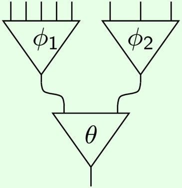
An algebraic view of entropy

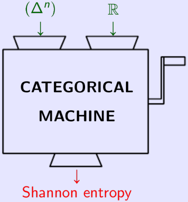
Tom Leinster

University of Edinburgh

Three talks

Sunday: Operads





Monday: An algebraic view of entropy

Thursday: Entropy modulo a prime

Trajectory of these three talks

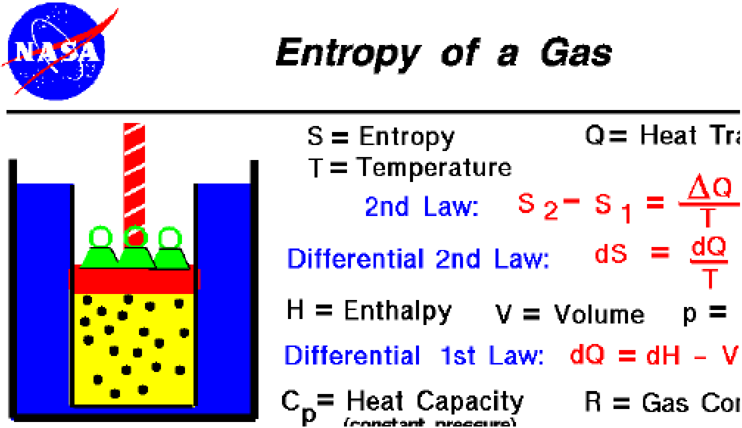
SUNDAY MONDAY

I will explain how operads—a cousin of algebraic theories—lead to the notion of entropy (which might seem to belong to other branches of science).

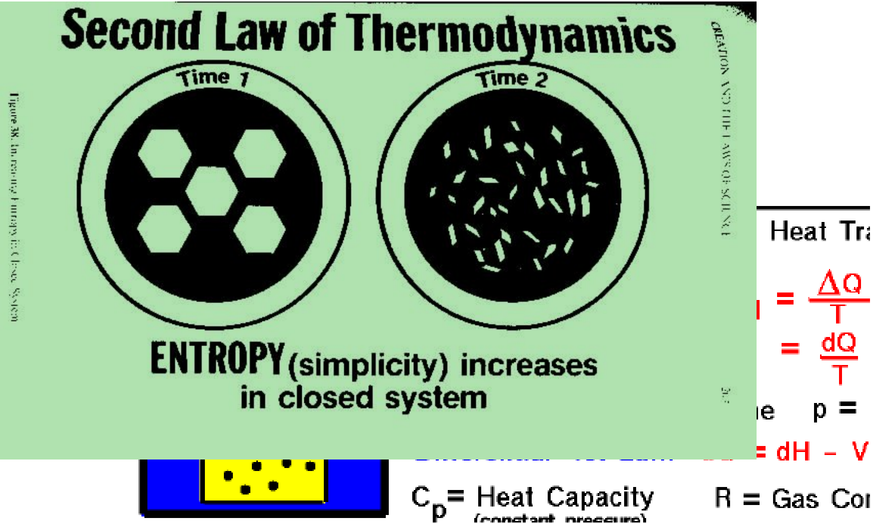
Then I will show how this story leads to a mysterious construction in number theory.

THURSDAY

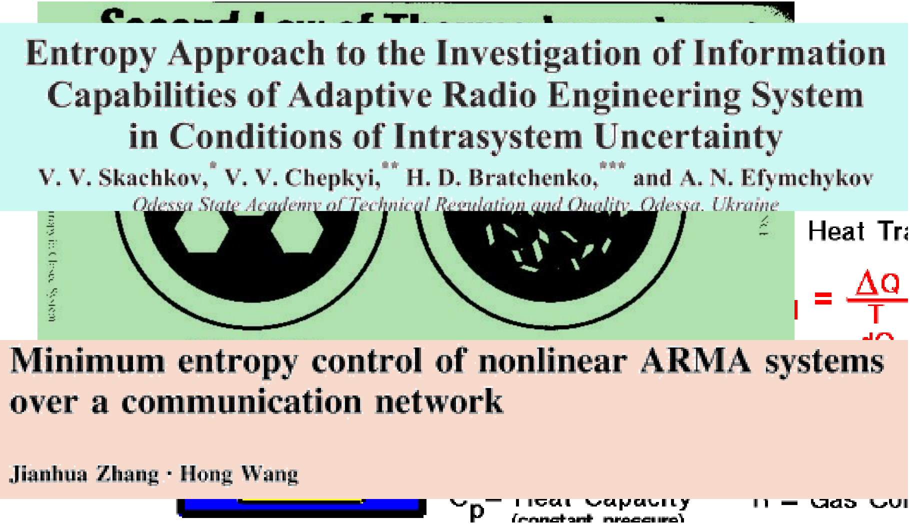
Entropy in the sciences



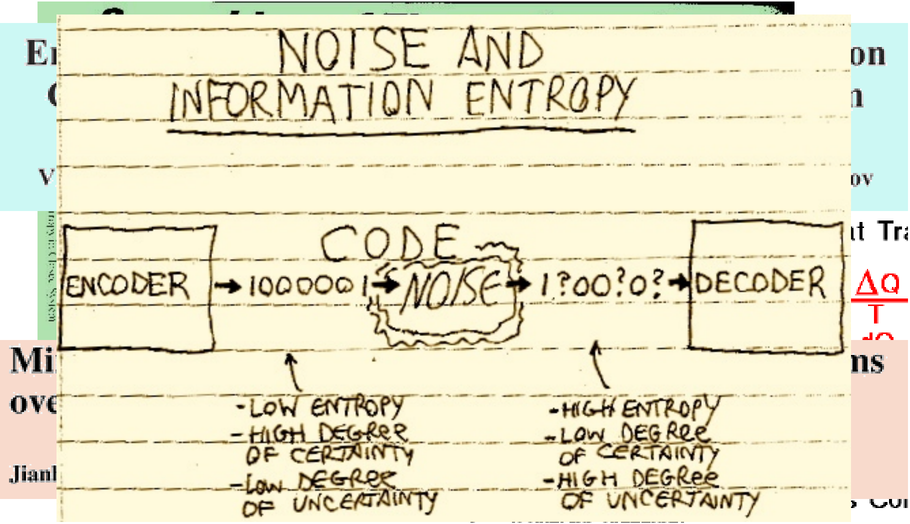
Entropy in the sciences



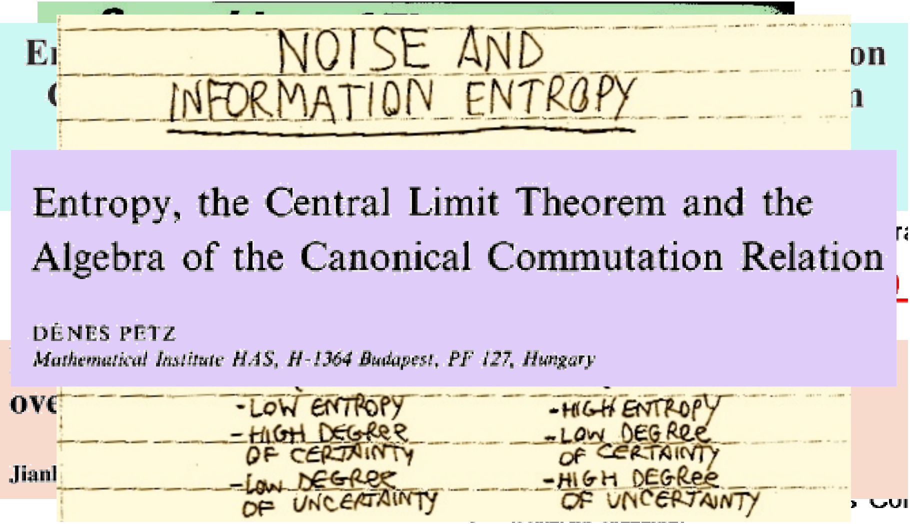
Entropy in the sciences



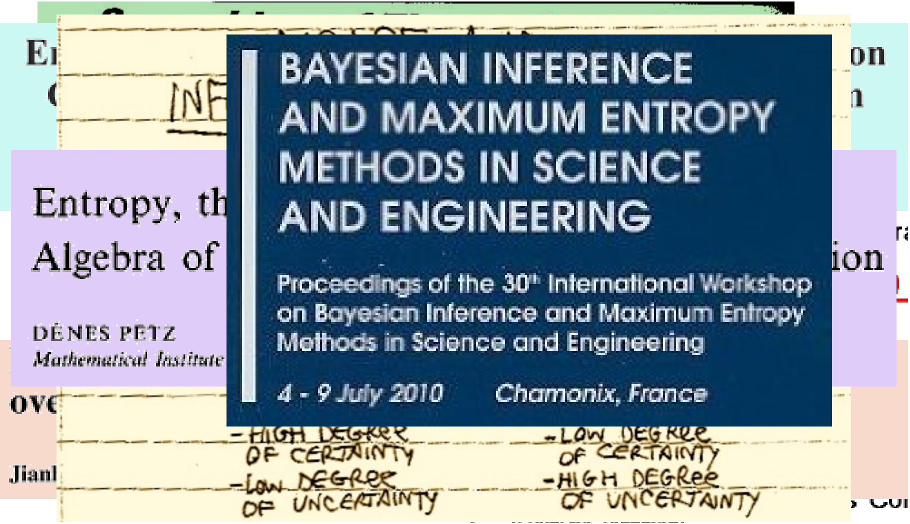
Entropy in the sciences



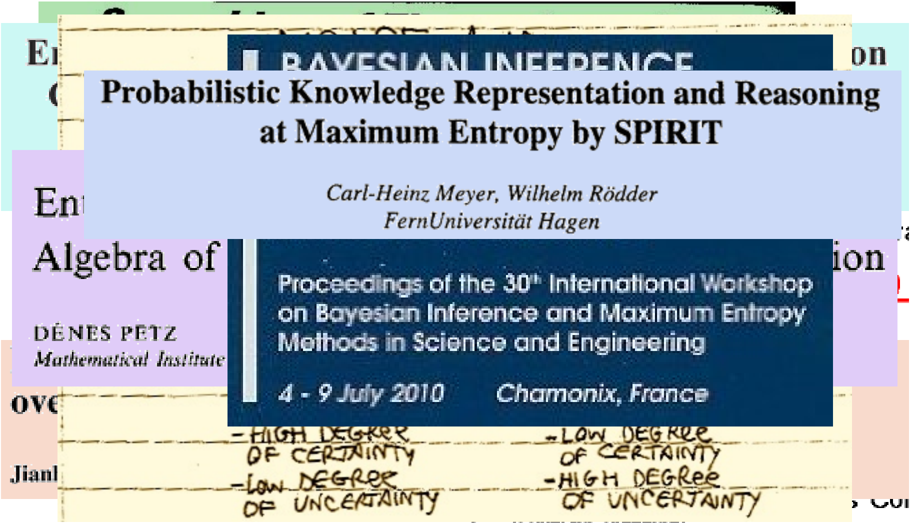
Entropy in the sciences



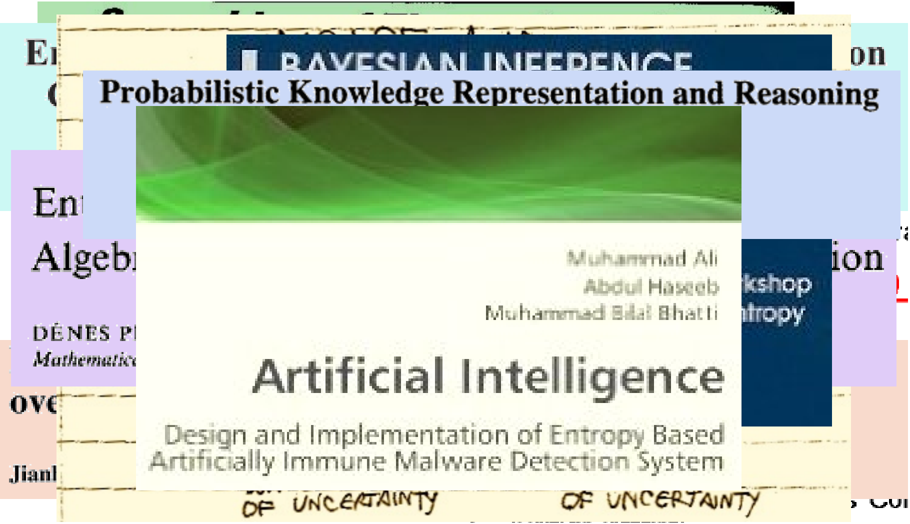
Entropy in the sciences



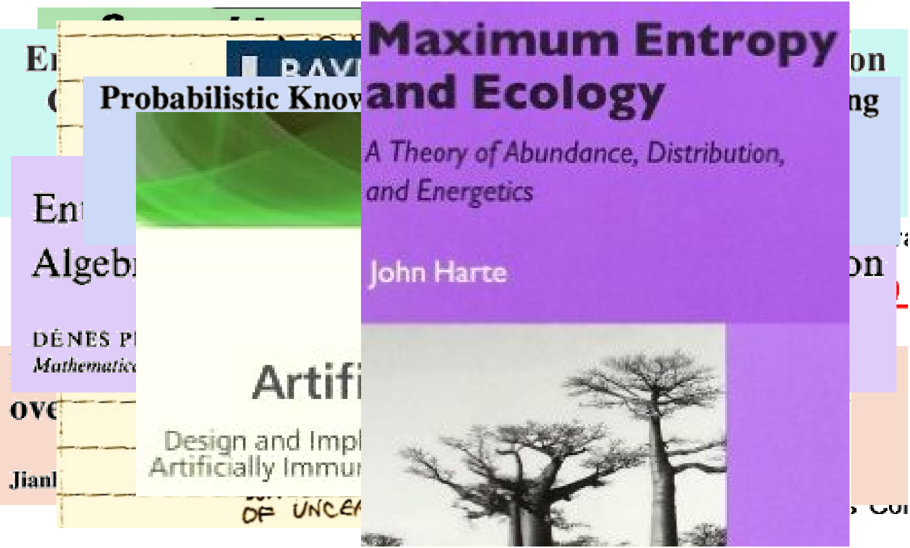
Entropy in the sciences



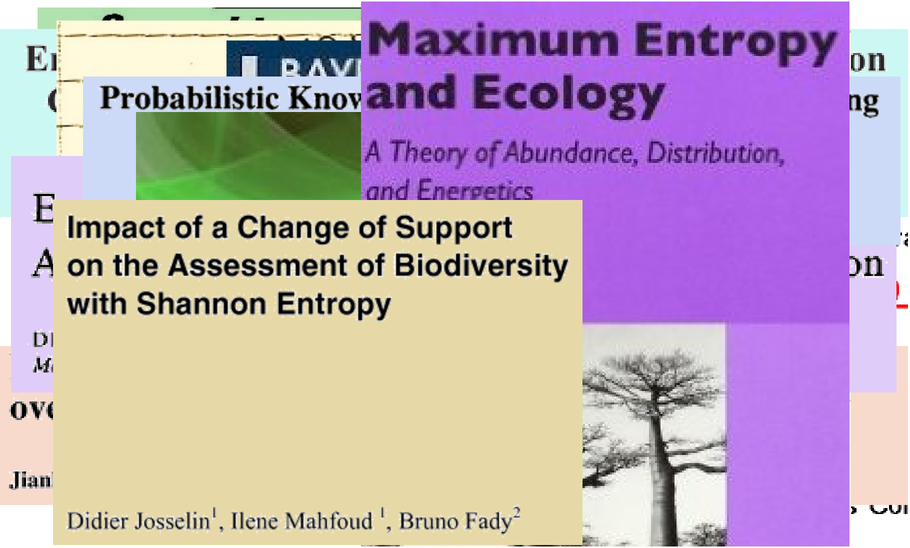
Entropy in the sciences



Entropy in the sciences



Entropy in the sciences



Entropy in the sciences

Roman F. Nalewajski

Reduced communication channels of molecular fragments and their entropy/information bond indices

An entropic characterization of protein interaction networks and cellular robustness

Thomas Manke, Lloyd Demetrius, Martin Vingron

Resilience and entropy as indices of robustness of water distribution networks

R. Greco, A. Di Nardo and G. Santonastaso

Entropy in the sciences

7 Algorithmic entropy and Kolmogorov complexity

Entropy and Quantum Kolmogorov Complexity: A Quantum Brudno's Theorem

Fabio Benatti[1], Tyll Krüger[2,3], Markus Müller[2], Rainer Siegmund-Schultze[2], Arleta Szkoła[2]

Entropy in the sciences

7 Algorithmic entropy and K

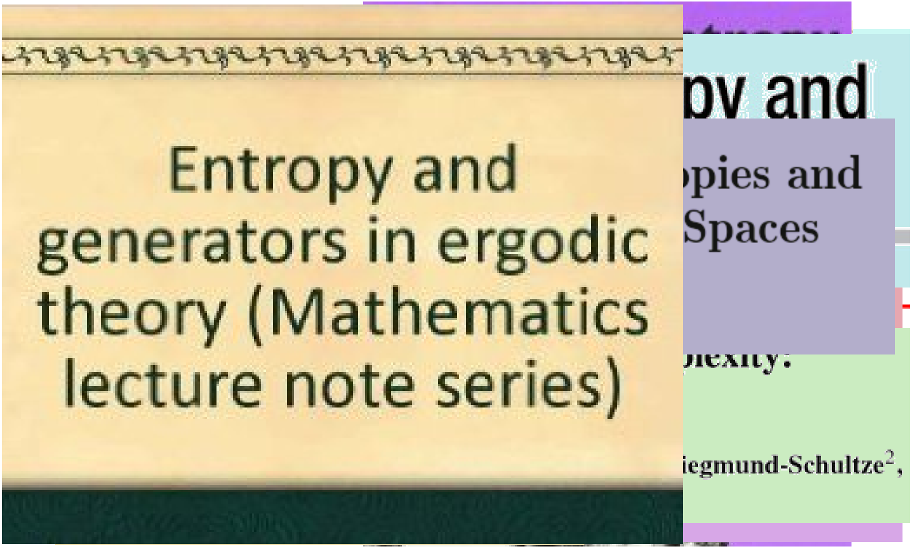
Rényi Relative Entropies and Noncommutative L_p-Spaces

Anna Jenčová

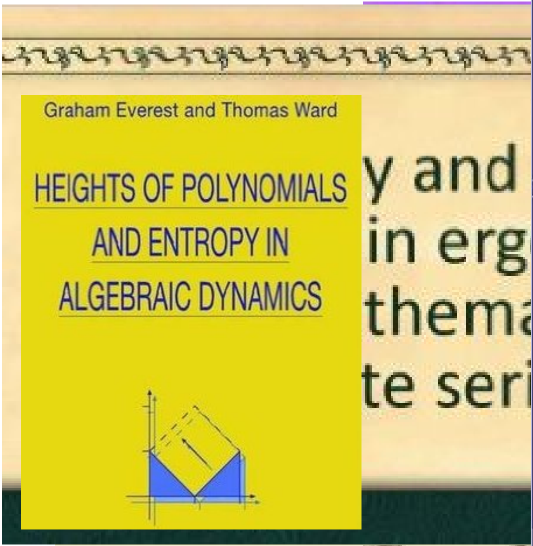
Entropy and Quantum Kolmogorov Complexity: A Quantum Brudno's Theorem

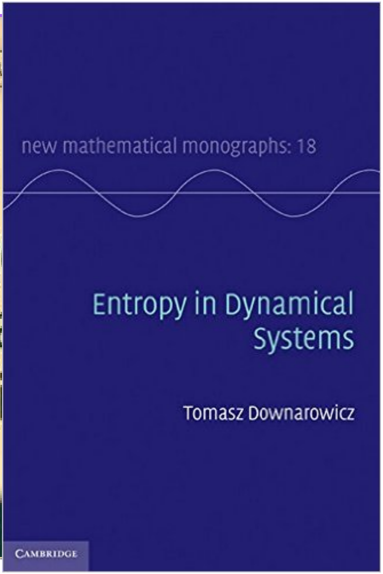
Fabio Benatti[1], Tyll Krüger[2,3], Markus Müller[2], Rainer Siegmund-Schultze[2], Arleta Szkoła[2]

Entropy in the sciences

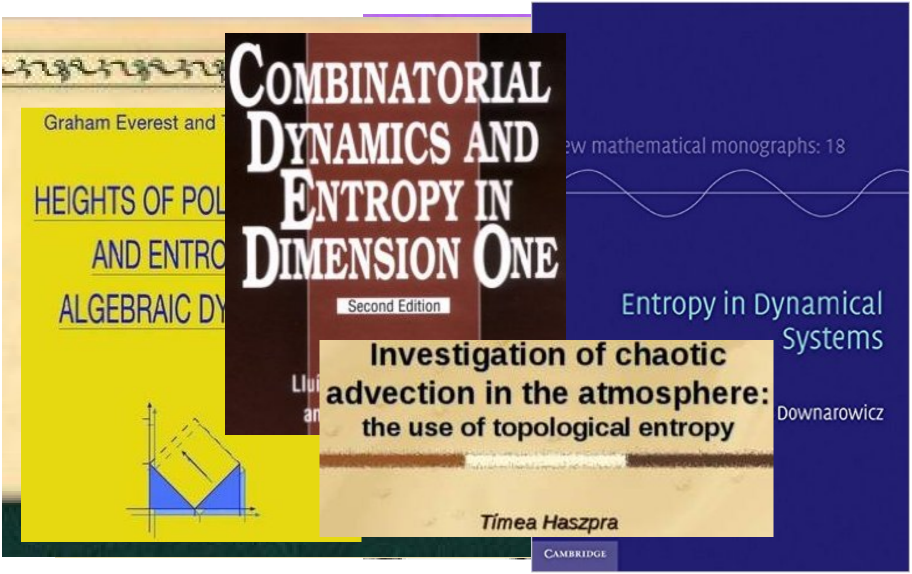


Entropy in the sciences

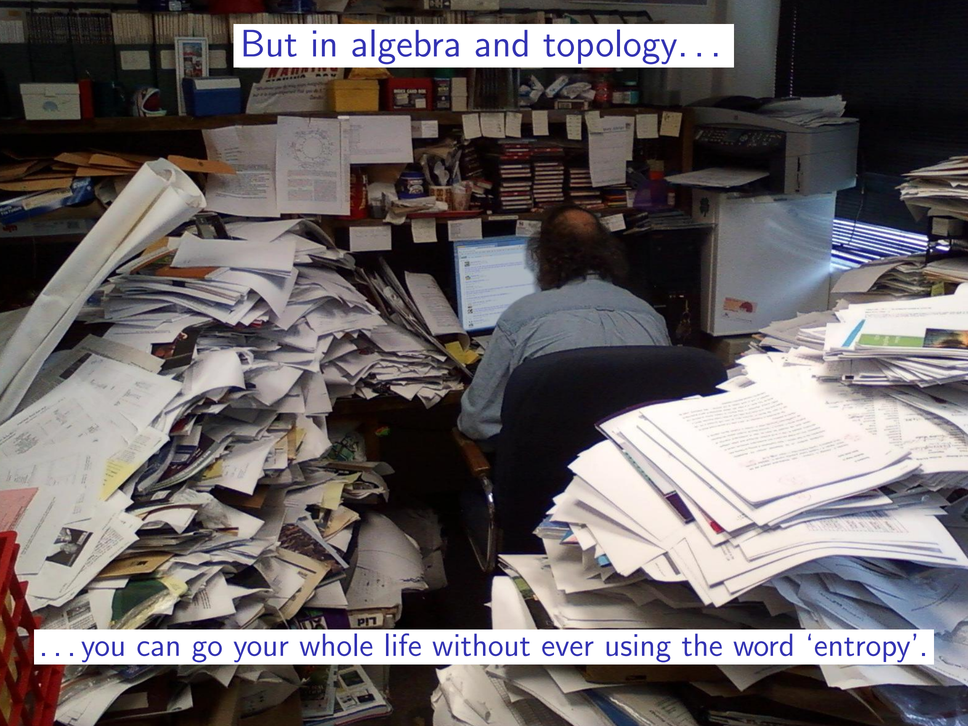




Entropy in the sciences



But in algebra and topology…



…you can go your whole life without ever using the word ‘entropy’.

The point of this talk

Entropy is notable by its relative absence from algebra and topology.

However, we will see that by considering general algebraic structures such as operads and categories—and with just a *tiny* bit of topological input—we naturally arrive at entropy.

It's there, whether we like it or not!

The point of this talk

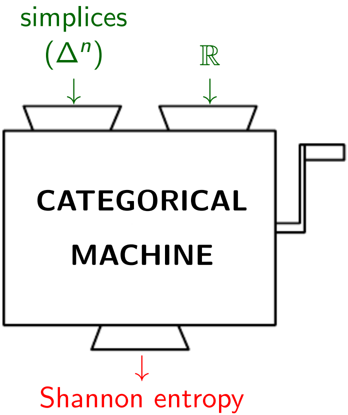


Image: J. Kock

The point of this talk

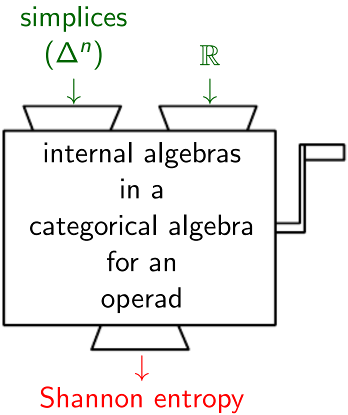


Image: J. Kock

Plan

1. What is entropy?
2. Review of yesterday
3. Categorical algebras for an operad
4. Internal algebras
5. The theorem: how entropy arises

I'll use a small amount of categorical vocabulary: category, functor, and (just once) natural transformation.

I'll also build on some of what I explained yesterday.

But even if you missed yesterday, come along for the ride. . .

1. What is entropy?

The definition of entropy

The simplest kind of entropy is the Shannon entropy of a finite probability distribution.

Let $\mathbf{p} = (p_1, \ldots, p_n)$ be a finite probability distribution: so $p_i \geq 0$ and $\sum p_i = 1$.

Its (Shannon) entropy is

$$H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \log p_i.$$

- When $p_i = 0$, interpret $0 \log 0$ as 0.
- Changing the base of the log only affects $H(\mathbf{p})$ up to a constant factor.

Entropy is the most important quantity associated with a probability distribution.

But what does it *mean*?

Entropy has a reputation for being mysterious...

> My greatest concern was what to call it. I thought of calling it "information", but the word was overly used, so I decided to call it "uncertainty". When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, "You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage."
>
> —*Claude Shannon*

A lot of mystique surrounds entropy.

But it's not so mysterious!

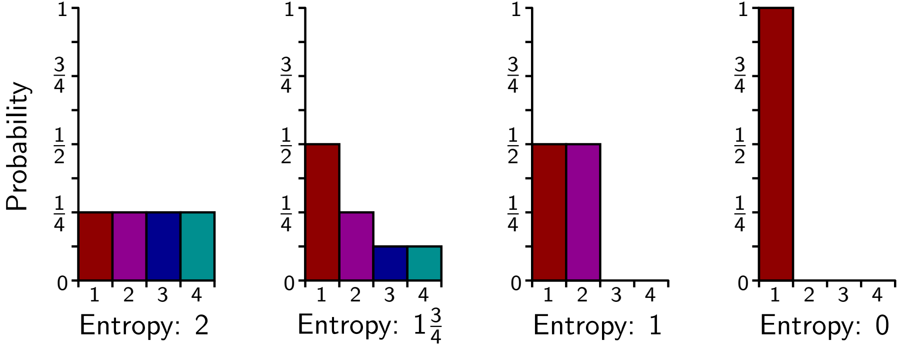
Entropy as uniformity

The entropy $H(\mathbf{p}) = -\sum p_i \log p_i$ can be understood as a measure of the *uniformity* of $\mathbf{p}$.

For distributions on an n-element set, its *maximum value* is $\log n$, achieved when $\mathbf{p} = (1/n, \ldots, 1/n)$.

Its *minimum value* is 0, achieved when $\mathbf{p} = (0, \ldots, 0, 1, 0, \ldots, 0)$.

Examples with $n = 4$, taking logarithms to base 2:

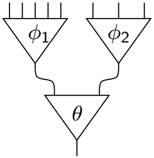


2. Review of yesterday

The definition of operad

An operad is like an abstract clone, but without the reindexing of variables: it's a sequence $(P_n)_{n \geq 0}$ of sets together with:

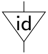
- a composition operator



$$\begin{array}{ccc} P_n \times P_{k_1} \times \cdots \times P_{k_n} & \to & P_{k_1 + \cdots + k_n} \\ (\theta, \phi_1, \ldots, \phi_n) & \mapsto & \theta \circ (\phi_1, \ldots, \phi_n) \end{array}$$

  for each $n, k_1, \ldots, k_n \geq 0$

- an identity element $\mathrm{id} \in P_1$,



satisfying associativity and identity axioms.

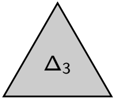
Examples:

- Terminal operad: $P_n = \{*_n\}$ for all n.
- For any monoid M, get operad P^M with $P_1^M = M$ and $P_n^M = \varnothing$ otherwise.

The operad of simplices

The operad Δ of simplices:

$$\Delta_n = \{(p_1, \ldots, p_n) \in \mathbb{R}^n \ : \ p_i \geq 0, \ \sum p_i = 1\}.$$



Composition is defined by thinking of $\mathbf{p} = (p_1, \ldots, p_n)$ as a probability distribution on $\{1, \ldots, n\}$. E.g. if

$\mathbf{p} =$  $= (\frac{1}{2}, \frac{1}{2})$, $\mathbf{q}_1 =$  $= (\frac{1}{6}, \ldots, \frac{1}{6})$, $\mathbf{q}_2 =$  $= (\frac{1}{52}, \ldots, \frac{1}{52})$

then

$$\mathbf{p} \circ (\mathbf{q}_1, \mathbf{q}_2) = (\underbrace{\tfrac{1}{12}, \ldots, \tfrac{1}{12}}_{6}, \underbrace{\tfrac{1}{104}, \ldots, \tfrac{1}{104}}_{52}) \in \Delta_{58}.$$

Generally, given

$$\mathbf{p} = (p_1, \ldots, p_n),$$
$$\mathbf{q}_1 = (q_1^1, \ldots, q_1^{k_1}), \ \ldots, \ \mathbf{q}_n = (q_n^1, \ldots, q_n^{k_n}),$$

define

$$\mathbf{p} \circ (\mathbf{q}_1, \ldots, \mathbf{q}_n) = (p_1 q_1^1, \ldots, p_1 q_1^{k_1}, \ \ldots, \ p_n q_n^1, \ldots, p_n q_n^{k_n}) \in \Delta_{k_1 + \cdots + k_n}.$$

Algebras for an operad

Let P be an operad.

A P-algebra is a set A together with a map

$$\overline{\theta}\colon A^n \to A$$

for each $n \geq 0$ and $\theta \in P_n$, satisfying action-like axioms:
(i) composition, (ii) identity.

Examples:

- When P is the terminal operad, a P-algebra is exactly a monoid.
- An P^M-algebra is exactly a set with an M-action.
- Let $A \subseteq \mathbb{R}^d$ be a convex set. Then A becomes a Δ-algebra as follows: given $\mathbf{p} \in \Delta_n$, define

$$\begin{array}{rccc} \overline{\mathbf{p}}\colon & A^n & \to & A \\ & (\mathbf{a}_1, \ldots, \mathbf{a}_n) & \longmapsto & \sum_i p_i \mathbf{a}_i. \end{array}$$

Algebras in categories other than **Set**

Let $\mathcal{M}$ be a category with some kind of product $\otimes$ and unit object I: it could be $(\mathbf{Set}, \times, \{*\})$, or something else.

Let P be an operad.

A P-algebra in $\mathcal{M}$ is an object A of $\mathcal{M}$ together with a map

$$\overline{\theta}\colon A^{\otimes n} \to A$$

for each $n \geq 0$ and $\theta \in P_n$, satisfying action-like axioms:
(i) composition, (ii) identity.

Today, we'll think about the case where $\mathcal{M}$ is **Cat**, the category of categories and functors.

3. Categorical algebras for operads

The definition of categorical algebra for an operad

Let P be an operad.

A categorical P-algebra is a P-algebra in **Cat**.

Explicitly: it's a category $\boldsymbol{A}$ together with a functor

$$\overline{\theta}\colon \boldsymbol{A}^n \to \boldsymbol{A}$$

for each $n \geq 0$ and $\theta \in P_n$, satisfying action-like axioms:
(i) composition, (ii) identity.

For $\overline{\theta}$ to be a *functor* $\boldsymbol{A}^n \to \boldsymbol{A}$ means:

- for all objects $a_1, \ldots, a_n$ of $\boldsymbol{A}$, we get an object $\overline{\theta}(a_1, \ldots, a_n)$ of $\boldsymbol{A}$
- for all maps $\begin{array}{c} a_1 \\ {\scriptstyle f_1}\downarrow \\ b_1 \end{array}, \ldots, \begin{array}{c} a_n \\ {\scriptstyle f_n}\downarrow \\ b_n \end{array}$ in $\boldsymbol{A}$, we get a map $\begin{array}{c} \overline{\theta}(a_1,\ldots,a_n) \\ {\scriptstyle \overline{\theta}(f_1,\ldots,f_n)}\downarrow \\ \overline{\theta}(b_1,\ldots,b_n) \end{array}$ in $\boldsymbol{A}$,

and that some axioms are satisfied.

Examples of categorical algebras

- Let P be the terminal operad: $P_n = \{*_n\}$ for all $n \geq 0$.

  By definition, a categorical P-algebra is a category $\boldsymbol{A}$ with a functor $\boldsymbol{A}^n \to \boldsymbol{A}$ for each $n \geq 0$, satisfying axioms.

  It's exactly a strict monoidal category: a category equipped with a strictly associative and unital product $\otimes$. The functor $\overline{*_n}\colon \boldsymbol{A}^n \to \boldsymbol{A}$ is

$$(a_1, \ldots, a_n) \mapsto a_1 \otimes \cdots \otimes a_n.$$

- Let M be a monoid and $P = P^M$ (so $P_1^M = M$ and $P_n^M = \varnothing$ otherwise).

  A categorical P^M-algebra is a category $\boldsymbol{A}$ with a functor $m \cdot -\colon \boldsymbol{A} \to \boldsymbol{A}$ for each $m \in M$, satisfying axioms.

  It's exactly a category with an M-action.

More examples of categorical algebras

Take the operad Δ of simplices:

$$\Delta_n = \{(p_1, \ldots, p_n) \in \mathbb{R}^n \ : \ p_i \geq 0, \ \sum p_i = 1\}.$$

We've already seen that the set $\mathbb{R}$ is a Δ-algebra: for $\mathbf{p} \in \Delta_n$, define

$$\begin{array}{rccc} \overline{\mathbf{p}}\colon & \mathbb{R}^n & \to & \mathbb{R} \\ & (a_1, \ldots, a_n) & \mapsto & p_1 a_1 + \cdots + p_n a_n. \end{array}$$

Crucial point A one-object category is the same thing as a monoid. The morphisms are the elements of the monoid, and composition is multiplication.

So, we can also view $\mathbb{R}$ as a category with only one object, with $\circ = +$.

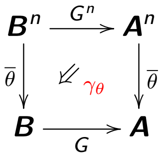
Each operation $\overline{\mathbf{p}}$ preserves addition. So $\overline{\mathbf{p}}$ is a *functor* $\mathbb{R}^n \to \mathbb{R}$.

It follows that $\mathbb{R}$, as a one-object category, is a categorical Δ-algebra.

Maps between categorical algebras for an operad

Fix an operad P and categorical P-algebras $\boldsymbol{B}$ and $\boldsymbol{A}$.

A lax map $\boldsymbol{B} \to \boldsymbol{A}$ is a functor $G\colon \boldsymbol{B} \to \boldsymbol{A}$ together with a natural transformation



for each $n \geq 0$ and $\theta \in P_n$, satisfying axioms.

Explicitly: it's a functor G together with a map

$$\gamma_{\theta, b^1, \ldots, b^n} \colon \overline{\theta}\big(Gb^1, \ldots, Gb^n\big) \to G\big(\overline{\theta}(b^1, \ldots, b^n)\big)$$

for each $\theta \in P_n$ and $b^1, \ldots, b^n \in \boldsymbol{B}$, satisfying naturality and axioms on:

(i) composition, (ii) identity.

4. Internal algebras

Internal algebras in a categorical algebra for an operad

Fix an operad P and a categorical P-algebra $\boldsymbol{A}$.

Write $\mathbf{1}$ for the categorical P-algebra with one object and only the identity morphism. (This category is a categorical P-algebra in a unique way.)

Definition: An internal algebra in $\boldsymbol{A}$ is a lax map $\mathbf{1} \to \boldsymbol{A}$.

Explicitly: it's an object $a \in \boldsymbol{A}$ together with a map

$$\gamma_\theta \colon \overline{\theta}(a, \ldots, a) \to a$$

for each $n \geq 0$ and $\theta \in P_n$, satisfying axioms on
(i) composition, (ii) identity.

Examples:

- Let P be the terminal operad. Let $\boldsymbol{A}$ be a strict monoidal category. An internal P-algebra in $\boldsymbol{A}$ is just a monoid in $\boldsymbol{A}$.
- Let $P = P^M$. Let $\boldsymbol{A}$ be a category with an M-action. An internal P^M-algebra in $\boldsymbol{A}$ is an object $a \in \boldsymbol{A}$ with a map $\gamma_m \colon m \cdot a \to a$ for each $m \in M$, satisfying action-like axioms.

Internal algebras in a one-object categorical algebra

Fix an operad P and a categorical P-algebra $\boldsymbol{A}$.

We just saw: an internal P-algebra in $\boldsymbol{A}$ is an object $a \in \boldsymbol{A}$ with a map

$$\gamma_\theta \colon \overline{\theta}(a, \ldots, a) \to a$$

for each $n \geq 0$ and $\theta \in P_n$, satisfying axioms.

What if $\boldsymbol{A}$ has only one object?

That is, what if $\boldsymbol{A}$ is a monoid A?

An internal algebra in the one-object category corresponding to A is a sequence of functions

$$\big(\gamma \colon P_n \to A\big)_{n \geq 0},$$

satisfying axioms on

(i) composition, (ii) identity.

Topologizing everything

Everything so far can be done in the world of topological spaces instead of sets.

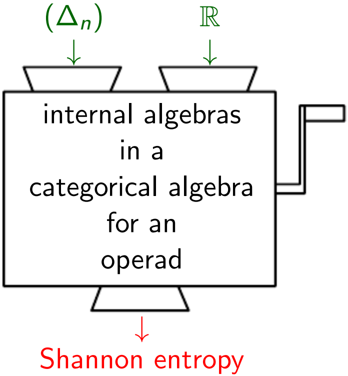
(Jargon: we work *internally* to the category **Top** instead of **Set**.)

Explicitly, this means that throughout, we add a condition

(iii) continuity

to the conditions (i) and (ii) that appear repeatedly.

5. The theorem: how entropy arises


(Δ_n)
$\mathbb{R}$
internal algebras
in a
categorical algebra
for an
operad
Shannon entropy

The theorem

Recall We have:

- the (topological) operad $\Delta = (\Delta_n)_{n \geq 0}$ of simplices
- the one-object (topological) category $\mathbb{R}$: morphisms are real numbers, $\circ = +$
- the one-object (topological) categorical Δ-algebra $\mathbb{R}$.

Recall For an operad P, an internal algebra in a one-object categorical P-algebra A is a sequence of functions $(P_n \to A)_{n \geq 0}$, satisfying axioms.

So, an internal algebra in the categorical Δ-algebra $\mathbb{R}$ is a sequence of functions $(\Delta_n \to \mathbb{R})_{n \geq 0}$, satisfying axioms.

One famous sequence of functions $(\Delta_n \to \mathbb{R})_{n \geq 0}$ is Shannon entropy:

$$H : \mathbf{p} \mapsto -\sum p_i \log p_i.$$

Theorem

The internal algebras in the categorical Δ-algebra $\mathbb{R}$ are precisely the scalar multiples of Shannon entropy.

The theorem

Theorem

The internal algebras in the categorical Δ-algebra $\mathbb{R}$ are precisely the scalar multiples of Shannon entropy.

Explicit version of the theorem (no categorical jargon)

Take a sequence of functions $\big(\gamma\colon \Delta_n \to \mathbb{R}\big)_{n\geq 0}$.
Then $\gamma = cH$ for some $c \in \mathbb{R}$ if and only if γ satisfies:

(i) composition: $\gamma\big(\mathbf{p} \circ (\mathbf{q}_1, \ldots, \mathbf{q}_n)\big) = \gamma(\mathbf{p}) + \sum_i p_i \gamma(\mathbf{q}_i)$

(ii) identity: $\gamma\big((1)\big) = 0$

(iii) continuity: each function γ is continuous.

Proof: This explicit form is almost equivalent to a 1956 theorem of Faddeev, except that he imposed a symmetry axiom that turns out to be redundant. $\square$

Summary

We have met various very general concepts:

- operads (a cousin of clones/algebraic theories)
- algebras for an operad: both set-based and categorical algebras
- internal algebras in a categorical algebra for an operad.

The simplest example:

- for the terminal operad ($P_n = \{*_n\}$ for all n), a categorical algebra is a strict monoidal category $\mathcal{M}$, and the internal algebras in $\mathcal{M}$ are the monoids in $\mathcal{M}$.

Another fundamental example:

- for the operad of simplices $(\Delta_n)_{n \geq 0}$, one categorical algebra is the one-object category $(\mathbb{R}, +)$, *and the internal algebras in it are the multiples of Shannon entropy.*

In short:

Entropy is inevitable.

Preview of Thursday

On Thursday...

We will follow this algebraic, axiomatic approach to entropy and use it to do something new, involving:

- "probabilities" that are not real numbers but integers modulo a prime
- an answer to the question: why is it reasonable to say that

$$\log \sqrt{8} \equiv 3 \pmod{7}?$$