



DEPARTMENT OF MATHEMATICAL
SCIENCES

TECHNICAL REPORT SERIES

Phenotype Spaces

by

Frédéric Mynard
Department of Mathematical Sciences
Georgia Southern University, Statesboro, GA 30460-8093

Gavin J. Seal
School of Computer Science
McGill University
Montreal, QC, Canada H3A 247

Number 2007-009
Submitted: June 25, 2007
© 2007

PHENOTYPE SPACES

FRÉDÉRIC MYNARD AND GAVIN J. SEAL

1. INTRODUCTION

In [12], Fontana, Stadler, Stadler, and Wagner present different aspects of a theory of evolutionary change, coined here as the GP-map model. In this context, RNA folding into secondary structures is the simplified model for evolutionary change the authors focus on. Indeed, despite its limitations, it is the only experimentally tractable example. The authors argue that “continuous” changes in the RNA sequences can lead to apparent “jumps” at the phenotypic level, so that evolution is not solely influenced by natural selection and genetic drift, but is also directed by internal dynamics. In order to analyze the “jumps” in the shapes of the RNA sequences, the authors equip the genotype set with a pretopology (via a not necessarily idempotent closure operator), which is then used to structure the phenotype set, and leads the authors to describe the apparent jumps as an essentially continuous process, interrupted at intervals by a *discontinuous* event.

The intent of these notes is to briefly recall the mathematical aspects of the GP-map model, as well as to discuss a number of concepts pertaining to the pretopological structure that Fontana, Stadler, Stadler, and Wagner endow both the genotype and phenotype spaces with. We begin by taking advantage of the theory of random walks on weighted digraphs to investigate the dynamics of the GP-map model from a more classical point of view. As suggested in [12], the quantitative information that this analysis exploits can be better represented by replacing the “topological” approach by a “metric” one (see 4.2 below). In the course of our analysis, we clarify the “topological” interpretation proposed in *op.cit.*, and put forth a number of its weaknesses, among which the rather non-topological nature of the model, and the difficulty to analyze the system’s evolution in terms of a satisfying notion of continuity.

Standard references for the mathematical themes we tackle can be found in most introductory probability books for random walks and Markov processes (for example [5]), and [3] for topological-like structures (filters, neighborhoods, pretopologies, etc.) used herein.

1.1. The GP-map model. Let us briefly recall the essentials of the *genotype-phenotype-map model* (or *GP-map model*) that we will use in the following (for more details, we refer to [12]). In the example of RNA folding, the GP-map model is given by a map $f : G \rightarrow P$, where the “genotype set” G is a set whose elements are RNA sequences of fixed length, referred to hereafter as *sequences*, and the “phenotype set” P is a set whose elements are the minimum free energy shapes, referred to simply as *shapes*, or *phenotypes*; the *GP-map* f then simply represents the “folding” process of a sequence into its shape. Each sequence of the genotype

Date: July 22, 2007.

set is a word of length l on the alphabet $\{a, c, g, u\}$ ⁽¹⁾, which can “mutate” into any of its $3l$ *one-mutants*, i.e., into any of the sequences that differ in exactly one coordinate from it. The genotype set becomes the set of vertices of a Hamming graph (in which two sequences are adjacent iff they differ in one coordinate), with a loop added at each vertex, so that a sequence is allowed to avoid mutation. Each edge is endowed with a probability $\frac{1}{3l+1}$, and a random walk on this graph therefore represents a series of mutations occurring—or not—at regular time intervals. For simplicity, we denote this weighted graph by G again. The phenotype set is then the vertex set of the quotient graph P induced by f . To describe this graph, let us denote the probability that a sequence $x \in G$ mutates into a sequence $y \in G$ by $\text{Prb}(y \curvearrowright x)$, so one obviously has

$$\text{Prb}(y \curvearrowright x) = \begin{cases} \frac{1}{3l+1} & \text{if } d_G(x, y) \leq 1 \\ 0 & \text{else,} \end{cases}$$

where $d_G : G \times G \rightarrow \mathbb{N}$ denotes the Hamming distance in G between two sequences (so that $d_G(x, y) \leq 1$ iff $x = y$ or y is a one-mutant of x). Therefore, for disjoint subsets $A, B \subseteq G$, the probability $\text{Prb}(B \curvearrowright A)$ that any element in B is the result of a one-mutation of an element in A is given by

$$(*) \quad \text{Prb}(B \curvearrowright A) = \frac{\sum_{y \in B} \sum_{x \in A} \text{Prb}(y \curvearrowright x)}{\sum_{y \in G} \sum_{x \in A} \text{Prb}(y \curvearrowright x)} = \frac{\text{Edg}(B, A)}{(3l+1)|A|},$$

where

$$\text{Edg}(B, A) = |\{(y, x) \mid y \in B, x \in A, d_G(y, x) = 1\}|.$$

Therefore, the probability $\text{Prb}(\beta \curvearrowright \alpha)$ that a mutation in G will lead to a change from shape α to a shape β in P is given by

$$(\dagger) \quad \text{Prb}(\beta \curvearrowright \alpha) = \frac{\text{Edg}(f^{-1}(\beta), f^{-1}(\alpha))}{(3l+1)|f^{-1}(\alpha)|}.$$

This probability is essentially the *occurrence frequency* described in [4]. Note however that the *frequency ratio* $A(\beta \curvearrowright \alpha)$ defined in [12] is not directly related to $\text{Prb}(\beta \curvearrowright \alpha)$, since it only takes into account the behavior of mutations near the border of a fibre $f^{-1}(\alpha)$.

The phenotype set P is endowed with a graph structure by associating each pair of shapes (α, β) with the probability $\text{Prb}(\beta \curvearrowright \alpha)$ (this is the quotient graph of the graph on G via $f : G \rightarrow P$), and a random walk on P therefore represents the change in shape induced by the mutations of the RNA sequences in G . This model puts forth the idea that while mutations can occur regularly in a sequence x of RNA, the shape that x takes on can remain the same while the mutating sequence moves in the same fibre $f^{-1}(\alpha)$, and then “suddenly” jumps into another shape when x changes fibre. The term *neutral drift* is used to describe a random walk in G that remains in a constant fibre.

2. DYNAMICAL ASPECTS AND MARKOV CHAINS

As is the case for many other evolutionary models (see [13]), directed graphs (or digraphs) provide a convenient conceptual framework in which to study the evolution of the phenotype space equipped with the dynamics described by the

¹ a stands for adenine, c for cytosine, g for guanine and u for uracil. Recall that a can pair with u only, and that g can pair with c or with u .

GP-map model. As suggested by the GP-map model, an evolutionary trajectory can be represented by a random walk on the graph whose vertices are the elements of the phenotype space, and whose directed edges carry a probability corresponding to the accessibility ratio:

$$a_{\alpha\beta} := \text{Prb}(\beta \curvearrowright \alpha) ,$$

for all phenotype shapes α and β . These probabilities provide the entries of the transition matrix of a Markov process with state space P :

$$M = (a_{\alpha\beta})_{\alpha, \beta \in P} .$$

The information pertaining to the probabilistic evolution of the phenotype space is then contained in the successive powers M, M^2, M^3, \dots of the transition matrix. In the context of the GP-map model, this matrix is strongly constrained: it is not symmetric, a fact that reflects the asymmetry in the direction of mutations, which in turn provides a preferred direction for evolution (see [12]); it also has strictly positive diagonal elements, a fact that allows the drift along the neutral network.

Let us analyze this matrix further. Although M is not symmetric, if an entry $a_{\beta\alpha}$ is distinct from 0, then so is $a_{\alpha\beta}$: indeed, according to the GP-map model, the mutations in the genotype space are following a random walk on an undirected graph whose edges all carry the same probability; thus, the mutation in the genotype space that accesses β from α is reversible, and the corresponding probability $a_{\beta\alpha}$ nonzero. By reducing if necessary the matrix into irreducible components, one can also suppose that M is irreducible, so that any phenotype in the corresponding state space P is accessible from any other. Better yet, since M is finite, and the diagonal entries are all distinct from 0, the matrix is regular: given a state α , there is a random walk to any state β and that subsequently stays there indefinitely (even though the probability of this event is small, it is non-zero), so by finiteness of P , one can take the maximum of the lengths of these walks to conclude that the row vectors $(a_{\alpha\beta}^{(n)})_{\beta \in P}$ of M^n all have strictly positive components for large enough n .

One can therefore conclude that the powers of a regular transition matrix converge to a limit matrix L (see [5]):

$$\lim_{n \rightarrow \infty} M^n = L ,$$

with constant columns. The values in each column then determine the fitness landscape of the phenotype space, and pinpoint which phenotypes will be more likely to prevail. The results described in [4] seem to indicate that a row vector of L will reveal a small number of “almost absorbing” states that occur among a wide majority of transient states (the column indices whose constant value in L will be close to 0).

3. THE FONTANA-STADLER-STADLER-WAGNER APPROACH

3.1. Pretopologies. In [12], Fontana, Stadler, Stadler, and Wagner endow a set of phenotypes P with a pretopology defined in terms of the GP-map $f : G \rightarrow P$. Recall that a *pretopology* ξ on a set X assigns to each point x of X a filter $\mathcal{N}_\xi(x)$ such that $x \in \bigcap \mathcal{N}(x)$, called the neighborhood filter of x . Unlike in the topological case, pretopological neighborhood filters do not need to have a filter-base composed of open sets (that is, sets which are neighborhoods of all of their points). The pretopology can alternatively be described in terms of a closure operator $\text{adh}_\xi : 2^X \rightarrow 2^X$ which is

- (1) grounded: $\text{adh}_\xi \emptyset = \emptyset$
- (2) expansive: $A \subseteq \text{adh}_\xi A$
- (3) isotone : $A \subseteq B \implies \text{adh}_\xi A \subseteq \text{adh}_\xi B$
- (4) additive: $\text{adh}_\xi(A \cup B) = \text{adh}_\xi A \cup \text{adh}_\xi B$,

but may fail to be idempotent ($\text{adh}_\xi \text{adh}_\xi A = \text{adh}_\xi A$). Note that

$$x \in \text{adh}_\xi A \iff \forall N \in \mathcal{N}(x) (A \cap N \neq \emptyset) .$$

The purpose of considering pretopologies on spaces of phenotypes is to introduce a notion of continuity for evolutionary trajectories. A map $h : (X, \xi) \rightarrow (Y, \tau)$ is *continuous* if

$$h(\text{adh}_\xi A) \subseteq \text{adh}_\tau h(A)$$

for each $A \subseteq X$, or equivalently, if

$$h(\mathcal{N}_\xi(x)) \supseteq \mathcal{N}_\tau(h(x))$$

for all $x \in X$. Recall that if ξ and τ are two pretopologies on the same set X , then ξ is *finer* than τ (or τ is *coarser* than ξ), in symbols $\xi \geq \tau$ if $\text{id}_X : (X, \xi) \rightarrow (X, \tau)$ is continuous, that is, if $\mathcal{N}_\xi(x) \supseteq \mathcal{N}_\tau(x)$ for every $x \in X$.

3.2. Finite pretopologies and the GP-map model. In the context of the GP-map model, one defines a pretopology ξ on G via

$$\text{adh}_\xi A := \{x \in G \mid \exists y \in A (d_G(x, y) \leq 1)\} ,$$

for all $A \subseteq G$, where $d_G : G \times G \rightarrow \mathbb{N}$ denotes the Hamming distance in G , as in 1.1. Since only finite sets are considered, every neighborhood filter $\mathcal{N}_\xi(x)$ is principal, that is, has a smallest element $N_\xi(x)$, and we have

$$\text{adh}_\xi A = \bigcup_{x \in A} N_\xi(x) = N_\xi(A) .$$

The above formula is true for the structure induced by the Hamming graph, because it is symmetric. It is not true in general for finite pretopological spaces. Note that a finite pretopological space (X, ξ) defines a directed graph whose set of vertices is X , and in which an edge from x to y exists iff $y \in N_\xi(x)$. Conversely, a directed graph with a loop at each vertex defines a pretopological space ξ whose underlying set is the set of its vertices and in which $y \in N_\xi(x)$ iff there is an edge from x to y . Since the previous correspondences are inverse of one another, finite pretopological spaces can be identified with directed graphs (with a loop at each vertex), and we will denote by $N_G(x)$ the smallest neighborhood of a point x of a digraph G in its natural pretopology. In this context, a map $h : G_1 \rightarrow G_2$ between two digraphs is *continuous at $x \in G_1$* if

$$h(N_{G_1}(x)) \subseteq N_{G_2}(h(x)) .$$

Since the genotype set G can be equipped with the pretopology induced by the Hamming graph structure, and the dynamics on P is inherited from G via the folding map $f : G \rightarrow P$ (see 1.1), it is natural to consider the quotient pretopology on P induced by f , that is, the finest pretopology ξ_a on P making f continuous. In [12], this structure is called the *accessibility pretopology* on P . In this case, a phenotype α is in the ξ_a -neighborhood of β when $\alpha = f(x)$, for a one-error mutant x of a sequence y such that $f(y) = \beta$. In other words, $N_{\xi_a}(\beta)$ is the set of phenotypes that are possibly accessible from β by a single mutation. However, as observed in [12], accessibility is a weak condition to be a neighbor, which doesn't distinguish between neighbors likely to be realized and those unlikely to be. In particular,

accessibility is symmetric, whereas the probability of transition is not, as shows (*).

On the other side of the spectrum, the authors of [12] also consider the *shadow pretopology* ξ_s , for which α is in the ξ_s -neighborhood of β if *every* sequence y folding into β admits a one-error mutant x folding into α :

$$N_{\xi_s}(\beta) = \bigcap_{y \in f^{-1}(\beta)} f(N_G(y)) .$$

This is now too strong a requirement. A more meaningful structure should reflect the likelihood of phenotypic change $\text{Prb}(\beta \curvearrowright \alpha)$ of (\alpha that have G -neighbors with phenotype β . A pretopology on P defined in these terms would have to use a ‘‘cut-off value’’ $\delta \in \mathbb{R}_+$. More specifically, the structure ξ_δ would be given by the sets

$$(\ddagger) \quad N_{\xi_\delta}(\alpha) = \{\beta \in P : \text{Prb}(\alpha \curvearrowright \beta) \geq \delta\} .$$

However, if δ is too large, this set $N_{\xi_\delta}(\alpha)$ will not contain α anymore, and can therefore hardly be considered to be a neighborhood of α ! In other words, if

$$\delta > \inf_{\alpha \in P} \text{Prb}(\alpha \curvearrowright \alpha) ,$$

then, as P is finite, there will be an $\alpha_0 \in P$, such that $\alpha_0 \notin N_{\xi_\delta}(\alpha_0)$, so the sets $N_{\xi_\delta}(\alpha)$ (for $\alpha \in P$) will not define a pretopology. Of course, one could add the element α to each $N_{\xi_\delta}(\alpha)$, but this would mean that the information pertaining to $\text{Prb}(\alpha \curvearrowright \alpha)$ is explicitly ignored, a feature that one would not wish for in a faithful model of the phenotype space. This issue was overlooked in [12]. We present a corrected model in the next Section, but the use of structures more general than pretopologies constitutes a step further away from a topological model.

Note also that contrary to the claim in [12], even if $\delta \leq \inf_{\alpha \in P} \text{Prb}(\alpha \curvearrowright \alpha)$, the accessibility and shadow pretopologies are not the limiting cases for pretopologies ξ_δ . More specifically:

3.2.1. Proposition. *If ξ_δ is defined by (P then:*

- (1) $\xi_\delta \geq \xi_a$ for every $\delta > 0$;
- (2) $\xi_\delta = \xi_a$ for every $\delta \in (0, \inf_{\alpha \in P} \frac{1}{(3l+1)|f^{-1}(\alpha)|}]$;
- (3) ξ_s may not be comparable with ξ_δ .

Proof.

(1). If $\beta \in N_{\xi_\delta}(\alpha)$ then $\text{Prb}(\alpha \curvearrowright \beta) > 0$, so that, in view of (f^{-1}(\beta) \cap N_G(f^{-1}(\alpha)) \neq \emptyset. Hence,

$$\beta \in f(N_G(f^{-1}(\alpha))) = N_{\xi_a}(\alpha) .$$

(2). Assume that $\delta \leq \inf_{\alpha \in P} \frac{1}{(3l+1)|f^{-1}(\alpha)|}$ and that $\beta \in N_{\xi_a}(\alpha)$. The latter condition means that $f^{-1}(\beta) \cap N_G(f^{-1}(\alpha)) \neq \emptyset$. Therefore, in view of (\text{Prb}(\alpha \curvearrowright \beta) \geq \frac{1}{(3l+1)|f^{-1}(\beta)|}. Hence $\text{Prb}(\alpha \curvearrowright \beta) \geq \delta$ so that $\beta \in N_{\xi_\delta}(\alpha)$.

(3). Consider an example in which $f^{-1}(\beta) \cap N_G(f^{-1}(\alpha))$ has several elements, but with a proper subset A with the property that $N(y) \cap A \neq \emptyset$ for every $y \in f^{-1}(\alpha)$. Then, $\beta \in N_{\xi_s}(\alpha)$ but $\beta \notin N_{\xi_\delta}(\alpha)$ if $\delta > \frac{|f^{-1}(\beta) \cap N_G(f^{-1}(\alpha))|}{(3l+1)|f^{-1}(\beta)|}$. On the other hand, it is clear that we may have $\beta \in N_{\xi_\delta}(\alpha)$ without every $y \in f^{-1}(\alpha)$ having a neighbor in $f^{-1}(\beta)$, that is, such that $\beta \notin N_{\xi_s}(\alpha)$. \square

It was pointed out in [12] that the need for an ad-hoc choice of a cut-off point δ is not satisfying and calls for the introduction of a probabilistic or fuzzy version of a pretopology on P .

4. QUANTITATIVE ATOTOPOLOGIES

4.1. Probabilistic atopologies. Since the sets $N_{\xi_\delta}(\alpha)$ that play the role of neighborhoods need not contain the element α , we introduce the following terminology, based in part on the *probabilistic pretopologies* of [10]. An *atopology* ξ on a set X is the assignment of a filter $\mathcal{N}_\xi(x)$ to each $x \in X$. Although we do not require anymore that $x \in \bigcap \mathcal{N}(x)$, we continue to call $\mathcal{N}_\xi(x)$ the neighborhood filter of x . A *probabilistic atopology* Ξ is a family $\Xi = (\xi_\delta)_{\delta \in [0,1]}$ of atopologies on X where ξ_0 is the indiscrete topology, and

$$\lambda \leq \mu \implies \xi_\lambda \leq \xi_\mu .$$

Note that if $(\xi_\alpha)_\alpha$ is a family of atopologies on X , then the atopology $\bigvee_\alpha \xi_\alpha$ defined by $\mathcal{N}_{\bigvee_\alpha \xi_\alpha}(x) = \bigvee_\alpha \mathcal{N}_{\xi_\alpha}(x)$ is the supremum of all the ξ_α , that is, the coarsest atopology finer than each ξ_α . A probabilistic atopology Ξ is called *left-continuous* if

$$\xi_\lambda = \bigvee_{\delta < \lambda} \xi_\delta$$

for each $\lambda \in (0, 1]$ (see [10]). As the probabilistic atopologies that we will consider are left-continuous, we assume from now on that this property is part of the definition of a probabilistic atopology. A map $h : (X, \Xi) \rightarrow (Y, T)$ between two probabilistic atopological spaces is *continuous* if $h : (X, \xi_\delta) \rightarrow (Y, \tau_\delta)$ is continuous for every $\delta \in [0, 1]$:

$$h(\mathcal{N}_{\xi_\delta}(x)) \supseteq \mathcal{N}_{\tau_\delta}(h(x))$$

for all $x \in X$.

We endow the phenotype space P described in the previous Section with the probabilistic atopology $\Xi = (\xi_\delta)_{\delta \in [0,1]}$ where ξ_δ is defined by (‡) when $\delta > 0$. Note that in view of Proposition 3.2.1, the probabilistic atopology Ξ contains a “constant interval” of copies of the accessibility pretopology for $\delta \in (0, \inf_{\alpha \in P} \frac{1}{(3l+1)l^{f^{-1}(\alpha)}}]$.

While the finite atopologies ξ_δ define digraph structures on the set of vertices P , the probabilistic atopology Ξ puts different weights on each edge, and can therefore be interpreted as a Markov chain on P . Indeed, given two phenotypes α and β in P , we define

$$a_{\alpha\beta} := \bigvee \{ \delta \in [0, 1] : \alpha \in N_{\xi_\delta}(\beta) \} ,$$

so that

$$\text{Prb}(\beta \curvearrowright \alpha) := \frac{a_{\alpha\beta}}{\sum_{\beta' \in N_{\xi_\delta}(\alpha)} a_{\alpha\beta'}}$$

uniquely defines a Markov chain on P . Conversely, the Markov chain given by the collection $\{\text{Prb}(\beta \curvearrowright \alpha) : \alpha, \beta \in P\}$ defines Ξ , and these operations are inverse of each other if one starts with a Markov chain. Thus, a Markov chain on a finite set defines a unique probabilistic atopology on that set. Hence, the probabilistic atopologies can reproduce in a “topological-like” setting the information conveyed by the Markov chain discussed in Section 2. The only advantage of the interpretation in terms of probabilistic atopologies is that it makes a notion of continuity readily available. We discuss this aspect in the next Section.

4.2. The metric alternative. As observed in [2], the category of probabilistic pretopological spaces and continuous maps is equivalent to that of pre-approach spaces and contractions (in the sense of [7]). One can extend this equivalence to the category of probabilistic atotopological spaces, and certain weak approach spaces, which then provide an alternative description of the structures at hand. Given a set X together with a map $\lambda : FX \rightarrow [0, \infty]^X$, where FX denotes the set of all filters on X , we say that a pair (X, λ) is an *ametric space* if λ satisfies

$$\lambda(\bigwedge_{\alpha} \mathcal{F}_{\alpha})(x) = \bigvee_{\alpha} \lambda(\mathcal{F}_{\alpha})(x) .$$

Let us mention that for (X, λ) to form a pre-approach space, the structure must also satisfy

$$\lambda(\dot{x})(x) = 0$$

Intuitively, the *limit function* $\lambda(\mathcal{F})$ measures the default of convergence of the filter \mathcal{F} at each point x . If $\lambda(\mathcal{F})(x) = 0$, the filter fully converges to x , while $\lambda(\mathcal{F})(x) = \infty$ means that \mathcal{F} is as far as possible from converging to x . Morphisms between ametric spaces are contractions: a map $f : (X, \lambda_X) \rightarrow (Y, \lambda_Y)$ between two pre-approach spaces is a *contraction* if

$$\lambda_Y(f(\mathcal{F}))(f(x)) \leq \lambda_X(\mathcal{F})(x) .$$

Roughly put, a probabilistic atotology assigns a probability of convergence to each pair of filter and point, while an ametric structure assigns to such a pair a measure of the default of convergence in $[0, \infty]$. The map $-\ln : [0, 1] \rightarrow [0, \infty]$ establishes a one-to-one correspondence between these two measures that also transforms sums into products (although this last feature only becomes relevant when one considers the “diagonal” condition used to define *approach* spaces). There are several other equivalent descriptions of pre-approach spaces that the interested reader may find in, e.g., [6], [7], [8], [9].

Ametric structures on a finite set are all *finitely generated* [7], and can simply be interpreted as sets equipped with a map $d : X \times X \rightarrow [0, \infty]$. Indeed, if (X, λ) is a finite ametric space then

$$d_{\lambda}(x, y) := \lambda(\dot{x})(y)$$

defines such a map on X . On the other hand, every filter is principal so that each map $d : X \times X \rightarrow [0, \infty]$ determines an ametric structure λ_d via

$$\lambda_d(A)(x) := \bigvee_{a \in A} d(a, x) .$$

Finally, it is easy to see that $\lambda_{d_{\lambda}} = \lambda$ and $d_{\lambda_d} = d$.

4.3. Ametric spaces via the unit interval. As mentioned in 4.2 above, the order-reversing map $-\ln : [0, 1] \rightarrow [0, \infty]$ yields a bijection between the probabilistic and metric views of the spaces that reflect the Markov structure of the phenotype space P . Because the link of the latter with the probabilistic presentation is more direct, it will be useful to translate the previous description of ametric spaces into its probabilistic counterpart. Thus, from now on, an *ametric space* (X, λ) will be a set X equipped with a map $\lambda : FX \rightarrow [0, 1]^X$ such that

$$\lambda(\bigwedge_{\alpha} \mathcal{F}_{\alpha})(x) = \bigwedge_{\alpha} \lambda(\mathcal{F}_{\alpha})(x)$$

In this case, the *limit function* $\lambda : FX \rightarrow [0, 1]^X$ measures the “probability” that a filter \mathcal{F} converges to a point $x \in X$. A *contraction* in this setting is a map $f : (X, \lambda_X) \rightarrow (Y, \lambda_Y)$ between ametric spaces satisfying

$$\lambda_X(\mathcal{F})(x) \leq \lambda_Y(f(\mathcal{F}))(f(x)) .$$

In other words, a contraction increases the probability of convergence of filters to points.

In this context, the finitely generated ametric spaces are those that can be described by a map $d' : X \times X \rightarrow [0, 1]$, and a map $f : (X, d'_X) \rightarrow (Y, d'_Y)$ corresponds to a contraction iff

$$d'_X(x, y) \leq d'_Y(f(x), f(y))$$

for all $x, y \in X$. The relation with the weighted digraph structure on P is immediate via the adequate normalization at each point $\alpha \in P$, as described previously in 4.2. In particular, one defines

$$d'_P(\alpha, \beta) := \text{Prb}(\beta \curvearrowright \alpha) ,$$

for all $\alpha, \beta \in P$, so that $d'_P(\alpha, \beta)$ represents the weight of each directed edge of the graph on P .

5. CONTINUITY, AND EVOLUTIONARY TRAJECTORIES

5.1. Continuity of the GP-map. As pointed out before, the accessibility pretopology is the finest pretopology on P making the GP-map everywhere continuous. Hence, the GP-map is *not* continuous into the finer atopologies ξ_δ as soon as

$$\delta > \inf\{\text{Prb}(\beta \curvearrowright \alpha) \mid \alpha, \beta \in P : \text{Prb}(\beta \curvearrowright \alpha) > 0\} .$$

Although the weighted digraph structure on the set P can faithfully be reproduced by a probabilistic atopological space, this structure is *not* the probabilistic atopological quotient structure on P induced by the GP-map $f : G \rightarrow P$ (if G is the metric space obtained thanks to the Hamming distance d_G). This can easily be seen by considering the quotient ametric structure on P :

$$d_P(\alpha, \beta) := \bigwedge_{\substack{x \in f^{-1}(\alpha) \\ y \in f^{-1}(\beta)}} d_G(x, y) ,$$

for all $\alpha, \beta \in P$. In the $[0, 1]$ setting, this structure is described by

$$d'_P(\alpha, \beta) = \bigvee_{\substack{x \in f^{-1}(\alpha) \\ y \in f^{-1}(\beta)}} \exp(-d_G(x, y)) .$$

The main obstacle to obtaining the required structure on P is that a contraction is in no way related to the sums used in a probabilistic distribution, as can be seen in the definition of a contraction in the unit interval setting. Thus, although probabilistic atopologies can completely reflect the initial data on a weighted digraph, the corresponding continuous maps do not preserve their structure conveniently. In particular, the quotient graph on P does not appear as a quotient structure of an atopological one. In the next Subsection, we briefly investigate whether maps between phenotype spaces can benefit from the point of view provided by probabilistic atopologies.

5.2. Evolutionary trajectories. The notion of continuity for evolutionary trajectories proposed in [12] is not very satisfying, because the structure on the space of phenotypes involves a choice of an ad-hoc cut-off point. Using the probabilistic atotology on P described in the previous Section should provide a useful alternative. However, the notion of trajectory needs to be re-interpreted. In [12], an evolutionary trajectory is a map $\tau : T \rightarrow P$ where T is a discrete time space (essentially, the natural numbers) and P the space of phenotypes. When T is endowed with its natural pretopology $N(t) = \{t-1, t\}$ (so that $t \rightarrow t+1$), and P with a pretopology (or an atotology), it is possible to consider whether trajectories are continuous or not. Here it should be noted that each increment in discrete time corresponds to a potential mutation. Hence, if τ represents an evolutionary trajectory, then $\tau : T \rightarrow (P, \xi_a)$ should be continuous. Indeed, τ can be factored through G as $\tau = f \circ h$ where $h : T \rightarrow G$ describes the state of a sequence at time t and $f : G \rightarrow P$ is the folding map. Here h is continuous if $h(t+1)$ is obtained from $h(t)$ by a single mutation, which in our view is an assumption to be made on an evolutionary trajectory in the context of a discrete time. Since $f : G \rightarrow (P, \xi_a)$ is continuous, so is τ .

In [12], the authors suggest the possibility of discontinuous genotypic change h , but we see the possibility of several mutations between time t and $t+1$ as inconsistent with the choice of a discrete space time: otherwise T is merely a discrete selection of instants in a continuous time space, and there is no reason not to consider the continuous time space instead. In this context, we think that an evolutionary trajectory should be defined as a *continuous* map $\tau : T \rightarrow (P, \xi_a)$. In contrast, for other atotologies ξ_δ , there will be discontinuities as observed in [12]. However, the biological interpretation of such discontinuities is uneasy, because a discontinuity for one δ may not be one for a smaller δ . If one endows P with a specific atotology ξ_{δ_0} , the discontinuities depend on the choice of the threshold value δ_0 . Replacing an atotology ξ_δ by a probabilistic version Ξ does not lead to a meaningful notion of continuity, because the time T is not endowed with a probabilistic pretopology. Indeed, time flows *necessarily* from t to $t+1$, so an atotology σ can be identified with a probabilistic atotology $\Sigma = (\sigma_t)_{t \in I}$ where $\sigma_t = \sigma$ for every $t \in (0, 1]$, but then a map $\tau : (T, \Sigma) \rightarrow (P, \Xi)$ would only be continuous if all transitions in P are certain, too.

Another approach is to consider the orbits $\{h^n(\alpha) : n \in \mathbb{N}\}$ in the dynamic of a map $h : P \rightarrow P$, where each iteration of h represents an increment in discrete time along an evolutionary trajectory. It seems natural to restrict ourselves to maps satisfying $h(\alpha) \in \text{adh}_{\xi_a} \alpha$, that is, maps for which $h(\alpha)$ is accessible from α . In this case, $h^n(\alpha) \in \text{adh}_{\xi_a}^n \alpha = \text{adh}_{\xi_a} \left(\text{adh}_{\xi_a}^{n-1} \alpha \right)$ for every n , so that each trajectory stays in the smallest ξ_a -open neighborhood of the initial condition α for the topological modification of ξ_a . This, however, doesn't provide much information, as this open neighborhood is simply the connected component of α . Note that for a single trajectory, that is, an orbit $\{h^n(\alpha) : n \in \mathbb{N}\}$ of a given phenotype α , we are only concerned by continuity at points of the orbit. However, even in restriction to an orbit, continuity of $h : (P, \Xi) \rightarrow (P, \Xi)$ is far too stringent a property. Indeed, in the ametric representation of P , global continuity of h reads as

$$d'_P(\alpha, \beta) \leq d'_P(h(\alpha), h(\beta)) ,$$

for all $\alpha, \beta \in P$. In particular, if at least one of these inequalities is strict, then h cannot be surjective, or the sum of the probabilities on the images of all β at $h(\alpha)$ will be strictly greater than 1. For the trajectory of a phenotype α , the condition of continuity (at $h^n(\alpha)$) reads as

$$d'_P(h^n(\alpha), h^p(\alpha)) \leq d'_P(h^{n+1}(\alpha), h^{p+1}(\alpha)) .$$

In particular, $d'_P(h^n(\alpha), h^{n+1}(\alpha)) \leq d'_P(h^{n+1}(\alpha), h^{n+2}(\alpha))$. Hence a continuous trajectory follows a path of increasing probability of phenotypic change. However, we also have $d'_P(h^n(\alpha), h^{n-1}(\alpha)) \leq d'_P(h^{n+1}(\alpha), h^n(\alpha))$. In other words, the probability of “moving back” also increases along the trajectory. This does not fit the expected behavior. Hence, continuity of $h : (P, \Xi) \rightarrow (P, \Xi)$, even in restriction to an orbit, doesn't yield a meaningful notion of “continuous trajectory”.

The main motivation for a topological approach to evolutionary change was to interpret some biological phenomenon like punctuated equilibrium as a topological discontinuity. It turns out that even if probabilistic atologies can reflect all the statistical information necessary to study the dynamic, the associated notion of continuity appears useless.

6. MUTATION, RECOMBINATION AND GENERAL CLOSURE SPACES

We have seen that the structure induced on the set of genotypes by simple mutation is that of a directed graph, which can be interpreted as a pretopology. The structure on the phenotype space that describes accessibility is then the quotient pretopology. If recombination (crossover) is considered, the adjacency relation of the graph is replaced by recombination sets $r(x, y)$ of all possible recombinants of two “parents” x and y , as discussed in e.g. [14], [11]. Minimal assumptions on recombination sets are that x and y are in $r(x, y)$, that is, replication is possible; and that $r(x, y) = r(y, x)$, because the role of the two parents is interchangeable. We may assume that $r(x, x) = \{x\}$ only if unequal crossover is ruled out. In a general model, this last property may not be satisfied. Also, we may include in $r(x, y)$ both recombinants and one-error mutants of x and y , in which case $r(x, x)$ would contain at least one-error mutants of x . The recombinations sets induce a closure operator on the set of genotype via

$$\text{cl } A = \bigcup_{(x,y) \in A \times A} r(x, y).$$

By definition, this closure operator is grounded, and isotone. Because replication is possible, it is also expansive. However, it needs not be additive nor idempotent. Sets equipped with such a closure operator have been considered under various names in the literature. We will call them *preclosure spaces*. Hence a pretopological space is an additive preclosure space and a topological space is a preclosure space that is both additive and idempotent. Idempotent preclosure spaces are usually called *closure spaces*. Hence, in a general theory of evolutionary accessibility finite preclosure spaces form a natural model of genotype spaces. The notion of continuity introduced for pretopological spaces generalizes to preclosure spaces: a map $f : (X, \text{cl}_X) \rightarrow (Y, \text{cl}_Y)$ between two preclosure spaces is *continuous* if

$$f(\text{cl}_X A) \subseteq \text{cl}_Y (f(A)) ,$$

for every $A \subset X$. Note that, as in the case of mutation, phenotypic accessibility is adequately described in topological terms. Indeed, if the set of genotype G is

endowed with a preclosure cl_G , the GP-map $f : (G, \text{cl}_G) \rightarrow P$ induces on P the *quotient preclosure* cl_P in which

$$\text{cl}_P(A) = f(\text{cl}_G(f^{-1}(A))) ,$$

for every $A \subset P$. This is the finest preclosure structure on P making f continuous. Note that $\beta \in \text{cl}_P(\alpha)$ if and only if $f^{-1}(\beta) \cap \text{cl}_G(f^{-1}(\alpha)) \neq \emptyset$, that is, if the potential offsprings of genotypes with phenotype α include a sequence with phenotype β . Hence, this structure describes phenotypic accessibility, as noticed in e.g. [11]. However, we have seen in the previous Section that even in the simplest case (mutation/pretopology), a topological model carrying the statistical information needed to make meaningful predictions will not yield a useful notion of continuity. In the next Section, we propose an alternative model, which is intrinsically non-topological. This makes any parallel with pre-existing closure operations difficult, if not impossible.

7. A NON-TOPOLOGICAL MODEL

The discussions in Section 5 point to the fact that the continuity condition for probabilistic pretopological spaces does not convey the pertinent information for the study of the GP-map model. We propose here an alternative point of view from which the model could be studied, but we insist on the fact that it is not topological in nature.

The emphasis in the GP-map model is on “accessibility” of a subset of G by a one-mutation of a sequence of RNA. Hence, an abstract structure on a finite set X reflecting this notion is an operation $\text{cl} : 2^X \rightarrow [0, 1]^X$ that measures the probability $\text{cl} A(x)$ that an element x is accessed from a subset $A \subseteq X$. This operation should therefore satisfy

- (1) $\text{cl} \emptyset = 0$;
- (2) $\forall x \in A, \text{cl} A(x) \neq 0$;
- (3) $\sum_{x \in X} \text{cl} A(x) = 1$.

The last condition ensures that $\text{cl} A$ is indeed a probability distribution for all $A \subseteq X$. To be meaningful, such an operation should not be monotone ($A \subseteq B \implies \text{cl} A \leq \text{cl} B$) in order to avoid a contradiction in the probabilistic condition. This contributes to the “non-topological” nature of the operation cl . We will call cl a *probabilistic closure* even though it differs in essential ways from probabilistic pretopologies or atopologies. A map $f : (X, \text{cl}_X) \rightarrow (Y, \text{cl}_Y)$ between two such spaces is *continuous* if

$$\sum_{x \in f^{-1}(y)} \text{cl}_X A(x) \leq \text{cl}_Y [f(A)](y) ,$$

for all $y \in Y, A \in 2^X$. Note that continuity of a map between probabilistic pretopological (or atopological) spaces can be described by a similar formula, in which the sum is replaced by a supremum. This is, however, an essential difference: in the probabilistic topological like models, continuity means that the probability of accessing a phenotype β from $f(A)$ is at least the largest probability of accessing a genotype with phenotype β from A , but it does not reflect how many times such genotypes are accessed from A . This drawback is corrected in this new model. Even better, the GP-map induces a quotient structure which is exactly the desired one.

Indeed, if f is a *surjective* map $f : (X, \text{cl}_X) \rightarrow Y$, then the quotient structure on Y is given by

$$\text{cl}_Y B(y) := \sum_{x \in f^{-1}(y)} \text{cl}_X[f^{-1}(B)](x),$$

for all $B \in 2^Y$, $y \in Y$. For example, when the genotype set G is equipped with its looped Hamming graph structure, then we define

$$\text{cl}_G A(x) := \text{Prb}(x \curvearrowright A),$$

where $\text{Prb}(x \curvearrowright A)$ is given by (*). Then the quotient structure on P is precisely the quotient graph structure in which

$$\text{cl}_P \alpha(\beta) = \text{Prb}(\beta \curvearrowright \alpha)$$

as in (†).

More generally, the genetic operations acting on G in the case of mutation or of recombination can be seen as maps of the form $c : G \rightarrow \{0, 1\}^G$ in the first case and $c : G \times G \rightarrow \{0, 1\}^G$ in the second. In the first case $c(x)(y) = 1$ if y is a one-error mutant of x . In the second case, $c(x, y)(z) = 1$ if z is a potential offspring of the pair (x, y) . We can identify $c : G \rightarrow \{0, 1\}^G$ to a map of the second kind $c' : G \times G \rightarrow \{0, 1\}^G$ by $c'(x, y) \equiv 0$ if $x \neq y$ and $c'(x, x)(z) = c(x)(z)$. Hence we can treat both cases simultaneously, and associate to such a map c a probabilistic closure via

$$\text{cl}_c A(x) = \frac{\sum_{(a,b) \in A \times A} c(a, b)(x)}{\sum_{y \in G} \sum_{(a,b) \in A \times A} c(a, b)(y)}.$$

Then $\text{cl}_c A(x)$ represents the probability to access the genotype x from the collection of genotypes A . Note that the same formula applies if $c : G \times G \rightarrow [0, 1]^G$ carries some statistical information on the likelihood to access a specific offspring, or mutant, among the potential offsprings, or mutants respectively. Once the genotype space is endowed with its probabilistic closure structure induced by the genetic operator c , the GP-map $f : G \rightarrow P$ induces the quotient probabilistic closure on P (the finest making f continuous) given by

$$\text{cl}_P B(\alpha) = \sum_{x \in f^{-1}(\alpha)} \text{cl}_c[f^{-1}(B)](x) = \sum_{x \in f^{-1}(\alpha)} \frac{\sum_{(y,z) \in f^{-1}(B) \times f^{-1}(B)} c(y, z)(x)}{\sum_{t \in G} \sum_{(y,z) \in f^{-1}(B) \times f^{-1}(B)} c(y, z)(t)}.$$

This operation adequately reflects the likelihood of accessing the phenotype α from the collection of phenotypes B .

Although final structures exist for surjective maps, they do not in general. Indeed, the substitution of a supremum by a sum in the definition of continuity is essential to adequately reflect the statistical information, but it makes the model non-topological in an essential way: the resulting category of the structured objects (X, cl_X) and their continuous maps is *not* topological over **Set** (in the sense of [1]). Topological categories share the following important property with the category of topological spaces and continuous maps: if $(f_i : X_i \rightarrow Y)_{i \in I}$ is a family of maps from topological spaces X_i to a set Y , there exists the finest topology on Y making each f_i continuous. This is the final structure. Similarly, if $(f_i : X \rightarrow Y_i)_{i \in I}$ is a family of maps from a set X to topological spaces Y_i , there exists the coarsest topology on Y making each f_i continuous. This is the initial structure. Final structures allow to define quotient maps while initial structures allow to define

subspaces and products, among others. With the new definition of continuity, final and initial structures do not always exist. In particular, the canonical notion of a product structure $\text{cl}_{X \times Y} : 2^{X \times Y} \rightarrow [0, 1]^{X \times Y}$ obtained from the structures $\text{cl}_X : 2^X \rightarrow [0, 1]^X$ and $\text{cl}_Y : 2^Y \rightarrow [0, 1]^Y$ by initial construction with respect to the projections is not available. Among other problems, this is a major obstacle to developing a theory of characters for these structures parallel to that developed in [14], which is based on the representation of a phenotype space or a region thereof, as a product space where each factor represents a character.

REFERENCES

- [1] J. Adámek, H. Herrlich, and E. Strecker. *Abstract and Concrete Categories*. John Wiley and Sons, Inc., 1990.
- [2] Paul Brock and D. C. Kent. Approach spaces, limit tower spaces, and probabilistic convergence spaces. *Appl. Categ. Structures*, 5(2):99–110, 1997.
- [3] S. Dolecki. *An Initiation Into Convergence Theory*, volume Beyond Topology. A.M.S., 2008. http://math.u-bourgogne.fr/topo/dolecki/Page/init_X.pdf.
- [4] W. Fontana and P. Schuster. Shaping space: The possible and the attainable in RNA genotype-phenotype mapping. *J. Theor. Biol.*, 194:491–515, 1998.
- [5] C.M. Grinstead and J.L. Snell. *Introduction to Probability*. American Mathematical Society, 1997.
- [6] E. Lowen and R. Lowen. Topological quasitopos hulls of categories containing topological and metric objects. *Cahiers de Topologies et Géométrie différentielle Catégorique*, 30:213–228, 1989.
- [7] E. Lowen, R. Lowen, and C. Verbeeck. Exponential objects in PRAP. *Cahiers de Topologies et Géométrie différentielle Catégorique*, 38:259–276, 1997.
- [8] R. Lowen. *Approach Spaces: The Missing Link in Topology-Uniformity-Metric Triad*. Oxford University Press, 1997.
- [9] F. Mynard. Measures of compactness for filters in the approach setting. *Quest. Math.*, to appear, 2007.
- [10] G. D. Richardson and D. C. Kent. Probabilistic convergence spaces. *J. Austral. Math. Soc. (Series A)*, 61:400–420, 1996.
- [11] B. Stadler and P. Stadler. The topology of evolutionary biology. In G. Ciobanu and G. Rozenberg, editors, *Modeling in Molecular Biology*, pages 267–286. Springer Verlag, 2004.
- [12] B. Stadler, P. Stadler, G. Wagner, and W. Fontana. The topology of the possible: Formal spaces underlying patterns of evolutionary change. *J. Theor. Biol.*, 213:241–274, 2001.
- [13] P. Stadler. Spectral landscape theory. In J. P. Crutchfield and P. Schuster, editors, *Evolutionary Dynamics. Exploring the Interplay of Selection, Neutrality, Accident and Function.*, pages 231–272. Oxford University Press, 2002.
- [14] Günter P. Wagner and Peter F. Stadler. Quasi-independence, homology and the unity of type: a topological theory of characters. *J. Theoret. Biol.*, 220(4):505–527, 2003.