Category theory for cognitive science

Some formal and conceptual analogies (abridged version)

Steven Phillips

Human Informatics and Interaction Research Institute (HIIRI)
National Institute of Advanced Industrial Science and Technology (AIST)
Tsukuba, Japan

(CogSci2022, Toronto)

July 27th, 2022

Categories and cognition

1. Introduction
   - Objectives: What/Why/How of category theory
   - Perspective: Category theory as theory of formal analogies
   - Where category theory *meets* cognitive science: *compositionality*
2. Basic correspondences
   - Categories and composition
   - Functors and representation
   - Natural transformations and re-representation (computation)
3. Universal cognition
   - Universal morphisms (limits) and compositionality
   - Universals and *systematicity*
4. Discussion
   - The *universal mapping principle* for cognitive science

Outline

1. Introduction
   - Objectives: What/Why/How of category theory
   - Perspective: Category theory as theory of formal analogies
   - Where category theory *meets* cognitive science: *compositionality*
2. Basic correspondences
   - Categories and composition
   - Functors and representation
   - Natural transformations and re-representation (computation)
3. Universal cognition
   - Universal morphisms (limits) and compositionality
   - Universals and *systematicity*
4. Discussion
   - The *universal mapping principle* for cognitive science

Objectives: What/Why/How of categories

Objectives: basic answers to three basic questions:

1. What is category theory?
   - A theory of structure
   - A theory of analogy
   - A theory of universal (mapping) properties
2. Why is category theory important (to cognitive science)?
   - A *lingua franca* for cognitive science
   - Explanations without *ad hoc* assumptions
3. How is category theory applied (in cognitive science)?
   - A *universal mapping principle* for cognitive science (Phillips, 2021a)

Learning objectives: schedule

Schedule (category theory concepts):

1. Session 1: Basics – categories, functors, natural transformations
   - categories and compositionality
   - functors and representation
   - natural transformations and re-representation (computation)
2. Session 2: Universal constructions
   - universal morphisms
   - limits/colimits
   - recursion/corecursion (iteration)

✍ Linking concept: (commutative) square of relations, □

Outline

1. Introduction
   - Objectives: What/Why/How of category theory
   - Perspective: Category theory as theory of formal analogies
   - Where category theory *meets* cognitive science: *compositionality*
2. Basic correspondences
   - Categories and composition
   - Functors and representation
   - Natural transformations and re-representation (computation)
3. Universal cognition
   - Universal morphisms (limits) and compositionality
   - Universals and *systematicity*
4. Discussion
   - The *universal mapping principle* for cognitive science

Perspective: Category theory as formal analogies

Conceptual analogies:

- Pen is to paper as chisel is to stone
- Category Theory is to Mathematics as Analogy Theory is to Cognitive Science

Compare:

- Structure Mapping Theory (Gentner, 1983) – analogy ($map : Source \rightarrow Target$)

$$\begin{tikzcd} pen \arrow[r, "map"] \arrow[d, "writes\ on"'] & \text{chisel} \arrow[d, "engraves\ on"] \\ paper \arrow[r, "map"'] & \text{stone} \end{tikzcd}$$

- Category Theory (Eilenberg & Mac Lane, 1945) – natural transformation ($\eta : F \dot{\rightarrow} G$)

$$\begin{tikzcd} F(A) \arrow[r, "\eta_A"] \arrow[d, "F(f)"'] & G(A) \arrow[d, "G(g)"] \\ F(B) \arrow[r, "\eta_B"'] & G(B) \end{tikzcd}$$

Outline

Where category theory *meets* cognitive science

Compositionality

Category theory as a formal framework for compositionality, e.g.,

$$\begin{array}{ccc} \textit{John} & \xrightarrow{\rho} & \text{John} \\ {\scriptstyle \textit{loves}}\big\downarrow & & \big\downarrow{\scriptstyle \text{loves}} \\ \textit{Mary} & \xrightarrow[\rho]{} & \text{Mary} \end{array}$$

Dimensions of compositionality

Three dimensions of composition:

- vertical: composition of arrows (within category)
- horizontal: composition of natural transformations
- out-of-plane (not shown): composition of functors

$$
\begin{tikzcd}
A \arrow[d, "f"] \arrow[dd, bend right, "g\circ f"'] \arrow[dr, dotted, mapsto] \arrow[drr, dotted, mapsto] \arrow[drrr, dotted, mapsto] & & & \\
B \arrow[d, "g"] \arrow[dr, dotted, mapsto] \arrow[drr, dotted, mapsto] \arrow[drrr, dotted, mapsto] & F(A) \arrow[r, "\alpha_A"] \arrow[d, "F(f)"] \arrow[dd, bend right, "F(g\circ f)"'] \arrow[rr, bend left, "(\beta\circ\alpha)_A"] & G(A) \arrow[r, "\beta_A"] \arrow[d, "G(f)"] & H(A) \arrow[d, "H(f)"'] \arrow[dd, bend left, "H(g\circ f)"] \\
C \arrow[dr, dotted, mapsto] \arrow[drr, dotted, mapsto] \arrow[drrr, dotted, mapsto] & F(B) \arrow[r, "\alpha_B"'] \arrow[d, "F(g)"] & G(B) \arrow[r, "\beta_B"'] \arrow[d, "G(g)"] & H(B) \arrow[d, "H(g)"'] \\
& F(C) \arrow[r, "\alpha_C"'] \arrow[rr, bend right, "(\beta\circ\alpha)_C"'] & G(C) \arrow[r, "\beta_C"'] & H(C)
\end{tikzcd}
$$

Commutative triangles and universal structures

$$
\begin{tikzcd}
A \arrow[d, "f"'] \arrow[dr, dotted] \arrow[drr, dotted] & & \\
B \arrow[dr, dotted] \arrow[drr, dotted] & X \arrow[r, "\eta_A"] \arrow[d, "1_X"'] & G(A) \arrow[d, "G(f)"] \\
 & X \arrow[r, "\eta_B"'] & G(B)
\end{tikzcd}
$$

triangles are special “squares” – cf. diagram for a *universal mapping property*:

$$
\begin{tikzcd}
X \arrow[r, "\phi"] \arrow[dr, "f"'] & G(A) \arrow[d, dashed, "G(u)"] & A \arrow[d, dashed, "u"] \\
 & G(B) & B
\end{tikzcd}
$$

every object X has an identity arrow 1_X (usually omitted from diagrams)

Outline

Categories and compositionality

Compositionality (basic): method for putting two things together to form something

In cognitive science (classical version):

- Principle: *composite symbol* built up from *constituent symbols* and their *relations*
- Example: *red circle* built from *red* and *circle* and their order ($\neq$ *circle red*)
- Motivation: *systematicity* and *productivity* properties of cognition

In category theory (basic form):

- Principle: *composite arrow* built up from *arrows* and *composition operation* ($\circ$)
- Example: $g \circ f$ built from f and g applied to $\circ$ ($\neq f \circ g$)
- Motivation: complex "structure" as composition of arrows

Category (definition)

A category $\mathbf{C} = (\mathbf{C}_0, \mathbf{C}_1, dom, cod, id, \circ)$ consists of:

- *objects*, $\mathbf{C}_0 = \{A, B, C, \dots\}$
- *arrows*, $\mathbf{C}_1 = \{f, g, h, \dots\}$—arrow $f : A \to B$ from *domain* A to *codomain* B
  - including *identity arrow* for each object: $id : A \mapsto (1_A : A \to A)$
- *domain/codomain maps*: $dom(f) = A$, $cod(f) = B$
- *composition operation*, i.e. $f : A \to B$ composed with $g : B \to C$ is $g \circ f : A \to C$
  - $1_B \circ f = f = f \circ 1_A$ – compare $1 \cdot x = x = x \cdot 1$
  - $h \circ (g \circ f) = (h \circ g) \circ f$ – compare $x \cdot (y \cdot z) = (x \cdot y) \cdot z$

✍ Identity arrows are usually omitted from diagrams

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "g \circ f"'] & B \arrow[d, "g"] \arrow[dl, phantom, "\Swarrow"] \\ & C \end{tikzcd} \quad \text{same as} \quad \begin{tikzcd} A \arrow[r, "f"] \arrow[d, "1_A"'] & B \arrow[d, "g"] \\ A \arrow[r, "g \circ f"'] & C \end{tikzcd}$$

Categories and composition (example: orders)

An ordered set is a category:

- *Ann is shorter (not taller) than Ben*—transitive

$$\begin{tikzcd} \text{Ann} \arrow[r, "\text{shorter}"] \arrow[rr, bend right, "\text{shorter}"'] & \text{Ben} \arrow[r, "\text{shorter}"] & \text{Caz} \end{tikzcd}$$

- *Ann is not taller than Ann*—reflexive: Ann $\leq$ Ann

Formally, an ordered set $(P, \leq)$ corresponds to a category:

- objects: $p \in P$
- arrows: $p \to q$ whenever $p \leq q$
  - identities: $p \leq p$, i.e. $id : p \mapsto (\leq_p : p \to p)$
- domain, codomain: $dom(\leq_{pq}) = p$, $cod(\leq_{pq}) = q$
- composition: $p \leq q$ composed with $q \leq r$ is $p \leq r$

Natural numbers $\mathbb{N} = \{0, 1, 2, \dots\}$ constitute category $(\mathbb{N}, \leq)$.

Categories and composition (example: operations)

Addition of natural numbers $(\mathbb{N}, +, 0)$—called a *monoid*—as a category

- $1 + 2 = 3$ (composition)

$$\begin{tikzcd} * \arrow[r, "1"] \arrow[rr, bend right, "3"'] & * \arrow[r, "2"] & * \end{tikzcd}$$

- $0 + x = x$ (identity)

$$\begin{tikzcd} * \arrow[r, "0"] \arrow[rr, bend right, "x"'] & * \arrow[r, "x"] & * \end{tikzcd}$$

The category has

- one object: $*$ (name unimportant)
- one arrow: $x : * \to *$ for each number $x \in \mathbb{N}$ (identity arrow is $0 : * \to *$)
- domain, codomain: $dom(x) = *$, $cod(x) = *$
- composition is addition: $n \circ m$ is $m + n$, which is
  - *associative*: $x + (y + z) = (x + y) + z$
  - *unital*: $x + 0 = x = 0 + x$

Categories and composition (e.g., sets and functions)

Sets and functions make up a category: **Set**

- objects: sets, $A = \{a, a', \dots\}$, $B = \{b, b', \dots\}$
- arrows: functions, $f : A \to B; a \mapsto b, a' \mapsto b', \dots$
  - identities are identity functions, $1_A : a \mapsto a$
- domain, codomain: $dom(f) = A$, $cod(f) = B$
- composition: composition of functions, $g \circ f : a \mapsto g(f(a))$

✍ Objects and arrows have *internal* structure: the actions on elements, $a \mapsto f(a)$

Contrast categories: set (S) *as* category vs. category *of* sets

- set (S): objects are elements $a \in S$; arrows are (only) identities $1_a : a \to a$
- **Set**: objects are sets; arrows are functions, $f : S \to T; a \mapsto f(a)$

✍ Notational difference: $\to$ (arrow) vs. $\mapsto$ (action)

Look ahead: monoid *as* category vs. category *of* monoids

Categories and structures as arrows (e.g., monoids)

Equivalently, a monoid (M, μ, η) is a set M and a pair of arrows:

- a *binary* function, $\mu : M \times M \to M$ and
- a *nullary* function $\eta : 1 \to M$ picking out the unit

expressed as the diagram

$$M \times M \xrightarrow{\mu} M \xleftarrow{\eta} 1$$

or, the single arrow

$$M \times M + 1 \xrightarrow{\mu+\eta} M$$

For example, addition of natural numbers is the monoid $(\mathbb{N}, \mu, \eta)$, where

- $\mu : (m, n) \mapsto m + n$
- $\eta : * \mapsto 0.$

Category structure as arrow

A category as the diagram of arrows ($d = id \circ dom$, $c = id \circ cod$):

$$\begin{tikzcd} \mathbf{C}_1 \times_{\mathbf{C}_0} \mathbf{C}_1 \arrow[r, "\circ_{\mathbf{C}}"] & \mathbf{C}_1 & \mathbf{C}_0 \arrow[l, "id"'] \\ & \mathbf{C}_1 \arrow[u, shift left, "d"] \arrow[u, shift right, "c"'] \arrow[ur, shift left, "dom"] \arrow[ur, shift right, "cod"'] & \end{tikzcd}$$

Category structure as one arrow ($\mathbf{C}_{\bullet} = \mathbf{C}_1 \times_{\mathbf{C}_0} \mathbf{C}_1 + 2\mathbf{C}_1 + \mathbf{C}_0$):

$$\mathbf{C}_{\bullet} \xrightarrow{\circ_{\mathbf{C}} + d + c + id} \mathbf{C}_1$$

Category of arrows *as* objects

A category of arrows and squares:

- objects are arrows: $\alpha : A \to A'$
- arrows are squares: pairs (f, f') making the squares commute:

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "\alpha"'] & B \arrow[d, "\beta"] \\
A' \arrow[r, "f'"'] & B'
\end{tikzcd}$$

- composition is "pasting" of squares:

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "\alpha"'] \arrow[rr, bend left, "g \circ f"] & B \arrow[r, "g"] \arrow[d, "\beta"] & C \arrow[d, "\gamma"] \\
A' \arrow[r, "f'"'] \arrow[rr, bend right, "g' \circ f'"'] & B' \arrow[r, "g'"'] & C'
\end{tikzcd}$$

Categories of structures (e.g., monoids)

Recall monoid structure as arrow (now an object):

$$M \times M + 1 \xrightarrow{\mu+\eta} M$$

object is arrow (structure), arrow is square (*structure-preserving* map):

$$\begin{tikzcd} M \times M + 1 \arrow[r, "h \times h + 1"] \arrow[d, "\mu+\eta"'] & N \times N + 1 \arrow[d, "\mu+\eta"] \\ M \arrow[r, "h"'] & N \end{tikzcd}$$

i.e. a monoid homomorphism $h : M \to N$ (cf. algebraic definition).

A *monoid homomorphism* is a function $h : M \to N$ such that for all $a, b \in M$

- $h(a \cdot b) = h(a) \cdot h(b)$
- $h(e_M) = e_N$

Look ahead: category as structure

Outline

Functors and representation

Representation (basic): one thing standing in for another thing

In cognitive science (classical version):

- Principle: semantic relations between parts “mirrored” by syntactic relations between corresponding symbols—*partial algebra homomorphism*
- Example: *circle is left of square* corresponds to left-of(circle, square)
- Motivation: *systematicity of representation* properties

In category theory (functor version):

- Principle: structure-preserving map (functor), *category homomorphism*, $F(g \circ f) = F(g) \circ F(f)$
- Example: composite arrow $g \circ f$ is represented as $F(g \circ f)$ by the composite of constituent arrow representations $F(f)$ and $F(g)$
- Motivation: represent objects A and arrows $f : A \to B$ in one domain as objects $F(A)$ and arrows $F(f) : F(A) \to F(B)$ in another (co)domain

Functor (definition)

A functor $F : \mathbf{C} \to \mathbf{D}$ is a (structure-preserving) map sending the objects and arrows in $\mathbf{C}$ to objects and arrows in $\mathbf{D}$ that preserves:

- domains, codomains: $F(dom(f)) = dom(F(f))$, $F(cod(f)) = cod(F(f))$
- identities: $F(1_A) = 1_{F(A)}$
- composition: $F(g \circ f) = F(g) \circ F(f)$

$$
\begin{tikzcd}
A \arrow[r, dotted] \arrow[d, "f"] \arrow[dd, bend right, "g \circ f"'] & F(A) \arrow[d, "F(f)"'] \arrow[dd, bend left, "F(g \circ f)"] \\
B \arrow[r, dotted] \arrow[d, "g"] & F(B) \arrow[d, "F(g)"'] \\
C \arrow[r, dotted] & F(C)
\end{tikzcd}
$$

Functors are structure-preserving maps

A functor $F : \mathbf{C} \to \mathbf{D}$ is a pair of maps $(F_0, F_1) : (\mathbf{C}_0, \mathbf{C}_1) \to (\mathbf{D}_0, \mathbf{D}_1)$ preserving:
domains/codomains:

$$\begin{tikzcd} \mathbf{C}_1 \arrow[r, "F_1"] \arrow[d, shift right, "dom"'] \arrow[d, shift left, "cod"] & \mathbf{D}_1 \arrow[d, shift right, "dom"'] \arrow[d, shift left, "cod"] \\ \mathbf{C}_0 \arrow[r, "F_0"'] & \mathbf{D}_0 \end{tikzcd}$$

identities:

$$\begin{tikzcd} \mathbf{C}_0 \arrow[r, "F_0"] \arrow[d, "id"'] & \mathbf{D}_0 \arrow[d, "id"] \\ \mathbf{C}_1 \arrow[r, "F_1"'] & \mathbf{D}_1 \end{tikzcd}$$

compositions:

$$\begin{tikzcd} \mathbf{C}_1 \times_{\mathbf{C}_0} \mathbf{C}_1 \arrow[r, "F_1 \times F_1"] \arrow[d, "\circ_{\mathbf{C}}"'] & \mathbf{D}_1 \times_{\mathbf{D}_0} \mathbf{D}_1 \arrow[d, "\circ_{\mathbf{D}}"] \\ \mathbf{C}_1 \arrow[r, "F_1"'] & \mathbf{D}_1 \end{tikzcd}$$

Functors are category homomorphisms

Recall category structure as arrow (now an object):

$$\mathbf{C}_\bullet \xrightarrow{\circ_{\mathbf{C}}+d+c+id} \mathbf{C}_1$$

object is arrow (structure), arrow is square (*structure-preserving* map):

$$\begin{tikzcd}
\mathbf{C}_\bullet \arrow[r, "F_1\times F_1+2F_1+F_0"] \arrow[d, "\circ_{\mathbf{C}}+d+c+id"'] & \mathbf{D}_\bullet \arrow[d, "\circ_{\mathbf{D}}+d+c+id"] \\
\mathbf{C}_1 \arrow[r, "F_1"'] & \mathbf{D}_1
\end{tikzcd}$$

i.e. a category homomorphism $F : \mathbf{C} \to \mathbf{D}$ (cf. algebraic definition).

A functor $F : \mathbf{C} \to \mathbf{D}$ is a (structure-preserving) map sending the objects and arrows in $\mathbf{C}$ to objects and arrows in $\mathbf{D}$ that preserves:

- domains, codomains: $F(dom(f)) = dom(F(f))$, $F(cod(f)) = cod(F(f))$
- identities: $F(1_A) = 1_{F(A)}$
- composition: $F(g \circ f) = F(g) \circ F(f)$

Interim: arrow $\rightarrow$ object

Internalization/chunking: enclosing (hiding) structure

$$
\begin{tikzcd}
\text{arrows} \arrow[rr, bend left, "\textit{objectify}"] & & \text{objects} \arrow[ll, bend left, dotted]
\end{tikzcd}
$$

Compare:

- arrows as objects
- verbs as nouns
- actions as states
- lines as points
- squares as lines (between lines)

Outline

Natural transformations and re-representation

Computation (basic): transformations of representations

In cognitive science (classical version):

- Principle: symbolic representations and processes
- Example: What is to the left of circle and square? circle
- Motivation: *systematicity of inference* properties

In category theory (natural transformation version):

- Principle: commutative square (representation homomorphism)
- Example: composites to constituents, $A \times B \to A$
- Motivation: naturality, generality (transcend specifics)

Natural transformation (definition)

A natural transformation $\eta : F \stackrel{\cdot}{\to} G$ is a family of arrows, $\eta_A : F(A) \to G(A)$, making the following diagram (square) commute:

$$
\begin{tikzcd}
A \arrow[d, "f"'] \arrow[dr, dotted] \arrow[drr, dotted] & & \\
B \arrow[dr, dotted] \arrow[drr, dotted] & F(A) \arrow[r, "\eta_A"] \arrow[d, "F(f)"'] & G(A) \arrow[d, "G(f)"] \\
 & F(B) \arrow[r, "\eta_B"'] & G(B)
\end{tikzcd}
$$

✍ Functors $F : \mathbf{C} \to \mathbf{D}$ and $G : \mathbf{C} \to \mathbf{D}$ are from the *same domain* to the *same codomain*

Nat. trans. (example): approximation

The *ceiling function* sends each real number x to the smallest integer not less than x: as a functor, $\lceil\cdot\rceil : \mathbb{R} \to \mathbb{Z}; x \mapsto \lceil x \rceil, (x \leq y) \mapsto (\lceil x \rceil \leq \lceil y \rceil)$.

Example: $2.1 \mapsto 3$ and $(2.9 \leq 3.1) \mapsto (3 \leq 4)$

The ceiling function as a natural transformation: $1_{\mathbb{R}} \dot{\to} \iota \circ \lceil\cdot\rceil$, where $1_{\mathbb{R}}$ is the identity functor on the category of ordered reals and ι is the injection of integers, $\iota : \mathbb{Z} \to \mathbb{R}$

$$
\begin{tikzcd}
x \arrow[d, "\leq"'] \arrow[dr, dotted] \arrow[drr, dotted] & & \\
y \arrow[dr, dotted] \arrow[drr, dotted] & x \arrow[r, "\leq"] \arrow[d, "\leq"'] & \lceil x \rceil \arrow[d, "\leq"] \\
& y \arrow[r, "\leq"'] & \lceil y \rceil
\end{tikzcd}
$$

Nat. trans. (example): parts and wholes

Functors: product – $\Pi : (A, B) \mapsto A \times B$; projection – $\acute{\Pi} : (A, B) \mapsto A$

Natural transformation, $\acute{\pi} : \Pi \dot{\rightarrow} \acute{\Pi}$ (cf. $A \wedge B \Rightarrow A$)

$$\begin{tikzcd} (A, B) \arrow[d, "{(f,g)}"'] \\ (C, D) \end{tikzcd} \qquad \begin{tikzcd} A \times B \arrow[r, "\acute{\pi}_{A,B}"] \arrow[d, "f \times g"'] & A \arrow[d, "f"] \\ C \times D \arrow[r, "\acute{\pi}_{C,D}"'] & C \end{tikzcd}$$

✍ Compare with *injection*; coproduct – $\amalg : (A, B) \mapsto A + B$

Natural transformation, $\acute{\iota} : \acute{\Pi} \dot{\rightarrow} \amalg$ (cf. $A \Rightarrow A \vee B$)

$$\begin{tikzcd} (A, B) \arrow[d, "{(f,g)}"'] \\ (C, D) \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r, "\acute{\iota}_{A,B}"] \arrow[d, "f"'] & A + B \arrow[d, "f+g"] \\ C \arrow[r, "\acute{\iota}_{C,D}"'] & C + D \end{tikzcd}$$

Natural transformation: commutative triangles

Constant functor – X : $A \mapsto X, f \mapsto 1_X$

$$\begin{tikzcd}
A \arrow[d, "f"'] \arrow[dr, dotted] \arrow[drr, dotted] & & \\
B \arrow[dr, dotted] \arrow[drr, dotted] & X \arrow[r, "\eta_A"] \arrow[d, "1_X"'] & G(A) \arrow[d, "G(f)"] \\
& X \arrow[r, "\eta_B"'] & G(B)
\end{tikzcd}$$

simplifies to

$$\begin{tikzcd}
A \arrow[d, "f"'] \arrow[dr, dotted] \arrow[drr, dotted] & & \\
B \arrow[r, dotted] \arrow[drr, dotted] & X \arrow[r, "\eta_A"] \arrow[dr, "\eta_B"'] & G(A) \arrow[d, "G(f)"] \\
& & G(B)
\end{tikzcd}$$

Natural transformation (example): least element

Zero (0) is the *least* natural number ($\mathbb{N}$) – as natural transformation between functors:

- Zero – $0 : \mathbb{N} \to \mathbb{N}; x \mapsto 0$
- Identity – $1_{\mathbb{N}} : \mathbb{N} \to \mathbb{N}; x \mapsto x$

Natural transformation $0 \leq 1_{\mathbb{N}}$ (cf. $0 \leq x$ for all $x \in \mathbb{N}$)

$$\begin{tikzcd} 0 \arrow[r, "\leq"] \arrow[dr, "\leq"'] & x \arrow[d, "\leq"] \\ & y \end{tikzcd}$$

Outline

Universal morphisms and limits

Category theory constructions are typically defined in terms of *universal (mapping) properties*, i.e. a property shared by all instances of some type.

A *universal mapping property*:

- a map that is *common (universal)* to *all* instances in given context and
- a map that is *unique (specific)* to *each* instance

Some intuition: travelling to the conference

Universal morphism (definitions)

<u>Primal form</u>: A *universal morphism* from object X to functor $F : \mathbf{B} \leftarrow \mathbf{A}$ is a pair (A, α) making the following diagram (triangle) commute (i.e. the *unique-existence* condition):

$$\begin{tikzcd} X \arrow[r, "\alpha"] \arrow[dr, "g"'] & F(A) \arrow[d, dashed, "F(u)"] & A \arrow[d, dashed, "u"] \\ & F(Y) & Y \end{tikzcd}$$

✍ unique-existence: for every Y and every g the exists a unique (dashed) arrow u

<u>Dual form</u>: A *universal morphism* from functor $F : \mathbf{B} \rightarrow \mathbf{A}$ to object Y is a pair (B, β) making the following diagram (triangle) commute (i.e. the *unique-existence* condition):

$$\begin{tikzcd} X \arrow[d, dashed, "u"'] & F(X) \arrow[d, dashed, "F(u)"'] \arrow[dr, "f"] & \\ B & F(B) \arrow[r, "\beta"'] & Y \end{tikzcd}$$

Universal morphism (examples): closest element

Primal form:

The *ceiling* of x (e.g., 2.1) is the smallest integer not less than x (i.e. 3)

$$\begin{tikzcd} 2.1 \arrow[r, "\leq"] \arrow[dr, "\leq"'] & 3 \arrow[d, dashed, "\leq"] \\ & y \end{tikzcd} \qquad \begin{tikzcd} 3 \arrow[d, dashed, "\leq"] \\ y \end{tikzcd}$$

i.e. the pair $(3, \leq)$ from 2.1 to the inclusion functor $\iota : (\mathbb{R}, \leq) \leftarrow (\mathbb{Z}, \leq)$.

Dual form:

The *floor* of y (e.g., 4.9) is the largest integer not greater than y (i.e. 4)

$$\begin{tikzcd} x \arrow[d, dashed, "\leq"'] \\ 4 \end{tikzcd} \qquad \begin{tikzcd} x \arrow[d, dashed, "\leq"'] \arrow[dr, "\leq"] & \\ 4 \arrow[r, "\leq"'] & 4.9 \end{tikzcd}$$

i.e. the pair $(4, \leq)$ from the inclusion functor $\iota : (\mathbb{Z}, \leq) \rightarrow (\mathbb{R}, \leq)$ to 4.9.

Limits and composition (of arrows *and* objects)

Limits constitute a general class of universal morphisms, seen here as another form of compositionality, e.g. a limit of A and B is their product, $A \times B$, and two maps retrieving A and B from the product, i.e. $A \times B \rightarrow A$ and $A \times B \rightarrow B$

A *limit* is the "best" way to pick out (reference) an arrangement of objects and arrows in some category:

- a map (cone) that is *common (universal)* to *all* references in given context and
- a map that is *unique (specific)* to *each* reference

Example: all pairs of maps $f_A : Z \rightarrow A$ and $f_B : Z \rightarrow B$ pass through product P (written $A \times B$) and maps $\pi_A : A \times B \rightarrow A$ and $\pi_B : A \times B \rightarrow B$, i.e. $f = \pi \circ u$, where $f = (f_A, f_B)$

$$\begin{tikzcd} & Z \arrow[dl, "f_A"'] \arrow[d, dashed, "u"] \arrow[dr, "f_B"] & \\ A & A \times B \arrow[l, "\pi_A"] \arrow[r, "\pi_B"'] & B \end{tikzcd}$$

Limits: Compositionality as universal construction

A *limit* is like an optimal focus of attention: the “best” way to pick out a collection of objects and morphisms of shape J in a category **C**:

Example: a *product* is the best way to pick out a *pair* of objects A and B in **C**

Construction	Category theory	Concept (attention)	Product
shape	category	window of attention	$2 = (\cdot, \cdot)$
diagram	functor	contents of attention	$(A, B) : 2 \to \mathbf{C}$
cone	(vertex, legs/nt.)	spotlight of attention	$(Z, (f_A, f_B))$
cone homomorph.	map (of cones)	shift of perspective	$u : Z \to A \times B$
limit (univ. cone)	univ. morphism	optimal perspective	$(A \times B, (\pi_A, \pi_B))$

Some intuition: spotlight of attention

✍ All limits have this form: universal cone to a J-shaped diagram in **C**

Universals and limits (diagrams)

Diagrams are formal constructions used to reference a part of a category

A *diagram* D of *shape* J in a category $\mathbf{C}$ is a functor $D : J \to \mathbf{C}$.

Some examples of diagrams are:

- *empty*, $D_\emptyset : 0 \to \mathbf{C}$, i.e. no objects or arrows
- *point*, $X : (\cdot) \to \mathbf{C}$, the object X and its identity 1_X
- *pair*, $(A, B) : (\cdot, \cdot) \to \mathbf{C}$, the pair of objects A and B (and their identities)
- arrow, $f : (\downarrow) \to \mathbf{C}$, the arrow $f : A \to B$

$\mathbf{C}^J$ is the category of J-shaped diagrams; e.g., $\mathbf{C}^\downarrow$ is the category of arrows

The diagonal functor $\Delta : \mathbf{C} \to \mathbf{C}^J$ sends each object A to a J-shaped diagram of A

Example: $\Delta : \mathbf{C} \to \mathbf{C}^\downarrow$ sends A to its identity arrow, i.e. $\Delta(A) = (1_A : A \to A)$

Universals and limits (definition)

Diagrams (functors, $D : J \to \mathbf{C}$), cones and cone homomorphisms:

A *cone* from a vertex V to a base D (i.e. a J-shaped *diagram*) is a pair (V, ϕ), where ϕ is a natural transformation; a *cone homomorphism* is a map $h : V \to W$ such that

$$\begin{tikzcd} V \arrow[r, "\phi_i"] \arrow[d, "h"'] \arrow[dr, "\phi_j"] & D_i \arrow[d, "\delta_{ij}"] \\ W \arrow[ur, "\psi_i"] \arrow[r, "\psi_j"'] & D_j \end{tikzcd}$$

Limits to diagrams (D) are *universal* cones:

A *limit* to D is a univeral morphism (L, ϕ) from *diagonal* functor $\Delta : \mathbf{C} \to \mathbf{C}^J$ to D:

$$\begin{tikzcd} V \arrow[d, dashed, "u"'] & \Delta(V) \arrow[d, dashed, "\Delta(u)"'] \arrow[dr, "\phi"] & \\ L & \Delta(L) \arrow[r, "\psi"'] & D \end{tikzcd}$$

Universals and limits (example: products)

A product of objects A and B is a limit to a pair diagram, $(A, B) : 2 \to \mathbf{C}$

A *product* of objects A and B is a product object $A \times B$ and two arrows retrieving A and B, i.e. $\pi_A : A \times B \to A$ and $\pi_B : A \times B \to B$

$$\begin{tikzcd} V \arrow[d, dashed, "\langle f,g \rangle"'] \\ A \times B \end{tikzcd} \qquad \begin{tikzcd} \Delta(V) \arrow[d, dashed, "\Delta\langle f,g \rangle"'] \arrow[dr, "(f,g)"] & \\ \Delta(A \times B) \arrow[r, "(\pi_A,\pi_B)"'] & (A, B) \end{tikzcd}$$

✍ Specific products depend on specific categories

Specific products (examples)

In **Set** (category of sets and functions), a product of sets A and B is their Cartesian product and two projections:

- Cartesian product: $A \times B = \{(a, b) | a \in A, b \in B\}$
- projections: $\pi_A : (a, b) \mapsto a$ and $\pi_B : (a, b) \mapsto b$

In **Set**$^{\subseteq}$ (category of sets and inclusions), a product of sets A and B is their intersection and two insertions:

- Intersection: $A \times B = \{c | c \in A, c \in B\}$
- insertions: $\iota_A : c \mapsto c$ and $\iota_B : c \mapsto c$ (also written $A \cap B \subseteq A$ and $A \cap B \subseteq B$)

Universals and limits (example: terminals)

A terminal (final) object is a limit to an *empty diagram*, $D_\emptyset : 0 \to \mathbf{C}$

A *terminal object* (denoted 1) has unique arrow to it from every object in **C**:

$$\begin{tikzcd} V \arrow[d, dashed, "u"'] & \Delta(V) = D_\emptyset \arrow[d, dashed, "1"'] \arrow[dr, dashed, "1"] & \\ 1 & \Delta(1) = D_\emptyset \arrow[r, dashed, "1"'] & D_\emptyset \end{tikzcd}$$

✍ $D_\emptyset \in \mathbf{C}^0 \cong \mathbf{1}$ (one-object category); unique arrow 1 is empty family of arrows ($\emptyset$); terminal object is pair $(1, \emptyset)$, usually just denoted 1 with triangle ignored

Specific terminal objects depend on specific categories. Contrast

- sets and functions (**Set**): a terminal object is any *singleton set*, $\{*\}$, where $*$ is the only (unnamed) element; the unique arrow is the constant function
- $(P, \leq)$: the terminal object is the *top element* ($\top$), i.e. $p \leq \top$ for all $p \in P$

Universals and colimits (definition)

Diagrams ($D : J \to \mathbf{C}$), cocones and cocone homomorphisms:

A *cocone* from a vertex V to a base D is a pair (V, ϕ), where ϕ is a natural transformation; a *cocone homomorphism* is a map $h : W \to V$ such that

$$
\begin{tikzcd}
D_i \arrow[r, "\psi_i"] \arrow[d, "\delta_{ij}"'] \arrow[dr, "\phi_i"' near end] & W \arrow[d, "h"] \\
D_j \arrow[r, "\phi_j"'] \arrow[ur, "\psi_j" near end] & V
\end{tikzcd}
$$

Colimits to diagrams (D) are *universal* cocones:

A *colimit* to D is a univeral morphism (L, ϕ) from D to diagonal functor $\Delta : \mathbf{C} \to \mathbf{C}^J$:

$$
\begin{tikzcd}
D \arrow[r, "\phi"] \arrow[dr, "\psi"'] & \Delta(L) \arrow[d, dashed, "\Delta(u)"] & L \arrow[d, dashed, "u"] \\
& \Delta(V) & V
\end{tikzcd}
$$

Universals and colimits (example: coproducts)

A coproduct (sum) of objects A and B is a colimit to a pair diagram, $(A, B) : 2 \to \mathbf{C}$

A coproduct of objects A and B is a coproduct object $A + B$ and two arrows inserting A and B, i.e. $\iota_A : A \to A + B$ and $\iota_B : B \to A + B$

$$\begin{tikzcd} (A, B) \arrow[r, "(\iota_A, \iota_B)"] \arrow[dr, "(f, g)"'] & \Delta(A + B) \arrow[d, dashed, "\Delta[f, g]"] \\ & \Delta(V) \end{tikzcd} \qquad \begin{tikzcd} A + B \arrow[d, dashed, "{[f, g]}"] \\ V \end{tikzcd}$$

Specific coproducts depend on specific categories. Contrast

- sets and functions (**Set**): *disjoint union*, i.e. $\{(1, a)|a \in A\} \cup \{(2, b)|b \in B\}$, and *injections*, i.e. $\iota_A : a \mapsto (1, a)$ and $\iota_B : b \mapsto (2, b)$
- sets and inclusions: set union, i.e. $A \cup B$, and *insertions*, i.e. $A \subseteq A \cup B$ and $B \subseteq A \cup B$)

Universals and colimits (example: initials)

An initial (cofinal) object is a colimit to an *empty diagram*, $D_\emptyset : 0 \to \mathbf{C}$

A *initial object* (denoted 0) has unique arrow from it to every object in $\mathbf{C}$:

$$\begin{tikzcd} 0 \arrow[d, dashed, "u"'] \\ V \end{tikzcd}$$

Specific initial objects depend on specific categories. Contrast

- sets and functions (**Set**): the initial object is the empty set ($\emptyset$); the unique arrow is the empty function
- $(P, \leq)$: the initial object is the *bottom element* ($\perp$), i.e. $\perp \leq p$ for all $p \in P$

Outline

Universals and *systematicity*

Systematicity problem: explain *why* having certain cognitive abilities implies having certain other cognitive abilities, e.g., the ability to understand the expression

- "John loves Mary" *if and only if*
- "Mary loves John"

or, conversely, *why not* the case of having the ability to understand the expression

- "John loves Mary" *but not*
- "Mary loves John"

i.e. why cognitive abilities cluster as *equivalence classes* in a certain way.

Classical explanation: combinatorial syntax and semantics (Fodor & Pylyshyn, 1988):

- a *grammar* common to all instances
- an *instantiation* specific to each instance

But, why *that* particular grammar (beyond fitting data)?

- explains how (possible), but not why (necessarily follows)
- grammar chosen to fit data—*ad hoc* (Aizawa, 2003)

Systematicity and universal morphism

Categorical explanation: underlying a systematicity property is a universal morphism (Phillips & Wilson, 2010). Each capacity is composed from

- a *mediating arrow* common to all instances and
- a *unique arrow* specific to each instance

Some intuition: capacities cluster around a universal morphism (common structure)

✍ Every universal morphism is an initial/terminal object in a *comma category*—the "best" one can do for the given situation

Systematicity (example: coloured shapes)

Inferring colours and shapes:

- Colours: red, green, blue; Shapes: circle, triangle, square
- Coloured shapes: ◯, ◯, △, □, . . .
- Colour projection: ◯ $\mapsto$ red, ◯ $\mapsto$ green, △ $\mapsto$ blue, □ $\mapsto$ blue, ...
- Shape projection: ◯ $\mapsto$ circle, ◯ $\mapsto$ circle, △ $\mapsto$ triangle, □ $\mapsto$ square, ...

In **Set**:

- Colours: $C = \{\text{red}, \text{green}, \text{blue}\}$; Shapes: $S = \{\text{circle}, \text{triangle}, \text{square}\}$
- Coloured-shapes: $CS = C \times S$
- Colour projection: $\pi_c : CS \rightarrow C$
- Colour projection: $\pi_s : CS \rightarrow S$
- Product: $(C \times S, (\pi_c, \pi_s))$

Systematicity (empirical test: product map)

Product map ($f \times g$): dashed arrow in the following commutative diagram

$$\begin{tikzcd} A & A \times B \arrow[l, "\pi_A"'] \arrow[r, "\pi_B"] \arrow[d, dashed, "f \times g"] & B \arrow[d, "g"] \\ C & C \times D \arrow[l, "\pi_C"] \arrow[r, "\pi_D"'] & D \end{tikzcd} \qquad \begin{tikzcd} Cue \arrow[d, "map"] \\ Target \end{tikzcd}$$

Empirical test: product of cue-target maps;

- Alphabet (letters): $A = \{k, m, p\}$; $B = \{g, r, v\}$
- Cues (letter pairs): $A \times B$; Targets (coloured shapes): $C \times S$
- Maps: AlphaA-to-Colour: $a2c : A \to C$; AlphaB-to-Shape: $b2s : B \to S$
- Product map: $a2c \times b2s : A \times B \to C \times S$
- Systematicity (?): training map implies testing map (generalization)

Some intuition: from partial map to complete map (Phillips, Takeda, & Sugimoto, 2016)

Categorical vs classical products

Classical compositionality: tokening principle (Fodor & Pylyshyn, 1988)

Constituent tokened whenever complex host is tokened: e.g.,

- $A \times B = \{(a, b) | a \in A, b \in B\}$, $\pi_A : (a, b) \mapsto a$; $\pi_B : (a, b) \mapsto b$

Categorical compositionality: unique-existence condition

Constituent symbol need not be tokened whenever complex host symbol is tokened: e.g.,

- $P = \{n | 1 \leq n \leq |A| \cdot |B|\}$, $\pi'_A : n \mapsto a_i, \ldots$; $\pi'_B : n \mapsto b_j, \ldots$

✍ classical (canonical) product as a special case of categorical product

Comma category: linking naturals and universals

Comma categories are constructed from functors with the same *codomain* category

Suppose functors $S : \mathbf{A} \to \mathbf{C}$ and $T : \mathbf{B} \to \mathbf{C}$. A comma category $(S \downarrow T)$ has:

- objects – a triple (A, B, ϕ) for each object A in $\mathbf{A}$, each object B in $\mathbf{B}$ and each arrow $\phi : S(A) \to T(B)$ in $\mathbf{C}$
- arrows – a pair (f, g) for each arrow $f : A \to A'$ in $\mathbf{A}$ and each arrow $g : B \to B'$ in $\mathbf{B}$ such that the diagram

$$\begin{tikzcd} S(A) \arrow[r, "\phi"] \arrow[d, "S(f)"'] & T(B) \arrow[d, "T(g)"] \\ S(A') \arrow[r, "\phi'"'] & T(B') \end{tikzcd} \qquad \begin{tikzcd} (A, B, \phi) \arrow[d, "(f,g)"] \\ (A', B', \phi') \end{tikzcd}$$

  in $\mathbf{C}$ is a commutative square, and
- composition – (vertical) pasting of compatible commutative squares

Univ. morphism: from X to F is *initial* in $(X \downarrow F)$; from F to X is *terminal* in $(F \downarrow X)$

Comma category (example) – products

A product of A and B is the terminal object $(A \times B, \pi)$ in the comma category $(\Delta \downarrow (A, B))$:

$$\begin{tikzcd} \Delta(V) \arrow[d, dashed, "\Delta\langle f,g\rangle"'] \arrow[dr, "(f,g)"] & \\ \Delta(A \times B) \arrow[r, "\pi"'] & (A, B) \end{tikzcd} \qquad \begin{tikzcd} (V, (f,g)) \arrow[d, dashed, "\langle f,g\rangle"] \\ (A \times B, \pi) \end{tikzcd}$$

$\pi = (\pi_A : A \times B \to A, \pi_B : A \times B \to B)$

For initials/terminals (hence, universal constructions generally):

- *All roads lead to Rome!* – “global optima”

✍ All universals are *unique up to unique isomorphism*: e.g., $(B \times A, \pi')$ is also a product, but only one arrow $\phi : (A \times B, \pi) \cong (B \times A, \pi') : \phi^{-1}$

Outline

Universals and cognition

A *universal mapping principle* for cognitive science (Phillips, 2021a):

- the "best" one can do in the given context
- compositionality, systematicity, productivity as universal (mapping) properties

Some other applications of principle:

- cognitive complexity (Phillips, Wilson, & Halford, 2009)
- learning and generalization (Phillips et al., 2016; Phillips, 2018)
- compositionality (Phillips, 2020)
- relational schema induction (Phillips, 2021b)

Further reading

Some introductions to category theory:

- conceptual (Lawvere & Schanuel, 2009; Simmons, 2011)
- formal (Leinster, 2014; Mac Lane, 1998)
- applied (Fong & Spivak, 2018; Spivak, 2014)
- philosophical (Kromer, 2007; Marquis, 2009)
- computational (Bird & De Moor, 1997; Walters, 1991) and many others ...

References

Aizawa, K. (2003). *The systematicity arguments*. New York, NY: Kluwer Academic. doi: 10.1007/978-1-4615-0275-3

Bird, R., & De Moor, O. (1997). *Algebra of programming*. London, UK: Prentice Hall. doi: 10.5555/248932

Eilenberg, S., & Mac Lane, S. (1945). General theory of natural equivalences. *Transactions of the American Mathematical Society*, *58*(2), 231–294. doi: 10.2307/1990284

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*(1–2), 3–71. doi: 10.1016/0010-0277(88)90031-5

Fong, B., & Spivak, D. I. (2018, October). *Seven sketches in compositionality: An invitation to applied category theory.* (Preprint: arXiv:1803.05316v3) doi: 10.48550/arXiv.1803.05316

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170. doi: 10.1016/S0364-0213(83)80009-3

Kromer, R. (2007). *Tool and object: A history and philosophy of category theory* (Hardcover ed., Vol. 32). Berlin, Germany: Birkhauser. doi: 10.1007/978-3-7643-7524-9

Lawvere, F. W., & Schanuel, S. H. (2009). *Conceptual mathematics: a first introduction to categories*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511804199

Leinster, T. (2014). *Basic category theory* (Vol. 143). Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9781107360068

Mac Lane, S. (1998). *Categories for the working mathematician* (2nd ed.). New York, NY: Springer. doi: 10.1007/978-1-4757-4721-8

Marquis, J.-P. (2009). *From a geometrical point of view: A study of the history and philosophy of category theory* (Hardcover ed.). Cambridge, UK: Springer. doi: 10.1093/philmat/nkq006

Phillips, S. (2018). Going beyond the data as the patching (sheaving) of local knowledge. *Frontiers in Psychology*, *9*, 1926. doi: 10.3389/fpsyg.2018.01926

Phillips, S. (2020). Sheaving—a universal construction for semantic compositionality. *Philosophical Transactions of the Royal Society B*, *375*(1791), 20190303. doi: 10.1098/rstb.2019.0303

Phillips, S. (2021a). A category theory principle for cognitive science: Cognition as universal construction. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, *28*(1), 11–24.

Phillips, S. (2021b). A reconstruction theory of relational schema induction. *PLoS Computational Biology*, *17*(1), e1008641. doi: 10.1371/journal.pcbi.1008641

Phillips, S., Takeda, Y., & Sugimoto, F. (2016). Why are there failures of systematicity? the empirical costs and benefits of inducing universal constructions. *Frontiers in psychology*, *7*, 1310. doi: 10.3389/fpsyg.2016.01310

Phillips, S., & Wilson, W. H. (2010). Categorial compositionality: A category theory explanation for the systematicity of human cognition. *PLoS Computational Biology*, *6*(7), e1000858. doi: 10.1371/journal.pcbi.1000858

Phillips, S., Wilson, W. H., & Halford, G. S. (2009). What do transitive inference and class inclusion have in common? categorical (co)products and cognitive development. *PLoS Computational Biology*, *5*(12), e1000599. doi: 10.1371/journal.pcbi.1000599

Simmons, H. (2011). *An introduction to category theory*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511863226

Spivak, D. I. (2014). *Category theory for the sciences*. Cambridge, MA: MIT Press. doi: 10.5555/2628001

Walters, R. F. C. (1991). *Categories and computer science* (Vol. 28). Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511608872