Categorical dynamics: from Hopfield to Pareto

Matilde Marcolli

Algebra, Geometry and Physics: a mathematical mosaic
Yuri Manin 85th birthday conference
MPI Bonn, March 8, 2022

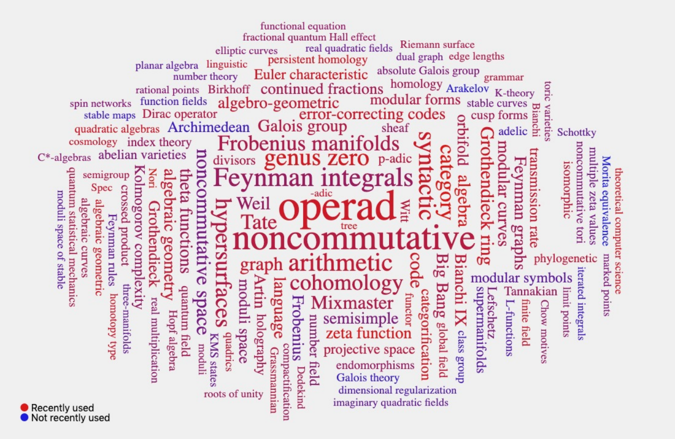
A Personal Introduction

Once upon a time... I moved from MIT to the MPI in 2000, I started collaborating with Yuri shortly after moving to Bonn

- Yuri I. Manin, Matilde Marcolli, *Continued fractions, modular symbols, and noncommutative geometry*. Selecta Math. (N.S.) 8 (2002), no. 3, 475–521.
- Yuri I. Manin, Matilde Marcolli, *Holography principle and arithmetic of algebraic curves*. Adv. Theor. Math. Phys. 5 (2001), no. 3, 617–650.

the first generated many interesting connections between arithmetic and noncommutative geometry, the second is connected to recent developments in p-adic AdS/CFT holography

Over the years, we coauthored 20 more papers: noncommutative geometry and arithmetic, $\mathbb{F}_1$-geometry, string theory, classical and quantum codes, cosmology, semantics, motives, persistent homology, dessins, neural networks, homotopy theory spectra, information



“Word Cloud” from the arXiv

Some things I learned:

1. math (a lot!)
2. how nice it is to look for unexpected connections between seemingly different things
3. the value of following one's own imagination
(math is a way of expressing the imagination)

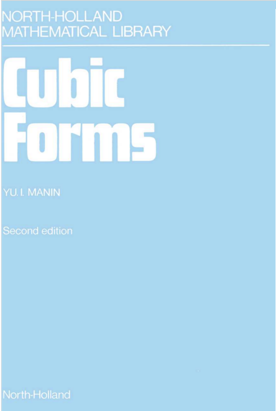
Ёжик в тумане

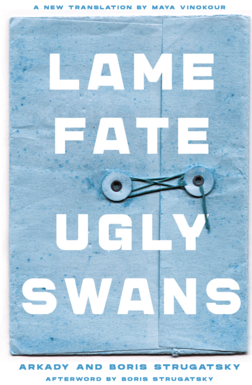
4. the value of continuously learning new things: what we half-see in the fog when exploring what we don't yet know
5. how nice it is to have an amazing friend!

... plus lots of fun, shared thoughts, and surprising coincidences

Recently... while working on one of our latest papers

- Noemie Combe, Yuri I. Manin, Matilde Marcolli, *Moufang Patterns and Geometry of Information*, arXiv:2107.07486 (for the birthday of our friend Don Zagier) ... I was reading:





"Мальчик-вундеркинд, почитывает «Кубические формы» Ю. Манина, очкарик; когда моет посуду, любит петь Высоцкого."

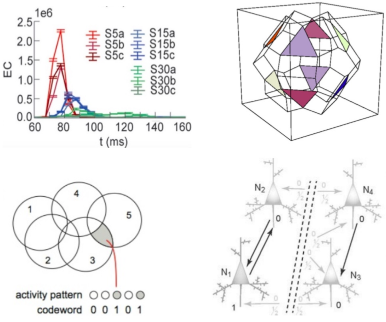
About this talk:

I will talk about shapes in the fog, not yet fully formed recognizable structures

The motivation comes from our recent work

- Yuri I. Manin, M. Marcolli, *Homotopy Theoretic and Categorical Models of Neural Information Networks*, arXiv:2006.15136

Original motivation: some observation from neuroscience and the theory of computation



(a) neocortical microcircuitry: formation of nontrivial Betti numbers in clique complex of network activated in response to stimuli; (b) in distributed computing enough nontrivial homology is necessary for successfully complete their task; (c) neural codes encode stimulus space up to homotopy; (d) integrated information and other information measures assigned to networks (not reducible to separate subsystems); also relation between informational and metabolic resources

Categories of Resources

- mathematical theory of resources
    - B. Coecke, T. Fritz, R.W. Spekkens, *A mathematical theory of resources*, Information and Computation 250 (2016), 59–86. [arXiv:1409.5531]
- Resources modelled by a symmetric monoidal category $(\mathcal{R}, \circ, \otimes, \mathbb{I})$ (or written "additively" as $(\mathcal{R}, \oplus, 0)$)
- objects $A \in \mathrm{Obj}(\mathcal{R})$ represent resources, product $A \otimes B$ represents combination of resources, unit object $\mathbb{I}$ empty resource
- morphisms $f : A \to B$ in $\mathrm{Mor}_{\mathcal{R}}(A, B)$ represent possible conversions of resource A into resource B
- convertibility of resources when $\mathrm{Mor}_{\mathcal{R}}(A, B) \neq \emptyset$

Measuring semigroups of categories of resources (Coecke, Fritz, Spekkens)

- preordered abelian semigroup $(R, +, \succeq, 0)$ on set R of isomorphism classes of $\mathrm{Obj}(\mathcal{R})$ with $A + B$ the class of $A \otimes B$ with unit 0 given by the unit object $\mathbb{I}$ and with $A \succeq B$ iff $\mathrm{Mor}_{\mathcal{R}}(A, B) \neq \emptyset$
- maximal conversion rate $\rho_{A \to B}$ of resources

$$\rho_{A \to B} := \sup\{ \frac{m}{n} \mid n \cdot A \succeq m \cdot B, \;\; m, n \in \mathbb{N} \}$$

  number of copies of resource A are needed on average to produce B
- measuring semigroup: abelian semigroup with partial ordering and semigroup homomorphism $M : (R, +) \to (S, *)$ with $M(A) \geq M(B)$ in S when $A \succeq B$ in R
- satisfy $\rho_{A \to B} \cdot M(B) \leq M(A)$

Summing functors (Segal's Gamma spaces formalism, 1973)

- $\mathcal{C}$ a category with sum and zero-object (binary codes, transition systems, resources, etc)
- (X, x_0) a pointed finite set and $\mathcal{P}(X)$ a category with objects the pointed subsets $A \subseteq X$ and morphisms the inclusions $j : A \subseteq A'$
- a functor $\Phi_X : \mathcal{P}(X) \to \mathcal{C}$ summing functor if

$$\Phi_X(A \cup A') = \Phi_X(A) \oplus \Phi_X(A') \quad \text{when} \quad A \cap A' = \{x_0\}$$

  and $\Phi_X(\{x_0\})$ is zero-object of $\mathcal{C}$
- $\Sigma_{\mathcal{C}}(X)$ category of summing functors $\Phi_X : \mathcal{P}(X) \to \mathcal{C}$, morphisms are *invertible* natural transformations

• Key idea: a summing functor is a *consistent assignment* of *resources* of type $\mathcal{C}$ to *all subsystems* of X so that a combination of independent subsystems corresponds to combined resources
• $\Sigma_{\mathcal{C}}(X)$ is the configuration space that parameterizes all possible such assignments

Category of summing functors

- original Segal formulation for $\mathcal{C}$ with sum and zero-object, extended by Thomason to $\mathcal{C}$ unital symmetric monoidal
- a summing functor $\Phi_X : P(X) \to \mathcal{C}$ completely determined by values $\Phi_X(x) := \Phi_X(A_x)$ on $A_x = \{x, \star\}$ for $x \in X \smallsetminus \{\star\}$
- $\hat{\mathcal{C}}$ category with same objects as $\mathcal{C}$ and the invertible morphisms of $\mathcal{C}$
- X finite pointed set with $\#X = n+1$: category $\Sigma_{\mathcal{C}}(X)$ of summing functors equivalent to $\hat{\mathcal{C}}^n$
- similarly for unital symmetric monoidal case: summing functors specified by

$$\Phi_X := \{\Phi_X(x)\}_{x \in X \smallsetminus \star}$$

in $\hat{\mathcal{C}}^n$ with morphisms in $\hat{\mathcal{C}}^n$

Meaning in homotopy theory (loop-deloop)

- take case where $\mathcal{C}$ is an abelian category, then (Quillen) the higher K-theory $K(\mathcal{C})$ is the K-theory of an infinite loop space
- the category of summing functors $\Sigma_{\mathcal{C}}(X)$ provides a delooping of this infinite loop space (Carlsson)
- a Gamma space defines an associated spectrum, by extending the functor $\Gamma : \mathcal{F} \to \Delta$ with $\mathcal{F}$ finite (pointed) sets and Δ (pointed) simplicial sets to an endofunctor $\Gamma : \Delta \to \Delta$ and applying it to spheres
- when $\mathcal{C} = \mathcal{F}$ with $\Gamma_{\mathcal{F}} : \mathcal{F} \to \Delta$ get the sphere spectrum
- all connective spectra are obtained through this construction for $\mathcal{C}$ a symmetric monoidal category (Thomason)
- hence nerves $\mathcal{N}(\Sigma_{\mathcal{C}}(X))$ are topologically very nontrivial

Meaning in our setting

- nerve $\mathcal{N}(\Sigma_{\mathcal{C}}(X))$ of category of summing functors organizes *all assignments of $\mathcal{C}$-resources to X-subsystems and their transformations* into a single topological structure that keeps track of equivalence relations between them (invertible natural transformations as morphisms of $\Sigma_{\mathcal{C}}(X)$ and their compositions become simplexes of the nerve)
- view $\mathcal{N}(\Sigma_{\mathcal{C}}(X))$ as a topological parameterizing space for all such consistent assignments of resources of type $\mathcal{C}$ to subsets of X

From finite sets to networks: directed graphs

- category $\mathbf{2}$ has two objects V, E and two morphisms $s, t \in \mathrm{Mor}(E, V)$
- $\mathcal{F}$ category of finite sets: objects finite sets, morphisms functions between finite sets
- a directed graph is a functor $G : \mathbf{2} \to \mathcal{F}$
  - $G(E)$ is the set of edges of the directed graph
  - $G(V)$ is the set of vertices of the directed graph
  - $G(s) : G(E) \to G(V)$ and $G(t) : G(E) \to G(V)$ are the usual source and target maps of the directed graph
- category of directed graphs $\mathrm{Func}(\mathbf{2}, \mathcal{F})$ objects are functors and morphisms are natural transformations

Network summing functors

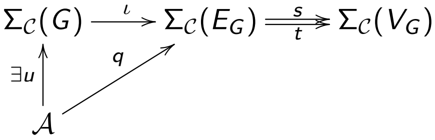
- want to replace finite set X with directed graph G (functor $G : \mathbf{2} \to \mathcal{F}$, from category with objects V, E morphisms s, t to finite sets)
- basic additivity rule functor $\Phi : P(G) \to \mathcal{C}$ from subgraphs with inclusions

$$\Phi(G' \sqcup G'') = \Phi(G') \oplus \Phi(G'')$$

- then want additional compositionality rules
  1. conservation laws at vertices (Kirchhoff): equalizer and coequalizer constructions for source and target
  2. grafting operation (when target category $\mathcal{C}$ has operad or properad structure)
  3. inclusion exclusion (as exact sequences) when $\mathcal{C}$ is abelian
  4. etc

Example: Conservation laws at vertices

- source and target functors $s, t : \Sigma_{\mathcal{C}}(E_G) \rightrightarrows \Sigma_{\mathcal{C}}(V_G)$
- equalizer category $\Sigma_{\mathcal{C}}(G)$ with functor $\iota : \Sigma_{\mathcal{C}}(G) \to \Sigma_{\mathcal{C}}(E_G)$ such that $s \circ \iota = t \circ \iota$ with universal property



- this is category of summing functors $\Phi_E : P(E_G) \to \mathcal{C}$ with conservation law at vertives: for all $A \in P(V_G)$

$$\Phi_E(s^{-1}(A)) = \Phi_E(t^{-1}(A))$$

  in particular for all $v \in V_G$ have inflow of $\mathcal{C}$-resources equal outflow

$$\oplus_{e:s(e)=v} \Phi_E(e) = \oplus_{e:t(e)=v} \Phi_E(e)$$

- another kind of conservation law expressed by coequalizer

Example: Grafting

- $\mathcal{C}$ symmetric monoidal category, full subcategories $\mathcal{C}(n,m)$
  - $\mathrm{Obj}(\mathcal{C}) = \cup_{n,m\in\mathbb{N}} \mathrm{Obj}(\mathcal{C}(n,m))$;
  - the monoidal structure $(\otimes, \mathbb{I})$ satisfies

$$\otimes : \mathcal{C}(n,k) \times \mathcal{C}(n,r) \to \mathcal{C}(m+n, k+r)\,;$$

  - the family $\{\mathcal{C}(n,m)\}_{n,m\in\mathbb{N}}$ is a properad in Cat.
- then $\Sigma^{prop}_{\mathcal{C}}(G) \subset \Sigma_{\mathcal{C}}(G)$ summing functors $\Phi : P(G) \to \mathcal{C}$:
  1. $\Phi(G') \in \mathrm{Obj}(\mathcal{C}(\deg^{in}(G'), \deg^{out}(G'))$
  2. $\Phi(\{v\}) = \Phi(C(v))$ corolla $C(v)$
  3. $\Phi(G' \star G'') = \Phi(G') \circ_{E(G',G'')} \Phi(G'')$ with $E(G',G'') \subset E_G$ edges one en in $V_{G'}$ one in V_G and $\circ_{E(G',G'')}$ properad composition

$$\circ_{E(G',G'')} : \mathcal{C}(\deg^{in}(G'), \deg^{out}(G')) \times \mathcal{C}(\deg^{in}(G''), \deg^{out}(G''))$$

$$\to \mathcal{C}(\deg^{in}(G' \star G''), \deg^{out}(G' \star G''))$$

- network summing functor $\Phi \in \Sigma^{prop}_{\mathcal{C}}(G)$ completely determined by value on corollas

Discrete and continuous Hopfield dynamics

- discrete version (binary neurons)

$$\nu_j(n+1) = \begin{cases} 1 & \text{if } \sum_k T_{jk}\nu_k(n) + \eta_j > 0 \\ 0 & \text{otherwise} \end{cases}$$

- continuous version (neuron firing rates as variables and threshold-linear dynamics)

$$\frac{dx_j}{dt} = -x_j + \left(\sum_k W_{jk} x_j + \theta_j\right)_+$$

W_{jk} real-valued connection strengths, θ_j constant external inputs, and $(\cdot)_+ = \max\{0, \cdot\}$ threshold function

- finite difference version

$$\frac{x_j(t+\Delta t) - x_j(t)}{\Delta t} = -x_j + (\sum_k W_{jk} x_k(t) + \theta_j)_+$$

(versions with or without "leak term" $-x_j$ on r.h.s.)

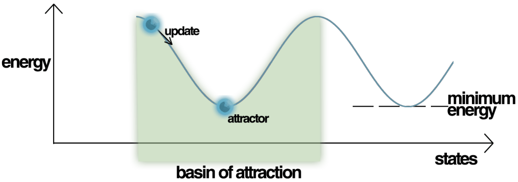
Hopfield Network (original binary case)

- historical connection between statistical physics of spin glass models and neural networks

- nodes variables $s_i = \pm 1$, update

$$s_i = \begin{cases} +1 & \sum_j w_{ij} s_j \geq \theta_i \\ -1 & \text{otherwise} \end{cases}$$

$$E = -\frac{1}{2} \sum_{i,j} w_{ij} s_i s_j - \sum_i \theta_i s_i$$

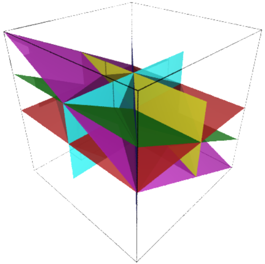


Energy landscape of the Hopfield network

convergence to fixed point attractor: model of associative memory

Hopfield Networks with threshold nonlinearity (recent work by Carina Curto and collaborators)

- Hopfield equations (continuous version) with threshold nonlinearity: patchwork of linear systems across wall crossings given by a hyperplane arrangement
- fixed points arise when fixed points of the linear system lie in the correct chamber of the hyperplane arrangement
- dynamical behavior (fixed points support) from combinatorial structures of (sub)graphs "motifs"



• Here we follow a different (but related) idea: Hopfield equations where variables are not the real x_i but assignments of resources to the network (summing functors Φ): equation on the configuration space $\Sigma_{\mathcal{C}}(G)$: key idea: all levels of structure associated to network (functorially related) evolve consistently in the dynamics

Categorical Hopfield dynamics

- as above the configuration space where the dynamics takes place is a category $\Sigma_{\mathcal{C}}(G)$ of network summing functors, for a given network G and category of resources $\mathcal{C}$
- two main examples $\Sigma_{\mathcal{C}}^{eq}(G)$ and $\Sigma_{\mathcal{C}}^{prop}(G)$
- possibly $\rho : \mathcal{C} \to \mathcal{R}$ functor to another category of resources with respect to which dynamics is measured (can take $\mathcal{C} = \mathcal{R}$ for simplicity)
- Two main ingredients for a categorical form of the Hopfield equations:
  1. threshold dynamics (using $(R, +, \succeq)$ preordered semigroup of category $\mathcal{R}$)
  2. weights matrix replaced by endofunctors

Threshold and endofunctors

- threshold endofunctor $(\cdot)_+ : \hat{\mathcal{C}} \to \hat{\mathcal{C}}$

$$(C)_+ = \begin{cases} C & \text{if } [\rho(C)] \succeq 0 \text{ in } (R, +, \succeq) \\ 0 & \text{otherwise,} \end{cases}$$

- not an endofunctor of $\mathcal{C}$ also $(\cdot)_+$ not monoidal
- $(R, +, \succeq)$ preordered semigroup of category $\mathcal{R}$
- summing functor in terms of $\{\Phi(x)\}_{x \in X \setminus \{\star\}}$ $\Rightarrow$ new summing functor $(\Phi)_+$ by values $(\Phi(x))_+$ in $\mathcal{C}$ for $x \in X$
- $\mathcal{E}(\mathcal{C}) = \mathrm{Func}(\mathcal{C}, \mathcal{C})$ category of monoidal endofunctors of $\mathcal{C}$, morphisms natural transformations
- sum of endofunctors defined pointwise $(F \oplus F')(C) = F(C) \oplus F'(C)$ for all $C \in \mathrm{Obj}(\mathcal{C})$

Categorical Hopfield dynamics for $\Sigma_{\mathcal{C}}^{eq}(G)$ and $\Sigma_{\mathcal{C}}^{prop}(G)$

- $\Sigma^{(2)}_{\mathcal{E}(\mathcal{C})}(E)$ category of *bisumming functors* $T : \mathcal{P}(E) \times \mathcal{P}(E) \to \mathcal{E}(\mathcal{C})$ summing in both arguments
- *coordinates*: $T_{ee'}$ with $T_{A,B} = \oplus_{e\in A, e'\in B} T_{ee'}$
- $\Sigma^{(2)}_{\mathcal{E}(\mathcal{C})}(G)$ equalizer of $s, t : \Sigma^{(2)}_{\mathcal{E}(\mathcal{C})}(E) \rightrightarrows \Sigma^{(2)}_{\mathcal{E}(\mathcal{C})}(V)$
- equation

$$X_e(n+1) = X_e(n) \oplus (\oplus_{e'\in E} T_{ee'}(X_{e'}(n)) \oplus \Theta_e)_+$$

or variant $X_e(n+1) = (\oplus_{e'\in E} T_{ee'}(X_{e'}(n)) \oplus \Theta_e)_+$ (leaking term or not)

- initial condition $\Phi_0 \in \Sigma_{\mathcal{C}}^{eq}(G)$ with $X_e(0) := \Phi_0(e)$
- fixed summing functor $\Psi \in \Sigma_{\mathcal{C}}^{eq}(G)$ with $\Theta_e = \Psi(e)$
- for case of $\Sigma_{\mathcal{C}}^{prop}(G)$ better use equation at vertices (where operad/properad composition happens)

$$X_v(n+1) = X_v(n) \oplus (\oplus_{v'\in V} T_{vv'}(X_{v'}(n)) \oplus \Theta_v)_+$$

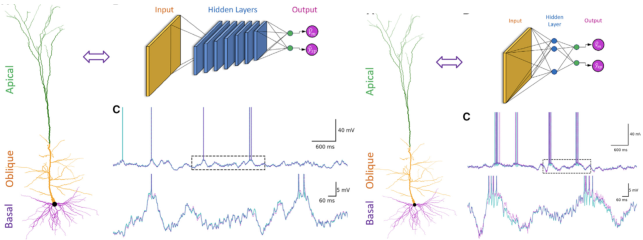
or $X_v(n+1) = (\oplus_{v'\in E} T_{vv'}(X_{v'}(n)) \oplus \Theta_v)_+$

Some properties of the dynamics

- $X_A(n) =: \Phi_n(A)$ sequence of summing functors in $\Sigma_{\mathcal{C}}(G)$
- assignment $\mathcal{T} : \Phi_n \mapsto \Phi_{n+1}$ defined by solution defines endofunctor $\mathcal{T} : \Sigma_{\mathcal{C}}(G) \to \Sigma_{\mathcal{C}}(G)$
- induced discrete topological dynamical system τ on realization $|\mathcal{N}(\Sigma_{\mathcal{C}}(G))| = B\Sigma_{\mathcal{C}}(G)$ (Segal's Gamma space)
- for $\mathcal{C}$ a category of weighted codes, categorical Hopfield dynamics induces usual (finite difference) Hopfield equation on the weights
- Question: general results in categorical setting about existence of solutions and behavior?

Example N.1: modeling neurons (nodes of network) with DNNs

• D. Beniaguev, I. Segev, M. London, *Single cortical neurons as deep artificial neural networks*, Neuron 109 (2021) 2727–2739.



L5 cortical pyramidal neuron and L5PC neuron modeled by DNNs (seven hidden layers and one hidden layer of 128 nodes)

Category of DNNs (deep neural networks)

- interfaces directed acyclic graphs (IDAGs): finite acyclic graphs $G = (V, E)$ with vertex set $V = I \sqcup H \sqcup O$ (inputs, hidden nodes, outputs) source and target maps $s : E \to I \sqcup H$ and $t : E \to H \sqcup O$
- category IDAG: objects IDAGs $G = (I, H, O, E)$ and morphisms $\varphi : G \to G'$ of directed graphs (mapping inputs to inputs, outputs to outputs)
- properad structure: $\mathcal{P}(n, m) = \mathrm{IDAG}(n, m)$ full subcategory with n inputs and m outputs: composition rules $\circ^{i_1,\dots,i_\ell}_{j_1,\dots,j_\ell}$ graft set of outputs $\{i_1, \dots, i_\ell\}$ to set of inputs $\{j_1, \dots, j_\ell\}$
- operad: single output case (case revelant for model) $\mathcal{O}(n) = \mathrm{IDAG}^o(n) = \mathrm{IDAG}(n, 1)$
- monoidal structure on IDAG^o: grafting at the output $G = (I, H, \{o\}, E)$, $G' = (I', H', \{o'\}, E')$

$$G \oplus G' = (I \times \{o'\} \sqcup \{o\} \times I', H \times \{o'\} \sqcup \{o\} \times H', \{(o, o')\}, E \sqcup E')$$

- this is the underlying graph architecture of the DNNs, then need weights

- category WIDAG^o of weighted single-output IDAGs
- objects pairs (G, W) with $G = (I, H, \{o\}, E)$ an IDAG with single output and W (weight matrix) function $W : V \times V \to \mathbb{R}$ with $V = I \sqcup H \sqcup \{o\}$ vertices, satisfying $W(u, v) = 0$ if $\{e \in E \,|\, s(e) = u, , \, t(e) = v\} = \emptyset$ (no looping edges so $W(v, v) = 0$ for all $v \in V$)
- $\varphi \in \mathrm{Mor}_{\mathrm{WIDAG}^o}((G, W), (G', W'))$ morphisms $\varphi : G \to G'$ of directed graphs with $W' = \varphi_*(W)$, pushforward of weight matrix

$$\varphi_*(W)(u', v') = \sum_{\substack{u : \varphi(u) = u' \\ v : \varphi(v) = v'}} W(u, v)$$

- monoidal structure $(G, W) \oplus (G', W') = (G \oplus G', W \oplus W')$ with $(W \oplus W')((u, o'), (v, o')) = W(u, v)$ and $(W \oplus W')((o, u'), (o, v')) = W'(u', v')$
- unit $(0, W_0) = ((\emptyset, \emptyset, \{o\}, \emptyset), W_0(o, o) = 0)$
- operad structure as in IDAG^o with composition on weights $(W \circ_j W')(u, v) = W(u, v)$ when $u, v \in V$, $(W \circ_j W')(u', v') = W'(u', v')$ when $u', v' \in V'$, and if output o grafted to input v'_j weights $(W \circ_j W')(u, v'_j) = W(u, o)$ and $(W \circ_j W')(o, v') = W'(v'_j, v')$, zero otherwise

inhibitory-excitatory balance

- unit $0 := (0, W_0)$ neither initial nor final
- condition $\exists$ morphism $\psi : (0, W_0) \to (G, W)$ implies $W : V \times V \to \mathbb{R}$ is the trivial map (zero weights)
- condition $\exists$ morphism $\varphi : (G, W) \to (0, W_0)$: the map collapsing graph G single output vertex o satisfies $\varphi_*(W)(o, o) = W_0(o, o) = 0$ so

$$\sum_{(u,v)\in V\times V} W(u, v) = 0$$

inhibitory-excitatory balance condition over whole network G

Dynamics case with leak term

- endofunctor $\mathcal{T}$ of the category $\Sigma_{\mathrm{WIDAG}^o}(G)$ that maps $\mathcal{T} : \Phi \mapsto (T(\Phi) \oplus \tilde{\Phi})_+$ with $\tilde{\Phi}$ a summing functor defined by local data $(\tilde{G}_v, \tilde{W}_v)$ and T endofunctor with local data $T_{vv'}$
- stationary solutions $\Phi_v = (G_v, W_v)$ fixed points of endofunctor $\mathcal{T}$ objects with isomorphism

$$\eta_v : (G_v, W_v) \stackrel{\simeq}{\to} (G_v, W_v) \oplus \left(\bigoplus_{v'} T_{vv'}(G_{v'}, W_{v'}) \oplus (\tilde{G}_v, \tilde{W}_v) \right)_+$$

- underlying directed graphs with isomorphism

$$\eta_v : G_v \stackrel{\simeq}{\to} G_v \oplus \mathbb{G}_v$$

for $(\mathbb{G}_v, \mathbb{W}_v) := \oplus_{v' \in V(G)} T_{vv'}(G_{v'}(n), W_{v'}(n)) \oplus (\tilde{G}_v, \tilde{W}_v)$
- by $\oplus$ in $\mathrm{WIDAG}^o \Rightarrow \mathbb{G}_V = \{o'\}$ and $(\mathbb{G}_v, \mathbb{W}_v) = (0, W_0)$
- regardless of endofunctors $T_{vv'}$, solutions converge to a fixed point (in finite time) iff inhibitory-excitatory balance on $(\mathbb{G}_v, \mathbb{W}_v)$ violated at some step; otherwise $\# V \to \infty$ no limit

Dynamics case without leak term (more interesting)

- evolution equation

$$(G_v(n+1), W_v(n+1)) = \left(\oplus_{v' \in V(G)} T_{vv'}(G_{v'}(n), W_{v'}(n)) \oplus (\tilde{G}_v, \tilde{W}_v) \right)_+$$

- stationary solutions pairs (G_v, W_v) where the right-hand-side satisfies balance condition and

$$(G_v, W_v) = \oplus_{v' \in V(G)} T_{vv'}(G_{v'}, W_{v'}) \oplus (\tilde{G}_v, \tilde{W}_v)$$

- if balanced condition is violated at some step solution stabilizes at $(0, W_0)$
- now behavior depends on specific choice of endofunctors $T_{vv'}$
- example $T_{vv'} = T_v \delta_{v,v'}$ with $T_v(G_v, W_v) = (G_v, T_v(W_v))$
- functor T_v as implementing a backpropagation mechanism on the DNN machine (G_v, W_v)

- example: updating weights by gradient descent

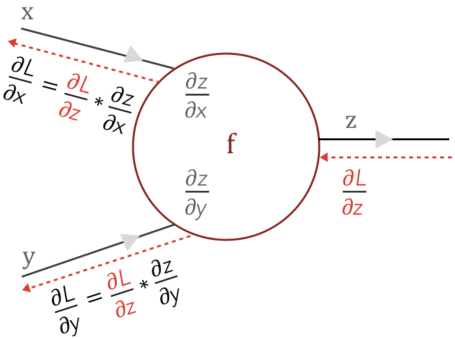
$$T_v(W_v) := W_v - \epsilon \nabla_{W_v} F_v$$

  some cost function F_v for the specific DNN (G_v, W_v) and a scale $\epsilon > 0$ (learning rate)

- then fixed point $T(W_v) = W_v$ is critical point of the cost function (minimum, depending on F_v shape)
- consistency condition $T(W') = \varphi_* T(W)$ holds since

$$(\varphi_* T(W))(a,b) = \sum_{\substack{u:\varphi(u)=a \\ v:\varphi(v)=b}} T(W)(u,v) = \sum_{u,v} W(u,v) - \epsilon \sum_{u,v} \nabla_{W(u,v)} F$$

gives $W'(a,b) - \epsilon \nabla_{W'(a,b)} F = T(W')(a,b)$



- Note: oversimplistic toy model as it assumes updating mechanism of weights of each DNN machine (G_v, W_v) uncoupled to the other machines at other nodes

- these just consistency checks: get usual Hopfield equations, get DNN backpropagation

- other related cases: more general automata (transition systems, models of concurrent and distributed computing) in our paper

Example N.2: a game of invisible varieties

inspired by some ideas discussed in our recent paper:

- Yuri I. Manin, Matilde Marcolli, *Homotopy Spectra and Diophantine Equations*, arXiv:2101.00197

- invisible varieties: number field K, varieties X, Y ovr K, say X is Y-invisible if $\mathrm{Hom}_{Var_K}(Y, X) = \emptyset$
- $\mathrm{Spec}(K)$-invisible: K-varieties with no K-points, $X(K) = \emptyset$
- local-to-global principle and Brauer-Manin obstruction: $X(K) \subseteq X(\mathbb{A}_K)^{\mathrm{Br}} \subseteq X(\mathbb{A}_K)$, ring of adèles $\mathbb{A}_K$ of K
- more general obstruction data (Corwin–Schlank 2020): very strong approximation $X(K) = X(\mathbb{A}_{K,S})^{\omega}$, conditions under which it holds for a finite affine open cover of X

- Spec(K)-invisibility gives an example of the threshold dynamics:

$$(X)_+ := \begin{cases} X & \mathrm{Hom}_{Var_K}(\mathrm{Spec}(K), X) \neq \emptyset \\ \mathrm{Spec}(K) & \mathrm{Hom}_{Var_K}(\mathrm{Spec}(K), X) = \emptyset \, . \end{cases}$$

- Var_K with monoidal structure given by fibered product $X \oplus Y = X \times_{\mathrm{Spec}(K)} Y$
- building varieties along graphs: G finite directed graph, $\Phi(e) = X_e$ assignment of K-varieties $X_e \in Var_K$ at edges of G
- "Kirchhoff conservation law" at vertices:

$$\oplus_{s(e)=v} X_e = \oplus_{t(e')=v} X_{e'}$$

Note this balance condition has "cancellation" subtleties like Zariski cancellation (for $char(K) > 0$ and $n \geq 3$ $\exists$ X with $X \times \mathbb{A}^1 \simeq \mathbb{A}^{n+1}$ but $X \not\simeq \mathbb{A}^n$, N.Gupta, 2014)

Where to get endofunctors?

Example that does not stabilize the dynamics:

- Denef-Loeser arc spaces $\mathcal{L}_m(X)$ algebraic variety whose K-points are $\mathcal{L}_m(X)(K) = X(K[u]/u^{m+1})$
- Intuition: points of $\mathcal{L}_1(X)$ are points x of X and tangent vectors v in $T_x(X)$
- for each pair $e, e' \in E(G)$ assign a choice of endofunctor $T_{ee'} = \mathcal{L}_{m_{ee'}}$ (can take all $\mathcal{L}_1$ for example)
- Dynamics (example without leak and forcing terms)

$$X_e(n+1) = (\oplus_{e'} T_{ee'}(X_{e'}(n)))_+$$

if initialized $X_e(0) = X$ with $X(K) \neq \emptyset$ dynamics just keeps building products of iterated arc spaces

Example that *does* stabilize the dynamics:

- use relation between torsors and K-points

Torsors (see Poonen's Rational Points book)

- $\mathbb{G}$-torsor over K: a K-variety X with a right action of an algebraic group $\mathbb{G}$ with $X_{\bar{K}}$ with $\mathbb{G}_{\bar{K}}$ action isomorphic to trivial torsor
- these torsors classified by $H^1(K, \mathbb{G})$
- $\mathbb{G}$-torsor over base scheme S: S-scheme X with right action $X \times_S \mathbb{G} \to X$ of a group scheme $\mathbb{G} \to S$ such that (after a base change $S' \to S$) $X_{S'}$ isom to $\mathbb{G}_{S'}$ with action on itself by translations, and

$$X \times_S \mathbb{G} \to X \times_S X \quad (x, g) \mapsto (x, xg) \quad \text{isomorphism}$$

- classification (via torsor sheaves and descent) by appropriate cohomology $H^1(S, \mathbb{G})$

Torsors and K-points (see Skorobogatov's book)

- $X \in \mathrm{Var}_K$ and $\mathbb{G}$ smooth alg group over K; G-torsor $f : Z \to X$ with class $\zeta \in H^1(X, \mathbb{G})$
- rational points $x \in X(K)$ give $\mathbb{G}$-torsor over K taking fiber $Z_x \to \{x\}$, class $\zeta(x) \in H^1(K, \mathbb{G})$

$$x : \mathrm{Spec}(K) \to X, \quad x^* : H^1(X, \mathbb{G}) \to H^1(K, \mathbb{G}), \quad \zeta \mapsto \zeta(x)$$

  evaluation $X(K) \to H^1(K, \mathbb{G})$, $x \mapsto \zeta(x)$

- Partition of K-points

$$X(K) = \bigsqcup_{\tau \in H^1(K,\mathbb{G})} \{x \in X(K) \,|\, \zeta(x) = \tau\} = \bigsqcup_{\tau \in H^1(K,\mathbb{G})} f^\tau(Z^\tau(K))$$

  $f^\tau : Z^\tau \to X$ twisted torsor: $Z \times_K^G T$ with $T \to \mathrm{Spec}(K)$ with class τ

- a $\mathbb{G}$-torsor over X determines a partition of the K-points of X

back to our game of invisible varieties

- network G with single output at each node
- Fix a choice of a group scheme $\mathbb{G}$ over a base S
- initialize the dynamics by assigning to egdes objects X_e in Var_S given by G-torsors $h_e : X_e \to S$ over S with balance condition
$$X_e = \prod_{e' \,:\, t(e')=s(e)} X_{e'}$$
with corresponding classes $\zeta_e = \sum_{e'} \zeta_{e'}$
- transform X_e by taking $T_{ee'}$ supported on e' with $t(e') = s(e)$ (just write $T_{e'}$ as unique output e)
$$T_e(X_e) = Z^{\tau_e}(X_e)$$
with $f^{\tau_e} : Z^{\tau_e} \to X_e$ is twisted $\mathbb{G}$-torsor, with τ_e still satisfying same balance condition
- equation: $X_e(n+1) = (\oplus_{e'} T_{ee'}(X_{e'}(n)))_+$ with threshold given by visibility

- above threshold (existence of K-points) have dynamics

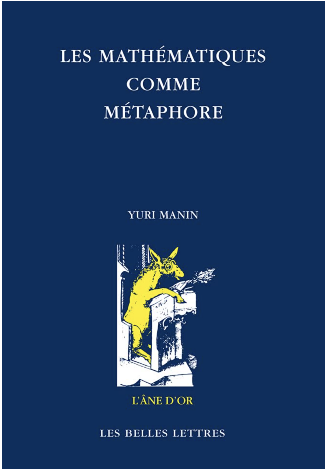
$$X_e(n+1) = \oplus_{e'} T_{ee'}(X_{e'}(n) = \prod_{e' : t(e')=s(e)} Z^{\tau_{e'}}(X_{e'}(n)) = Z^{\tau_e}(X_e(n))$$

- if $Z^{\tau_e}(X_e(n))$ has no more K-points dynamics collapses to a point $X_e(n+1) = \mathrm{Spec}(K)$ contributing trivially to product at following vertex
- initialize with X_e's with finite sets of K-points
- function $\sum_e \mathcal{L}(X_e) = \sum_e \# h_e(X_e(K))$ is a Lyapunov function of the dynamics $\sum_e \mathcal{L}(X_e(n+1)) \leq \sum_e \mathcal{L}(X_e(n))$ because of partition of K-points $X_e(K)$ by $f^{\tau_e}(Z^{\tau_e}(X_e)(K))$
- fixed points occur when dynamics collapses to a point (when the $Z^{\tau_e}(X_e)$ are K-invisible), as otherwise fixed point condition would require $X \simeq Z^{\tau}(X)$

Some questions

- can more interesting examples of categorical Hopfield dynamics be constructed in categories of varieties and schemes?
- what about other categories of resources?
- what about other models of categorical dynamics (related to Hopfield network dynamics)?
- about this last question: searching for equilibria of Hopfield networks is related to searching for Pareto optimization solutions
- Next question: is there a categorical Pareto optimization?

A background story of my ongoing Pareto musings
Yuri sent me last November a copy of his new book
(by the way, read it: it's great!)

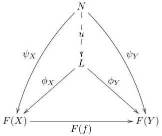


> La contradiction interne qui se trouve au cœur de l'idée de marché (y compris de la scandaleuse expression « libre marché des idées ») est la suivante : nous projetons le monde multidimensionnel des degrés incomparables et incompatibles de liberté, sur le monde à une dimension des prix monétaires. Par principe, on ne peut pas le rendre compatible avec des relations d'ordre, même basiques, sur ces axes, et encore moins le rendre compatible avec des valeurs de différentes sortes, des valeurs qui n'ont pas d'existence ou qui sont incomparables.

- first thought: I agree completely with this! ...
- second thought: what if objectives and values incompatible and irreducible to a single real variable would be imagined instead like *objects in categories*?

Note: categories and universal properties are naturally expressing constraints and optimization

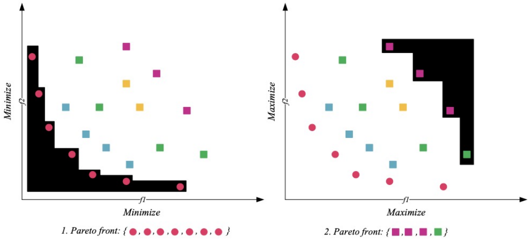
- Example: limits and colimits in categories
    - diagram $F : \mathcal{J} \to \mathcal{C}$ and cone N, limit is "optimal cone" (dual version for colimits)



    - special cases of limits and colimits: equalizers, coequalizers
- Example: thin categories $(S, \leq)$ set of objects S and one morphism $s \to s'$ when $s \leq s'$
    - diagram in $(S, \leq)$ is selection of a subet $A \subset S$
    - limits and colimits greatest lower bounds and least upper bounds for subsets $A \subseteq S$
- functors compatible with limits and colimits describe constrained optimization

Pareto optimization problem (multi-objective programming)

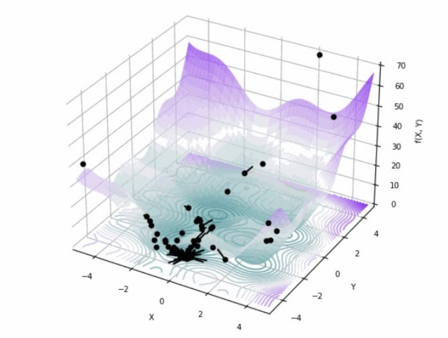
- several simultaneous objective functions, valued in cones inside a real Euclidean space $\mathbb{R}^n$ need to be optimized
- subject to constraints: cannot individually maximize each function
- Pareto optimal solution (non-unique): none of the objective functions can be improved without worsening some of the others
- Pareto frontier: set of Pareto optimal solutions
- possible solution S_1 Pareto dominates S_2 if all objective valuations f_i satisfy $f_i(S_1) \geq f_i(S_2)$ and for at least one of strict inequality
- Pareto optimal solutions: not Pareto dominated by any other



Swarm Intelligence and the Pareto frontier

- swarm intelligence approach: find solutions as close as possible to the Pareto frontier and as diverse as possible (map out different regions of the Pareto frontier)
- virtual swarm of "particles" that moves according to some dynamical rules across the landscape of possibilities (the configuration space)
- structure of the swarm intelligence algorithm:
  1. initializes the swarm by random distribution of positions and momenta with uniform measure over configuration space
  2. each individual particle in the swarm can memorize its best solution up to the present time
  3. each particle in the swarm tends to search near its best position obtained so far
  4. each particle can see the positions of the other particles of the swarm at the same time and evaluate the best position achieved by the swarm at that moment;
  5. each individual particle tends to move towards the best position achieved within the swarm at that time

Swarm Intelligence dynamics



swarm of particle flowing towards energy minimum
(image by Alex Thevenot)

Update rules of the swarm dynamics (discrete time)

- Update velocities:

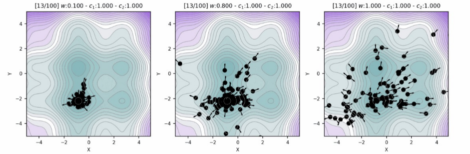
$$V_i(n+1) = \lambda_3\, V_i(n) + \lambda_1\, G_1(X_i(n) - X_{i,best}(n)) + \lambda_2\, G_2(X_i(n) - X_{best}(n))$$

G_i Gaussians, λ_i tunable parameters; $X_{i,best}(n)$ best position among $\{X_i(0), \ldots, X_i(n)\}$ and $X_{best}(n)$ best position among $\{X_1(n), \ldots, X_N(n)\}$ (swarm of N particles)

- Update positions:

$$X_i(n+1) = X_i(n) + V_i(n+1)$$

Known results: general conditions under which for large swarm size N and many iterations n positions of the swarm draw out the Pareto frontier



role of parameters in swarm dynamics (image by A.Thevenot)

Next step: instead of being optimization of functions in cones in Euclidean spaces imagine the *desirable objectives* are just *objects in categories*, no direct convertibility to real valued evaluations

1. need a notion of objective valuations in the setting of categories of resources
2. need the appropriate notion of Pareto frontier
3. need an analog of the swarm intelligence dynamics
4. categories need to be enriched with probabilistic data since the swarm intelligence approach is intrinsically probabilistic

• as before want to think of this as a problem of optimization of *resources assigned to a network*

Objective valuation functors

- G a network, $\Sigma_{\mathcal{C}}(G)$ a category of summing functors (say $\Sigma^{eq}_{\mathcal{C}}(G)$ or $\Sigma^{prop}_{\mathcal{C}}(G)$)
- a finite family $\mathcal{V}_\alpha$ of categories that describes possible objectives for optimization with functors $F_\alpha : \Sigma_{\mathcal{C}}(G) \to \mathcal{V}_\alpha$ (valuations) and $X_\alpha \in \mathrm{Obj}(\mathcal{V}_\alpha)$ (goals)
- $(F, X) = (F_\alpha, X_\alpha)_{\alpha \in \mathcal{I}}$ a choice of valuation functors and target objects
- Note: valuation functors may factor through the target category of resources $\mathcal{C}$ (not necessarily)
- a summing functor $\Phi \in \Sigma_{\mathcal{C}}(G)$ is F-minorized by another $\Psi \in \Sigma_{\mathcal{C}}(G)$ if

  $$\mathrm{Hom}_{\mathcal{V}_\alpha}(F_\alpha(\Phi), F_\alpha(\Psi)) \neq \emptyset \quad \forall \alpha$$

  strictly F-minorized if also $\exists \alpha$ with $F_\alpha(\Phi)$ and $F_\alpha(\Psi)$ not isomorphic in $\mathcal{V}_\alpha$ (*majorization* similarly defined)
- this means $F_\alpha(\Psi)$ is obtainable from $F_\alpha(\Phi)$ by an admissible "conversion of resources" in $\mathcal{V}_\alpha$

Pareto frontier

- a summing functor $\Phi \in \Sigma_{\mathcal{C}}(G)$ is (F,X)-minorized by another $\Psi \in \Sigma_{\mathcal{C}}(G)$ if

$$\mathrm{Hom}_{\mathcal{V}_\alpha}(F_\alpha(\Phi), F_\alpha(\Psi)) \neq \emptyset \quad \forall \alpha$$

$$\mathrm{Hom}_{\mathcal{V}_\alpha}(F_\alpha(\Phi), X_\alpha) \neq \emptyset \quad \forall \alpha$$

$$\mathrm{Hom}_{\mathcal{V}_\alpha}(F_\alpha(\Psi), X_\alpha) \neq \emptyset \quad \forall \alpha$$

- that is, $F_\alpha(\Psi)$ is obtainable from $F_\alpha(\Phi)$, while both are good enough to achieve the goals X_α
- $\Phi \in \Sigma_{\mathcal{C}}(G)$ is on the (F,X)-Pareto upper frontier if

$$\mathrm{Hom}_{\mathcal{V}_\alpha}(F_\alpha(\Phi), X_\alpha) \neq \emptyset \quad \forall \alpha$$

  but there is no $\Psi \in \Sigma_{\mathcal{C}}(G)$ not isomorphic to Φ that is a strict (F,X)-minorization of Φ
- "upper frontier" as "above the goals": case of lower frontier is analogous with $\mathrm{Hom}_{\mathcal{V}_\alpha}(X_\alpha, F_\alpha(\Phi)) \neq \emptyset$ conditions and majorizations (visibility conditions for varieties, for instance)

- valuation functors $F_\alpha : \Sigma_{\mathcal{C}}(G) \to \mathcal{V}_\alpha$ are in general *not* fully faithful (not faithful, also not full)
- $\mathcal{D}_\alpha$ full subcategory of $\mathcal{V}_\alpha$ on $F_\alpha(\Phi)$ objects
- can think of (F, X)-minorizations as a subcategory $\mathcal{C}(\mathcal{D}_\alpha, X_\alpha)$ of co-cones with tip X_α (and $\mathcal{D}_\alpha$ diagrams)

$$\begin{array}{ccc} F_\alpha(\Phi) & \longrightarrow & F_\alpha(\Psi) \\ & \searrow \quad \swarrow & \\ & X_\alpha & \end{array}$$

- if $\mathcal{V}_\alpha$ has colimits $L_\alpha = \mathrm{colim}\, \mathcal{C}(\mathcal{D}_\alpha, X_\alpha)$ and valuations F_α are *essentially surjective* then Ψ on Pareto frontier iff $F_\alpha(\Psi) \simeq L_\alpha$ for some α
- so can get Pareto frontier as subcategory
- Question: is there a swarm dynamics associated to it?

Preliminaries: probabilistic categories

- category $\mathcal{C} \Rightarrow$ probabilistic version $\mathcal{PC}$: like a wreath product $\mathcal{FP} \wr \mathcal{C}$ of the category $\mathcal{C}$ with the category $\mathcal{FP}$ of finite probabilities
- $\mathcal{FP}$ finite probabilities: objects (X, P) of a finite set X with a probability measure P, morphisms $S \in \mathrm{Hom}_{\mathcal{FP}}((X, P), (Y, Q))$ given by stochastic $(\#Y \times \#X)$-matrices S with
  - $S_{yx} \geq 0$, for all $x \in X$, $y \in Y$;
  - $\sum_{y \in Y} S_{yx} = 1$ for all $x \in X$;
  - the probability measures are related by $Q = S\, P$.

- objects of $\mathcal{PC}$: formal finite convex combinations

$$\Lambda C = \sum_i \lambda_i C_i$$

  with $\Lambda = (\lambda_i)$ finite probability and $C_i \in \mathrm{Obj}(\mathcal{C})$
- morphisms $\mathrm{Hom}_{\mathcal{PC}}(\Lambda C, \Lambda' C')$ are pairs $(S, f) : \Lambda C \to \Lambda' C'$ with
  - S stochastic matrix with $S\Lambda = \Lambda'$;
  - $f = \{f_{ab,r}\}$ finite collection of morphisms $f_{ab,r} : C_b \to C'_a$ with assigned probabilities μ_r^{ab};
  - probabilities satisfy $\sum_r \mu_r^{ab} = S_{ab}$.
- interpret morphism (S, f) as mapping C_b to C'_a by randomly choosing a morphism from the set $\{f_{ab,r}\}$, with probability μ_r^{ab} of choosing $f_{ab,r}$

Observations

- objects $\Lambda X = \sum_i \lambda_{i=1}^n X_i$ and $\Lambda' X' = \sum_{j=1}^m \lambda'_j X'_j$ in a probabilistic category $\mathcal{PC}$ are isomorphic if and only if $n = m$ with $X_i \simeq X'_{\sigma(i)}$ (isomorphic in $\mathcal{C}$) for some permutation σ and $\lambda_i = \lambda'_{\sigma(i)}$ (stochastic matrices with stochastic inverse are permutations)
- sequence $\Lambda^{(n)} X^{(n)} = \sum_{k=1}^{r_n} \lambda_k^{(n)} X_k^{(n)}$ of objects in $\mathcal{PC}$ converges to $\Lambda X = \sum_{k=1}^r \lambda_r X_r$ if for all $n \geq n_0$, have $r_n \geq r$ and a subset $\mathcal{I}_n \subset \{1, \dots, r_n\}$ of size $\#\mathcal{I}_n = r$

$$X^{(n)}_{\sigma_n(k)} \simeq X_k \quad \text{for some permutation } \sigma_n \in S_r$$

$$\sigma_n(\Lambda_n|_{\mathcal{I}_n}) \to \Lambda \quad \text{in } \Delta_r \,, \quad \text{and } \Lambda_n|_{\mathcal{I}_n^c} \to 0$$

Other probabilistic conditions

- category $\mathcal{C}$ or resources is a *small* category with a probability distribution $\mathbb{P}$ on the set $\mathrm{Obj}(\mathcal{C})$ (can represent relative abundance or scarcity of resources)
- identification of summing functors with objects in $\hat{\mathcal{C}}^n \Rightarrow$ induced probability $\mathbb{P}$ on $\mathrm{Obj}(\Sigma_{\mathcal{C}}(G))$
- (X, F)-admissible objects

$$\mathrm{Obj}^{adm}_{(X,F)}(\Sigma_{\mathcal{C}}(G)) := \{\Phi \in \Sigma_{\mathcal{C}}(G) \,|\, \mathrm{Hom}_{\mathcal{V}_\alpha}(F_\alpha(\Phi), X_\alpha) \neq \emptyset \, \forall \alpha\}$$

- condition that choice of goals and valuations not incompatible with resource availability:

$$\mathbb{P}(\mathrm{Obj}^{adm}_{(X,F)}(\Sigma_{\mathcal{C}}(G))) > 0$$

- given $\Phi \in \Sigma_{\mathcal{C}}(G)$ set of all (F, X)-minorizations

$$\mathcal{M}^{adm}_{(X,F)}(\Phi) = \{\Psi \in \Sigma_{\mathcal{C}}(G) \,|\, \Psi \text{ is an } (F, X)\text{-minorization of } \Phi\}$$

$$= \{\Psi \not\simeq \Phi \,|\, \mathrm{Hom}_{\mathcal{V}_\alpha}(F_\alpha(\Phi), F_\alpha(\Psi)) \neq \emptyset \, \forall \alpha\} \cap \mathrm{Obj}^{adm}_{(X,F)}(\Sigma_{\mathcal{C}}(G))$$

- $\lambda(\Phi) = \mathbb{P}(\mathcal{M}^{adm}_{(X,F)}(\Phi))$
- if Φ on the Pareto frontier: $\mathcal{M}^{adm}_{(X,F)}(\Phi) = \emptyset$
- if $\mathbb{P}$ has no non-empty sets of measure zero $\lambda(\Phi) = 0$ iff on Pareto frontier

Step 1: single particle

- initialize: X_0 is drawn from $\mathrm{Obj}^{adm}_{(X,F)}(\Sigma_{\mathcal{C}}(G))$ uniformly at random with respect to $\mathbb{P}$
- equivalently: drawn from all $\mathrm{Obj}(\Sigma_{\mathcal{C}}(G))$ and rejected if not admissible (with new draw)
- the dynamics proceeds by making new random steps and comparing them: "velocities" are (probabilistic) jumps to a new position,

$$(\Lambda X)_{n+1} = T_{n+1}\,(\Lambda X)_n$$

- first step: new draw of point X_1 with probability $\lambda_0 = \mathbb{P}(\mathcal{M}^{adm}_{(X,F)}(X_0))$ improves on X_0, if not keep X_0:

$$(\Lambda X)_1 = (1-\lambda_0)X_0 + \lambda_0 X_1$$

- second step: second draw X_2 and $\lambda_1 = \mathbb{P}(\mathcal{M}^{adm}_{(X,F)}(X_1))$

$$(\Lambda X)_2 = (1-\lambda_0)((1-\lambda_0)X_0+\lambda_0 X_2)+\lambda_0((1-\lambda_1)X_1+\lambda_1 X_2)$$
$$= (1-\lambda_0)^2 X_0 + \lambda_0(1-\lambda_1)X_1 + \lambda_0(1-(\lambda_0-\lambda_1))X_2$$

- Recursion for the single particle coefficients $(\Lambda X)_n = \sum_{k=0}^{n} c_n^k X_k$

$$\begin{cases} c_n^k = c_k^k (1-\lambda_k)^{n-k} & 0 \leq k \leq n-1 \\ c_n^n = \sum_{k=0}^{n-1} \lambda_k (1-\lambda_k)^{n-1-k} c_k^k \end{cases}$$

- probability: satisfy $\sum_k c_n^k = 1$
- coefficients c_n^k are polynomials in the λ_i so depend on $X_0, \ldots, X_n$
- $c_n^k =$ the probability of having X_k as the "best position" of the particle during the first n steps
- Problem: this probability distribution is very spread out, it peaks somewhere (not always at the end) but can be very non-concentrated (inefficient search)

Step N.2: swarm

- Have N particles behaving as above, initialized as $X_0^{(i)}$ and with $(\Lambda^{(i)} X^{(i)})_n = \sum_{k=0}^n c_n^k(i)\, X_k^{(i)}$ with $X_k^{(i)}$ the k-th random draw for the i-th particle, best position with probability $c_n^k(i)$
- At each step, the N particles can compare positions and select also a "best position of the swarm" at time n
- $(\Lambda_s X_s)_n = \sum_{\ell=1}^N \pi_n^\ell\, X_n^{(\ell)}$ with coefficients the probability π_n^ℓ of $X_n^{(\ell)}$ being the best position of the swarm at time n
- new update rule for particles in the swarm

$$(\Lambda^{(i)} X^{(i)})_{n+1} = \mu\, T_{n+1}^{(i)}\, (\Lambda^{(i)} X^{(i)})_n \,+\, (1-\mu)\, (\Lambda_s X_s)_n$$

here $T_{n+1}^{(i)}$ differs as new draw $X_{n+1}^{(i)}$ compared not only to the previous $X_k^{(i)}$ but also to all the previous $X_k^{(\ell)}$ and probabilities $c_n^k(i)$ now also functions of μ parameter and of the π_k^ℓ

- swarm approaches the Pareto frontier (best particle does)
- this happens when $\inf_{n,i} \lambda_n^{(i)} = 0$
- if nonzero lower bound $\lambda_{inf} > 0$ then would have uniform nonzero lower bound for all the $c_n^k(i)$, incompatible with maintaining normalization $\sum_{k=1}^n c_n^k(i) = 1$ for $n \to \infty$
- but even with the comparison terms between swarm particles very inefficient

Optimization questions:

- can limit/colimit constructions in categories be realized by algorithms in corresponding probabilistic categories?
- swarm algorithms concentrating probability around best particle?

More general kind of question: interesting examples of threshold dynamics in categories? what is it good for?

preview of work in progress...

Mathematics is like a voyage of exploration: sometimes it leads to unexpected landscapes sometimes to a mirage. What is most important is to build somewhere a solid home to which one knows one can return to watch the beauty of the stars in the night sky: in the past 20 years I've always known where to find that place

HAPPY 85th BIRTHDAY YURI !!

and Happy March 8 to everybody!