Sense-making: accounting for intelligibility

David I. Spivak

TOPOS
INSTITUTE

Mathematics of Collective Intelligence
2022 February 19

Outline

Why am I here?

I want to know what's going on here. What's this thing we're doing?

- We've engaged in something here at IPAM, a sense-making activity.
- We're improving our understanding of collective intelligence.
- I want a mathematical account of how collective sense-making works.

Why am I here?

I want to know what's going on here. What's this thing we're doing?

- We've engaged in something here at IPAM, a sense-making activity.
- We're improving our understanding of collective intelligence.
- I want a mathematical account of how collective sense-making works.

There are two reasons: it's fun and it's important.

- Finding the right abstractions to explain something is exhilarating.
- I love discussing how intelligence is collectively generated.
- Simultaneously, the world is distressed.
- New complexity challenges our collective sense-making capacity.

Why am I here?

I want to know what's going on here. What's this thing we're doing?

- We've engaged in something here at IPAM, a sense-making activity.
- We're improving our understanding of collective intelligence.
- I want a mathematical account of how collective sense-making works.

There are two reasons: it's fun and it's important.

- Finding the right abstractions to explain something is exhilarating.
- I love discussing how intelligence is collectively generated.
- Simultaneously, the world is distressed.
- New complexity challenges our collective sense-making capacity.

A constructive account of sense-making may suggest new ways forward.

Mathematical fields as accounting systems

I think of mathematical fields as accounting systems.

- Arithmetic accounts for the flow of quantities, as in finance.
- Hilbert spaces account for the states of elementary particles, as in QM.
- Probability distributions account for likelihoods, as in game theory.

Mathematical fields as accounting systems

I think of mathematical fields as accounting systems.

- Arithmetic accounts for the flow of quantities, as in finance.
- Hilbert spaces account for the states of elementary particles, as in QM.
- Probability distributions account for likelihoods, as in game theory.

An account of phenomena needs to be written in high-fidelity language.

- To understand the phenomena requires that certain aspects be tracked.
- The language must articulate the relevant type-differences ...
- ... and provide operations that correspond with their interactions.

Mathematical fields as accounting systems

I think of mathematical fields as accounting systems.

- Arithmetic accounts for the flow of quantities, as in finance.
- Hilbert spaces account for the states of elementary particles, as in QM.
- Probability distributions account for likelihoods, as in game theory.

An account of phenomena needs to be written in high-fidelity language.

- To understand the phenomena requires that certain aspects be tracked.
- The language must articulate the relevant type-differences ...
- ... and provide operations that correspond with their interactions.

Category theory is the accounting system for interlocking structures.

- Mathematical definitions are composed of interlocking structures.
- Category theory tracks the layers of structure and their connections.
- This makes analogies—similarities of structure—into formal objects.

The morphology of collective intelligence

Collective intelligence—the product of culture—is all around us.

- It's in our science, our technology, our governance, our morality.
- Each of these is the product of our work over millennia.
- Each body is a collective of cells whose individual intelligences...
- ... work harmoniously to create the intelligence at our level.

The morphology of collective intelligence

Collective intelligence—the product of culture—is all around us.

- It's in our science, our technology, our governance, our morality.
- Each of these is the product of our work over millennia.
- Each body is a collective of cells whose individual intelligences...
- ... work harmoniously to create the intelligence at our level.

I want a language and logic for the *shape* of collective intelligence.

- In particular, I want to be able to talk about this leveling up.
- Rather than understanding the lowest level physics...
- ...and relying on "emergence" to get us to human intelligence, ...
- ... I prefer to look for construction principles that are *compositional*.

The morphology of collective intelligence

Collective intelligence—the product of culture—is all around us.

- It's in our science, our technology, our governance, our morality.
- Each of these is the product of our work over millennia.
- Each body is a collective of cells whose individual intelligences...
- ... work harmoniously to create the intelligence at our level.

I want a language and logic for the *shape* of collective intelligence.

- In particular, I want to be able to talk about this leveling up.
- Rather than understanding the lowest level physics...
- ...and relying on "emergence" to get us to human intelligence, ...
- ... I prefer to look for construction principles that are *compositional*.

Wanted: an algebra in which interacting intelligences form an intelligence.

- The category-theoretic notion of *operad* seems appropriate.
- An operad lets you create arbitrary—e.g. geometric—syntax.
- You design the operad so that its combination-rules make sense...
- ... in this case, make sense for collectivizing intelligences.
- It'd be a custom accounting system for how intelligences combine.
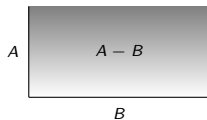
3 / 14

Outline

Getting a sense

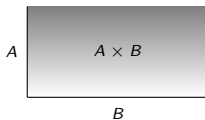How does this make you feel?
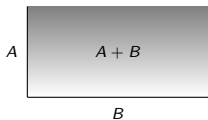
Getting a sense

How does this make you feel? Like trying on a friend's glasses?
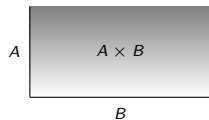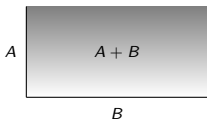
Getting a sense

How does this make you feel?   Ah... that's better.

What separates the good math student—or tennis player—from the bad?

Getting a sense

How does this make you feel?

What separates the good math student—or tennis player—from the bad?

- The bad math student memorizes formulas.
- The good student gets a sense. They track the ideas with the symbols.
- Our senses are the entirety of our connection to the matter at hand.
- The good tennis player senses distances, speeds, angles.
- 100 trillion atoms are involved, but our senses track the "right" ones.

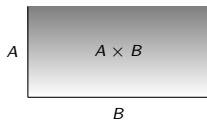# Getting a sense

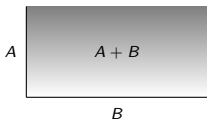How does this make you feel?

What separates the good math student—or tennis player—from the bad?

- The bad math student memorizes formulas.
- The good student gets a sense. They track the ideas with the symbols.
- Our senses are the entirety of our connection to the matter at hand.
- The good tennis player senses distances, speeds, angles.
- 100 trillion atoms are involved, but our senses track the "right" ones.

How did we get our senses? How is it that we *sense* the situation?

- $x^2 + 3 = 7$; have a sense of what to do before you know the answer?
- Do you have a sense of which way the exit is?
- So much to track. But we do it! Can we make sense of *that*?

Sense-making: the pun that wasn't

I'm using the term sense-making in two ways. First, we make sense of xyz.

- Sometimes we shake our head and say "but that doesn't make sense".
- There's no point to accepting something if it doesn't make sense.
- But when they explain it a different way, "ohh, that makes sense!"
- We sit there with the math or social problem until it clicks.
- What's the clicking? Can we make sense of *that*?

Sense-making: the pun that wasn't

I'm using the term sense-making in two ways. First, we make sense of xyz.

- Sometimes we shake our head and say "but that doesn't make sense".
- There's no point to accepting something if it doesn't make sense.
- But when they explain it a different way, "ohh, that makes sense!"
- We sit there with the math or social problem until it clicks.
- What's the clicking? Can we make sense of *that*?

Second, we are considering what has made our senses.

- We sense the color and taste of the strawberry.
- We track the right aspects in the tennis match or the math problem.
- What made these senses for us? How do we play so beautifully?

Sense-making: the pun that wasn't

I'm using the term sense-making in two ways. First, we make sense of xyz.

- Sometimes we shake our head and say "but that doesn't make sense".
- There's no point to accepting something if it doesn't make sense.
- But when they explain it a different way, "ohh, that makes sense!"
- We sit there with the math or social problem until it clicks.
- What's the clicking? Can we make sense of *that*?

Second, we are considering what has made our senses.

- We sense the color and taste of the strawberry.
- We track the right aspects in the tennis match or the math problem.
- What made these senses for us? How do we play so beautifully?

I hypothesize that these two meanings are the same, past and present.

- We produce the senses we later enjoy through our sense-making work.
- We install what we know into the deepest structures we can find:
    - We write our sense of how-to into computer code and books, ...
    - ... our sense of beauty & good into DNA with mate selection.

Settling accounts

What do you think of this hypothesis. Does it make sense?

- Could past sense-making activity, installed into deep structures...
- ... account for the senses we have today?
- Does that *account* feel settled enough that we can go with it?

Settling accounts

What do you think of this hypothesis. Does it make sense?

- Could past sense-making activity, installed into deep structures...
- ... account for the senses we have today?
- Does that *account* feel settled enough that we can go with it?
- If so, the real work is to understand what sense-making is.
- How is it produced? What are the specs of sense?

Settling accounts

What do you think of this hypothesis. Does it make sense?

- Could past sense-making activity, installed into deep structures...
- ... account for the senses we have today?
- Does that *account* feel settled enough that we can go with it?
- If so, the real work is to understand what sense-making is.
- How is it produced? What are the specs of sense?

I hypothesize that sense-production has to do with proper accounting.

- When we shake our head and say "that doesn't make sense"...
- ... we're saying it doesn't settle the accounts. Something is left over.
- We jiggle the pieces, try different arrangements until click.
- Something settles. The energy level drops. Delightful!

Settling accounts

What do you think of this hypothesis. Does it make sense?

- Could past sense-making activity, installed into deep structures...
- ... account for the senses we have today?
- Does that *account* feel settled enough that we can go with it?
- If so, the real work is to understand what sense-making is.
- How is it produced? What are the specs of sense?

I hypothesize that sense-production has to do with proper accounting.

- When we shake our head and say "that doesn't make sense"...
- ... we're saying it doesn't settle the accounts. Something is left over.
- We jiggle the pieces, try different arrangements until click.
- Something settles. The energy level drops. Delightful!

I don't think this is merely Bayesian; it's too much of a phase change.

- Bayesian update, free energy principle stuff doesn't feel right.
- Sense-making has a groping-in-the-dark feel, followed by a click.
- At this point we can *build* on it. The edifice of sense.
- I haven't seen anything account for the delight of the click-into-place.

Outline

Where we are, and where we're going

Luckily, mathematical fields are accounting systems!

- So if we want to account for the click, we could try to do it in math.
- I propose that category theory is a great language for this.
- Don't get your hopes up: I don't have an account of sense-making!
- The remainder of the talk is about an accounting system, …
- … which I think an account of sense-making given be told within.

Where we are, and where we're going

Luckily, mathematical fields are accounting systems!

- So if we want to account for the click, we could try to do it in math.
- I propose that category theory is a great language for this.
- Don't get your hopes up: I don't have an account of sense-making!
- The remainder of the talk is about an accounting system, ...
- ... which I think an account of sense-making given be told within.

So let me tell you very briefly about the operadic approach.

- Please ask for details if you're interested.
- Discussing this and its relevance to your thinking is what I'm here for.

The operad idea

An operad is an *e pluribus unum* system.

- You specify a set of possible *interfaces*.
- You specify how interfaces can be *arranged* within any interface.
- For example, maybe interfaces are sets and arrangements are functions:
$$\varphi \colon S_1 \times \cdots \times S_n \to S'$$
- This φ builds one element $\varphi(s_1, \ldots, s_n) \in S'$ out of n-many elements.
- You also specify how nesting works.

The operad idea

An operad is an *e pluribus unum* system.

- You specify a set of possible *interfaces*.
- You specify how interfaces can be *arranged* within any interface.
- For example, maybe interfaces are sets and arrangements are functions:
$$\varphi \colon S_1 \times \cdots \times S_n \to S'$$
- This φ builds one element $\varphi(s_1, \ldots, s_n) \in S'$ out of n-many elements.
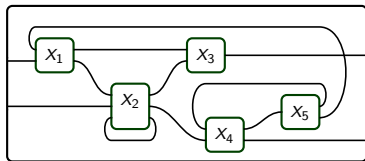- You also specify how nesting works.

The original operad $\mathcal{B}$ was "boxes positioned within boxes":
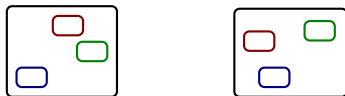
- The operad $\mathcal{B}$ only has one object, "the box", but for every $n \in \mathbb{N}$ ...
- ... $\mathcal{B}$ has a whole space of arrangements, two of which are shown.
- This nests: you can put tiny boxes inside the small boxes.
- Turtles can go all the way down, as deep as you want.

Interaction patterns

Here's a more-relevant operad: wiring diagrams.

The nesting allows you to build ever-higher levels of abstraction:

- transistors in logic gates, in adder circuits, in CPUs, in server farms, ...
- You can stop at any top or bottom, but you don't have to.

Interaction patterns
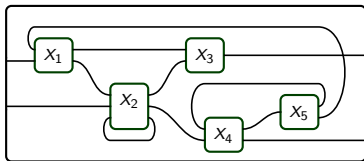
Here's a more-relevant operad: wiring diagrams.

The nesting allows you to build ever-higher levels of abstraction:

- transistors in logic gates, in adder circuits, in CPUs, in server farms, ...
- You can stop at any top or bottom, but you don't have to.

There's an even-more-relevant operad that's much more dynamic.

- The boxes can change their number of ports, and ...
- ... the wiring diagram itself can change in time, ...
- ... based on what flows on the wires. The boxes can run around.

I'm calling this *operadic morphology*.

Interaction patterns

Here's a more-relevant operad: wiring diagrams.

The nesting allows you to build ever-higher levels of abstraction:
- transistors in logic gates, in adder circuits, in CPUs, in server farms, ...
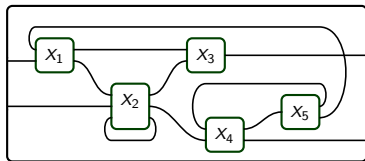- You can stop at any top or bottom, but you don't have to.

There's an even-more-relevant operad that's much more dynamic.
- The boxes can change their number of ports, and ...
- ... the wiring diagram itself can change in time, ...
- ... based on what flows on the wires. The boxes can run around.

I'm calling this *operadic morphology*.
- It's difficult to distinguish between morphology and behavior.
- Our behavior in the company changes the company's morphology.   9/14

Interacting dynamical systems

Operads are designed to give algebraic theories.

- When $\mathcal{O}$ consists of these interfaces and those arrangements...
- ... then an $\mathcal{O}$-algebra says what can *fill* the interfaces ...
- ... and how arrangements of fillers form a composite filler.
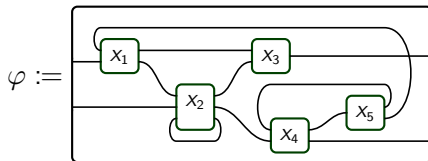
Interacting dynamical systems

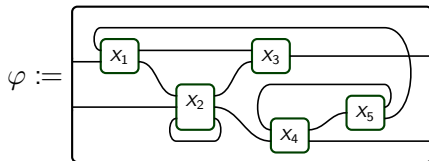Operads are designed to give algebraic theories.

- When $\mathcal{O}$ consists of these interfaces and those arrangements...
- ... then an $\mathcal{O}$-algebra says what can *fill* the interfaces ...
- ... and how arrangements of fillers form a composite filler.

$$\varphi :=$$

Say you fill each X_i with an open dynamical system (automaton or ODE).

- Each dynamical system takes inputs and produces outputs,...
- ...updating (or flowing) according to some formula you specify.

The arrangement φ tells us how to form a composite open dyn'l system.

- We can make the connection pattern adjust itself based on what flows.
- This can be formalized *very cleanly* using polynomial functors.

What algorithm works?

The above is worked out, and the math is elegant.

- It's a container for little machines that adjust their configuration.
- This "container" holds both *electrical circuits* and *deep learning*.
- There's tons of room between these two: it's an empty continent.

What algorithm works?

The above is worked out, and the math is elegant.

- It's a container for little machines that adjust their configuration.
- This "container" holds both *electrical circuits* and *deep learning*.
- There's tons of room between these two: it's an empty continent.

But I think deep learning and circuits are missing a key idea.

- Ok, they're amazing, but they don't seem like enough to explain life.
- Training DNNs doesn't result in an *explanation* of anything.
- Explainability—sense-making—is the thing we're trying to track.

What algorithm works?

The above is worked out, and the math is elegant.

- It's a container for little machines that adjust their configuration.
- This "container" holds both *electrical circuits* and *deep learning*.
- There's tons of room between these two: it's an empty continent.

But I think deep learning and circuits are missing a key idea.

- Ok, they're amazing, but they don't seem like enough to explain life.
- Training DNNs doesn't result in an *explanation* of anything.
- Explainability—sense-making—is the thing we're trying to track.

Question: by what algorithmic strategy could we build up sense-making?

- If each of the little boxes is a sense-maker, by what adjustments...
- ...would the collective itself be a sense-maker?
- If every box could announce "here's what problems I make sense of"...
- ...could the adjusting collective arrange them to solve higher problems?

Governance, accountability, and sense-making

The operadic approach says that we just need the "inductive step".

- Can we say how a collective of sense-makers constitute a sense-maker?
- The following is hypothetical. I'm groping for the right idea.

Governance, accountability, and sense-making

The operadic approach says that we just need the "inductive step".

- Can we say how a collective of sense-makers constitute a sense-maker?
- The following is hypothetical. I'm groping for the right idea.

What if sense-making is just proper accounting?

- We make sense by accounting for what's happening, putting it away.
- If each small box says "I can account for *this* aspect of my input"...
- ...then when the adjusting interaction pattern gets things right...
- ...maybe it would "click" and accounts would settle.

Governance, accountability, and sense-making

The operadic approach says that we just need the "inductive step".

- Can we say how a collective of sense-makers constitute a sense-maker?
- The following is hypothetical. I'm groping for the right idea.

What if sense-making is just proper accounting?

- We make sense by accounting for what's happening, putting it away.
- If each small box says "I can account for *this* aspect of my input"...
- ...then when the adjusting interaction pattern gets things right...
- ...maybe it would "click" and accounts would settle.

What does proper accounting do?

- The claim was that our senses constitute the totality of our ability.
- So if accounting explains sense-making, it must explain ability. (Let's try..)
- Proper accounting creates intra-level and inter-level coherence.
- Cohering structures align high-level "decisions" to low-level actions.
- And it makes the low-level activity intelligible.
- So good accounting creates a thread of intelligibility, top-to-bottom.

Outline

The sixth great extinction

The sixth great extinction is nipping at our toes.

- The activity that's killing our animal friends won't stop there.
- My guess is that the civil unrest is caused by people feeling it coming.
- $100k / year jobs, e.g. mortgage broker, have largely gone extinct.
- People are trying to make sense of it. They're turning to Q-anon.

The sixth great extinction

The sixth great extinction is nipping at our toes.

- The activity that's killing our animal friends won't stop there.
- My guess is that the civil unrest is caused by people feeling it coming.
- $100k / year jobs, e.g. mortgage broker, have largely gone extinct.
- People are trying to make sense of it. They're turning to Q-anon.

Everyone naturally knows how to sense-make, but many are disoriented.

- Sense-making is nutritious for everyone, *always*; right? That's rare.

The sixth great extinction

The sixth great extinction is nipping at our toes.

- The activity that's killing our animal friends won't stop there.
- My guess is that the civil unrest is caused by people feeling it coming.
- $100k / year jobs, e.g. mortgage broker, have largely gone extinct.
- People are trying to make sense of it. They're turning to Q-anon.

Everyone naturally knows how to sense-make, but many are disoriented.

- Sense-making is nutritious for everyone, *always*; right? That's rare.
- Understanding sense-making should be fun, useful, and safe.
- We need to understand how sense is produced & tuned, but also how...
- ...it's distinguished from just-so stories in constant need of shoring up.
- To keep it grounded, generalizable, and uninfected by agenda,...
- ... the description should be as formal and elegant as possible.

Summary

Sense-making is an activity we naturally do.

- Trying to survive isn't as effective for survival as trying to sense-make.
- Maybe all the sense we have was made & installed by our predecessors.

Summary

Sense-making is an activity we naturally do.

- Trying to survive isn't as effective for survival as trying to sense-make.
- Maybe all the sense we have was made & installed by our predecessors.

Category theory is an accounting system for interlocking structures.

- Operads help us make sense of hierarchical systems.
- There's an operad for dynamical systems that rearrange connectivity.

Summary

Sense-making is an activity we naturally do.

- Trying to survive isn't as effective for survival as trying to sense-make.
- Maybe all the sense we have was made & installed by our predecessors.

Category theory is an accounting system for interlocking structures.

- Operads help us make sense of hierarchical systems.
- There's an operad for dynamical systems that rearrange connectivity.

It's possible that sense-making is produced by proper accounting.

- So to make sense of sense-making we simply account for it.
- Operadic account: sense is made by properly arranging sense-makers.
- But the "proper arranging" algorithm—the seed—is missing. Ideas?

Summary

Sense-making is an activity we naturally do.

- Trying to survive isn't as effective for survival as trying to sense-make.
- Maybe all the sense we have was made & installed by our predecessors.

Category theory is an accounting system for interlocking structures.

- Operads help us make sense of hierarchical systems.
- There's an operad for dynamical systems that rearrange connectivity.

It's possible that sense-making is produced by proper accounting.

- So to make sense of sense-making we simply account for it.
- Operadic account: sense is made by properly arranging sense-makers.
- But the "proper arranging" algorithm—the seed—is missing. Ideas?

Our world needs better sense-making; luckily, it's fun, profitable, and good.

Thanks! Comments and questions welcome...