Stochastic Techniques in Concurrency

Prakash Panangaden
School of Computer Science
McGill University
Montréal, Québec
CANADA

New directions in semantics of concurrency

- Real-time - continuously varying quantities

- Hybrid Systems - control theory, differential equations, dynamical systems

- Probabilistic systems - probability theory, queueing theory

- Probabilistic systems with continuous state spaces - measure theory

Inspiration and Sources:

- Kozen - Probabilistic Semantics and Probabilistic PDL
- Giry (Lawvere) - Categorical Approach to Probability
- Saheb-Djharomi, Jones-Plotkin Probabilistic Powerdomains
- Larsen-Skou - Probabilistic bisimulation
- Hillston PEPA
- Edalat - Integration on Domains
- Abramsky CONCUR96 Talk

This is not an exhaustive list.

Outline of this Course

- **Day 1**
  - Introduction and Review of Larsen-Skou
  - Review of Probability Theory
- **Day 2**
  - Measure Theory
  - Integration - The Monotone Convergence Mantra
- **Day 3**
  - The Radon-Nikodym Theorem
  - Conditional Probability - Stochastic Kernels

- **Day 4**
  - The Category of Stochastic Relations
  - Iteration in **SRel**

- **Day 5**
  - Labelled Markov Processes - Bisimulation
  - The Logical characterization of bisimulation

Elementary Probability Theory

A *process* has a set of outcomes (final states), say S - suppose finite. Assume that each outcome is *equally likely*. Then the probability of observing the final state to be in $A \subseteq S$ is

$$Pr(A) = |A|/|S|$$

We call A an **event**.

Everything is essentially based on Boolean algebra, thus, e.g.

$$Pr(A \wedge B) = |A \cap B|/|S|$$

From this foundation we build up to probabilities for compound events.

Probability Spaces

We can have a countable set S with a real number, $Pr(x)$, between 0 and 1 assigned to each x in S such that

$$\sum_{x \in S} Pr(x) = 1.$$

Now for $A \subseteq S$ we define

$$Pr(A) = \sum_{x \in A} Pr(x).$$

Expectation Value

A **random variable** is a function from a probability space to the reals.

The **expectation value** of such an f is

$$E[f] =_{df} \sum_{x \in S} f(x) Pr(x).$$

Conditional Probability

We know the result is in B what is the probability that it is in A?

$$Pr(A|B) = Pr(A \cap B)/Pr(B)$$

Undefined if $Pr(B)$ is 0.

Suppose that n men throw their hats into a bag and then randomly select a hat from the bag what is the probability that none of them get their hat back?

Let E_i mean "the ith man gets his hat back" then

for all i : $Pr(E_i) = \frac{1}{n}$

for $i \neq j$:

$Pr(E_i E_j) = Pr(E_i)Pr(E_j|E_i) = \frac{1}{n(n-1)}$

for $i \neq j \neq k$:

$Pr(E_i E_j E_k) = Pr(E_i E_j)Pr(E_k|E_i E_j) = \frac{1}{n(n-1)(n-2)}$

. . . .

$$Pr(\cup_i E_i) = \sum_i Pr(E_i) - \sum_{i \neq j} Pr(E_i E_j) + - \ldots$$

Hence the probability that no one gets his hat back is

$$\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} \cdots$$

As $n \longrightarrow \infty$ we get $\frac{1}{e}$!

Conditional Expectation

Just as for conditional probability we define

$$E[f|B] = \sum_x f(x)Pr(x|B).$$

Process in your favourite process algebra

$$Q = a.NIL + b.Q + c.Q$$

with the following information: a takes 1 unit of time, b takes two units and c takes 3 units. The environment offers one action at a time with equal probability. What is the expected time to termination?

Let t be the time to termination, it is a random variable.

$$E[t] = E[t|a]Pr(a) + E[t|b]Pr(b) + E[t|c]Pr(c)$$

From the recursive nature of the process we deduce

$$E[t|a] = 1,\ E[t|b] = 2 + E[t],\ E[t|c] = 3 + E[t]$$

and from the fact that the probabilities are equal we get

$$E[t] = \frac{1}{3} + \frac{2}{3} + 1 + \frac{2}{3}E[t]$$

or $E[t] = 6$.

This solution avoids irritating infinite summations.

Consider an infnite sequence of coin tosses. We cannot assign probabilities to each outcome in a useful way. We cannot expect that we work with *arbitrary* unions but - Kolmogorov's insight - we can work with *countable* unions.

The singletons are not the basic building blocks but we can work with the "open sets" - sets with a given initial sequence. Now closure under countable union and complementation gives us all the probabilities we can reasonably expect.

In an infinite sequence of coin tosses we expect the number of heads and tails to be equal. If we work with 0s and 1s we can express this quantitatively. Let s be a sequence of 0s and 1s generated by a fair coin. Let s_i be the ith element of the sequence.

The **weak law of large numbers**:

$$\lim_{n \longrightarrow \infty} Pr\{s : |\frac{1}{n}\sum_{i=1}^{n} s_i - \frac{1}{2}| \geq \epsilon\} = 0$$

for any $\epsilon \geq 0$.

$$N = \{s : \lim_{n \longrightarrow \infty} [\frac{1}{n}\sum_{i=1}^{n} s_i] = \frac{1}{2}\}.$$

This is a condition that is stated on the entire sequence.

The **strong law of large numbers** says N has measure 1. This makes no sense unless we can define measures on sets like N.

Measure Theory

Definition 1 *A* **measurable space** (X, Σ) *is a set* X *together with a family of subsets of* X*, called a* σ-**field** *satisfying the following axioms:*
(1) $\emptyset \in \Sigma$,
(2) $A \in \Sigma$ *implies that* $A^c \in \Sigma$, *and*
(3) if $\{A_i \in \Sigma | i \in I\}$ *is a countable family then* $\cup_{i \in I} A_i \in \Sigma$.

If we require only finite additivity rather than countable additivity we get a **field**.

Proposition 2 *The intersection of an arbitrary collection of σ-fields on a set X is a σ-field on X.*

Corollary 3 *Given any subset $\mathcal{B}$ of $\mathcal{P}(X)$ there is a least σ-field containing $\mathcal{B}$.*

We often refer to the least σ-field containing $\mathcal{B}$ as the σ-field **generated** by $\mathcal{B}$.

If we have a nested family of sets

$$A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \ldots$$

with $\cup_n A_n = A$ we write $A_n \uparrow A$. Similarly if

$$A_1 \supseteq A_2 \supseteq \ldots \supseteq A_n \ldots$$

with $\cap_n A_n = A$ we write $A_n \downarrow A$.

Definition 4 *A collection of sets $\mathcal{M}$ is called a* **monotone class** *if whenever $A_n \uparrow A$ with all $A_n \in \mathcal{M}$ then $A \in \mathcal{M}$ and also if $A_n \downarrow A$ with all $A_n \in \mathcal{M}$ then $A \in \mathcal{M}$.*

A field that is also a monotone class must be a σ-field.

Proposition 5 *If $\mathcal{F}$ is a field of subsets of X then the monotone class, $\mathcal{M}$, generated by $\mathcal{F}$ is a σ-field on X.*

Definition 6 *A function f from a σ-field (X, Σ_X) to a σ-field (Y, Σ_Y) is said to be* **measurable** *if $f^{-1}(B) \in \Sigma_X$ whenever $B \in \Sigma_Y$.*

Definition 7 *The category $\mathbf{Mes}$ has as objects sets equipped with σ-fields and as morphisms measurable functions.*

Proposition 8 *The category $\mathbf{Mes}$ has binary products and coproducts.*

A very important class of examples: Separable metric spaces.

A metric space has a topology. The σ-field generated is called the **Borel algebra**, $\mathcal{B}$. If the metric space has a countable basis it is called **separable**. Anything you are ever likely to meet is a separable metric space. If the space is also complete it is called **Polish**.

Question: If we take two topological spaces we get two Borel algebras, $\mathcal{B}_1, \mathcal{B}_2$. Is the product σ-field $\mathcal{B}_1 \otimes \mathcal{B}_2$ the same as the Borel algebra generated by the product? Without knowing this we do not know that the product of two measurable functions is measurable. Answer: No! However for separable metric spaces it is.

These spaces dominate the subject.

Definition 9 *Given a family of functions* $\{f_n : X \longrightarrow Y | n \in \mathbb{N}\}$ *we say that the family* **converges pointwise** *to* f *if* $\forall x \in X. \lim_{n\to\infty} f_n(x) = f(x)$.

Theorem 10 *If a family of measurable functions* $\{f_n : X \longrightarrow Y | n \in \mathbb{N}\}$ *converges pointwise to* f *then* f *is also measurable.*

Definition 11 *A measurable function is called* **simple** *if its range is finite.*

For real valued functions we have the following very important result.

Theorem 12 *Given a non-negative measurable function $f : X \longrightarrow \mathbf{R}$ there is a family of simple functions s_i such that $\forall i. s_i \leq s_{i+1} \leq f$ and the s_i converge pointwise to f.*

For $n \in \mathbf{N}$ and $1 \leq i \leq n2^n$ define
$E_{n,i} \overset{\text{def}}{=} f^{-1}([\frac{i-1}{2^n}, \frac{i}{2^n}])$
and $F_n \overset{\text{def}}{=} f^{-1}([n, \infty])$.

$$s_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{n,i}} + n\chi_{F_n}.$$

Definition 13 *A* **measure** *(***probability measure***)* μ *on a measurable space* (X, Σ) *is a function from* Σ *to* $[0, \infty]$ *(*$[0, 1]$*), such that if* $\{A_i | i \in I\}$ *is a countable family of pairwise disjoint sets then*

$$\mu(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i).$$

Also

$$\mu(\emptyset) = 0.$$

A set equipped with a σ*-field and a measure defined on it is called a* **measure space***.*

Proposition 14 (Monotonicity and Continuity)

1. *If* $A \subseteq B$ *then* $\mu(A) \leq \mu(B)$.

2. *If* $A_1 \subseteq A_2 \subseteq \ldots A_n \subseteq \ldots$ *and* $\cup_i A_i = A$ *then* $\lim_{i\to\infty} \mu(A_i) = \mu(A)$.

3. *If* $A_1 \supseteq A_2 \supseteq \ldots A_n \supseteq \ldots$ *and* $\cap_i A_i = A$ *then* $\lim_{i\to\infty} \mu(A_i) = \mu(A)$, *if* $\mu(A_1)$ *is finite.*

Corollary 15 (Convexity) *For any countable family of sets* B_i *we have* $\mu(\cup_i B_i) \leq \sum_i \mu(B_i)$.

Fix a set X and a point x of X. We define a measure, in fact a probability measure, on the σ-field of all subsets of X as follows. We use the slightly peculiar notation $\delta(x, A)$ to emphasize that x is a parameter in the definition.

$$\delta(x, A) = \begin{array}{ll} 1 & \text{if } \ x \in A, \\ 0 & \text{if } \ x \notin A. \end{array}$$

This measure is called the *Dirac delta measure*. Note that we can fix the set A and view this as the definition of a (measurable) function on X. What we get is the characteristic function of the set A, χ_A.

Definition 16 *A measure space (X, Σ, μ) is said to be* **complete** *if every subset of a negligible set is in Σ.*

Theorem 17 *Given (X, Σ, μ), there is a σ-field, $\Sigma' \supseteq \Sigma$ and a measure μ' on Σ' such that (X, Σ', μ') is complete and $\mu(A) = \mu'(A)$ for any A in Σ.*

We define the extended σ-field as follows

$$\Sigma' = \{Z \in \mathcal{P}(X) | \exists A, B \in \Sigma. A \subseteq Z \subseteq B \wedge \mu(B - A) = 0\}.$$

Theorem 18 (Vitali) *Assuming the axiom of choice, there is no measure on the real numbers which is translation-invariant, assigns a measure to all sets and which assigns a nonzero value to the interval* $[0,1]$.

Lebesgue Measure I

We begin as follows

$$m((a,b)) = m((a,b]) = m([a,b)) = m([a,b]) = b - a$$

where a and b are real numbers. If we have a pairwise disjoint collection of intervals $\{I_k | k \in \mathcal{K}\}$ we define

$$m(\bigcup_{k \in \mathcal{K}} I_k) = \sum_{k \in \mathcal{K}} m(I_k)$$

If A is arbitrary we define $m^*(A)$ to be the **inf** of m over all families of intervals that cover A. Clearly (?) if A is countable $m^*(A)$ is zero.

What properties does μ^* have? It is clearly monotone but not countably additive. We claim that it satisfies the convexity property

$$\mu^*(\cup_i A_i) \leq \sum_i \mu^*(A_i)$$

where $\{A_i\}$ is a countable family not necessarily pairwise disjoint.

Definition 19 *An* **outer measure** *on a set X is a set function $\mu^* : \mathcal{P}(X) \longrightarrow [0, \infty]$, defined on all subsets of X, such that*

1. $\mu^*(\emptyset) = 0$,

2. $A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B)$ *and*

3. *for any countable family of subsets of X, say $\{A_i\}$, we have*

$$\mu^*(\cup_i A_i) \leq \sum_i \mu^*(A_i).$$

Theorem 20 *Let X be a set and let μ^* be an outer measure defined on X. Denote by Σ the collection of all subsets, say A, such that for every subset E of X we have*

$$\mu^*(E) = \mu^*(A \cap E) + \mu^*(A^c \cap E).$$

For all A in Σ define $\mu(A) = \mu^*(A)$. Then (X, Σ, μ) is a measure space.

Integration

Two approaches:

- **Riemann:** Divide up the domain, functions have to be "well-behaved"; continuous (Lebesgue)-almost everywhere. The integral has poor convergence properties. The basis of all numerical algorithms; much more constructive in flavour.

- **Lebesgue:** Divide up the range, functions can be perverse (measurable), very good convergence properties. Hard to see how to make it constructive.

The strategy:

1. First define the integral of characteristic functions
$$\int \chi_A \mu =_{df} \mu(A).$$

2. Using linearity, define the integral of simple functions
$$\int s\mu = \sum_{i=1}^{n} a_i \mu(A_i)$$
where
$$s = \sum_{i=1}^{n} a_i \chi_{A_i}.$$

3. Using the fact that all positive measurable functions are sups of *increasing* sequences of simple functions we define the integral of a positive measurable function as a sup:
$$\int f\mu = \text{sup}\{\int s\mu : s \text{ is simple and } s \leq f\}.$$

It might happen that the integral is infinite: one uses the phrase **integrable function** to mean a measurable function with finite integral. I will be sloppy in the lectures but I am more careful in the notes.

If A is a measurable set we write

$$\int_A f\mu$$

for integration just over the subset A. The definition is

$$\int_A f\mu = \int f \cdot \chi_A \mu.$$

Arithmetic with ∞

We adopt the following conventions:

$$a + \infty = \infty = \infty + a$$

and

$$a \cdot \infty = \infty$$

UNLESS $a = 0$, in which case

$$0 \cdot \infty = 0.$$

Two examples

$$\int_{\mathbf{R}} f(x)\delta_{x_0} = f(x_0)$$

where f is a real-valued function on $\mathbf{R}$ and δ_{x_0} is the Dirac measure concentrated at the point x_0.

Let $f : [0,1] \longrightarrow \mathbf{R}$ be the function that takes the value 1 if x is irrational and 0 if x is rational. Then

$$\int f dx = 1.$$

This is trivial! Our f is a simple function.

Basic Properties

1. $0 \leq f \leq g$ implies $\int_A f\mu \leq \int_A g\mu$
2. $c \cdot \int_A f\mu = \int_A c \cdot f\mu$
3. If $f(x) = 0$ then $\int_A f\mu = 0$ even if $\mu(A) = \infty$
4. If $\mu(A) = 0$ then $\int_A f\mu = 0$ even if f is ∞ everywhere on A
5. $\int (f+g)\mu = \int f\mu + \int g\mu$

The Monotone Convergence Theorem

Theorem 21 *Let $\{f_n\}$ be a sequence of measurable functions on X and suppose that*

1. $\forall x \in X. 0 \leq f_1(x) \leq f_2(x) \leq \ldots \leq \infty$

2. $\forall x \in X. \sqcup_n f_n(x) = f(x)$.

Then f is measurable, and $\sqcup_n \int_X f_n \mu = \int_X f \mu$.

Theorem 22 *If* $\{f_i : i \in I\}$ *is a countable family of measurable functions from* X *to* $[0, \infty]$.

$$\int_X (\sum_i f_i)\mu = \sum_i \int f_i \mu$$

A typical application of the monotone convergence theorem.

The Monotone Convergence Mantra

Want to prove that some property P - involving integration - holds for all measurable functions.

First prove P for characteristic functions only - usually trivial.

Then P holds for all simple functions s by linearity.

Now by the monotone convergence theorem the result follows for all positive measurable functions and by a simple case analysis (usually!) for all functions.

Let $T : (X, \Sigma_X) \longrightarrow (Y, \Sigma_Y)$ be a measurable function. A measure μ on X induces a measure ν on Y by the formula $\nu = \mu \circ T^{-1}$. If $X = Y$, $\Sigma_X = \Sigma_Y$ and $\mu = \nu$ we call T an *ergodic* transformation.

Proposition 23 *If we have X, Y, T as above and $f : Y \longrightarrow \mathbf{R}$ is measurable and $B \in \Sigma_Y$ then*

$$\int_{T^{-1}(B)} f \circ T \mu = \int_B f \nu$$

Theorem 24 *Let* (X, Σ_x, μ) *and* (Y, Σ_Y, ν) *be* σ*-finite measure spaces. Then* ρ *extends uniquely to a measure on* $\Sigma_x \otimes \Sigma_Y$ *such that for all* $E \in \Sigma_x \otimes \Sigma_Y$

$$\rho(E) = \int [\int \chi_E(x,y)\mu]\nu = \int [\int \chi_E(x,y)\nu]\mu.$$

In the above ρ is defined on rectangles by

$$\rho(A \times B) = \mu(A) \times \nu(B).$$

This uses the monotone convergence theorem.

Theorem 25 (Fubini-Tonelli) *Let* (X, Σ_x, μ) *and* (Y, Σ_Y, ν) *be* σ*-finite measure spaces and let* f *be an measurable function with respect to* $\Sigma_X \otimes \Sigma_Y$ *from* $X \times Y$ *to* $[0, \infty]$*. Then*

$$\int f(\mu \times \nu) = \int [\int f\mu]\nu = \int [\int f\nu]\mu.$$

The same holds for not necessarily positive integrable functions.

The proof uses the monotone convergence theorem in several places.

Conditional Probability

$$P(A|B) \stackrel{\text{def}}{=} \frac{P(A \cap B)}{P(B)}$$

Undefined if $P(B)$ is 0.

What to do in the continuous case?

Heuristic approach (Feller):

$$B_1 \supseteq B_2 \supseteq \ldots \supseteq B_n \supseteq \ldots$$

such that

$$\cap_i B_i = \{y\}.$$

$$P(A|y) = \lim_{i \longrightarrow \infty} \frac{P(A \times B_i)}{P(X \times B_i)}.$$

We need some notion of “differentiation of measures” in order to make sense of the last idea.

This is what the **Radon-Nikodym** theorem does.

$$F(x) = \int_{\infty}^{x} f(x)dx$$

If F has a finite jump then f will have a δ-function singularity.

Conjecture 26 *If λ is a measure on $\mathbf{R}$ that has the property $\lambda(\{x\}) = 0$ for any x, then there is some measurable function f such that*

$$\lambda(B) = \int_{B} f(x)dx$$

for any measurable set B.

False BUT if we say $\lambda(B) = 0$ whenever B has Lebesgue measure 0 then it is true.

Definition 27 *If* (X, Σ) *is a measurable space and* $\lambda : \Sigma \longrightarrow [-\infty, \infty]$ *is an extended-real valued function we say that* λ *is* **countably additive (finitely additive***) if, whenever* $\{A_i\}_{i\in I}$ *is a countable (finite) family of* Σ*-measurable sets that are pairwise disjoint then*

$$\lambda(\cup_{i\in I} A_i) = \Sigma_{i\in I}\lambda(A_i).$$

The following easy facts are proved just like the analogous results for measures.

Lemma 28 *Let* λ *be a countably additive set function on* (X, Σ)*. All the sets mentioned in this lemma are assumed* Σ*-measurable.*

1. *If* $A_i \uparrow A$ *then* $\lim_{i\longrightarrow\infty} \lambda(A_i) = \lambda(A)$*.*

2. *If* $A_i \downarrow A$ *and* $\lambda(A_1)$ *is finite, then* $\lim_{i\longrightarrow\infty} \lambda(A_i) = \lambda(A)$*.*

Hahn-Jordan Decomposition

Theorem 29 *Let* λ *be a countably additive set function on* (X, Σ). *Then there are sets* $C, D \in \Sigma$ *such that* $\lambda(C) = sup\{\lambda(A) : A \in \Sigma\}$ *and* $\lambda(D) = inf\{\lambda(A) : A \in \Sigma\}$.

Remark: This theorem says that λ actually attains a maximum and a minimum value. This is obviously trivial for a measure.

Definition 30 *Let λ be a set function defined on the measurable space (X, Σ). A set A is called* **positive** *with respect to λ if $\forall E \subseteq A . \lambda(E) \geq 0$; similarly we define* **negative**.

Theorem 31 (Hahn decomposition) *For any countably additive set function λ defined on the measurable space (X, Σ), there are disjoint sets A^+ and A^- such that $X = A^+ \cup A^-$ where A^+ (A^-)is positive (negative) with respect to λ.*

Corollary 32 (Jordan decomposition) *If λ is countably additive set function defined on a measurable space (X, Σ), then there are two positive measures λ^+ and λ^- such that for all $S \in \Sigma$ we have*

$$\lambda(S) = \lambda^+(S) - \lambda^-(S).$$

Definition 33 *Two measures,* μ, ν *on a measurable space* (X, Σ) *are* **mutually singular**, *written* $\mu \perp \nu$, *if there are disjoint measurable sets* A, B *with* $\mu(X \setminus A) = 0$ *and* $\nu(X \setminus B) = 0$.

Definition 34 *Suppose that* μ *and* ν *are measures defined on a measurable space* (X, Σ). *We say that* ν *is* **absolutely continuous** *with respect to* μ, *written* $\nu << \mu$, *if* $\forall A \in \Sigma . \mu(A) = 0 \Rightarrow \nu(A) = 0$.

Theorem 35 (Radon-Nikodym-Lebesgue) *If μ and ν are both σ-finite measures on a measurable space (X, Σ) then:*

1. *ν can be written as $\nu_a + \nu_s$ where $\nu_a << \mu$ and $\nu_s \perp \mu$.*

2. *there is a non-negative measurable function f such that*

$$\forall A \in \Sigma. \int_A f\mu = \nu_a(A).$$

If g is another function satisfying the same property as f then the set of points where f and g differ have μ measure 0.

Remark 36 Part (1) of the theorem is usually called the Lebesgue decomposition while part (2) is usually called the Radon-Nikodym theorem. The function f (unique μ-almost everywhere) is often called the Radon-Nikodym derivative and is written $\frac{d\nu_a}{d\mu}$

Suppose we have (X, Σ_X) and (Y, Σ_Y) with a probability measure P on $X \times Y$. We get two probability measures

$$P_X(A) = P(A \times Y), P_Y(B) = P_Y(X \times B)$$

called the *marginals*.

How do we define a reasonable notion of conditional probability in this case?

Answer: Work with densities of the form $P(x, B)$ and $P(y, A)$.

Proposition 37 *Given the product space $X \times Y$ of the previous slide with a joint probability distribution P there are, for each fixed subset $B \in \Sigma_Y$ and for each fixed subset $A \in \Sigma_X$ measurable functions $P(x, B)$ and $P(y, A)$, such that*

$$P(A \times B) = \int P(x, B) dP_X = \int P(y, A) dP_Y.$$

These functions are unique P_X-almost (or P_Y-almost) everywhere.

Proposition 38 *Let $f : X \to Y$ be a measurable function and P a probabilistic distribution on X. Let P_f be the probability measure $P \circ f^{-1}$ on Y induced by f. Then, for each $A \in \Sigma_X$, there is a P_f-almost everywhere unique measurable function, written $P(A|y)$, such that*

$$\forall B \in \Sigma_Y . P(f^{-1}(B) \cap A) = \int_B P(A|y) dP_f.$$

More general situation:
Conditioning with respect to a σ-field.

Theorem 39 *Let* (X, Σ, P) *be a probability space and let* Γ *be a subfield of* Σ*. For each* A*, a member of* Σ*, there is a measurable function,* P*-almost everywhere unique,* $P[A|\Gamma] : (X, \Gamma) \longrightarrow (\mathbf{R}, \mathcal{B})$*, called the conditional probability of* B *given* Γ*, such that*

$$\forall C \in \Gamma . P(A \cap C) = \int_C P[A|\Gamma] dP.$$

The object $P[A|\Gamma]$ is a measurable function with a dependence on A. How does $P[\cdot|\Gamma](x)$ behave viewed as a set function with x fixed?

It is not always a measure - certainly Radon-Nikodym by itself does not do it.

Definition 40 *Let f be a random variable on (X, Σ, P) and let Γ be a subfield of Σ. We say that a function $F : X \times \mathbf{R} \longrightarrow [0, 1]$ is a* **regular conditional distribution function** *for f given Γ iff the following two conditions are satisfied:*

1. *for each x, $F(x, \cdot)$ is a distribution function, i.e. increasing, right-continuous and* $\lim_{r \longrightarrow \infty} F(x, r) = 1$ *and* $\lim_{r \longrightarrow -\infty} F(x, r) = 0$,

2. $\forall r \in \mathbf{R} F(x, r) = P[f^{-1}((-\infty, r])|\Gamma](x)$.

Polish and Analytic spaces

A **Polish** space is the topological space underlying a complete separable, metric space.

Examples: $\mathbf{R}$ and $\mathbf{N}^\infty$

An analytic space is the continuous or measurable image of a Polish space under a map of Polish spaces. Every analytic space is the continuous image (or measurable image) of $\mathbf{N}^\infty$.

Theorem 41 *Let $f : (X, \Sigma_X, P) \longrightarrow (Y, \Sigma_Y)$ be a measurable function and Y be a analytic space with Σ_Y its σ-field of Borel sets. Let Γ be a sub σ-field of Σ_X. Then a regular conditional probability for f given Γ exists.*

The gadget $P(x, B)$ is called a "stochastic kernel" or "Markov kernel".

Stochastic Relations - **SRel**

Objects: (X, Σ_X), sets with a σ-field.
Morphisms: Stochastic kernels:

$$h : (X, \Sigma_X) \longrightarrow (Y, \Sigma_Y)$$

is a function $h : X \times \Sigma_Y \longrightarrow [0, 1]$ such that

1. $\forall B \in \Sigma_Y . \lambda x \in X . h(x, B)$ is a bounded measurable function,

2. $\forall x \in X . \lambda B \in \Sigma_Y . h(x, B)$ is a subprobability measure on Σ_Y.

Composition in **SRel**

$$h : (X, \Sigma_X) \longrightarrow (Y, \Sigma_Y) \; h : X \times \Sigma_Y \longrightarrow [0, 1]$$

$$k : (Y, \Sigma_Y) \longrightarrow (Z, \Sigma_Z) \; k : Y \times \Sigma_Z \longrightarrow [0, 1]$$

$$(k \circ h)(x, C) = \int_Y k(y, C) h(x, dy)$$

The identity morphisms:

$$Id_X : X \longrightarrow X \text{ is } \delta(x, A)$$

Proposition 42 *With composition defined as above* $\mathbf{SRel}$ *is a category.*

$$\int_Y [\int_Z p(z, D) k(y, dz)] h(x, dy) =$$

$$\int_Z p(z, D) [\int_Y k(y, dz) h(x, dy)].$$

Use the monotone convergence mantra.

SRel as a Kleisli category

Let **Mes** be the category of measurable spaces and measurable functions.

$\Pi(X) =_{df} \{\nu | \nu$ is a subprobability measure on$X\}$.

Given $A \in \Sigma_X$ we have $p_A : \Pi(X) \longrightarrow [0,1]$ given by $p_A(\nu) = \nu(A)$. Choose the "smallest" σ-field on $\Pi(X)$ that makes these maps measurable.

If we have $f : X \longrightarrow Y$ then

$$\Pi(f)(\nu) = \nu \circ f^{-1}.$$

Π is a monad on **Mes**

$\eta : I \longrightarrow \Pi$ is given by

$$\eta_X(x) = \delta(x, \cdot)$$

and
$\mu : \Pi^2 \longrightarrow \Pi$ is

$$\mu_X(\Omega) = \lambda B \in \Sigma_X . \int_{\Pi(X)} p_B \Omega.$$

Verification uses the monotone convergence theorem very often.

Why is the Kleisli category of Π the same as **SRel**?

$f : X \longrightarrow Y$ in $\mathcal{K}_{\Pi}$ is
$f : X \longrightarrow \Pi(Y)$ in **Mes**
or
$f : X \longrightarrow (\Sigma_Y \longrightarrow [0, 1])$
or uncurrying
$f : X \times \Sigma_Y \longrightarrow [0, 1]$

SRel has countable coproducts

Given $\{(X_i, \Sigma_I) | i \in I\}$ in **SRel** we define

$$X = \biguplus_i X_i$$

with evident σ-field. A measurable set will look like

$$\biguplus_{i \in I} A_i \times \{i\}$$

The injections $\iota_i : X_i \longrightarrow X$ are

$$\iota_i(x, \biguplus_{k \in I} A_k) = \delta((x, i), \biguplus_{k \in I} A_k) = \delta(x, A_i).$$

Given $f_j : X_j \longrightarrow (Y, \Sigma_Y)$ of $\mathbf{SRel}$ morphisms the mediating morphism $f : X \longrightarrow Y$ is $f((x, i), B) = f_i(x, B)$.

Rel is a compact-closed category, what about **SRel**?

It has a tensor but no easy closed structure we could find. It (actually a closely related category) has a related structure – but that is another story.

Rel has biproducts but the **SRel** coproduct is not a product - though it really wants to be!

Partially Additive Structure

One can add **SRel** morphisms - sometimes!

Given $f, g : X \longrightarrow Y$ we define

$$(f+g)(x,B) = f(x,B) + g(x,B)$$

if $f(x,Y) + g(x,Y)$ is bounded by 1 for all x.

Definition 43 *A* **partially additive monoid** *is a pair* $(M, \sum)$ *where* M *is a nonempty set and* $\sum$ *is a partial function which maps* some *countable subsets of* M *to* M. *We say that* $\{x_i | i \in I\}$ *is* **summable** *if* $\sum_{i \in I} x_i$ *is defined. The following axioms are obeyed.*

1. **Partition-Associativity:** *Suppose that* $\{x_i | i \in I\}$ *is a countable family and* $\{I_j | j \in J\}$ *is a countable partition of* I. *Then* $\{x_i | i \in I\}$ *is summable iff for every* $j \in J$ $\{x_i | i \in I_j\}$ *is summable and* $\{\sum_{i \in I_j} x_i | j \in J\}$ *is summable. In this case we require*
$$\sum_{i \in I} x_i = \sum_{j \in J} \sum_{i \in I_j} x_i.$$

2. **Unary-sum:** *A singleton family is always summable.*

3. **Limit:** *If* $\{x_i | i \in I\}$ *is countable and* every finite subfamily *is summable then the whole family is summable.*

Definition 44 *Let $\mathcal{C}$ be a category. A* **partially additive structure** *on $\mathcal{C}$ is a partially additive monoid structure on the homsets of $\mathcal{C}$ such that if $\{f_i : X \longrightarrow Y | i \in I\}$ is summable, then $\forall W, Z, g : W \longrightarrow X, h : Y \longrightarrow Z$ we have that $\{h \circ f_i | i \in I\}$ and $\{f_i \circ g | i \in I\}$ are summable and, furthermore, the equations*

$$h \circ \sum_{i \in I} f_i = \sum_{i \in I} h \circ f_i, (\sum_{i \in I} f_i) \circ g = \sum_{i \in I} f_i \circ g$$

hold.

Proposition 45 *The sum of the empty family exists, call it* 0*. It is the identity for* $\sum$*.*

Definition 46 *A category has* **zero morphisms** *if there is a distinguished morphism in every homset, we write* 0_{XY} *for the distinguished member of* $hom(X,Y)$*, such that* $\forall W,X,Y,Z,f : W \longrightarrow X, g : Z \longrightarrow Y$ *we have* $g \circ 0_{WZ} = 0_{XY} \circ f$*.*

Proposition 47 *If a category has a partially additive structure it has zero morphisms.*

Proposition 48 *The category* $\mathbf{SRel}$ *has a partially additive structure.*

How do we prove that composition distributes over addition?

The monotone convergence mantra of course!

Definition 49 *Let $\mathcal{C}$ be a category with countable coproducts and zero morphisms and let $\{X_i | i \in I\}$ be a countable family of objects of $\mathcal{C}$.*

1. *For any $J \subset I$ we define the* **quasi-projection** $PR_J : \coprod_{i \in I} X_i \longrightarrow \coprod_{j \in J} X_j$ *by*

$$PR_J \circ \iota_i = \begin{cases} \iota_i & i \in J \\ 0 & i \notin J \end{cases}$$

2. *We write $I \cdot X$ for the coproduct of $|I|$ copies of X. We define the* **diagonal-injection** Δ *by couniversality:*

$$\begin{tikzcd}
\coprod(X_i | i \in I) \arrow[r, "\Delta"] & I \cdot \coprod(X_i | i \in I) \\
X_j \arrow[u, "in_j"] \arrow[r, "in_j"] & \coprod(X_i | i \in I) \arrow[u, "in_j"]
\end{tikzcd}$$

3. *We have a morphism σ from $I \cdot X$ to X given by:*

$$
\begin{tikzcd}
I \cdot X \arrow[r, "\sigma"] & X \\
X \arrow[u, "in_j"] \arrow[ur, "id_X"'] &
\end{tikzcd}
$$

$$PR_J((x,k),\uplus_{j\in J}) = \begin{cases} \delta(x,A_k) & k\in J \\ 0 & k\notin J \end{cases}$$

The Δ morphism in $\mathbf{SRel}$ is

$$\Delta((x,k),\uplus_{i\in I}(\uplus_{j\in I}A^i_j)) = \delta(x,A^k_k).$$

The analogous map in **Set** is $\Delta((x,k)) = ((x,k),k)$. Finally

$$\sigma((x,k),A) = \delta(x,A)$$

in $\mathbf{SRel}$ while in **Set** we have $\sigma((x,k)) = x$.

Definition 50 *A* **partially additive category**, *$\mathcal{C}$ is a category with countable coproducts and a partially additive structure satisfying the following two axioms.*

1. **Compatible sum axiom**: *If $\{f_i | i \in I\}$ is a countable set of morphisms in $\mathcal{C}(X, Y)$ and there is a morphism $f : X \longrightarrow I \cdot Y$ with $PR_i \circ f = f_i$ then $\{f_i | i \in I\}$ is summable.*

2. **Untying axiom**: *If $f, g : X \longrightarrow Y$ are summable then $\iota_1 \circ f$ and $\iota_2 \circ g$ are summable as morphisms from X to $Y + Y$.*

Theorem 51 (Arbib-Manes) *Given* $f : X \longrightarrow X + Y$ *in a partially additive category, we can find a unique* $f_1 : X \longrightarrow X$ *and* $f_2 : X \longrightarrow Y$ *such that* $f = \iota_1 \circ f_1 + \iota_2 \circ f_2$. *Furthermore there is a morphism* $\dagger f =_{df} \sum_{n=0}^{\infty} f_2 \circ f_1^n : X \longrightarrow Y$. *The morphism* $\dagger f$ *is called the* **iterate** *of* f.

$$S ::== x_i := f(\vec{x}) | S_1 ; S_2 |$$

$$if \; \mathbf{B} \; then \; S_1 \; else \; S_2 | while \; \mathbf{B} \; do \; S.$$

We assume that the program has n variables $\vec{x}$, and they each take values in (X, Σ). We model statements as $\mathbf{SRel}$ morphisms of type $(X^n, \Sigma^n) \longrightarrow (X^n, \Sigma^n)$. We write $\vec{A}$ for the product $A_1 \times \ldots \times A_n$

Assignment: $x := f(\vec{x})$

$$[\![\, x_i := f(\vec{x}) \,]\!](\vec{x}, \vec{A}) =$$

$$\delta(x_1, A_1) \ldots \delta(x_{i-1}, A_{i-1})$$

$$\delta(f(\vec{x}), A_i)$$

$$\delta(x_{i+1}, A_{i+1}) \ldots \delta(x_n, A_n)$$

Sequential Composition: $S_1; S_2$

$$[\![\, S_1; S_2 \,]\!] = [\![\, S_2 \,]\!] \circ [\![\, S_1 \,]\!]$$

Conditionals: $if\ \mathrm{B}\ then\ S_1\ else\ S_2$

$$[\![if\ \mathrm{B}\ then\ S_1\ else\ S_2]\!](\vec{x}, \vec{A}) =$$

$$\delta(\vec{x}, [\![\mathbf{B}]\!])[\![S_1]\!](\vec{x}, \vec{A}) + \delta(\vec{x}, [\![\mathbf{B}]\!]^c)[\![S_2]\!](\vec{x}, \vec{A})$$

where U^c is the complement of U.

While Loops: $\mathit{while}\ \mathbf{B}\ \mathit{do}\ S$

$$[\![\mathit{while}\ \mathbf{B}\ \mathit{do}\ S]\!] = h^{\dagger}$$

where we are using the $\dagger$ in $\mathbf{SRel}$ and the morphism $h : (X^n, \Sigma^n) \longrightarrow (X^n, \Sigma^n) + (X^n, \Sigma^n)$ is given by

$$h(\vec{x}, \vec{A_1} \uplus \vec{A_2}) =$$

$$\delta(\vec{x}, [\![\mathbf{B}]\!]) [\![S]\!] (\vec{x}, \vec{A_1}) + \delta(\vec{x}, [\![\mathbf{B}]\!]^c) \delta(\vec{x}, \vec{A_2}).$$

Stochastic Processes

Definition 52 *A* **stochastic process** *is an indexed family of random variables* $X_t : \Omega \longrightarrow \mathbf{R}$ *where* $(\Omega, \mathcal{B}, P)$ *is a probability space and* $t \in T$ *is the indexing set.*

Alternatively: a functor from ω to **SRel**.

Joint distributions can be defined:

$$P_{t_1 \ldots t_n}(B) = P(\{x | (X_{t_1}(x), \ldots, X_{t_n}(x)) \in B\}).$$

First Kolmogorov consistency requirement

$$P_{t_1 \dots t_n t_{n+1}}(B \times \mathbf{R}) = P_{t_1 \dots t_n}(B).$$

The second reqirement says that the distributions behave the way they should under permutation of the variables.

Fundamental Theorem: Any family of finite-dimensional probability distributions satisfying these requirements can be realized as a stochastic process.

Markov Processes

We write $P(A_{n+1}|x_1,\ldots,x_n)$ for the conditional probability that the system is in the set A_{n+1} given that at time t_1 it was at x_1 etc.

In a Markov Process we only depend on the last state:

$$P(A_{n+1}|x_1,\ldots,x_n) = P(A_{n+1}|x_n).$$

These are precisely what can be described as matrices or stochastic kernels.

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.

Discrete Probabilistic Systems

Definition 53 *A* **labelled Markov chain** *is a quadruple* $(S, \mathcal{L}, C_l, P_l)$*, where* S *is a countable set of states,* $\mathcal{L}$ *is a set of labels, and for each* $l \in \mathcal{L}$*, we have a subset* C_l *of* S*, a function,* P_l*, called a* **transition probability matrix**,

$$P_l : C_l \times S \longrightarrow [0, 1]$$

satisfying the normalization condition

$$\forall l \in \mathcal{L}, s \in C_l.\Sigma_{s' \in S} P_l(s, s') = 1.$$

If we have the weaker property,

$$\forall l \in \mathcal{L}, s \in C_l.\Sigma_{s' \in S} P_l(s, s') \leq 1$$

we call the system a **partial** *labelled Markov chain.*

Larsen-Skou Bisimulation

Definition 54 *Let* $T = (S, P_l)$ *be a labelled Markov chain. Then a* **probabilistic bisimulation** $\equiv_p$, *is an equivalence on* S *such that, whenever* $s \equiv_p t$, *the following holds:*

$$\forall l \in \mathcal{L}. \forall A \in S/\equiv_p,\ \Sigma_{s' \in A} P_l(s, s') = \Sigma_{s' \in A} P_l(t, s').$$

Two states s *and* t *are said to be* **probabilistically bisimilar** *(*$s \sim_{LS} t$*) in case* (s, t) *is contained in some probabilistic bisimulation.*

Defining Labelled Markov Processes

Definition 55 *A* **partial labelled Markov process** *with label set* $\mathcal{L}$ *is a structure* $(S, \Sigma, \{k_l \mid l \in \mathcal{L}\})$*, where* S *is the set of states, which is assumed to be an analytic space, and* Σ *is the* σ*-field on* S*, and*

$$\forall l \in \mathcal{L}, k_l : S \times \Sigma \longrightarrow [0, 1]$$

is a transition sub-probability function.

Analytic spaces guarantee that we can find regular conditional probability densities.

Simulation

Definition 56 *A* **simulation morphism** f *between two partial labelled Markov processes, (S, Σ, k_l) and (S', Σ', k'_l) is a measurable function $f : (S, \Sigma) \longrightarrow (S', \Sigma')$ such that*

$$\forall l \in \mathcal{L}. \forall x \in S. \forall \sigma' \in \Sigma'. k_l(x, f^{-1}(\sigma')) \leq k'_l(f(x), \sigma').$$

Proposition 57 *Given two partial labelled Markov chains, a simulation morphism between them is also a simulation morphism between the associated labelled transition systems.*

Definition 58 *The objects of the category* $\mathbf{LMP}$ *are labelled Markov processes, having $\mathcal{L}$ as set of labels, with simulations as the morphisms. The category of labelled, Markov chains is written* $\mathbf{LMC}$ *and is the full subcategory of* $\mathbf{LMP}$ *that includes only the labelled, Markov chains as objects.*

Zigzag morphisms

Definition 59 *A morphism f from (S, Σ, k_l) to (S', Σ', k'_l) is* a zigzag morphism *if it satisfies the properties:*

1. *f is surjective;*

2. *f is measurable;*

3. *$\forall l \in \mathcal{L}, s \in S, \sigma' \in \Sigma', \;\; k_l(s, f^{-1}(\sigma')) = k'_l(f(s), \sigma')$.*

Bisimulation as span

Definition 60 *Let T and T' be two labelled Markov processes. T is* probabilistically bisimilar *to T' (written $T \sim T'$) if there is a span of zigzag morphisms between them, i.e. there exists a labelled Markov process S and zigzag morphisms f and g such that*

$$\begin{tikzcd} & S \arrow[dl, "f"'] \arrow[dr, "g"] & \\ T & & T' \end{tikzcd}$$

Notice that if there is a zigzag morphism between two systems, they are bisimilar since the identity is a zigzag morphism.

Theorem 61 *Bisimulation, defined as above on labelled Markov processes, is an equivalence relation.*

Composing spans requires that we can "complete" the square.

$$\begin{tikzcd}
S_1 \arrow[dr, "f_1"] & & S_2 \arrow[dl, "f_2"'] \\
& S &
\end{tikzcd}$$

We can always do this if we have zigzags.

Warning We earlier claimed that we have pullbacks but this is **FALSE**, we can, however, always complete the square.

Try to mimic the pullbacks in **Set**:
we take U to be all pairs (s_1, s_2) from $S_1 \times S_2$ such that $f_1(s_1) = f_2(s_2)$.

We are forced to have: $h((s_1, s_2), (\sigma_1 \times S_2) \cap U) = k_1(s_1, \sigma_1)$ and similarly for π_2.

Unfortunately, sets of the form $\sigma_1 \times S_2$ and $S_1 \times \sigma_2$ are not enough to determine a measure on all of Σ_U.

Given a set of the form

$$(\sigma_1 \times \sigma_2) \cap U$$

we do not have σ_1 and σ_2 unambiguously determined.

The Idea

Forget about U, define h on all of $S_1 \times S_2$ but using conditional probabilities to make the transitions agree. The resulting measure is **automatically** supported only on U.

We have the two conditional probability densities

$$P((s_i, \sigma_i)|\Sigma_i') : S_i \to [0, 1]$$

of the sub-probability measure $k_i(s_i, \cdot)$ given the sub-σ-field $\Sigma_i' \subseteq \Sigma_i$, where $\Sigma_i' = f_i^{-1}(\Sigma)$

$P((s_i, \sigma_i)|\Sigma_i')(\cdot)$ is a *regular* conditional probability distribution, i.e.

1. $P((s_i, \sigma_i)|\Sigma_i') : S_i \to [0, 1]$ is Σ_i' measurable and integrable.

2. For all $\gamma_i' \in \Sigma_i'$, we have

$$\int_{\gamma_i'} P((s_i, \sigma_i)|\Sigma_i')(\omega_i) k_i(s_i, d\omega_i) = k_i(s_i, \sigma_i \cap \gamma_i').$$

3. For $k_i(s_i, \cdot)$-almost all $\omega_i \in S_i$, $P((s_i, \cdot)|\Sigma_i')(\omega_i) : \Sigma_i \to [0, 1]$ is a sub-probability measure on S_i.

One should think of $P((s_i, \sigma_i)|\Sigma_i')(\omega_i)$ as the probability that s_i makes a transition to σ_i given that s_i makes a transition to $f_i^{-1} f_i \omega_i$, or equivalently as the probability that s_i makes a transition to σ_i given that $f_i s_i$ makes a transition to $f_i \omega_i$.

We can construct (using Radon-Nikodym)

$$P((s_i, \sigma_i)|\Sigma) : S \to [0, 1].$$

Here, $P((s_i, \sigma_i)|\Sigma)(s)$ gives the probability that s_i makes a transition to σ_i given that s_i makes a transition to some ω_i with $f_i(\omega_i) = s$ or equivalently the probability that s_i makes a transition to σ_i given that $f_i s_i$ makes a transition to s.

1. $P((s_i, \sigma_i)|\Sigma)$ is Σ-measurable and integrable.

2. Since $k_i(s_i, \cdot) \circ f_i^{-1} = k(f_i s_i, \cdot)$, we have, for all $\sigma \in \Sigma$:

$$\int_\sigma P((s_i, \sigma_i)|\Sigma) k(f_i s_i, ds)$$
$$= \int_{f_i^{-1}\sigma} P((s_i, \sigma_i)|\Sigma_i') k_i(s_i, d\omega_i)$$
$$= k_i(s_i, \sigma_i \cap f_i^{-1}\sigma).$$

3. For $k(f_i s_i, \cdot)$-almost all s, we have a sub-probability measure

$$P((s_i, \cdot)|\Sigma)(s) : \Sigma_i \to [0, 1]$$

on S_i.

Let $\mathcal{R} = \Sigma_1 \times \Sigma_2$ be the collection of all "rectangles" $\sigma_1 \times \sigma_2 \in \Sigma_{S_1 \times S_2}$; $\mathcal{R}$ is a semi-ring. For $\sigma_1 \times \sigma_2 \in \mathcal{R}$ and $s \in S$, put

$$\rho((s_1, s_2), \sigma_1 \times \sigma_2)(s) =$$

$$P((s_1, \sigma_1)|\Sigma)(s) P((s_2, \sigma_2)|\Sigma)(s).$$

Proposition 62 *For $k(f_1 s_1, \cdot)$-almost all $s \in S$, $\rho((s_1, s_2), \cdot)(s)$ is countably additive on $\mathcal{R}$.*

For $\sigma_1 \times \sigma_2 \in \mathcal{R}$, put

$$g((s_1, s_2), \sigma_1 \times \sigma_2) = \tag{1}$$
$$\int_S P((s_1, \sigma_1)|\Sigma)(s) P((s_2, \sigma_2)|\Sigma)(s) k(f_1 s_1, ds).$$

Proposition 63 *$g((s_1, s_2), \cdot)$ has a unique extension to a sub-probability measure on $(S_1 \times S_2, \Sigma_{S_1 \times S_2})$.*

Proposition 64 *For any $\sigma_1 \in \Sigma_1$ and $\sigma_2 \in \Sigma_2$ with $U \cap (\sigma_1 \times \sigma_2) = \emptyset$ we have $g((s_1, s_2), \sigma_1 \times \sigma_2) = 0$.*

Corollary 65 *$g((s_1, s_2), (S_1 \times S_2) \setminus U) = 0$.*

Theorem 66 $(U, \Sigma_U, h : U \times \Sigma_U \to [0,1])$ *is an object of the category and* $\pi_i : U \to S_i$ *are morphisms.*

A Modal Logic for Bisimulation

$$\mathsf{T} | \phi_1 \wedge \phi_2 | \langle a \rangle_q \phi$$

We say $s \models \langle a \rangle_q \phi$ iff $\exists A \in \Sigma.(\forall s' \in A.s' \models \phi) \wedge (k_a(s, A) \geq q)$.

Proposition 67 *For all formulas ϕ, we have $[\![\phi]\!] \in \Sigma$.*

Proposition 68 *If f is a simulation morphism from S to S' then for all state $s \in S$ and all formulas ϕ,*

$$s \models \phi \Rightarrow f(s) \models \phi.$$

If, moreover, f is a zigzag morphism, the converse is also true.

We say that S and S' satisfy *all the same formulas* if $\forall s \in S \exists s' \in S'$ such that s and s' satisfy all the same formulas (written $s \approx s'$) and the same with s and s' interchanged.

Corollary 69 *If S and S' are bisimilar, then they satisfy the same formulas.*

Given (S, Σ) the quotient $(S/\approx, \Sigma_\approx)$ has an evident σ-field.

Proposition 70 *Let* (S, Σ, k_a) *be a* **LMP**. *We can define* h_a *so that the canonical projection* q *from* (S, Σ, k_a) *to* $(S/\approx, \Sigma_\approx, h_a)$ *is a zigzag morphism.*

Theorem 71 *Two systems are bisimilar iff they obey the same formulas of our logic.*

$$
\begin{tikzcd}
S \arrow[dr, "\iota"] & & S' \arrow[dl, "\iota'"'] \\
& T \arrow[d, "r"'] & \\
& T/{\approx} &
\end{tikzcd}
$$