given by

$$\Gamma((X, Sigma_X)) = \{\nu | \nu \text{ is a probability measure on } \Sigma_X\}$$

and given $f : (X, \Sigma_X) \longrightarrow (Y, \Sigma_Y)$

$$\Gamma(f)(\nu : \Gamma(X)) = \lambda B : \Sigma_Y . \nu(f^{-1}(B)).$$

LMPs are coalgebras for this monad.

LMPs as Coalgebras

There is a monad defined by Gìry in 1981:

$$\Gamma : \mathbf{Mes} \longrightarrow \mathbf{Mes}$$

given by

$$\Gamma((X, Sigma_X)) = \{\nu | \nu \text{ is a probability measure on } \Sigma_X\}$$

and given $f : (X, \Sigma_X) \longrightarrow (Y, \Sigma_Y)$

$$\Gamma(f)(\nu : \Gamma(X)) = \lambda B : \Sigma_Y . \nu(f^{-1}(B)).$$

LMPs are coalgebras for this monad.

Bisimulation as a Span

Define a *zig-zag* to be a measurable function between LMPs (X, Σ_X, τ_a) and (Y, Σ_Y, ρ_a) such that

$$\tau_a(x, f^{-1}(B)) = \rho_a(f(x), B).$$

This is exactly the notion of co-algebra homomorphism.

We say two systems are bisimilar if there is a span of zig-zags connecting them.

Bisimulation as a Span

Define a *zig-zag* to be a measurable function between LMPs (X, Σ_X, τ_a) and (Y, Σ_Y, ρ_a) such that

$$\tau_a(x, f^{-1}(B)) = \rho_a(f(x), B).$$

This is exactly the notion of co-algebra homomorphism.

We say two systems are bisimilar if there is a span of zig-zags connecting them.

Is bisimulation transitive?

- Ideally we would like to be able to construct pullbacks.
- Unfortunately, they do not exist in general.
- Weak pullbacks will do (works for example in ultrametric spaces).
- Unfortunately even weak pullbacks do not exist!
- Edalat showed how to construct semi-pullbacks (with great pain!)
- and Doberkat improved and generalized the construction.

Is bisimulation transitive?

- Ideally we would like to be able to construct pullbacks.
- Unfortunately, they do not exist in general.
- Weak pullbacks will do (works for example in ultrametric spaces).
- Unfortunately even weak pullbacks do not exist!
- Edalat showed how to construct semi-pullbacks (with great pain!)
- and Doberkat improved and generalized the construction.

Is bisimulation transitive?

- Ideally we would like to be able to construct pullbacks.
- Unfortunately, they do not exist in general.
- Weak pullbacks will do (works for example in ultrametric spaces).
- Unfortunately even weak pullbacks do not exist!
- Edalat showed how to construct semi-pullbacks (with great pain!)
- and Doberkat improved and generalized the construction.

Is bisimulation transitive?

- Ideally we would like to be able to construct pullbacks.
- Unfortunately, they do not exist in general.
- Weak pullbacks will do (works for example in ultrametric spaces).
- Unfortunately even weak pullbacks do not exist!
- Edalat showed how to construct semi-pullbacks (with great pain!)
- and Doberkat improved and generalized the construction.

Is bisimulation transitive?

- Ideally we would like to be able to construct pullbacks.
- Unfortunately, they do not exist in general.
- Weak pullbacks will do (works for example in ultrametric spaces).
- Unfortunately even weak pullbacks do not exist!
- Edalat showed how to construct semi-pullbacks (with great pain!)
- and Doberkat improved and generalized the construction.

Is bisimulation transitive?

- Ideally we would like to be able to construct pullbacks.
- Unfortunately, they do not exist in general.
- Weak pullbacks will do (works for example in ultrametric spaces).
- Unfortunately even weak pullbacks do not exist!
- Edalat showed how to construct semi-pullbacks (with great pain!)
- and Doberkat improved and generalized the construction.

Bisimulation à la Larsen and Skou

Let $\mathcal{S} = (S, i, \Sigma, \tau)$ be a labelled Markov process. An equivalence relation R on S is a **bisimulation** if whenever sRs', with $s, s' \in S$, we have that for all $a \in \mathcal{A}$ and every R-closed measurable set $A \in \Sigma$, $\tau_a(s, A) = \tau_a(s', A)$.
Two states are bisimilar if they are related by a bisimulation relation.
Can be extended to bisimulation between two different LMPs.

Co-bisimulation

Define the dual of bisimulation using co-spans.

$$\begin{tikzcd}
X \arrow[dr, "f"'] & & Y \arrow[dl, "g"] \\
& Z &
\end{tikzcd}$$

This always yields an equivalence relation because pushouts exist by general abstract nonsense.

This seems to be independently due to Bartels, Sokolova and de Vink and Danos, Desharnais, Laviolette and P.

Co-bisimulation

Define the dual of bisimulation using co-spans.

$$\begin{tikzcd} X \arrow[dr, "f"'] & & Y \arrow[dl, "g"] \\ & Z & \end{tikzcd}$$

This always yields an equivalence relation because pushouts exist by general abstract nonsense.

This seems to be independently due to Bartels, Sokolova and de Vink and Danos, Desharnais, Laviolette and P.

A Modal Logic

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

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 $\mathcal{L}$.

This depends on properties of analytic spaces and quotients of such spaces under “nice” equivalence relations.

Modal Logic and Co-bisimulation

The theorem that the modal logic characterizes co-bisimulation is (relatively) easy and works for general measure spaces.

It does not require properties of analytic spaces.

For analytic spaces the two concepts coincide.

Modal Logic and Co-bisimulation

The theorem that the modal logic characterizes co-bisimulation is (relatively) easy and works for general measure spaces.

It does not require properties of analytic spaces.

For analytic spaces the two concepts coincide.

Provocative Slogan

Co-bisimulation is the *real* concept; it is only a coincidence that bisimulation works for discrete systems.

Finite Approximations 1

Our main result: A systematic approximation scheme for labelled Markov processes.

The set of LMPs is a Polish space. Furthermore, our approximation results allow us to approximate integrals of continuous functions by computing them on finite approximants.

Finite Approximations 2

- For any LMP, we explicitly provide a (countable) sequence of approximants to it such that:

For every logical 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.

Finite Approximations 2

- For any LMP, we explicitly provide a (countable) sequence of approximants to it such that:

For every logical 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.

Finite Approximations 2

- For any LMP, we explicitly provide a (countable) sequence of approximants to it such that:

For every logical 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.

Finite Approximations 2

- For any LMP, we explicitly provide a (countable) sequence of approximants to it such that:

For every logical 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.

Finite Approximations 2

- For any LMP, we explicitly provide a (countable) sequence of approximants to it such that:

For every logical 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.

The Approximation Construction

- Given a labelled Markov process $\mathcal{S} = (S, \Sigma, \tau)$, an integer n and a *rational* number $\epsilon > 0$, we define $\mathcal{S}(n, \epsilon)$ to be an n-step unfolding approximation of $\mathcal{S}$.
- Its state-space is divided into $n + 1$ levels which are numbered $0, 1, \ldots, n$.
- A state is a pair (X, l) where $X \in \Sigma$ and $l \in \{0, 1, \ldots, n\}$.
- *At each level*, the sets that define states form a partition of S. The initial state of $\mathcal{S}(n, \epsilon)$ is at level n and transitions only occur between a state of one level to a state of one lower level.

The Approximation Construction

- Given a labelled Markov process $\mathcal{S} = (S, \Sigma, \tau)$, an integer n and a *rational* number $\epsilon > 0$, we define $\mathcal{S}(n, \epsilon)$ to be an n-step unfolding approximation of $\mathcal{S}$.
- Its state-space is divided into $n + 1$ levels which are numbered $0, 1, \ldots, n$.
- A state is a pair (X, l) where $X \in \Sigma$ and $l \in \{0, 1, \ldots, n\}$.
- *At each level*, the sets that define states form a partition of S. The initial state of $\mathcal{S}(n, \epsilon)$ is at level n and transitions only occur between a state of one level to a state of one lower level.

The Approximation Construction

- Given a labelled Markov process $\mathcal{S} = (S, \Sigma, \tau)$, an integer n and a *rational* number $\epsilon > 0$, we define $\mathcal{S}(n, \epsilon)$ to be an n-step unfolding approximation of $\mathcal{S}$.
- Its state-space is divided into $n + 1$ levels which are numbered $0, 1, \ldots, n$.
- A state is a pair (X, l) where $X \in \Sigma$ and $l \in \{0, 1, \ldots, n\}$.
- *At each level*, the sets that define states form a partition of S. The initial state of $\mathcal{S}(n, \epsilon)$ is at level n and transitions only occur between a state of one level to a state of one lower level.

The Approximation Construction

- Given a labelled Markov process $\mathcal{S} = (S, \Sigma, \tau)$, an integer n and a *rational* number $\epsilon > 0$, we define $\mathcal{S}(n, \epsilon)$ to be an n-step unfolding approximation of $\mathcal{S}$.
- Its state-space is divided into $n + 1$ levels which are numbered $0, 1, \ldots, n$.
- A state is a pair (X, l) where $X \in \Sigma$ and $l \in \{0, 1, \ldots, n\}$.
- *At each level*, the sets that define states form a partition of S. The initial state of $\mathcal{S}(n, \epsilon)$ is at level n and transitions only occur between a state of one level to a state of one lower level.

Approximating the Transition Probabilities

What is the transition probability between A and B (sets of states of the real system)?

$$\rho(A, B) = \inf_{x \in A} \tau(x, B).$$

This is an under approximation.

Improvements

- Sometimes the approximation is “spectacularly dumb”; it unwinds loops that should not be unwound.
- Danos and Desharnais fixed this but their approximants had measures that were not additive.
- DDP fixed this by using averaging rather than under approximating.
- This required a very restrictive condition in order to get rid of the problem that in measure theory things are defined upto sets of measure 0.

Improvements

- Sometimes the approximation is “spectacularly dumb”; it unwinds loops that should not be unwound.
- Danos and Desharnais fixed this but their approximants had measures that were not additive.
- DDP fixed this by using averaging rather than under approximating.
- This required a very restrictive condition in order to get rid of the problem that in measure theory things are defined upto sets of measure 0.

Improvements

- Sometimes the approximation is “spectacularly dumb”; it unwinds loops that should not be unwound.
- Danos and Desharnais fixed this but their approximants had measures that were not additive.
- DDP fixed this by using averaging rather than under approximating.
- This required a very restrictive condition in order to get rid of the problem that in measure theory things are defined upto sets of measure 0.

Improvements

- Sometimes the approximation is “spectacularly dumb”; it unwinds loops that should not be unwound.
- Danos and Desharnais fixed this but their approximants had measures that were not additive.
- DDP fixed this by using averaging rather than under approximating.
- This required a very restrictive condition in order to get rid of the problem that in measure theory things are defined upto sets of measure 0.

Dualize Everything!

An LMP is not to be thought of not as $\tau : X \times \Sigma_X \longrightarrow [0,1]$ but, rather as a function $f \mapsto \tau(f)$ where

$$\tau(f)(x) = \int_X f(x')\tau(x, dx').$$

In other words as a "function" transformer:

the quantitative analogue of a "predicate transformer."

Functions as Formulas

A function on the state space describes partial information about the state of the system.

Example

The function 1_B says that the state is somewhere in B.

Kozen's Analogy

Logic	Probability
s state	P distribution
ϕ formula	χ random variable
$s \models \phi$	$\int \chi dP$

Functions as Formulas

A function on the state space describes partial information about the state of the system.

Example

The function 1_B says that the state is somewhere in B.

Kozen's Analogy

Logic	Probability
s state	P distribution
ϕ formula	χ random variable
$s \models \phi$	$\int \chi dP$

Functions as Formulas

A function on the state space describes partial information about the state of the system.

Example

The function 1_B says that the state is somewhere in B.

Kozen's Analogy

Logic	Probability
s state	P distribution
ϕ formula	χ random variable
$s \models \phi$	$\int \chi dP$

LMPs as Predicate Transformers

- Given a Markov kernel τ on X we define a linear operator $\hat{\tau}$ on bounded real-valued functions as

$$\hat{\tau}(f)(x) = \int_X f(y)\tau(x, dy).$$

- Given a probabilistic predicate ϕ on X we interpret $\phi(x)$ as the probability that x satisfies ϕ.
- Then $\hat{\tau}(\phi)(x)$ is the probability that *after a transition* x satisfies ϕ.
- In other words $\hat{\tau}$ is the weakest precondition.

LMPs as Predicate Transformers

- Given a Markov kernel τ on X we define a linear operator $\hat{\tau}$ on bounded real-valued functions as

$$\hat{\tau}(f)(x) = \int_X f(y)\tau(x, dy).$$

- Given a probabilistic predicate ϕ on X we interpret $\phi(x)$ as the probability that x satisfies ϕ.
- Then $\hat{\tau}(\phi)(x)$ is the probability that *after a transition* x satisfies ϕ.
- In other words $\hat{\tau}$ is the weakest precondition.

LMPs as Predicate Transformers

- Given a Markov kernel τ on X we define a linear operator $\hat{\tau}$ on bounded real-valued functions as

$$\hat{\tau}(f)(x) = \int_X f(y)\tau(x, dy).$$

- Given a probabilistic predicate ϕ on X we interpret $\phi(x)$ as the probability that x satisfies ϕ.
- Then $\hat{\tau}(\phi)(x)$ is the probability that *after a transition* x satisfies ϕ.
- In other words $\hat{\tau}$ is the weakest precondition.

LMPs as Predicate Transformers

- Given a Markov kernel τ on X we define a linear operator $\hat{\tau}$ on bounded real-valued functions as

$$\hat{\tau}(f)(x) = \int_X f(y)\tau(x, dy).$$

- Given a probabilistic predicate ϕ on X we interpret $\phi(x)$ as the probability that x satisfies ϕ.
- Then $\hat{\tau}(\phi)(x)$ is the probability that *after a transition* x satisfies ϕ.
- In other words $\hat{\tau}$ is the weakest precondition.

AMPs

- An abstract Markov process on a probability space is a linear operator on a space of almost-everywhere bounded real-valued functions.
- I am skipping the exact details but if you really want to know it is $L_\infty(X, P)$.
- Note that there is now an underlying measure on the state space.
- We can forget the points and just think of everything pointlessly!

AMPs

- An abstract Markov process on a probability space is a linear operator on a space of almost-everywhere bounded real-valued functions.
- I am skipping the exact details but if you really want to know it is $L_\infty(X, P)$.
- Note that there is now an underlying measure on the state space.
- We can forget the points and just think of everything pointlessly!

AMPs

- An abstract Markov process on a probability space is a linear operator on a space of almost-everywhere bounded real-valued functions.
- I am skipping the exact details but if you really want to know it is $L_\infty(X, P)$.
- Note that there is now an underlying measure on the state space.
- We can forget the points and just think of everything pointlessly!

AMPs

- An abstract Markov process on a probability space is a linear operator on a space of almost-everywhere bounded real-valued functions.
- I am skipping the exact details but if you really want to know it is $L_\infty(X, P)$.
- Note that there is now an underlying measure on the state space.
- We can forget the points and just think of everything pointlessly!

What is Averaging?

Given a real-valued function f defined on a probability space (X, Σ, P), we define the expectation (average) value of f to be

$$\langle f \rangle = \int_X f(x) dP.$$

Here P is a probability distribution on X and f is assumed to be measurable with respect to Σ.

What Measurable Really Means

- To say that $f : (X, \Sigma) \longrightarrow \mathbb{R}$ is measurable means that f does not vary "too fast."
- Imagine that there are some "minimal" measurable sets: f must be a constant on them.
- Of course Σ usually includes individual points but what if it did not?

What Measurable Really Means

- To say that $f : (X, \Sigma) \longrightarrow \mathbb{R}$ is measurable means that f does not vary "too fast."
- Imagine that there are some "minimal" measurable sets: f must be a constant on them.
- Of course Σ usually includes individual points but what if it did not?

What Measurable Really Means

- To say that $f : (X, \Sigma) \longrightarrow \mathbb{R}$ is measurable means that f does not vary "too fast."
- Imagine that there are some "minimal" measurable sets: f must be a constant on them.
- Of course Σ usually includes individual points but what if it did not?

Coarsening a σ-algebra

- Suppose that we have $\Lambda \subset \Sigma$. Then a Λ-measurable function has to be constant on minimal Λ sets.
- Thus a smaller σ-algebra means that we do not have such a refined view of the state space.
- Constructing approximations means making coarser σ-algebras rather than just clustering the points.

Coarsening a σ-algebra

- Suppose that we have $\Lambda \subset \Sigma$. Then a Λ-measurable function has to be constant on minimal Λ sets.
- Thus a smaller σ-algebra means that we do not have such a refined view of the state space.
- Constructing approximations means making coarser σ-algebras rather than just clustering the points.

Coarsening a σ-algebra

- Suppose that we have $\Lambda \subset \Sigma$. Then a Λ-measurable function has to be constant on minimal Λ sets.
- Thus a smaller σ-algebra means that we do not have such a refined view of the state space.
- Constructing approximations means making coarser σ-algebras rather than just clustering the points.

Conditional Expectation

Suppose we have (X, Σ, P) and $\Lambda \subset \Sigma$. Suppose that we are given f, real-valued and Σ-measurable.

Theorem

There exists a Λ-measurable function, written $\mathbb{E}(f|\Lambda)$ such that for any $B \in \Lambda$

$$\int_B \int f dP = \int_B \mathbb{E}(f|\Lambda) dP.$$

In other words, there is a smoothed-out version of f that is too crude to see the variations in Σ but is good enough for Λ.

Conditional Expectation

Suppose we have (X, Σ, P) and $\Lambda \subset \Sigma$. Suppose that we are given f, real-valued and Σ-measurable.

Theorem

There exists a Λ-measurable function, written $\mathbb{E}(f|\Lambda)$ such that for any $B \in \Lambda$

$$\int_B \int f dP = \int_B \mathbb{E}(f|\Lambda) dP.$$

In other words, there is a smoothed-out version of f that is too crude to see the variations in Σ but is good enough for Λ.

Conditional Expectation

Suppose we have (X, Σ, P) and $\Lambda \subset \Sigma$. Suppose that we are given f, real-valued and Σ-measurable.

Theorem

There exists a Λ-measurable function, written $\mathbb{E}(f|\Lambda)$ such that for any $B \in \Lambda$

$$\int_B \int f dP = \int_B \mathbb{E}(f|\Lambda) dP.$$

In other words, there is a smoothed-out version of f that is too crude to see the variations in Σ but is good enough for Λ.

Conditional Expectation

Suppose we have (X, Σ, P) and $\Lambda \subset \Sigma$. Suppose that we are given f, real-valued and Σ-measurable.

Theorem

There exists a Λ-measurable function, written $\mathbb{E}(f|\Lambda)$ such that for any $B \in \Lambda$

$$\int_B \int f dP = \int_B \mathbb{E}(f|\Lambda) dP.$$

In other words, there is a smoothed-out version of f that is too crude to see the variations in Σ but is good enough for Λ.

Co-spans Rule!

The definition of bisimulation naturally becomes dualized.

Bisimulation

Two AMPs are bisimilar if there is a cospan of zigzag morphisms relating them.

- It is fairly easy to show that bisimulation is transitive.
- Much easier than when using spans!
- Completely general: works for all measurable spaces.

Co-spans Rule!

The definition of bisimulation naturally becomes dualized.

Bisimulation

Two AMPs are bisimilar if there is a cospan of zigzag morphisms relating them.

- It is fairly easy to show that bisimulation is transitive.
- Much easier than when using spans!
- Completely general: works for all measurable spaces.

Co-spans Rule!

The definition of bisimulation naturally becomes dualized.

Bisimulation

Two AMPs are bisimilar if there is a cospan of zigzag morphisms relating them.

- It is fairly easy to show that bisimulation is transitive.
- Much easier than when using spans!
- Completely general: works for all measurable spaces.

Co-spans Rule!

The definition of bisimulation naturally becomes dualized.

Bisimulation

Two AMPs are bisimilar if there is a cospan of zigzag morphisms relating them.

- It is fairly easy to show that bisimulation is transitive.
- Much easier than when using spans!
- Completely general: works for all measurable spaces.

The Smallest Bisimulation

- Given an AMP X, one can show the existence of a smallest bisimilar process $\tilde{X}$.
- This is unique up to isomorphism.
- The σ-algebra can be obtained from the modal logic.

The Smallest Bisimulation

- Given an AMP X, one can show the existence of a smallest bisimilar process $\tilde{X}$.
- This is unique up to isomorphism.
- The σ-algebra can be obtained from the modal logic.

The Smallest Bisimulation

- Given an AMP X, one can show the existence of a smallest bisimilar process $\tilde{X}$.
- This is unique up to isomorphism.
- The σ-algebra can be obtained from the modal logic.

Approximations

Let τ be an AMP on (X, Σ) and we want to define an AMP $\Lambda(\tau)$ on (X, Λ).
The approximation scheme of DGJP (2000,2003) yields this diagram:

$$\begin{tikzcd} (X, \Sigma) \arrow[d, "i"] \\ (X, \Lambda) \end{tikzcd} \qquad \begin{tikzcd} L_\infty^+(X, \Sigma) \arrow[r, "\tau"] & L_\infty^+(X, \Sigma) \arrow[d, "\mathbb{E}_\Lambda"] \\ L_\infty^+(X, \Lambda) \arrow[u, "(\cdot)\circ i"] \arrow[r, dotted, "\Lambda(\tau)"'] & L_\infty^+(X, \Lambda) \end{tikzcd}$$

Our Scheme

We generalize the previous diagram to any measurable map α, by constructing a functor $\mathbb{E}_{(\cdot)}$.

$$\begin{array}{c} (X,\Sigma) \\ \Big\downarrow \alpha \\ (Y,\Lambda) \end{array} \qquad \begin{tikzcd} L_\infty^+(X,\Sigma) \arrow[r, "\tau"] & L_\infty^+(X,\Sigma) \arrow[d, "\mathbb{E}_\alpha"] \\ L_\infty^+(Y,\Lambda) \arrow[u, "(\cdot)\circ\alpha"] \arrow[r, dotted, "\alpha(\tau)"'] & L_\infty^+(Y,\Lambda) \end{tikzcd}$$

Finite Approximants from the Logic

- We use the logic as follows. Take a finite set $\mathcal{Q}$ of rationals in $[0, 1]$ and a natural number N.
- Consider formulas with nesting depth up to N and using only members of $\mathcal{Q}$.
- Take the sets denoted by these formulas and look at the σ-algebra generated. This gives a finite σ-algebra which is a sub-σ-algebra of Σ.
- Use conditional expectations as described above to produce the approximation.

Finite Approximants from the Logic

- We use the logic as follows. Take a finite set $\mathcal{Q}$ of rationals in $[0, 1]$ and a natural number N.
- Consider formulas with nesting depth up to N and using only members of $\mathcal{Q}$.
- Take the sets denoted by these formulas and look at the σ-algebra generated. This gives a finite σ-algebra which is a sub-σ-algebra of Σ.
- Use conditional expectations as described above to produce the approximation.

Finite Approximants from the Logic

- We use the logic as follows. Take a finite set $\mathcal{Q}$ of rationals in $[0, 1]$ and a natural number N.
- Consider formulas with nesting depth up to N and using only members of $\mathcal{Q}$.
- Take the sets denoted by these formulas and look at the σ-algebra generated. This gives a finite σ-algebra which is a sub-σ-algebra of Σ.
- Use conditional expectations as described above to produce the approximation.

Finite Approximants from the Logic

- We use the logic as follows. Take a finite set $\mathcal{Q}$ of rationals in $[0, 1]$ and a natural number N.
- Consider formulas with nesting depth up to N and using only members of $\mathcal{Q}$.
- Take the sets denoted by these formulas and look at the σ-algebra generated. This gives a finite σ-algebra which is a sub-σ-algebra of Σ.
- Use conditional expectations as described above to produce the approximation.

A Projective System

- As we vary over $\mathcal{Q}$ and N we get a projective system.
- Such systems have projective limits with nice properties [Choksi 1958].
- The projective limit is exactly the smallest bisimilar process. [Our main technical result]

A Projective System

- As we vary over $\mathcal{Q}$ and N we get a projective system.
- Such systems have projective limits with nice properties [Choksi 1958].
- The projective limit is exactly the smallest bisimilar process. [Our main technical result]

A Projective System

- As we vary over $\mathcal{Q}$ and N we get a projective system.
- Such systems have projective limits with nice properties [Choksi 1958].
- The projective limit is exactly the smallest bisimilar process. [Our main technical result]

What We Did

- We dualized LMPs to AMPs and defined a category of AMPs where the arrows behave as generalized projections.
- We defined a conditional expectation functor.
- Bisimulation is defined by a co-span and
- is characterized by a modal logic.
- There is a smallest bisimilar process for any given AMP.
- We gave an approximation scheme
- that reconstructs the smallest bisimilar process as a projective limit.

What We Did

- We dualized LMPs to AMPs and defined a category of AMPs where the arrows behave as generalized projections.
- We defined a conditional expectation functor.
- Bisimulation is defined by a co-span and
- is characterized by a modal logic.
- There is a smallest bisimilar process for any given AMP.
- We gave an approximation scheme
- that reconstructs the smallest bisimilar process as a projective limit.

What We Did

- We dualized LMPs to AMPs and defined a category of AMPs where the arrows behave as generalized projections.
- We defined a conditional expectation functor.
- Bisimulation is defined by a co-span and
- is characterized by a modal logic.
- There is a smallest bisimilar process for any given AMP.
- We gave an approximation scheme
- that reconstructs the smallest bisimilar process as a projective limit.

What We Did

- We dualized LMPs to AMPs and defined a category of AMPs where the arrows behave as generalized projections.
- We defined a conditional expectation functor.
- Bisimulation is defined by a co-span and
- is characterized by a modal logic.
- There is a smallest bisimilar process for any given AMP.
- We gave an approximation scheme
- that reconstructs the smallest bisimilar process as a projective limit.

What We Did

- We dualized LMPs to AMPs and defined a category of AMPs where the arrows behave as generalized projections.
- We defined a conditional expectation functor.
- Bisimulation is defined by a co-span and
- is characterized by a modal logic.
- There is a smallest bisimilar process for any given AMP.
- We gave an approximation scheme
- that reconstructs the smallest bisimilar process as a projective limit.

What We Did

- We dualized LMPs to AMPs and defined a category of AMPs where the arrows behave as generalized projections.
- We defined a conditional expectation functor.
- Bisimulation is defined by a co-span and
- is characterized by a modal logic.
- There is a smallest bisimilar process for any given AMP.
- We gave an approximation scheme
- that reconstructs the smallest bisimilar process as a projective limit.

What We Did

- We dualized LMPs to AMPs and defined a category of AMPs where the arrows behave as generalized projections.
- We defined a conditional expectation functor.
- Bisimulation is defined by a co-span and
- is characterized by a modal logic.
- There is a smallest bisimilar process for any given AMP.
- We gave an approximation scheme
- that reconstructs the smallest bisimilar process as a projective limit.

What Is To Be Done?

- We need to implement the approximation scheme. Actually Philippe has done this using a Monte Carlo scheme, but we do not yet have a proof that it works correctly.
- We would like to define metrics.
- I would like to push the dual view of bisimulation to (all) other settings.

What Is To Be Done?

- We need to implement the approximation scheme. Actually Philippe has done this using a Monte Carlo scheme, but we do not yet have a proof that it works correctly.
- We would like to define metrics.
- I would like to push the dual view of bisimulation to (all) other settings.

What Is To Be Done?

- We need to implement the approximation scheme. Actually Philippe has done this using a Monte Carlo scheme, but we do not yet have a proof that it works correctly.
- We would like to define metrics.
- I would like to push the dual view of bisimulation to (all) other settings.

What Is To Be Done?

- We need to implement the approximation scheme. Actually Philippe has done this using a Monte Carlo scheme, but we do not yet have a proof that it works correctly.
- We would like to define metrics.
- I would like to push the dual view of bisimulation to (all) other settings.

References

1. *Labelled Markov Processes*, recent book published by Imperial College Press.
2. Recent paper (2009) in ICALP.

References

1. *Labelled Markov Processes*, recent book published by Imperial College Press.
2. Recent paper (2009) in ICALP.