Applied Category Theory:
Towards a science of interdisciplinarity

David I. Spivak

Department of Mathematics
Massachusetts Institute of Technology

Outline

A road to true interdisciplinarity

- Scientific disciplines are conceptual analogies of the world.
  - Science: a schematic, conceptual account of phenomena.
  - Engineering is using these accounts to channel world events.
  - But how do different disciplines and accounts cohere?
  - To solve big problems, we need to connect different approaches.

A road to true interdisciplinarity

- Scientific disciplines are conceptual analogies of the world.
    - Science: a schematic, conceptual account of phenomena.
    - Engineering is using these accounts to channel world events.
    - But how do different disciplines and accounts cohere?
    - To solve big problems, we need to connect different approaches.
- We need a shared fabric, a substrate for interdisciplinarity.
    - Interdisciplinarity consists of effective analogy-making.
    - To go further, we need to formalize the analogies themselves.

A road to true interdisciplinarity

- Scientific disciplines are conceptual analogies of the world.
  - Science: a schematic, conceptual account of phenomena.
  - Engineering is using these accounts to channel world events.
  - But how do different disciplines and accounts cohere?
  - To solve big problems, we need to connect different approaches.
- We need a shared fabric, a substrate for interdisciplinarity.
  - Interdisciplinarity consists of effective analogy-making.
  - To go further, we need to formalize the analogies themselves.
- Better yet: we need a conceptual stem-cell.
  - Something that can differentiate into huge variety of forms.
  - Find the analogies between forms as aspects within the stem cell.

Hard science requires math

Consider as you watch:

- Is a hard science of interdisciplinarity & interoperability possible?

Hard science requires math

Consider as you watch:

- Is a hard science of interdisciplinarity & interoperability possible?
- What sort of mathematical theory would it require?

Hard science requires math

Consider as you watch:

- Is a hard science of interdisciplinarity & interoperability possible?
- What sort of mathematical theory would it require?
- Does Category Theory fit the bill?

Hard science requires math

Consider as you watch:

- Is a hard science of interdisciplinarity & interoperability possible?
- What sort of mathematical theory would it require?
- Does Category Theory fit the bill?

Category theory as conceptual stem-cell

Category theory (CT) can differentiate into many forms:

- All forms of pure math... (we'll briefly discuss this)

Category theory as conceptual stem-cell

Category theory (CT) can differentiate into many forms:

- All forms of pure math... (we'll briefly discuss this)
- Databases and knowledge representation (categories and functors)
- Functional programming languages (cartesian closed categories)
- Universal algebra (finite-product categories)
- Dynamical systems and fractals (operad-algebras, co-algebras)
- Shannon Entropy (operad of simplices)
- Partially-ordered sets and metric spaces (enriched categories)
- Higher order logic (toposes = categories of sheaves)
- Measurements of diversity in populations (magnitude of categories)
- Collaborative design (enriched categories and profunctors)
- Petri nets and chemical reaction networks (monoidal categories)
- Quantum processes and NLP (compact closed categories)

Popper's objection

"A theory that explains everything explains nothing." – Karl Popper

Popper's objection

"A theory that explains everything explains nothing." – Karl Popper

We counter this objection in two ways:

- Couldn't the same objection be made about mathematics?
  - Mathematics is the basis of hard science, used everywhere.
  - CT—like math—explains, models, formalizes many many things.
  - Conclude that math/CT explains everything and hence nothing?

Popper's objection

"A theory that explains everything explains nothing." – Karl Popper

We counter this objection in two ways:

- Couldn't the same objection be made about mathematics?
  - Mathematics is the basis of hard science, used everywhere.
  - CT—like math—explains, models, formalizes many many things.
  - Conclude that math/CT explains everything and hence nothing?
- Stem cells don't do work until they differentiate.
  - "Adult-level" work requires differentiation and optimization.
  - But the unified origins lead to impressive interoperability.

CT is the mathematics of mathematics.

You could also say: CT is the mathematics of mathematics.

- Designed to transport theorems from one area of math to another.
  - Example: from topology (shapes) to algebra (equations).
  - This isn't mere analogy, it's analogy made rigorous.

CT is the mathematics of mathematics.

You could also say: CT is the mathematics of mathematics.

- Designed to transport theorems from one area of math to another.
    - Example: from topology (shapes) to algebra (equations).
    - This isn't mere analogy, it's analogy made rigorous.
- It's revolutionized pure math since its inception in 1940s.
    - It's touched or greatly influenced all corners of mathematics.
    - It's become a gateway to learning mathematics.

CT is the mathematics of mathematics.

You could also say: CT is the mathematics of mathematics.

- Designed to transport theorems from one area of math to another.
  - Example: from topology (shapes) to algebra (equations).
  - This isn't mere analogy, it's analogy made rigorous.
- It's revolutionized pure math since its inception in 1940s.
  - It's touched or greatly influenced all corners of mathematics.
  - It's become a gateway to learning mathematics.
- And it's branched out from math in a big way.
  - Databases and knowledge representation (categories and functors)
  - Functional programming languages (cartesian closed categories)
  - Dynamical systems and fractals (operad-algebras, co-algebras)
  - Shannon Entropy (operad of simplices)
  - Measurements of diversity in populations (magnitude of categories)
  - Collaborative design (enriched categories and profunctors)
  - Petri nets and chemical reaction networks (monoidal categories)
  - Quantum processes and NLP (compact closed categories)

Getting to specifics

The category-theoretic stem cell is about compositional design patterns.

Getting to specifics

The category-theoretic stem cell is about compositional design patterns.

Let's focus on one.

- Operads:

Getting to specifics

The category-theoretic stem cell is about compositional design patterns.

Let's focus on one.

- Operads: the category-theoretic formalization of "operations".
- A branch of category theory that well represents the spirit.
- "Operadic": a *theme* of design patterns.

Continue to consider as you watch:

- Is a hard science of interdisciplinarity & interoperability possible?

Getting to specifics

The category-theoretic stem cell is about compositional design patterns.

Let's focus on one.

- Operads: the category-theoretic formalization of "operations".
- A branch of category theory that well represents the spirit.
- "Operadic": a *theme* of design patterns.

Continue to consider as you watch:

- Is a hard science of interdisciplinarity & interoperability possible?
- What sort of mathematical theory would it require?

Getting to specifics

The category-theoretic stem cell is about compositional design patterns.

Let's focus on one.

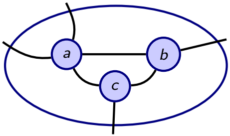
- Operads: the category-theoretic formalization of "operations".
- A branch of category theory that well represents the spirit.
- "Operadic": a *theme* of design patterns.

Continue to consider as you watch:

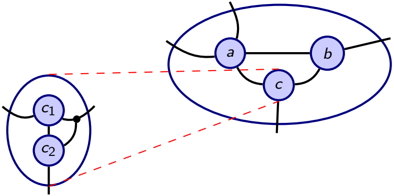
- Is a hard science of interdisciplinarity & interoperability possible?
- What sort of mathematical theory would it require?
- Does Category Theory fit the bill?

Outline

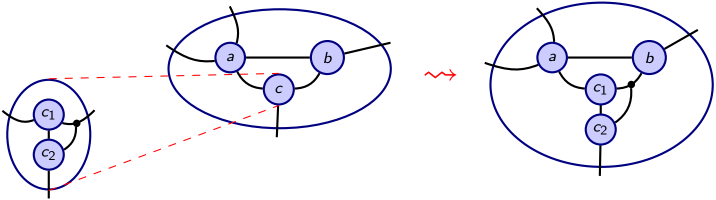
Compositionality: nestable arrangements



Compositionality: nestable arrangements



Compositionality: nestable arrangements



What are compositional operations?

An operad consists of:

- A collection of sorts $X, Y, \ldots$,
- And ways to arrange them, $\varphi\colon X_1, \ldots, X_k \to Y$,
- Such that arrangements can be nested inside each other.

What are compositional operations?

An operad consists of:

- A collection of sorts $X, Y, \ldots$,
- And ways to arrange them, $\varphi\colon X_1, \ldots, X_k \to Y$,
- Such that arrangements can be nested inside each other.
  (That last part is the compositionality.)

What are compositional operations?

An operad consists of:

- A collection of sorts $X, Y, \ldots,$
- And ways to arrange them, $\varphi\colon X_1, \ldots, X_k \to Y,$
- Such that arrangements can be nested inside each other. (That last part is the compositionality.)

Ask me now or later, if you want a more formal definition.

Operads are everywhere

Operads are used unconsciously in many fields.

- Electrical engineering: “wiring diagrams”
- Design: “set-based design”
- Computer programming: “data flow”
- Natural language processing: “grammars”
- Materials science: “hierarchical materials”
- Information theory: “Shannon entropy”

Operads are everywhere

Operads are used unconsciously in many fields.

- Electrical engineering: “wiring diagrams”
- Design: “set-based design”
- Computer programming: “data flow”
- Natural language processing: “grammars”
- Materials science: “hierarchical materials”
- Information theory: “Shannon entropy”

We want to bring operads to the fore.

- There’s a common theme in the way we think.
- Operads structure this sort of thinking.
- With mathematical structure, we can go much further.

Operads are everywhere

Operads are used unconsciously in many fields.

- Electrical engineering: "wiring diagrams"
- Design: "set-based design"
- Computer programming: "data flow"
- Natural language processing: "grammars"
- Materials science: "hierarchical materials"
- Information theory: "Shannon entropy"

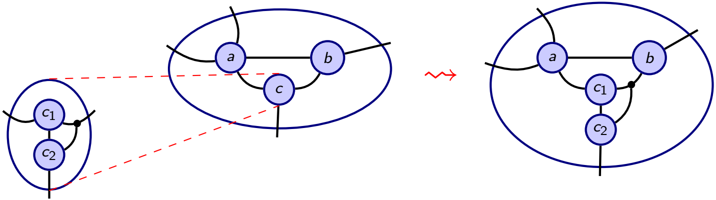
We want to bring operads to the fore.

- There's a common theme in the way we think.
- Operads structure this sort of thinking.
- With mathematical structure, we can go much further.

Let's look for sorts, arrangements, and nesting in some examples.

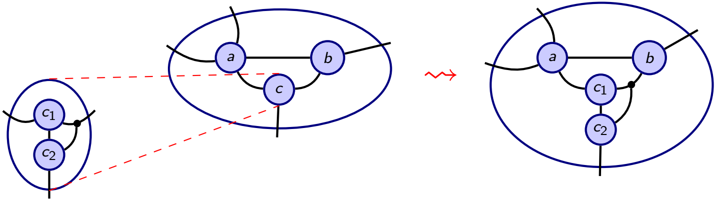
Operad 1: wiring diagrams

Sorts: circles with ports. Arrangements: wiring diagrams.



Operad 1: wiring diagrams

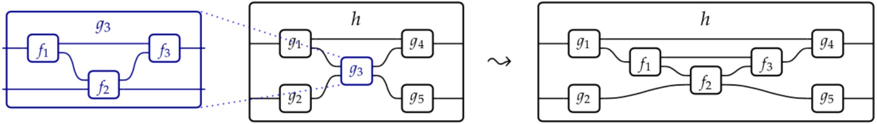
Sorts: circles with ports. Arrangements: wiring diagrams.



Applications: tensor networks, databases, contracts.

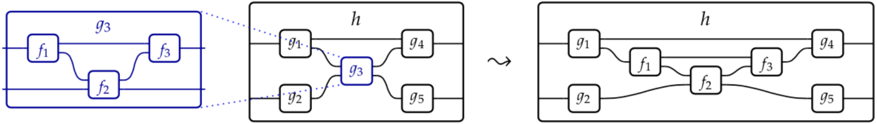
Operad 2: another kind of wiring diagram

Sorts: boxes with ports. Arrangements: a different kind of wiring diagrams.



Operad 2: another kind of wiring diagram

Sorts: boxes with ports. Arrangements: a different kind of wiring diagrams.



Applications: dynamical systems, control theory.

Operad 2: hierarchical protein materials

There is an operad $\mathcal{M}$ for composing hierarchical protein materials.

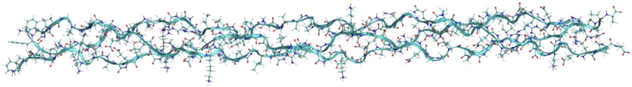
- A protein is an arrangement of simpler proteins.
  - There are "atomic" proteins: amino acids.
  - Protein materials include your skin: stretchable, breathable, waterproof.
  - Materials scientists would *love* to make materials like this.

[1] Giesa, T.; Jagadeesan, R.; Spivak, D.I.; Buehler, M.J. (2015) "Matriarch: a Python library for materials architecture." *ACS Biomaterials Science & Engineering*.

Operad 2: hierarchical protein materials

There is an operad $\mathcal{M}$ for composing hierarchical protein materials.

- A protein is an arrangement of simpler proteins.
  - There are "atomic" proteins: amino acids.
  - Protein materials include your skin: stretchable, breathable, waterproof.
  - Materials scientists would *love* to make materials like this.
- Assemble new proteins from old:
  - arrange in series or parallel (H-bonds), or
  - arrange in helices, double helices, any conceivable curve, etc.



- Collagen has a nested structure: it is an array, each fiber of which is a triple helix, each strand of which is a helix, each unit of which is an amino acid.[1]

[1] Giesa, T.; Jagadeesan, R.; Spivak, D.I.; Buehler, M.J. (2015) "Matriarch: a Python library for materials architecture." *ACS Biomaterials Science & Engineering*.

Operad 3: probabilities

Even an operad with one sort can be interesting.

- Call the operad $\mathcal{P}$ for "probabilities".
- Say sorts={event}, i.e. "event" is the only sort in $\mathcal{P}$.
- What defines the operad are its arrangements and nesting.

Operad 3: probabilities

Even an operad with one sort can be interesting.

- Call the operad $\mathcal{P}$ for "probabilities".
- Say sorts={event}, i.e. "event" is the only sort in $\mathcal{P}$.
- What defines the operad are its arrangements and nesting.

Example: think arrangements = probability distributions.

- Let $\mathcal{P}_k := \left\{(x_1, \ldots, x_k) \in \mathbb{R}^k_{\geq 0} \;\middle|\; x_1 + \cdots + x_k = 1\right\}$.

Operad 3: probabilities

Even an operad with one sort can be interesting.

- Call the operad $\mathcal{P}$ for "probabilities".
- Say sorts=\{event\}, i.e. "event" is the only sort in $\mathcal{P}$.
- What defines the operad are its arrangements and nesting.

Example: think arrangements = probability distributions.

- Let $\mathcal{P}_k := \left\{(x_1, \ldots, x_k) \in \mathbb{R}^k_{\geq 0} \;\middle|\; x_1 + \cdots + x_k = 1\right\}$.
- Examples:
  - flip coin: $f = (\frac{1}{2}, \frac{1}{2}) \in \mathcal{P}_2$.
    - The event "flip coin" as arrangement of two sub-events
  - roll die: $r = (\frac{1}{6}, \ldots, \frac{1}{6}) \in \mathcal{P}_6$.
  - pick card: $p = (\frac{1}{52}, \ldots, \frac{1}{52}) \in \mathcal{P}_{52}$.

Operad 3: probabilities

Even an operad with one sort can be interesting.

- Call the operad $\mathcal{P}$ for "probabilities".
- Say sorts={event}, i.e. "event" is the only sort in $\mathcal{P}$.
- What defines the operad are its arrangements and nesting.

Example: think arrangements = probability distributions.

- Let $\mathcal{P}_k := \left\{(x_1, \ldots, x_k) \in \mathbb{R}^k_{\geq 0} \;\middle|\; x_1 + \cdots + x_k = 1\right\}$.
- Examples:
  - flip coin: $f = (\frac{1}{2}, \frac{1}{2}) \in \mathcal{P}_2$.
    - The event "flip coin" as arrangement of two sub-events
  - roll die: $r = (\frac{1}{6}, \ldots, \frac{1}{6}) \in \mathcal{P}_6$.
  - pick card: $p = (\frac{1}{52}, \ldots, \frac{1}{52}) \in \mathcal{P}_{52}$.
- The nesting rule is multiplication:
  - Flip a coin: result decides whether to roll a die or pick a card.

$$f \circ (r, p) = \Bigg(\underbrace{\frac{1}{12}, \ldots, \frac{1}{12}}_{6 \text{ times}}, \underbrace{\frac{1}{104}, \ldots, \frac{1}{104}}_{52 \text{ times}}\Bigg) \in \mathcal{P}_{58}$$

A zoo of operads: Grammars

Any context-free grammar is an operad.

⟨sentence⟩	::=	⟨noun-phrase⟩ ⟨verb-phrase⟩
⟨noun-phrase⟩	::=	⟨pronoun⟩ | ⟨proper-noun⟩ | ⟨determiner⟩ ⟨nominal⟩
⟨nominal⟩	::=	⟨noun⟩ | ⟨nominal⟩ ⟨noun⟩
⟨verb-phrase⟩	::=	⟨verb⟩ | ⟨verb⟩ ⟨noun-phrase⟩ | ⟨verb⟩ ⟨prep-phrase⟩
⟨prep-phrase⟩	::=	⟨preposition⟩ ⟨noun-phrase⟩

Of course, it doesn't have to be English

- Grammars are used to design custom languages.
- Legal arguments, arithmetic expressions, programming languages.

And how is this an operad?

- The sorts are the parts of speech.
- The arrangements are the production rules.
- Nesting is given by "syntax trees".

The operad of sets

There is an operad **Set** of sets and functions.

- The sorts in **Set** are all the sets, e.g. $\{\mathtt{r}, \mathtt{g}, \mathtt{b}\}$, $\mathbb{N}$, $\mathbb{R}^2$, etc.
- An arrangement is a function $f\colon X_1 \times \cdots \times X_k \to Y$.
- Nesting is composition $f \circ (g_1, \ldots, g_k)$.

The operad of sets

There is an operad **Set** of sets and functions.

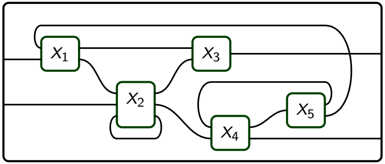
- The sorts in **Set** are all the sets, e.g. $\{\mathtt{r}, \mathtt{g}, \mathtt{b}\}$, $\mathbb{N}$, $\mathbb{R}^2$, etc.
- An arrangement is a function $f \colon X_1 \times \cdots \times X_k \to Y$.
- Nesting is composition $f \circ (g_1, \ldots, g_k)$.

Think of sets as containing features one can sort into.

- An arrangement is a clustering of a k-dimensional feature space.

Mode-dependent networks of dynamical systems

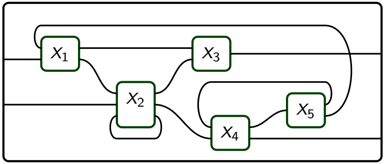
Recall the wiring diagrams operad, used in dynamical systems.



There's an operad where the arrangement changes in time.

Mode-dependent networks of dynamical systems

Recall the wiring diagrams operad, used in dynamical systems.

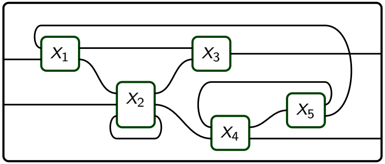


There's an operad where the arrangement changes in time.

- Each box has a bunch of "communicative modes".

Mode-dependent networks of dynamical systems

Recall the wiring diagrams operad, used in dynamical systems.

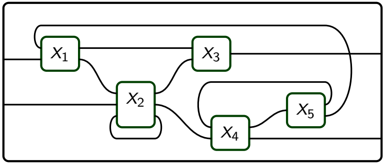


There's an operad where the arrangement changes in time.

- Each box has a bunch of “communicative modes”.
- Its communicative mode determines its number of inputs and outputs.

Mode-dependent networks of dynamical systems

Recall the wiring diagrams operad, used in dynamical systems.

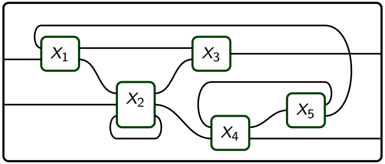


There's an operad where the arrangement changes in time.

- Each box has a bunch of "communicative modes".
- Its communicative mode determines its number of inputs and outputs.
- The modes of all the boxes together determine that of the outer box...
- ... and the wiring diagram by which they are arranged.

Mode-dependent networks of dynamical systems

Recall the wiring diagrams operad, used in dynamical systems.



There's an operad where the arrangement changes in time.

- Each box has a bunch of "communicative modes".
- Its communicative mode determines its number of inputs and outputs.
- The modes of all the boxes together determine that of the outer box...
- ... and the wiring diagram by which they are arranged.

As boxes send signals and update their states, their modes change, and hence their arrangement does too.

Outline

CT for a science of interdisciplinarity

Operads can be used for many things.

- People are already using them in many different fields.
- We can describe them all in a common language.
- We can also invent our own new compositional languages.

CT for a science of interdisciplinarity

Operads can be used for many things.

- People are already using them in many different fields.
- We can describe them all in a common language.
- We can also invent our own new compositional languages.

Functors = formal mappings between languages (didn't discuss: time)

- A richly interconnected network of languages.

CT for a science of interdisciplinarity

Operads can be used for many things.

- People are already using them in many different fields.
- We can describe them all in a common language.
- We can also invent our own new compositional languages.

Functors = formal mappings between languages (didn't discuss: time)

- A richly interconnected network of languages.
- Some languages are more expressive, others are easier to reason about.
- Think rounding: real numbers = more expressive, integers = easier.)

CT for a science of interdisciplinarity

Operads can be used for many things.

- People are already using them in many different fields.
- We can describe them all in a common language.
- We can also invent our own new compositional languages.

Functors = formal mappings between languages (didn't discuss: time)

- A richly interconnected network of languages.
- Some languages are more expressive, others are easier to reason about.
- Think rounding: real numbers = more expressive, integers = easier.)

What is the right mathematics to underlie a science of interdisciplinarity?

- Category theory has an impressive record inside of math and out.
- Recently highlighted by US agencies such as NIST and DARPA.
- Whether or not it can really do the job at scale remains to be seen.

CT for a science of interdisciplinarity

Operads can be used for many things.

- People are already using them in many different fields.
- We can describe them all in a common language.
- We can also invent our own new compositional languages.

Functors = formal mappings between languages (didn't discuss: time)

- A richly interconnected network of languages.
- Some languages are more expressive, others are easier to reason about.
- Think rounding: real numbers = more expressive, integers = easier.)

What is the right mathematics to underlie a science of interdisciplinarity?

- Category theory has an impressive record inside of math and out.
- Recently highlighted by US agencies such as NIST and DARPA.
- Whether or not it can really do the job at scale remains to be seen.

But if you want to work towards a science of interdisciplinarity,

CT for a science of interdisciplinarity

Operads can be used for many things.

- People are already using them in many different fields.
- We can describe them all in a common language.
- We can also invent our own new compositional languages.

Functors = formal mappings between languages (didn't discuss: time)

- A richly interconnected network of languages.
- Some languages are more expressive, others are easier to reason about.
- Think rounding: real numbers = more expressive, integers = easier.)

What is the right mathematics to underlie a science of interdisciplinarity?

- Category theory has an impressive record inside of math and out.
- Recently highlighted by US agencies such as NIST and DARPA.
- Whether or not it can really do the job at scale remains to be seen.

But if you want to work towards a science of interdisciplinarity,
Category Theory should at least be on your radar.

CT for a science of interdisciplinarity

Operads can be used for many things.

- People are already using them in many different fields.
- We can describe them all in a common language.
- We can also invent our own new compositional languages.

Functors = formal mappings between languages (didn't discuss: time)

- A richly interconnected network of languages.
- Some languages are more expressive, others are easier to reason about.
- Think rounding: real numbers = more expressive, integers = easier.)

What is the right mathematics to underlie a science of interdisciplinarity?

- Category theory has an impressive record inside of math and out.
- Recently highlighted by US agencies such as NIST and DARPA.
- Whether or not it can really do the job at scale remains to be seen.

But if you want to work towards a science of interdisciplinarity,
Category Theory should at least be on your radar.

Thanks! Questions and comments welcome.