Here is the highly referenced seminar handout notes by Lawvere, 1962. The pen markings are mine; in several places my initial thoughts are incorrect (e.g., $\mathcal{P}$ has products therefore... In fact $\mathcal{P}$ does not have products or equalizers -only weak products and weak equalizers). Rather than "correcting them" by more markings I left them incorrect; we have a rather detailed analysis of this category as we were trying to determine if it had equalizers (we proved it doesn't). Attached to his notes is a recent email exchange I had with him concerning "probabilistic relations" using this category.

G.-C. Rota

THE CATEGORY OF PROBABILISTIC MAPPINGS

— With Applications to Stochastic Processes, Statistics, and Pattern Recognition

- by F. W. Lawvere

I. Objects and Maps in the Category of Probabilistic Mappings

I.1 Measurable Spaces

I.1.1 The objects which we consider are measurable spaces Ω. That is, $\Omega = \langle S, \mathbb{B} \rangle$ will be an ordered pair in which S is any set and $\mathbb{B}$ is any σ-algebra of subsets of S. This means that:

A σ-algebra $\langle S, \mathbb{B} \rangle$ forms a Boolean algebra under $\vee \equiv \cup$, $B_1 \wedge B_2 = B_1^c \cup B_2^c$ and the partial order $\leq \equiv \subseteq$. The top element $\top = S$, bottom $\bot = \phi$

(0) Every member of $\mathbb{B}$ is a subset of S.

$\forall B_1, B_2 \in \mathbb{B}$ (1) The empty set φ and the "whole space" S are members of $\mathbb{B}$.

These imply $B_1 \cap B_2 \in \mathbb{B}$

(2) If $B \in \mathbb{B}$ (i.e., if B is a member of $\mathbb{B}$) then the complement $(S \sim B) \in \mathbb{B}$.

- A σ-algebra gives a topological space

(3) If B_i, $i = 0, 1, 2, \cdots$ is any countable family of members of $\mathbb{B}$, then the union $\bigcup_{i=0}^{\infty} B_i$ is also a member of $\mathbb{B}$.

We also say that $\mathbb{B}$ is the class of measurable sets of Ω.

A σ-alg. is (co)complete. Moreover, it has exponentials $x^y = \bigvee\{w \mid w \wedge y \leq x\}$ so a σ-alg. is a locale.

1.2 If $\Omega = \langle S, \mathbb{B} \rangle$ is any measurable space and if f is a function defined on S with values in a partially ordered set Λ, then f is said to be Ω-Λ measurable if for each $\lambda \in \Lambda$ we have $\{\omega \mid f(\omega) \leq \lambda\} \in \mathbb{B}$; that is, if the set of all $\omega \in \Omega$ whose value under f precedes a given λ is measurable for each λ. For example, we will use this notion when $\Lambda = R$, the real number.

$= \{\omega \mid \omega \in f^{-1}(\downarrow\lambda)\}$

$f^{-1}(\Omega)$ see

Choose the σ-alg. on Λ to be generated by $\{\downarrow\alpha\}_{\alpha \in \Lambda}$ where $\downarrow\alpha = \{\lambda \mid \lambda \leq \alpha\}$.

1.3 More generally, if $\Omega = \langle S, \mathbb{B} \rangle$ and $\Omega' = \langle S', \mathbb{B}' \rangle$ are any measurable spaces, and if f is any function defined on S with values in S', then f is said to be a measurable mapping if and only if $f^{-1}(B') \in \mathbb{B}$ for every $B' \in \mathbb{B}'$, where $f^{-1}(B')$ denotes the set of all $x \in S$ for which $f(x) \in B'$. The foregoing paragraph is seen to be a special case of this by considering $\Omega' = \langle \Lambda, \mathbb{B}(\Lambda) \rangle$ where $\mathbb{B}(\Lambda)$ is the smallest σ-algebra containing all sets of the form $\{\lambda' \mid \lambda' \leq \lambda\}$ for all $\lambda \in \Lambda$.

$S \xrightarrow{f} S'$

~~Isolate~~ $P : \mathbb{B} \to [0,1]$ ~~is ... that~~ ~~induces a mapping~~

Is a probability measure measurable? no, $\mathrm{dom}(P) = \mathbb{B}$ $(\neq S)$

1.4 If $\Omega = \langle S, \mathbb{B} \rangle$ is a measurable space, then by a probability measure on Ω is meant a function P which assigns to every measurable set $B \in \mathbb{B}$ a real number $P(B)$, in such a way that:

(0) $0 \leq P(B) \leq 1$ for every $B \in \mathbb{B}$

(1) $P(S) = 1$

(2) If $B_i \in \mathbb{B}$ for $i = 1, 2, \cdots$ and if $B_i \cap B_j = 0$ for $i \neq j$ (i.e., B_i are pair-wise disjoint measurable sets) then

$$P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i) .$$

1.5 In case S is a countable set and $\mathbb{B}$ consists of all subsets of S, then for any probability measure P on $\langle S, \mathbb{B} \rangle$ and any $B \in \mathbb{B}$, we have

$$P(B) = \sum_{x \in B} P(\{x\})$$

where $\{x\}$ is the "singleton" subset of S whose only member is x, for each $x \in B$. Thus, in this case, a probability measure is already determined by a function $p(x) = P(\{x\})$ of members of S; this function is arbitrary, save for the two conditions $0 \leq p(x) \leq 1$, $\sum_{x \in S} p(x) = 1$.

$p : S \to [0,1]$ $\quad$ $P : \mathbb{B} \to [0,1]$

$S \xrightarrow{p} [0,1]$, $\mathbb{B} = 2^S \xrightarrow{P} [0,1]$

If S is not countable, then probability measures on Ω are not determined by their values at singletons. For example, if $S = \{x \mid 0 \leq x \leq 1\}$ = the "unit interval", and if $\mathbb{B}$ = the smallest σ-algebra containing all closed subintervals = the class of "Borel sets", then there are a great many probability measures P on $\Omega = \langle S, \mathbb{B} \rangle$ for which $P(\{x\}) = 0$ for all x. For example, in this case P = Lebesque measure = (generalized) length is a probability measure but every singleton has zero probability.

1.6 If $\Omega = \langle S, \mathbb{B} \rangle$ and $\Omega' = \langle S', \mathbb{B}' \rangle$ are measurables spaces, if f is a measurable mapping (1.3) from Ω to Ω', and if P is a probability measure on Ω, then the probability measure Pf induced on Ω' by P via f is defined by

$$(Pf)(B') \triangleq P(f^{-1}(B'))$$

for every $B' \in \mathbb{B}'$. Note that in $\mathcal{P}$ = cat of prob. mappings, that Pf is the image of $f \circ P$: $1 \xrightarrow{P} (X,\mathcal{B}) \xrightarrow{f} (X',\mathcal{B}')$, $f \circ P$, $\mathrm{im}_{fP} (= Pf)$

Pf is the smallest monic thru which $f \circ P$ factors since if $1 \xrightarrow{f \circ P} (X',\mathcal{B}')$, $1 \xrightarrow{g} (Y,\mathcal{C}) \xrightarrow{m} (X',\mathcal{B}')$ then $Pf = m \circ g$ and so $Pf \leq m$ i.e., $1 \xrightarrow{Pf} (X',\mathcal{B}')$, $1 \xrightarrow{g} (Y,\mathcal{C}) \xrightarrow{m} (X',\mathcal{B}')$ so g is the required factorization.

3

To verify that Pf is probability measure (i.e., satisfies the conditions 0, 1, 2 of 1.4) note that the mapping f^{-1} from $\mathbb{B}'$ to $\mathbb{B}$ is a σ-homomorphism; i.e., that

$$f^{-1}(S' \sim B') = S \sim f^{-1}(B') \qquad \text{for } B' \in \mathbb{B}'$$

$$f^{-1}\left(\bigcup_{i=1}^{\infty} B_i'\right) = \bigcup_{i=1}^{\infty} f^{-1}(B_i') \qquad \text{for } B_i' \in \mathbb{B}'$$

$$f^{-1}(S') = S$$

From this it is obvious that Pf is a probability measure if P is, in fact, any mapping from $\mathbb{B}'$ to $\mathbb{B}$ which satisfies the above conditions (whether induced by a mapping from S to S' or not) will induce a mapping from probability measures on Ω to those on Ω'.

1.7 If $\Omega = \langle S, \mathbb{B} \rangle$ is a measurable space and, if $x \in S$, then P_x defined by

$$P_x(B) = \begin{cases} 1 \text{ if } x \in B \\ 0 \text{ if } x \notin B \end{cases}$$

for any $B \in \mathbb{B}$, is a probability measure on Ω, known as a "one-point" or "Dirac" measure.

1.8 Let $\Omega = \langle S, \mathbb{B} \rangle$ be a measurable space, P a probability measure on Ω, f a bounded measurable mapping from Ω to $\mathbb{R}$ = the space of real numbers with Borel sets as the measurable sets ("Bounded" means that for some positive real number M, $|f(x)| \leq M$ for all $x \in S$). Such an f is often called a bounded random variable. We now wish to define the P-expectation of f, also called the integral of f with respect to P, denoted by either

$$E_P(f), \int_\Omega f\, dP$$

or by $$\int_\Omega f(x)P(dx) \quad .$$

7

Expectation is a classical concept, defined only for "random variables" (real-valued)
We want to generalize it to the category of probabilistic mappings

This can be done by considering approximations to the integral based on doubly infinite increasing sequences

$X \xrightarrow{f} \mathbb{R}$

$1 \xrightarrow{P} \Omega \xrightarrow{\delta_f} \langle \mathbb{R}, \mathbb{B} \rangle$

Pick a seq. of real numbers

$$\cdots \le a_{-2} \le a_{-1} \le a_0 \le a_1 \le a_2 \le \cdots$$

$a_{-2}\ a_{-1}\ a_0\ a_1\ a_2\ \ldots$ The diagram

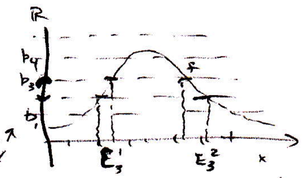
of real numbers. Given any such sequence a, define the upper approximation

Decomposing the range ..

$\int_\Omega f\, dP = \lim_{n\to\infty} \left\{ \sum_{i=1}^{n} a_i\, P(x, f^{-1}(a_i)) \right\}$

undefined!

$$\overline{J}\,(f, P, a) = \sum_{-\infty<n<\infty} \cancel{f}(a_n)\, Pf(a_{n-1}, a_n]$$



This is a bit awkward cuz $f(a_n)$ are not ordered ↗ or ↙

and the lower approximation

$$\underline{J}\,(f, P, a) = \sum_{-\infty<n<\infty} \cancel{f}(a_n) \circ Pf(a_n, a_{n+1}] \ .$$

define $E_3 = E_3^1 \cup E_3^2$

$\Rightarrow \sum b_n\, P(\underbrace{E_n}_{= f^{-1}[b_n, b_{n+1})})$

Here $Pf(a, b] = P\{x \mid a < f(x) \le b\}$ as defined in 1. The upper integral is defined by

$$\overline{I}\,(f, P) = \inf \overline{J}\,(f, P, a)$$

and the lower integral by

$$\underline{I}\,(f, P) = \sup \underline{J}\,(f, P, a)$$

where the infimum and supremum are taken over all doubly infinite increasing sequences a. If $\underline{I}\,(f, P) = \overline{I}\,(f, P)$, then the function f is said to be <u>integrable</u> with respect to P, and the integral is defined to be the common value

$$\underline{I}\,(f, P) = \int_\Omega f dP = \overline{I}\,(f, P)\ .$$

It can be shown that <u>every</u> bounded measurable function (on Ω) is integrable with respect to every probability measure (on Ω). For each individual P, there will ordinarily be many unbounded functions which are integrable with respect to P.

1.9 If S is a <u>countable</u> set, B the family of all subsets of S, f any bounded measurable function on $\Omega = \langle S, \mathbb{B} \rangle$, and P any probability measure on Ω, then

$$\int_\Omega f(x)P(dx) = \sum_{x \varepsilon S} f(x)p(x)$$

where $p(x) = P(\{x\})$ as defined in 1.5.

1.10 Let f, g be any two bounded measurable functions on a measurable space Ω, and let P be any probability measure on Ω. Then

$$\int_\Omega (f+g)\, dP = \int_\Omega f dP + \int_\Omega g dP \ .$$

If a is any real number, then

$$\int_\Omega af(x)\, P(dx) = a \int_\Omega f(x)\, P(dx) \ .$$

If f_n is any sequence of bounded measurable functions such that f_n is uniformly bounded ($|f_n(x)| \leq M$ for all x, n) and if $\lim_{n \to \infty} f_n(x) = f(x)$ for each $x \varepsilon S$,

then $$\lim_{n \to \infty} \int_\Omega f_n dP = \int_\Omega f dP.$$

1.11 If $0 \leq \theta \leq 1$ and if P_1, P_2 are any two probability measures on the measurable space Ω, then $P = \theta P_1 + (1-\theta)P_2$ is also a probability measure, and

$$\int_\Omega f dP = \theta \int_\Omega f dP_1 + (1-\theta) \int_\Omega f dP_2$$

for any bounded measurable function f on Ω.

I.2 Probabilistic Mappings

2.1 Let $\Omega = <S, \mathcal{B}>$ and $\Omega' = <S', \mathcal{B}'>$ be any measurable spaces. We say T is a probabilistic mapping from Ω to Ω' and write $\Omega \xrightarrow{T} \Omega'$ if and only if T assigns, to each point in Ω, a probability measure on Ω', and does so in a measurable way. More precisely, T is a function of two variables $x \varepsilon S$, $B' \varepsilon \mathcal{B}'$ having the properties

The first 3 conditions mean for each S $T(x,\cdot)$ is a measure on Ω'

(0) $0 \leq T(x, B') \leq 1$ for all $x \varepsilon S$, $B' \varepsilon \mathcal{B}'$

(1) $T(x, S') = 1$ for all $x \varepsilon S$

6

(2) $T(x, \bigcup_{i=1}^{\infty} B'_i) = \sum_{i=1}^{\infty} T(x, B'_i)$ for each $x \varepsilon S$ and for each disjoint sequence B'_i of measurable sets of Ω'.

$T(\cdot, B'): S \to [0,1]$

$T(\cdot, B'): (S, \mathbb{B}) \to ([0,1], \mathbb{B}_{\text{Borel}})$ (algebra)

(3) $\{x \mid T(x, B') \leq a\} \varepsilon \mathbb{B}$ for each $0 \leq a \leq 1$ and each $B' \varepsilon \mathbb{B}'$.

$T(\cdot, B')^{-1}([0,a]) \in \mathbb{B}$

$\therefore$ $T(\cdot, B')$ must be $(\mathbb{B}, \text{Borel})$-measurable

We will refer to $T(x, B')$ as the (conditional) T-probability of the event B' in Ω', given the elementary event x in Ω, or as the T-probability that x is mapped into B'. In case S' is countable and B' consists of all subsets of S', then a probabilistic mapping $\Omega \xrightarrow{T} \Omega'$ is entirely determined by a function t of two point variables $x \varepsilon S$, $x' \varepsilon S'$. (See 1.5)

i.e. $T(x, B') = T(x, \bigcup_{x'_j \in B'} \{x'_j\}) = \sum_{x'_j \in B'} T(x, \{x'_j\}) = \sum_{x'_j \in B'} t(x, x'_j)$ where $\Omega \times \Omega' \xrightarrow{t} \mathbb{R}$

2.2 Every measurable mapping f from Ω to Ω' (these being measurable spaces) may be regarded as a probabilistic mapping $\Omega \xrightarrow{T_f} \Omega'$ as follows:

$$T_f(x, B') = \begin{cases} 1 & f(x) \varepsilon B' \\ 0 & f(x) \not\varepsilon B' \end{cases} .$$

$1 \xrightarrow{x} X$ $\Downarrow$ $1 \xrightarrow{\delta_x} (X, \mathcal{B})$ where $\delta_x(*, B) = \begin{cases} 1 & \text{iff } x \in B \\ 0 & \end{cases}$

$X \xrightarrow{f} Y$ $\Downarrow$ $(X, \mathcal{B}) \xrightarrow{\delta_f} (Y, \mathcal{B}')$ where $\delta_f(x, B') = \begin{cases} 1 & \text{iff } f(x) \in B' \\ 0 & \text{otherwise} \end{cases}$

dirac (delta) measure

That is, T_f assigns to x the one-point measure (on Ω') which is concentrated at $f(x)$. Probabilistic mappings of this special sort we call deterministic.

2.3 Let $\Omega \xrightarrow{T} \Omega' \xrightarrow{U} \Omega''$ be probabilistic mappings. We define the composition $\Omega \xrightarrow{T \circ U} \Omega''$ to be the probabilistic mapping defined by

$$(T \circ U)(x, B'') = \int_{\Omega'} T(x, dx') \cdot U(x', B'') \text{, i.e., } \int_{\Omega'} \underbrace{U(\circ, B'')}_{\text{bounded measurable function on } \Omega'} d\underbrace{(T(x, \circ))}_{\text{a prob. measure on } \Omega'}$$

Note These are SAMPLING Distributions

That is, $(T \circ U)(x, B'')$ is the $T(x, \circ)$-expectation of $U(\circ, B'')$.

This is the correct law for composition of conditional probabilities in physical and other situations.

2.4 If Ω' is a countable space is 2.3, then $(T \circ U)(x, B'') = \sum_{x' \varepsilon S'} t(x, x') \cdot U(x', B'')$. If Ω'' is also countable, then

$$(T \circ U)(x, \{x''\}) = \sum_{x' \varepsilon S'} t(x, x') \cdot u(x', x'') .$$

$\Downarrow$

You can define $(t \circ u)(x, x'') = $ r.h.s.

$T_x(B') = T_x(\bigcup_{x' \in B'} \{x'\}) = \sum_{x' \in B'} T(x, \{x'\}) \stackrel{\text{def}}{=} \sum_{x' \in B'} t(x, x')$

7

2.5 If $\Omega \xrightarrow{f} \Omega' \xrightarrow{g} \Omega''$ are measurable mappings then

$$T_{f \circ g} = T_f \circ T_g$$

where $f \circ g$ is the usual composition of functions (thus the deterministic mappings constitute a subcategory (2.7) of the category of all probabilistic mappings. Note: This subcategory $\mathcal{P}_{det}$ has objects = measurable spaces, arrows = measurable functions. "$\underline{\text{Meas}}$"

2.6 A probabilistic mapping $1 \xrightarrow{P} \Omega$, where 1 is a one-point space, is just a probability measure on Ω. If $\Omega \xrightarrow{T} \Omega'$ is a probabilistic mapping, then $P \circ T$ is the induced distribution on Ω'. This is familiar in case Ω' is
This has to be the definition of induced distribution for a general probabilistic mapping T. When T is deterministic this reduces to the familiar def
Euclidean space and T a deterministic mapping (i.e., T is a "random variable"). Another special case is that where $\Omega = <S, \mathbb{B}>$, $\Omega' = <S, \mathbb{B}'>$, and $\mathbb{B}'$ is a sub-σ-algebra of B, while T is the "identity" mapping; then $P \circ T$ is the restriction of P from $\mathbb{B}$ to $\mathbb{B}'$!

$P \circ T_f(\star, B) = P(f^{-1}(B))$ i.e. $[P(\star, \cdot)](f^{-1}(B))$

2.7 If

$$\Omega \xrightarrow{T} \Omega' \xrightarrow{U} \Omega'' \xrightarrow{V} \Omega'''$$

then

$$T \circ (U \circ V) = (T \circ U) \circ V \, .$$

Also, if i_Ω denotes the probabilistic mapping defined by the (deterministic) identity map on Ω, then

$$i_\Omega \circ T = T = T \circ i_{\Omega'}$$

whenever $\Omega \xrightarrow{T} \Omega'$. Thus, the class $\mathcal{P}$ of all probabilistic mappings between measurable spaces, together with our notion of composition, is a category in the sense of Eilenberg-MacLane. Thus, the notions of functor, natural transformation, and adjoint functor have a well-defined meaning in connection with $\mathcal{P}$. The "objects" of $\mathcal{P}$ are arbitrary measurable spaces.

Exponentials of [0,1], i.e. $[0,1]^{\Omega}$ ... $a \in [0,1]$

2.8 Let, for each object Ω in $\mathcal{P}$, $\mathcal{D}(\Omega)$ = the set of all probability measures on Ω, equipped with the smallest σ-algebra such that for each ~~measurable~~

7

$\Omega = \langle S, \mathcal{B} \rangle$

$A \in \mathcal{B}$

~~$A \subseteq \Omega$~~, the evaluation $\mathcal{D}(\Omega) \longrightarrow [0,1]$ at A is measurable. Thus, $\mathcal{D}(\Omega)$ is also an object in $\mathcal{P}$. For any $\Omega \xrightarrow{T} \Omega'$ in $\mathcal{P}$, define the deterministic map $\mathcal{D}(\Omega) \xrightarrow{\mathcal{D}(T)} \mathcal{D}(\Omega')$ by

$$\mathcal{D}(T)(P)(A') = \int_{\Omega} P(d\omega)\, T(\omega, A)$$

for every $P \in \mathcal{D}(\Omega)$ and every measurable $A' \subseteq \Omega'$. Thus, $\mathcal{D}(T)(P) = P \circ T$ for $P \in \mathcal{D}(\Omega)$; i.e., viewed as a probabilistic mapping, $1 \xrightarrow{P} \Omega \xrightarrow{T} \Omega'$

$$\mathcal{D}(T)(P, \mathcal{A}') = \begin{cases} 1 & P \circ T \in \mathcal{A} \\ 0 & P \circ T \notin \mathcal{A} \end{cases}$$

for every element P of $\mathcal{D}(\Omega)$, and for every measurable set $\mathcal{A}'$ of probability measures on ~~Ω'~~ $\mathcal{D}(\Omega')$.

Define also the probabilistic mapping

$$\mathcal{D}(\Omega) \xrightarrow{\varphi_{\Omega}} \Omega$$

This is just the evaluation map. By construction it is measurable, i.e. condition (3) of 2.1 is satisfied

for each object Ω in $\mathcal{P}$ by the formula

$$\varphi_{\Omega}(P, A) = P(A) \quad \text{, i.e., } P(*, A) \qquad 1 \xrightarrow[P]{} \Omega$$

for each element P of $\mathcal{D}(\Omega)$ and each measurable $A \subseteq \Omega$. Then for any $\Omega \xrightarrow{T} \Omega'$ in $\mathcal{P}$, the diagram

$\mathcal{D}, id : \mathcal{P} \to \mathcal{P}$

φ is the N.T.

$\mathcal{D} \to id$

$$\begin{tikzcd} \mathcal{D}(\Omega) \arrow[r, "\varphi_{\Omega}"] \arrow[d, "\mathcal{D}(T)"'] & \Omega \arrow[d, "T"] \\ \mathcal{D}(\Omega') \arrow[r, "\varphi_{\Omega'}"'] & \Omega' \end{tikzcd}$$

in $\mathcal{P}$

is commutative, so that φ is a natural transformation of the functor $\mathcal{D}$ into the identity functor on $\mathcal{P}$.

2.9 Actually $\mathcal{D}$ is co-adjoint to the inclusion of the deterministic subcategory into $\mathcal{P}$; i.e., ~~each $\varphi_{\Omega'}$ is deterministic, and~~ if $\Omega \xrightarrow{T} \Omega'$ is any probabilistic mapping then there is a unique deterministic mapping f such that the

$i \dashv \mathcal{D}$ $\quad \mathcal{P}_{det} \underset{\mathcal{D}}{\overset{i}{\rightleftarrows}} \mathcal{P}$

The unit $\eta : Id_{\mathcal{P}_{det}} \to \mathcal{D} \circ i$

counit $\varepsilon : i \circ \mathcal{D} \to Id_{\mathcal{P}}$

For each $\Omega \in_{ob} \mathcal{P}$, $\varphi_{\Omega'}$ is a universal arrow from the functor i to Ω:

CARTESIAN Closed? No; but the map objects $[0,1]^{\Omega}$ exist

diagram

$$\begin{tikzcd} \mathcal{D}(\Omega') \\ \Omega \arrow[u, "\bar{T}"] \end{tikzcd} \quad \text{in } \mathcal{P}_{Def} \qquad \begin{tikzcd} i(\mathcal{D}(\Omega')) \arrow[r, "\varphi_{\Omega'}"] & \Omega' \\ i(\Omega) \arrow[u, "i(\bar{T})"] \arrow[ur, "T"'] \end{tikzcd} \quad \text{in } \mathcal{P}$$

$$\begin{tikzcd} \Omega \arrow[d, dashed, "f"'] \arrow[dr, "T"] & \\ \mathcal{D}(\Omega') \arrow[r, "\varphi_{\Omega'}"] & \Omega' \end{tikzcd}$$

$$\begin{tikzcd} \mathcal{D}(\Omega) \\ \Omega \arrow[u, "id"] \end{tikzcd} \qquad \begin{tikzcd} i(\mathcal{D}(\Omega)) \arrow[r, "\varphi_{\Omega}"] & \Omega \\ i(\Omega) \arrow[u, "i(id)"] \arrow[ur, "id"'] \end{tikzcd}$$

so $hom_{\mathcal{P}}(i(\Omega),\Omega') \cong hom_{\mathcal{P}_{Def}}(\Omega, \mathcal{D}(\Omega'))$

$\varphi_{\Omega'} \circ i(\bar{T}) \longleftarrow\!\mid \bar{T}$

is commutative. (In particular, there is a deterministic inclusion $\Omega \to \mathcal{D}(\Omega)$ and this is actually a retract with associated retraction $\varphi_{\Omega'}$.) It is expected that this adjointness observation will aid in the analysis of various methodological problems such as Bohm's questions about quantum mechanics.

I.3 Stochastic Processes and Decision Maps

3.1 A fairly general class of decision problems may be formulated as follows. There is a basic space Ω and a measurable partition Δ of Ω, elements of Δ being called "patterns" or "decisions". We denote the quotient (deterministic mapping) mapping $\Omega \to \Delta$ by f. (Actually, for the formulation of the problem we could allow f itself to be "fuzzy"; i.e., probabilistic.) There is also a space T of "observable states" and a probabilistic mapping $\Omega \xrightarrow{F} T$ expressing the conditional probability $F(\omega, A)$ that the observed state lies in any $A \subseteq T$, given that the basic state is $\omega \in \Omega$. The problem is then to find a "best" completion δ of the diagram

$$\begin{tikzcd} \Omega \arrow[r, "F"] \arrow[dr, "f"'] & T \arrow[d, dashed, "\delta"] \\ & \Delta \end{tikzcd}$$

One of the "virtues" of probability theory (and hence of the category $\mathcal{P}$) is that this general problem, when properly explicated, has a solution in many cases in which the corresponding deterministic problem does not; a basic reason for this is the possibility in $\mathcal{P}$ of forming convex combinations of maps, whereas there is no corresponding operation which produces deterministic maps. Of course, if there exists δ such that

If $\Omega \underset{T_2}{\overset{T_1}{\rightrightarrows}} \Omega'$ then $\Omega \xrightarrow{\theta T_1 + (1-\theta)T_2} \Omega'$

since $(\theta T_1 + (1-\theta)T_2)(x,\cdot) = \theta T_1(x,\cdot) + (1-\theta)T_2(x,\cdot) : \mathcal{B}' \to \mathbb{R}$

$F \circ \delta = f$, we would choose such δ as the solution to our problem; unfortunately, this is not possible for many F, f of interest. One popular scheme for making definite the criterion for choosing δ is to work with a given distribution $1 \xrightarrow{P} \Omega$ on Ω, and to choose δ so as to maximize the quantity

why should this even be an element of $\mathbb{B}'$?

$$\int_\Omega F\circ\delta(\cdot, [f(x)])\, d[P(x,\cdot)] \qquad \int_\Omega (F \circ \delta)(x, \{f(x)\})\, P(dx)$$

f is a partition so $\{f(x)\}$ probably means the equivalence class of $f(x)$, $[f(x)]$ which is taken to be an element of the algebra on Δ

which represents the average (with respect to P) of the probability of making the correct decision by first making the observation F and then following the decision rule δ. The probability measure P clearly expresses the relative importance attached to various basic states $x \in \Omega$ when evaluating the decision rule δ. In the absence of any such P, one could choose δ so as to maximize

$$\inf_{x \in \Omega} (F \circ \delta)(x, [f(x)]) .$$

The existence of solutions δ to these optimization problems can be established in very great generality by topological arguments.

3.2 We consider stochastic processes with discrete time. Let N be the category with countably many objects and no non-identity maps, and let $\mathcal{P}^N$ denote the category whose objects are sequences $\Omega_0, \Omega_1, \cdots$ of objects in $\mathcal{P}$. We define a functor

$$\mathcal{P}^N \xrightarrow{\Phi} \mathcal{P}^N$$

Products exist in $\mathcal{P}$ and hence $\mathcal{P}^N$

by

$$\Phi\{\Omega_n\} = \left\{\prod_{k<n} \Omega_k\right\}_n$$

for each sequence Ω of measurable spaces, where $\prod_{k<n} \Omega_k$ denotes the measurable space whose elements are all n-tuples $\langle x_0, \cdots, x_{n-1}\rangle$ with $x_i \in \Omega_i$, equipped with the smallest σ-algebra which makes each projection $\prod_{k<n} \Omega_k \to \Omega_j$ measurable. If Ω_n is thought of as the space of all possible

states of a system at time n, then $\Phi(\Omega)_n$ is the space of all possible histories of the system up to time n. We define a general temporally discrete stochastic process P in Ω to be any map

the stochastic process is the N.T. P

$$\Phi(\Omega) \xrightarrow{P} \Omega$$

here he is treating $\Omega \in_{ob} \mathcal{P}^N$

in $\mathcal{P}^N$. Given any two processes

SHOW THIS IS EQUIVALENT TO the CLASSICAL Def

$$\Phi(\Omega) \xrightarrow{P} \Omega, \quad \Phi(\Omega') \xrightarrow{P'} \Omega'$$

the general theory of categories indicates that a map $P \xrightarrow{f} P'$ of stochastic processes should be defined as a sequence

$$\Omega_n \xrightarrow{f_n} \Omega'_n$$

of maps in $\mathcal{P}$, such that for each time $n \in N$, the diagram

This is the unique arrow to the product space

He is claiming a (temporally discrete) stochastic process is a N.T. of the functors Φ, id:

$$\mathcal{P}^N \underset{id}{\overset{\Phi}{\rightrightarrows}} \mathcal{P}^N \quad ; \; P : \Phi \to id$$

$$\begin{array}{ccc} \Phi(\Omega)_n & \xrightarrow{\Phi(f)_n} & \Phi(\Omega')_n \\ {\scriptstyle P_n}\downarrow & & \downarrow{\scriptstyle P'_n} \\ \Omega_n & \xrightarrow[f_n]{} & \Omega'_n \end{array}$$

So given $\mathcal{F}, \mathcal{G} \in_{ob} \mathcal{P}^N$ and any arrow $\mathcal{F} \xrightarrow{f \in_a \mathcal{P}^N} \mathcal{G}$

$$\begin{array}{ccc} \Phi(\mathcal{F}) & \xrightarrow{P_{\mathcal{F}}} & \mathcal{F} \\ \downarrow{\scriptstyle \Phi(f)} & & \downarrow{\scriptstyle f} \\ \Phi(\mathcal{G}) & \xrightarrow[P_{\mathcal{G}}]{} & \mathcal{G} \end{array}$$

is commutative. Since there is also an obvious notion of composition for such maps, all stochastic processes and all maps of such determine a category

$$(\Phi, \mathcal{P}^N)$$

so that at component n

$$\begin{array}{ccc} \prod_{k \le n} \Omega_k & \xrightarrow{(P_{\mathcal{F}})_n} & \Omega_n \\ \downarrow & & \downarrow \\ \prod_{k \le n} \Omega_k & \xrightarrow[(P'_{\mathcal{G}})_n]{} & \Omega'_n \end{array}$$

which we call the category of temporally discrete stochastic processes. All the machinery developed in the general theory of categories, as well as that which can be developed for the particular category $\mathcal{P}$, can thus be applied to formulate, explicate, and solve many methodological problems within the category $(\Phi, \mathcal{P}^N)$.

3.3 If $\mathbb{N}$ denotes the additive monoid of non-negative integers, considered as a category with one object 0, then the functor category

12 $\mathcal{P}^{\mathbb{N}}$

7

is the <u>category of temporally discrete Markov processes.</u> Explicitly, an object in $\mathcal{P}^{\mathbb{N}}$ is just a measurable space Ω together with a probabilistic mapping $\Omega \xrightarrow{T} \Omega$, and maps f in $\mathcal{P}^{\mathbb{N}}$ satisfy a commutative diagram

Looks like the category ? $\mathcal{P}$

$$\begin{tikzcd} \Omega \arrow[r, "f"] \arrow[d, "T"'] & \Omega' \arrow[d, "T'"] \\ \Omega \arrow[r, "f"] & \Omega' \end{tikzcd}$$

If we are given a Markov process $<\Omega, T>$ together with an initial distribution $1 \xrightarrow{P_o} \Omega$, we can view our situation as a general stochastic process in which

1. $\Omega_n = \Omega$ for all $n \in \mathbb{N}$
2. $\Phi(\Omega)_o \to \Omega_o$ is just P_o
3. $\Phi(\Omega)_n \to \Omega_n$ is just the composition

$$\prod_{k<n} \Omega_k \to \Omega_{n-1} \xrightarrow{T} \Omega$$

where the first is the projection; i.e., the dependence on the past is really only on the preceding moment and, furthermore, the law of transition from one time to the next does not change with time.

If we denote by $(1, \mathcal{P}^{\mathbb{N}})$ the category of Markov processes augmented with initial distributions, then the foregoing discussion determines a functor

$$(1, \mathcal{P}^{\mathbb{N}}) \longrightarrow (\Phi, \mathcal{P}^{\mathbb{N}}).$$

This assertion carries the additional information that the various mappings match up properly, and also raises the question of whether the above functor has an adjoint (or co-adjoint). That is, is it possible to extend any process to a Markov process in a fashion which is universal with respect to maps to (or from) Markov processes??

From: wlawvere@buffalo.edu
Subject: Re: Probabilistic Relations
Date: 07/19/2011 08:23 AM
To: kirksturtz@universalmath.com

Dear Dr. Sturtz
For this and many other constructions,
for example an internal Hom, it seems
that one needs to consider the category of
all convex spaces and not just its full subcategory P.
That is, the whole Eilenberg-Moore category of
the commutative pr monad, not only the Kleisli
category of free algebras (="simplices" in this case).

In group theory one deals with actual groups not only
their presentations (= maps in the Kleisli category).

The most obvious property of this monad that most
do not have is that the free on 1 is 1, with the result
that the tensor product has projections ("marginals")
which however do not have the universal property of
the associated cartesian product.

The commutativity of the monad means that all sorts of
diagram categories arising in statistics can be enriched.

Thanks for your interest and I look forward to your
further comments.

Best wishes
Bill Lawvere

On Tue 07/19/11 9:39 AM , "kirksturtz@universalmath.com" kirksturtz@universalmath.com sent:
> Dear Prof. Lawvere, I have been trying to develop the concept of
> probabilistic relations using the Category of Probabilistic Mappings,
> P, via Rel(P). Such an approach requires P be regular - it is not.
> It only has weak equalizers; given a parallel pair f,g: X----> Y,
> the weak equalizer is the extreme set of the set of all probability
> measures P on X which satisfy f P = g P, along with the evaluation
> map. (Choquet Theory) In P arrows to 2 with the powerset
> algebra correspond to measurable functions, and seemingly the
> apparent alternative to the Rel(P) approach is to define a
> probabilistic relation as either a P map X x Y ----------> 2, which
> for finite spaces correspond directly to fuzzy relations, or as a P
> map X x Y -------> [0,1] . Defining composition is the challenge.
> I am familiar with the current literature - it falls short of
> capturing this critical concept. Any thoughts are greatly
> appreciated. Respectfully,
> Kirk Sturtz, Ph.D.
> Universal Mathematics
> Dayton, OH
> 937-610-8704
>