Stochastic category

Vladimir Voevodsky

Started February 3, 2004

1 Introduction

Let us look at the classical approach to model mathematically a deterministic process of some sort. One starts with a set (usually with an additional structure) whose points correspond to possible states of the system in question. A change in the state of the system is modeled as a map from this set to itself. A "process" is usually a family of such maps – one for each interval $[t_0, t_1]$ of the line representing time, which satisfy the obvious composition condition for intervals of the form $[t_0, t_1]$, $[t_1, t_2]$ and $[t_0, t_2]$. In particular any (deterministic) computer program which takes t_0, t_1, and the state of the system at time t_0 as an input and produces the state of the system at time t_1 as an output defines a "process" in the sense specified above.

If the program we use is not deterministic but uses a random number generator to compute new values of the variables from the old ones it does not define such a process.

Consider now the case when we have a process whose computer model is based on a randomized algorithm to produce the new values of the variables from the old ones. As an example we may look at a simple population dynamics model where the the state of the system is determined by the number of organisms currently alive, time is discrete and to produce the state at the next moment of time our algorithm uses a random number generator to determine whether a given organism survives (with probability p) or dies (with probability $1-p$).

The **stochastic category** described below allows one to repeat the same description in a randomized case simply by replacing the category of sets with the stochastic category.

Conventions . For a topological space X we will write simply X instead of the usual $(X, \mathcal{B})$ for the measure space with the underlying set X and the underlying σ-algebra the Borel σ-algebra on X. We will further consider sets as topological spaces with the discrete topology (all subsets are open). Combining these two conventions we will write X for the measure space with the underlying set X and the underlying σ-algebra of all subsets of X.

2 The value category

Value category V . Objects of $\boldsymbol{V}$ are (X, A) measurable spaces. **Meas**

Definition 2.1 *A value morphism* $f = f(x, U)$ *from* (X, A) *to* (Y, B) *is a function*

$$f(-,-) : X \times B \to [0, \infty]$$ transition kernel

such that for any $x \in X$ *the function*

$$U \mapsto f(x, U)$$

is a measure on (Y, B) *and for any* $U \in B$ *the function*

$$x \mapsto f(x, U)$$

is a measurable function on (X, A).

Let $f : (X, A) \to (Y, B)$, $g : (Y, B) \to (Z, C)$ be two value morphisms.
Composition of value morphisms:

$$\boxed{(x, W) \mapsto \int_Y g(y, W) f(x, dy)} \tag{1}$$

fg(x, dz) = ∫f(x,dy)g(y,dz)

For every (X, A) the value morphism Id: $x,U \to \delta_x(U)$

Lemma 2.3 *Let* f *be a value morphism* $(X, A) \to (Y, B)$ *and* $g : Y \to [0, \infty]$ *be a non-negative measurable function on* Y. *Then the function*

g(y,{*}) = g(y)

$$f^*(g) : x \mapsto \int_Y g df(x, -)$$ f o g

is a measurable function on (X, A).

Proof. Monotonic class thm.

X –> Y –> 1

Measure spaces, value morphisms and compositions (1) define a category. We denote this category by V and call the value category.

Lemma 2.4 *Let* μ *be a measure on* (X, A) *and* $f : (X, A) \to (Y, B)$ *a value morphism. Then the function* $f_*(\mu)$ *on* B *of the form*

$$U \mapsto \int f(x, U) d\mu$$ 1 –> X –> Y

is a measure on (Y, B).

μ(f{-1}[U]), f: deterministic

Lemma 2.5 *Let* $f : (X, A) \to (Y, B)$ *be a value morphism,* μ *a measure on* (X, A) *and* g *a measurable non-negative function on* (Y, B). *Then*

$$\int f^*(g) d\mu = \int g df_*(\mu)$$

$1 \xrightarrow{\mu} X \xrightarrow{f} Y \xrightarrow{g} 1$

Lemma 2.6 *The composition of value morphisms defined by (1) is associative.*

Example 2.7 For any (X, A) there is a unique morphism from $\emptyset$ to (X, A). Therefore $\emptyset$ is the initial object of the value category. Since there is a unique measure on $\emptyset$ there is also a unique morphism from any (X, A) to the empty set i.e. $\emptyset$ is also the final/terminal object. null object.

$1 = \{*\}$

Example 2.8 denote the object of the value category corresponding to the one element set by $\mathbf{1}$. A morphism from $\mathbf{1}$ to (X, A) is the same as a measure on (X, A). A morphism from (X, A) to $\mathbf{1}$ is a non-negative measurable function on (X, A).

$$Hom(\mathbf{1}, \mathbf{1}) = \mathbf{R}_{\geq 0} \cup \{\infty\} \qquad (2)$$

$\mu(*, \{*\}) = ?$
$\mu(*, \emptyset) = 0$

and for any (X, A) , $(\mu, f) \to \int f \mu$. $1 -> X -> 1$

$$Hom(\mathbf{1}, (X, A)) \times Hom((X, A), \mathbf{1}) \to Hom(\mathbf{1}, \mathbf{1})$$

Note the composition on (2) is of the form $(a, b) \mapsto ab$ where $0\infty = \infty 0 = 0$.

Example 2.9 Let $\mathbf{n}$ be the measure space with the underlying set $\{1, \ldots, n\}$ and the σ-algebra of all subsets. Then $Hom(\mathbf{n}, \mathbf{n})$ is the set of $n \times n$ matrices with entries from $[0, \infty]$. The composition is given by the product of matrices. $\mu(i,\{j\}) = M_{ij}$, M: $n \times n$

Lemma. gf(x,W) = f(x, g ^ {-1}[W]), where g: X— >Y measurable

Let $(X,A),(Y,B)$ be measurable spaces and let $f:X\to Y$ be a measurable map. Sending $x\in X$ to the measure $\delta_{f(x)}$ on Y defines a morphism from (X,A) to (X,B) in the value category.

δ: f — > δf(x, U) = δ{f(x)}(U)= δx(f{-1}[U]) = 1{f{-1}[U]}(x)

This construction defines a functor from the category of measurable spaces and measurable maps to the value category.

To distinguish morphisms in V which correspond to maps of measure spaces from the general morphisms we will call the former *deterministic* morphisms.

Example 2.10 Let $\mu:\mathbf{1}\to(X,A)$ be a measure on (X,A) and $f:(X,A)\to(Y,B)$ a measurable map considered as a morphism in the value category. Then $f\circ\mu=f_*(\mu)$ is the "direct image" of μ wrt f.

1— >X— >Y pushforward

p(x,-) ~ 0 if x ∉U
δx if x ∈U

Example 2.11 Let (X,A) be a measure set and (U,A_U) be a measurable subset of X considered with the induced σ-algebra. Then the embedding $(U,A_U)\to(X,A)$ can be split by a projection p where $p(x,-)$ is zero for $x\in X-U$ and is the measure concentrated in x for $x\in U$. Hence any measurable subset (including the empty one) of a measure space is canonically a retract of this space in V.

idU:
X— >U

id p
U — > X — > U

Lemma 2.14 *[11] Let G be a finite group of measurable automorphisms of a measure space (X,A). Then the measure space $(X/G,A^G)$ is the categorical quotient of (X,A) in V wrt the action of G.*

The functor from the category of measurable spaces to the value category does not reflect isomorphisms i.e. somme morphisms of measurable spaces may become isomorphisms when considered in the value category. Let (Y,B) be a measurable space and $f:X\to Y$ a be any surjection of sets. Let $f^{-1}(B)$ be the σ-algebra on X which consists of subsets of the form $f^{-1}(U)$ for $U\in B$. Then measures on $(X,f^{-1}(B))$ are in one-to-one correspondence with measures on (Y,B). In particular for each point $b\in Y$ we have a measure f_b on $(X,f^{-1}(B))$ corresponding to the delta measure δ_b on (Y,B). gives us a value morphism $(Y,B)\to(X,f^{-1}(B))$ and one verifies easily that it is inverse to the obvious morphism $(X,f^{-1}(B))\to(Y,B)$. Hence, from the point of view of the value category, the measurable spaces (Y,B) and $(X,f^{-1}(B))$ are indistinguishable. isomorphism

iso on Boolean algebra
f(b,f{-1}[U]): = fb(f{-1}[U])
= δb[U] = 1U(b), U: B

Why?

For measure spaces (X,A), (X',A') the measure space $(X\coprod X',A\coprod A')$ is easily seen to be both a product and a coproduct of (X,A) and (X',A') in V. V has both finite products and finite coproducts.

X ×{0} ∪ X'×{1}

For any two objects the set of morphisms between them is an abelian semi-group and moreover a "module" over $\mathbf{R}_+\cup\{\infty\}$. However (since we do not allow negative measures) morphisms can not be subtracted and therefore V is not an additive category. semi-module

V: semi-Ab-Category:
(f+g)(h+k) = fh+gh+fk+gk?

Let (X_α,A_α) be a family of measure spaces. Then there are two obvious ways to define a σ-algebra on $\coprod X_\alpha$. Let $A_\alpha^\cap$ denote the σ-algebra generated by elements of the form $U\subset X_\alpha\subset\coprod X_\alpha$ for all α and all U in A_α. Let $A_\alpha^\cup$ denote the set (σ-algebra) of subsets U in $\coprod X_\alpha$ such that for each α one has $U\cap X_\alpha\in A_\alpha$.

Lemma 2.12 *[prcopr] The measure space $(\coprod X_\alpha,A_\alpha^\cap)$ (resp. $(\coprod X_\alpha,A_\alpha^\cup)$) is the product (resp. the coproduct) of the family (X_α,A_α) in V.*

The families $A_\alpha^\cap$ and $A_\alpha^\cup$ coincide if our family is finite or countable but are different in general. In particular the countable products and coproducts in V coincide.

3 Bounded value category

A morphism $f : (X, A) \to (Y, B)$ is called bounded if the function

$$\boxed{\beta_f : x \mapsto f(x, Y)}$$

is a bounded function on X. Note that this condition means in particular that β_f takes only finite values i.e. that for any x the measure $f(x, -)$ on (Y, B) is finite. The composition of bounded morphisms is bounded and therefore measure spaces and bounded morphisms form a subcategory V_b in V called the bounded value category.

For (X, A), (Y, B) consider the measure space $(X \times Y, A \times B)$ where $A \times B$ is the σ-algebra generated by $U \times V$ with $U \in A$ and $V \in B$. If $f : (X, A) \to (Y, B)$ and $f' : (X', A') \to (Y', B')$ are (bounded) value morphisms define $f \times f'$ as the family which takes (x, x') to the product measure $f(x, -) \times f'(x', -)$ on $Y \times Y'$. Standard results about products of (finite) measures imply that $f \times f'$ is a morphism in the (bounded) value category. One can easily see that this construction defines a symmetric monoidal structure on V_b which we will denote by $\otimes$ instead of $\times$ to avoid confusion with the category product. The one element set is the unit of this monoidal structure which is why we denote it by $\mathbf{1}$.

Example 3.1 [**net1**] The standard example of trouble which one can get into if one tries to define the product of two measures one of which is not necessarily finite can be found in [**?**, p.78]. The source of the problem seems to lie in the fact that while all measures are continuous wrt countable filtered colimits (cf. [**?**, Lemma 1.10(a)]) only finite measures are continuous wrt countable filtered limits ([**?**, Lemma 1.10(b)]). Since limits are required to produce measurable subsets of the product of two measure spaces (e.g. the diagonal), a pair of measures on the factors can not be canonically extended to a measure on the product.

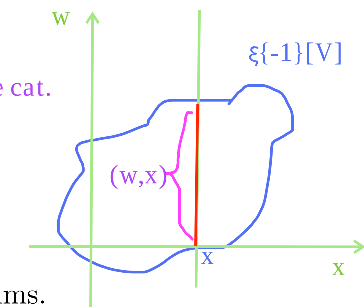
Remark 3.2 For each (X, A) the diagonal $(X, A) \to (X, A) \otimes (X, A)$ and the projection $(X, A) \to \mathbf{1}$ make (X, A) into a (commutative) comonoid in V_b.
Note. this structure is not canonical i.e. morphisms in V_b are not morphisms of comonoids.

Definition 3.3 *Let $f:(X,A)\to(Y,B)$ be a bounded value morphism. An* implementation *of f is $(\Omega,\mathcal{F},\mathbf{P},\xi)$ where $(\Omega,\mathcal{F})$ is a finite measure space with $\mathbf{P}:\mathbf{1}\to(\Omega,\mathcal{F})$ and $\xi:(\Omega,\mathcal{F})\times X\to(Y,B)$ is a deterministic morphism such that the diagram*

P×Id(U ×A) = P(U)1{x∈A}, for fixed x

$$\begin{array}{ccc} X & \xrightarrow{\mathbf{P}\otimes Id} & (\Omega,\mathcal{F})\otimes X \\ {\scriptstyle id}\downarrow & & \downarrow{\scriptstyle \xi} \\ (X,A) & \xrightarrow{f} & (Y,B) \end{array}$$

: Value cat.



commutes.

x, V — > P{ξx}V = = f(x, V) for x:X

ξx ~ P{ξx} = = fx

ξx(w) := ξ(w,x): Ω — > Y

ξ{-1}[V]x = ξx{-1}[V]

Remark 3.4 Explain relation to implementations of randomized algorithms.

A(x,V) = Y ×...× V ×...× Y: Cyl

x

Let X be a set and (Y,B) a measure space. Consider the set Y^X of all maps of sets from X to Y. For any V in B and any x in X let $A(x,V)$ be the set of all $g:X\to Y$ such that $g(x)\in V$. Let B^X be the σ-algebra on Y^X generated by the subsets $A(x,V)$. We will denote the measure space (Y^X,B^X) by $(Y,B)^X$.

Note it may be considered as the infinite product of as many copies of (Y,B).

{fx}, as a seq. of measures

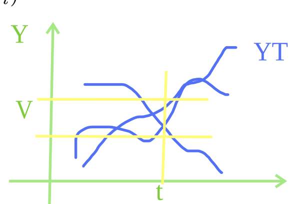
Lemma 3.5 (Kolmogorov) *Let $f:X\to(Y,B)$ be a bounded value morphism. Then there exists a unique measure μ_f on $(Y,B)^X$ such that for any finite set of pairwise distinct points $x_1,\ldots,x_n$ of X and any finite set $V_1,\ldots,V_n$ of elements of B one has*

$$\mu_f(\cap_{i=1}^n A_{(x_i,V_i)})=\prod_{i=1}^n f(x_i,V_i)$$

Y ×...V1...× V2 ×...Y...

value space of Y-valued process

Example 3.6 [**paths1**] Let $X=T$ be an interval of real line. Then Y^T is the space of paths in Y. An elementary measurable subset $A(t,V)$ in $(Y,B)^T$ is the subset of all paths γ such that $\gamma(t)\in V$. More generally $\cap_{i=1}^n A_{(t_i,V_i)}$ in Y^T is the subset of all paths which pass through V_i at time t_i. Lemma 3.5 asserts that any non-deterministic path $\phi:T\to(Y,B)$ defines a measure on $(Y,B)^T$ such that the "size" of $\cap_{i=1}^n A_{(t_i,V_i)}$ relative to this measure is the product of the probabilities (determined by ϕ) that t_i lands in V_i.



depict a stochastic process

Let $ev : (Y,B)^X \otimes X \to (Y,B)$ be the evaluation morphism $(g,x) \mapsto g(x)$. it is a measurable map. Consider μ_f as a morphism $\mathbf{1} \to (Y,B)^X$. Then the diagram

$$\begin{array}{ccc} X & \xrightarrow{\mu_f \otimes Id} & (Y,B)^X \otimes X \\ {\scriptstyle Id}\downarrow & & \downarrow{\scriptstyle ev} \\ X & \xrightarrow{f} & (Y,B) \end{array}$$

commutes and provides a canonical implementation of the morphism f. The obvious extension of this construction to bounded value morphisms $(X,A) \to (Y,B)$ implies the following result.

Lemma 3.7 *For any bounded value morphism $f : (X,A) \to (Y,B)$ the diagram*

$$\begin{array}{ccc} X & \xrightarrow{\mu_f \otimes Id} & \boxed{(Y,B)^X} \otimes X \\ {\scriptstyle Id}\downarrow & & \downarrow{\scriptstyle ev} \\ (X,A) & \xrightarrow{f} & (Y,B) \end{array}$$

= = P(A(x,V))

x, V — > P{evx}V = = f(x, V) for x:X

evx ~ P{evx} = = fx

evx(w): = w(x): Ω— > Y

where μ_f is the measure of Lemma 3.5, is an implementation of f.

Remark 3.8 Let $f_\alpha : (X_\alpha, A_\alpha) \to (Y,B)$ be a countable family of morphisms in V. Our definitions imply that $\coprod f_\alpha$ is a bounded morphism if and only if the functions β_{f_α} are uniformly bounded. This observation shows in particular that $(\coprod X_\alpha, \coprod A_\alpha)$ is not a coproduct of our family in V.

Similarly for $f_\alpha : (X,A) \to (Y_\alpha, A_\alpha)$, the family which sends x to the measure $\sum f_\alpha(x,-)$ is not a bounded morphism unless this measure is finite i.e. unless

$$\sum \beta_{f_\alpha} < \infty$$

everywhere on X, which shows that $(\coprod Y_\alpha, \coprod B_\alpha)$ is not a product of our family in V.

Note. (cf. 5.3 below) sending a family (X_α, A_α) to the coproduct space $(\coprod X_\alpha, \coprod A_\alpha)$ is not a functor from the category of families of objects in V to V. These properties make the bounded value category to be of limited use.

4 The stochastic category

A morphism $f:(X,A)\to(Y,B)$ is stochastic if for any x one has $f(x,Y)=1$ i.e. if the corresponding measures are probability measures. The subcategory is called the *stochastic category S.*

Remark 4.1 Stochastic morphisms from a measure space to itself are known in probability theory as stochastic /Markov kernels.

Note. for a non-empty (X,A) there are no stochastic morphisms from $(X,A)\to\emptyset$. Therefore, while $\emptyset$ is an initial object of the stochastic category it is not a finial object . On the other hand for any (X,A) there is exactly one stochastic morphism $(X,A)\to\mathbf{1}$. Therefore, $\mathbf{1}$ is the final object of the stochastic category (but not of the value category).

For (X,A) and (X',A') the coproduct $(X\coprod X', A\coprod A')$ is easily seen too be the coproduct of (X,A) and (X',A') in the stochastic category. However it is not the product of (X,A) and (X',A') in the stochastic category since the sum of two probability measures is not a probability measure.

For any measurable map of measure spaces $(X,A)\to(Y,B)$ the corresponding morphism in V is stochastic. Therefore the functor from measure spaces to the value category factors through the stochastic category.

Lemma 4.2 *[l3] Let (X_α, A_α) be a family of measure spaces. Then $(\coprod X_\alpha, A_\alpha^\cup)$ is a coproduct of this family in the stochastic category.*

In view of Lemma 4.2 we will write $\coprod(X_\alpha, A_\alpha)$ instead of $(\coprod X_\alpha, A_\alpha^\cup)$.

Note. the finite group quotients of Lemma 2.14 remain quotients in the stochastic category.

The tensor product of two stochastic morphisms is a stochastic morphism and therefore the symmetric monoidal structure defined above for the bounded value category gives a similar structure on S.

Example 4.3 [**markov2**]Let G be a set which is finite or countable. We consider G as a measure space with respect to the σ-algebra which contains all subsets of G. Then $Hom_{V_b}(G,G)$ is the set of matrices $(p_{ij})_{i,j\in G}$ such that $p_{ij}\geq 0$, for any i the sum $p_i=\sum_j p_{ij}$ is finite and the set of numbers p_i is bounded. The set $Hom_S(G,G)$ is the set of stochastic matrices with rows and columns numbered by elements of G. The composition of morphisms corresponds in this description to multiplication of matrices. If P is an element of this set and $f: G\to\mathbf{1}$ a morphism in V (corresponding to a rv by 2.8) then the sequence of rvs $f_n=f\circ P^n$ is the **Markov chain** generated by the stochastic matrix P.

For any (X,A) let

$$tr_n := \sum_{i=1}^{n} pr_i : (X,A)^{\otimes n}\to (X,A) \tag{3}$$

be the morphism which sends a point $(x_1,\ldots,x_n)$ to the measure $\sum_{i=0}^{n}\delta_{x_i}$. For $n=0$ we take tr_0 to be the zero morphism. The following lemma gives an important property of stochastic morphisms.

Lemma 4.4 *For any stochastic morphism $f : (X, A) \to (Y, B)$ and any $n \geq 0$ the diagram*

$$\begin{array}{ccc} (X,A)^{\otimes n} & \xrightarrow{f^{\otimes n}} & (Y,B)^{\otimes n} \\ {\scriptstyle tr_n}\downarrow & & \downarrow{\scriptstyle tr_n} \\ (X,A) & \xrightarrow{f} & (Y,B) \end{array}$$

commutes.

Proof: In view of the definition of tr_n it is sufficient to verify that $pr_i \circ f^{\otimes n} = f \circ pr_i$ for all i. More generally it is sufficient to see that for a morphism $f : X \to Y$ and a stochastic morphism $f' : X' \to Y'$ one has $pr_Y \circ (f \otimes f') = f \circ pr_X$ i.e. that the square

$$\begin{array}{ccc} X \otimes X' & \xrightarrow{f \otimes f'} & Y \otimes Y' \\ {\scriptstyle pr_X}\downarrow & & \downarrow{\scriptstyle pr_Y} \\ X & \xrightarrow{f} & Y \end{array}$$

commutes. Let e be the canonical stochastic morphism from an object to the point. We have

$$pr_Y \circ (f \otimes f') = (Id_Y \otimes e) \circ (f \otimes f') = f \otimes (e \circ f') = f \otimes e = f \circ pr_X$$

where the third equality holds since $e \circ f' = e$ eactly means that f' is stochastic.

5 Branching morphisms and branching category

For a measure space (X, A) let $S^n(X, A) = (X, A)^n/\Sigma_n$ be the n-th symmetric power of (X, A). For $n = 0$ we set $S^0(X, A) := \mathbf{1}$ for all (X, A) including the empty set. set

$$S^\bullet(X, A) = \coprod_{n \geq 0} S^n(X, A)$$

Example 5.1 We have:

$$S^\bullet(\emptyset) = \mathbf{1} \quad S^\bullet(\mathbf{1}) = \mathbf{N}$$

Lemma 2.14 shows that for each n, $S^n(-)$ is a functor from the bounded value category to itself. Since $S^\bullet(X, A)$ is the coproduct of $S^n(X, A)$ in V we conclude that $S^\bullet(-)$ is a functor from the bounded value category to the value category. Finally, since coproduct of stochastic morphisms is stochastic we conclude that both the individual symmetric powers $S^n(X, A)$ and the total symmetric power $S^\bullet(X, A)$ are functors from the stochastic category to itself.

Remark 5.2 For a sufficiently nice space (X, A) the space $S^\bullet(X, A)$ is isomorphic to the space of integer-valued measures $M((X, A), \mathbf{Z}_+)$ on (X, A). This interpretation of the total symmetric power appears in some probabilistic texts on branching processes (e.g. [**?**]). The theory of measure valued branching processes studies the analogs of branching processes with the integer-valued measures replaced by more general measures.

Remark 5.3 One can easily see that the total symmetric power $S^\bullet$ is not a functor from V_b to V_b. Indeed consider a morphism $a : \mathbf{1} \to \mathbf{1}$ where $a > 1$ (see (2)). Then $S^n(a) = a^n$ and $S_\bullet(a)$ is not bounded since the volumes of corresponding measures on $\mathbf{N}$ are $a, a^2, \ldots$ which is an unbounded function on $\mathbf{N}$.

Definition 5.4 *A branching morphism ϕ from (X, A) to (Y, B) is a morphism in S of the form $(X, A) \to S^\bullet(Y, B)$.*

The functor $S^\bullet(-)$ is an extension to S of a functor with the same notation and meaning on the category of measure spaces and measurable maps to itself. In particular the obvious monad structure

$$S^\bullet \circ S^\bullet \to S^\bullet$$

$$Id \to S^\bullet$$

of the total symmetric power functor on sets defines a monad structure on $S^\bullet$ on S. We define the *branching category* B as the category of free algebras over $S^\bullet$. The objects of B are again measure spaces (X, A) and morphisms from (X, A) to (Y, B) are the branching morphisms of Definition 5.4.

Remark 5.5 In view of Lemma 4.2 algebras over $S^\bullet$ are exactly commutative monoids in S wrt $\otimes$.

We will write $\phi : [X, A] \to [Y, B]$ for branching morphisms to distinguish them from morphisms in V and S. Let us describe the composition of branching morphisms more explicitly. Observe first that there is a measurable map of measure spaces

$$m : S^\bullet(Y, B) \times S^\bullet(Y, B) \to S^\bullet(Y, B)$$

which makes $S^\bullet(Y, B)$ into a commutative monoid. In view of Lemma 2.14 and the definition of the symmetric product it shows that any morphism ϕ from (X, A) to $S^\bullet(Y, B)$ in V_b defines a family of morphisms of the form

$$\phi_n : S^n(X, A) \to S^\bullet(Y, B)$$

(where we set ϕ_0 to be identically 1). If the original morphism is stochastic so are the morphisms ϕ_n and therefore by Lemma 4.2 they define a morphism

$$\phi_* = \coprod \phi_n : S^\bullet(X, A) \to S^\bullet(Y, B)$$

We can now define the composition of two branching morphisms by the rule:

$$\psi \circ_B \phi := \psi \circ \phi_*$$

Forgetting the $S^\bullet$ algebra structure defines a functor

$$F : B \to S$$

which takes (X, A) to $S^\bullet(X, A)$ and ϕ to the morphism ϕ_* defines above.

Example 5.6 *[***ex8***]*Consider morphisms in the branching category of the form $\phi : [\mathbf{1}] \to [\mathbf{1}]$. Since $S^\bullet(\mathbf{1}) = \mathbf{N}$ we may identify this set with the set of probability measures on $\mathbf{N}$. For any ϕ let $p_\phi = \sum p_i t^i$ be the generating function of this measure. This construction identifies $Hom_B([\mathbf{1}], [\mathbf{1}])$ with formal power series $\sum p_i t^i$ satisfying $p_i \geq 0$ and $\sum p_i = 1$. If ϕ, ψ two endomorphisms of $[\mathbf{1}]$ in B then one has

$$p_{\phi \circ \psi} = p_\psi(p_\phi(t)) \tag{4}$$

i.e. in this description the composition of morphisms corresponds to the composition of power series in the reverse order.

Example 5.7 *[***ex10***]*The previous example has an immediate generalization to branching morphisms of the form $\phi : [\mathbf{n}] \to [\mathbf{n}]$ where $\mathbf{n} := \coprod_{i=1}^n \mathbf{1}$ is the set of n elements considered as a measure space with respect to the maximal σ-algebra. Such morphism is a collection of n probability measures on $\mathbf{N}^n$. If we describe these measures through their generating functions we may identify $Hom_B([\mathbf{n}], [\mathbf{n}])$ with the set of n-tuples $(f_1, \ldots, f_n)$ where each f_i is a formal power series in n-variables with non-negative coefficients satisfying the condition $f_i(1, \ldots, 1) = 1$. The composition of morphisms corresponds to the substitution composition for such n-tuples.

The morphism (3) is clearly invariant under the action of the symmetric group and by Lemma 4.2 it defines a bounded value morphism

$$tr_n : S^n(X, A) \to (X, A)$$

which sends the point $x_1, \ldots, x_n$ to the sum of δ-measures $\delta_{x_1} + \ldots + \delta_{x_n}$ (for $n = 0$ our morphism is 0) and which we continue to denote by tr_n. The coproduct of tr_n's is a morphism $tr_* : S^\bullet(X, A) \to (X, A)$ in V. For a stochastic morphism $(X, A) \to S^\bullet(Y, B)$ (i.e. for a branching morphism $\phi : [X, A] \to [Y, B]$) define a value morphism

$$tr(\phi) : (X, A) \to (Y, B)$$

as the composition $tr_* \circ \phi$.

Proposition 5.8 *For any ϕ as above the diagram*

$$\begin{array}{ccc} S^\bullet(X, A) & \xrightarrow{\phi_*} & S^\bullet(Y, B) \\ {\scriptstyle tr_*}\downarrow & & {\scriptstyle tr_*}\downarrow \\ (X, A) & \xrightarrow{tr(\phi)} & (Y, B) \end{array}$$

commutes.

Proof: By definition of ϕ_* it is sufficient to verify that for any n the diagram

$$\begin{array}{ccccc} (X,A)^{\otimes n} & \xrightarrow{\phi^{\otimes n}} & S^\bullet(Y,B)^{\otimes n} & \xrightarrow{m} & S^\bullet(Y,B) \\ {\scriptstyle tr_n}\downarrow & & {\scriptstyle tr_n}\downarrow & & \downarrow{\scriptstyle tr_*} \\ (X,A) & \xrightarrow{\phi} & S^\bullet(Y,B) & \xrightarrow{tr_*} & (Y,B) \end{array}$$

commutes. The rhs square consists of morphisms which take a point to the sum of finitely many points and it is easy to verify its commutativity explicitly. The lhs square commutes by Lemma 4.4.

Corollary 5.9 *For a pair of branching morphisms $\phi : [X,A] \to [Y,B]$, $\psi : [Y,B] \to [Z,C]$ one has*

$$tr(\psi \circ \phi) = tr(\psi) \circ tr(\phi)$$

Example 5.10 *[***ex11***]*Consider a branching morphism $\phi : [\mathbf{1}] \to [\mathbf{1}]$ which we describe through the corresponding probability generating function $p_\phi = \sum p_i t^i$ as in Example 5.6. Then $tr(\phi)$ is a morphism $\mathbf{1} \to \mathbf{1}$ i.e. a non-negative number. One can easily see that

$$tr(\phi) = \sum i p_i = p'_\phi(1)$$

where p'_ϕ is the formal derivative of p_ϕ with respect to t. In other words, $tr(\phi)$ is in this case the expectation value of ϕ. For two morphisms ϕ, ψ of this form Corollary 5.9 asserts that

$$tr(\psi \circ \phi) = tr(\psi) tr(\phi).$$

In view of (4) this follows from the equation

$$(p_\phi \circ p_\psi)'(1) = p'_\psi(1) p'_\phi(p_\psi(1)) = p'_\psi(1) p'_\phi(1)$$

where the last equation holds since the $p_\psi(1) = 1$ because ψ is a stochastic morphism.

Example 5.11 *[***ex12***]*Consider now branching morphisms $[\mathbf{n}] \to [\mathbf{n}]$ as in Example 5.7. For a morphism ϕ of this form $tr(\phi)$ is a morphism $\mathbf{n} \to \mathbf{n}$ i.e. an $n \times n$-matrix (a_{ij}) with entries from $[0, \infty]$. If we represent ϕ a sequence of power series $(f_1, \dots, f_n)$ in variables $t_1, \dots, t_n$ then one gets

$$a_{ij} = \frac{\partial f_i}{\partial t_j}(1)$$

If $\psi = (g_1, \dots, g_n)$ is another such morphism then the statement of Corollary 5.9 is again equivalent to the formula for the differential of a composition combined with the fact that $g_i(1) = 1$ since ψ is stochastic.

6 Application A. The history of a non-aging haploid population.

The problem considered in this section is the most simple non-trivial example of the general reconstruction problem which I know of. The model we deal with here is very unrealistic and probably does not approximate any real biological phenomenon whatsoever. Much more realistic models will be considered in the following sections.

Here is the setup. We are given a times interval $[t_{min}, t_{max}]$, a constant $b \geq 0$ called the birth (division) rate and a function $d(t) \geq 0$ on $[t_{min}, t_{max}]$ called the death rate. At time t_{min} we have a population which consists of N identical creatures. The population starts to live according to the following rules:

1. all living creatures obey the same rules and their lives are independent from each other
2. if x is a creature alive at t then during the time interval $[t, t + \Delta t]$ one of the following will occur:
   (a) it will turn into 0 creatures (die) with probability $d(t)\Delta t + o(\Delta t)$
   (b) it will continue being 1 creature with probability $1 - b\Delta t - d(t)\Delta t - o(\Delta t)$
   (c) it will turn into 2 creatures (divide) with probability $b\Delta t + o(\Delta t)$
   (d) it will turn into n creatures where $n > 2$ with probability $o(\Delta t)$

An independent observer keeps a list of all creatures ever alive, of their birth and death times (when a creature divides it is considered to die with two new ones being born) and with the information on who is the descendant of whom.

We stop the process at t_{max}. At this time the observer provides us with the list of all creatures currently alive together with the following additional information. For every pair of creatures x, y we are given a number $d(x, y)$ which equals to the death time of their most recent common ancestor if such a common ancestor existed and ∞ otherwise.

Assume that we know the birth rate b and t_{min}, t_{max}. Our task is to do our best to reconstruct the history of the population i.e. the number of creatures $\alpha(t)$ alive at time t for all t in $[t_{min}, t_{max}]$. We would also like to be able to check our observer for mistakes and cheating.

Let us first describe the mathematical model corresponding to the algorithmic description given above. The state of our system at a given time t is completely determined by the number of creatures alive at this time. Therefore, the set of states of our system is $\mathbf{N}$. The evolution of the system from time t to time t' is given by a morphism $\phi(t, t') : \mathbf{N} \to \mathbf{N}$ in the stochastic category such that $n \mapsto \mu_n$ where $\mu_n(m)$ is the probability that we get m creatures at time t' assuming that we had n creatures at time t. Since the fates of creatures are independent from each other this is a branching morphism and it is determined by its value on 1.

For each $t \in [t_{min}, t_{max}]$ let $A(t)$ be the set of creatures alive at time t and let $\tilde{A}(t)$ be the subset in $A(t)$ which consists of those creatures who have at least one descendant alive at t_{max}. Let further $\alpha(t)$ and $\tilde{\alpha}(t)$ be the number of elements in $A(t)$ and $\tilde{A}(t)$ respectively.

Note that we know $A = A(t_{max})$ and that $\tilde{A}(t_{max}) = A$. For every $t \in [t_{min}, t_{max}]$ we have an equivalence relation $\cong_t$ on A where $x \cong_t y$ if $d(x, y) \leq t$. One verifies easily that

$$\tilde{A}(t) = A/\cong_t .$$

and since we know $d(-,-)$ we conclude that we know $\tilde{A}(t)$ for all t. Consider $\tilde{A}(t_{min}) = A/\cong_{t_{min}}$. Let $\tilde{N} = \tilde{\alpha}(t_{min})$ be the number of elements in this set. This number is a major parameter of our data. We first consider the case when $\tilde{N}$ is "sufficiently large". It is clear that at the opposite extreme when $\tilde{N} = 1$ we can say little about the history. In general it is natural to expect that a reliable inference about the history from the data is possible only for periods of time t such that $\tilde{\alpha}(t_{max} - t)$ is large.

The equivalence relation $\cong_{t_{min}}$ splits the set A into $\tilde{N}$ equivalence classes. Let us number those classes and denote them by A_i, $i = 1, \ldots, \tilde{N}$. Considering the number of elements