Title
A formal comparison/contrast of associative and relational learning: a case study of relational schema induction

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN
1069-7977

Author
Phillips, Steven

Publication Date
2021

Peer reviewed

A formal comparison/contrast of associative and relational learning: a case study of relational schema induction

Steven Phillips (steve@ni.aist.go.jp)
National Institute of Advanced Industrial Science and Technology (AIST)
Tsukuba, Ibaraki, 305-8566, Japan

Abstract

Relational schema induction involves a series of learning tasks conforming to a common (group-like) structure. The paradigm contrasts associative versus relational aspects of learning for cognitive, developmental and comparative psychology. Yet, a theory accounting for the relationship between these forms of learning has not been fully developed. We use (mathematical) category theory methods to redress this situation: both forms of learning involve a (universal) construction that differs in terms of "dimensionality", i.e. one-dimensional (associative) versus two-dimensional (relational). Accordingly, the development of relational learning pertains to changes in the dimensionality of the underlying relational schemas induced.

Keywords: associative learning; relational schema induction; supervised learning; category; functor; universal construction

Introduction

The *relational schema induction* paradigm, involving a series of learning tasks, was introduced to contrast associative versus relational models of learning (Halford, Bain, Maybery, & Andrews, 1998) on the basis of *learning transfer*, i.e. a faster rate of learning with tasks learned (Harlow, 1949). Associative models (Miller, Barnet, & Grahame, 1995) suppose stimulus-response co-occurrence drives learning. Relational models, by contrast, suppose some form of *structure mapping* (Gentner, 1983), whereby relations induced from prior tasks are applied in new tasks to correctly predict novel stimulus responses. (See the next section, Figure 1, for a characterization.) Evidently (Halford, Bain, et al., 1998; Halford & Busby, 2007), such *first-trial* performance ruled out associative models as the basis of induction, because novel stimuli were not previously paired with a target response to drive learning.

First-trial performance is a simple criterion for relational processes. However, methodological differences complicate comparisons across species and age-groups (Halford, Wilson, Andrews, & Phillips, 2014): e.g., whether the relations vary over trials. Such differences led some to argue that relational cognition is uniquely human (Penn, Holyoak, & Povinelli, 2008), while others maintain that relational processes are evident in other species, including insects (Giurfa, 2021).

The current state of affairs raises a conundrum for cognitive, comparative and developmental psychology. On one hand, if non-humans and younger age-groups do have a comparable capacity for relations, then why don't they show comparable performance to older humans; on the other hand, if relational cognitive processes are unique to older humans, then how does such a capacity evolve or develop?

As a way of addressing this conundrum, we revisit the relational schema induction paradigm to compare/contrast associative and relational learning. Relational schema induction is attractive in this regard because associative and relational aspects of learning are involved within the same paradigm. Yet, until recently (Phillips, 2021b), a theory of relational schema induction had not been proposed, beyond appealing to some form of structure mapping between tasks (Halford, Bain, et al., 1998). Recent work showed that induction follows from a particular kind of *category theory* (Leinster, 2014; Mac Lane, 1998) construction (Phillips, 2021b), detailed later. However, this work left open a comparison/contrast of the learning aspects of the paradigm, which we take up here.

We proceed by detailing a specific example of the relational schema induction paradigm and the formal methods used for comparison and contrast (Methods). Both forms of learning are different instances of a common (universal) construction, but associative and relational learning differ in terms of their *dimensionality* (Results), which we preview in the next section. The theory also provides the formal connection between these two forms of learning and hence a basis for development (Discussion). Additional details appear in the appendices.

Dimensionality: associative (one) vs. relational (two)

The difference between associative and relational processing is characterized as the dimensionality of maps: associative processes involve "one-dimensional" maps from stimuli to responses; relational processes involve two-dimensional maps between stimulus and response relations (Figure 1). Relation-based learning transfer is afforded because novel association a' obtains from known association a, stimulus relation s and response relation r (i.e. $a' \circ s = r \circ a$). This characterization is made formally precise in the rest of the paper.

$$S \xrightarrow{a} R \qquad\qquad \begin{tikzcd} S \arrow[r, "a"] \arrow[d, "s"'] & R \arrow[d, "r"] \\ S' \arrow[r, "a'"'] & R' \end{tikzcd}$$

Figure 1: Associative ($S \to R$) and relational ($s \to r$) maps.

Methods

We describe an example of relational schema induction and the background theory used for comparison and contrast of the associative and relational learning aspects of this paradigm.

Relational schema induction

Relational schema induction involves a series of cue-target learning tasks conforming to a group-like structure (Halford, Bain, et al., 1998). For example, one task instance consists of stimuli drawn from the set of shapes $Sh_1 = \{\triangle, \square, \heartsuit\}$ and the set of trigrams $Tri_1 = \{\text{BEH}, \text{FUT}, \text{PEJ}\}$ and the cue-target map, $\tau_1 : Sh_1 \times Tri_1 \to Tri_1$, is shown in Table 1 (left): e.g., $\tau_1 : (\square, \text{BEH}) \mapsto \text{FUT}$. The next task instance consists of new sets of stimuli and cue-target map, as shown in Table 1 (right): e.g., $\tau_2 : (\spadesuit, \text{NIZ}) \mapsto \text{HUQ}$. If each trigram locates on a unique vertex of an equilateral triangle, then each shape corresponds to a unique rotation sending trigrams to trigrams. For instance, in the first task, $\triangle$, $\square$ and $\heartsuit$ correspond to 0°, 120° and 240° rotations, respectively. Participants must demonstrate correct responses to each cue for a block of trials for a task instance before proceeding to the next learning task. They are not told the structure of the task, but are given feedback on each trial indicating the target (correct response) for the given cue.

Table 1: Relational schema induction tasks.

τ_1	BEH	FUT	PEJ	τ_2	HUQ	KES	NIZ
$\triangle$	BEH	FUT	PEJ	$\clubsuit$	HUQ	KES	NIZ
$\square$	FUT	PEJ	BEH	$\spadesuit$	KES	NIZ	HUQ
$\heartsuit$	PEJ	BEH	FUT	$\bigstar$	NIZ	HUQ	KES

The shapes can also be interpreted as permutation actions on a row of trigrams. For instance (τ_1), $\triangle$ is the "no change" permutation that leaves the relative positions of each trigram with a row unchanged, $\square$ is "shift left with wrap around" action and *heartsuit* is the "shift right with wrap around" action.

The relevant data pertain to first-trial responses (Halford, Bain, et al., 1998) as a measure of learning transfer across task instances. For the first task instance, first-trial responses are expected to be at chance level, since the relational schema has not been determined at this stage. However, first-trial responses on subsequent tasks are expected to be better than chance once the relational schema has been induced. Indeed, participants demonstrated learning transfer on this structure (relational schema) and other group-like structures (Halford, Bain, et al., 1998; Halford & Busby, 2007)

The crucial observation is that each shape acts on trigrams. By locating the trigrams at the vertices of an (imaginary) equilateral triangle we see that the shapes $\triangle$, $\square$ and $\heartsuit$ correspond to 0°, 120° and 240° degree rotations. Thus, learning transfer on a new task instance is afforded by aligning the shapes to actions on the new trigrams. In this way, feedback on correct responses for two (so-called information) trials is necessary and sufficient to determine the correct responses for the remaining six (shape, trigram) pairs. For example, suppose we learn from feedback that $(\clubsuit, \text{KES}) \mapsto \text{KES}$ and $(\spadesuit, \text{NIZ}) \mapsto \text{HUQ}$, hence $\clubsuit$ and $\spadesuit$ correspond to 0° and 120° rotation, respectively. Thus, we can infer that $(\bigstar, \text{NIZ}) \mapsto \text{KES}$ as the other rotation, and so on.

Background theory

A summary description is given here of the formal theory in the appendices. We proceed in three stages. The first stage formalizes relational schema induction in the more familiar terms of sets and functions (Appendix A). The second stage recasts this basic formalism in terms of categories and functors (Appendix B), providing the framework for our third stage, which is the comparison/contrast in terms of universal constructions as constrained optimization (Appendix C).

Sets. Formally, the target of each learning task is a function from a set of cues to a set of target responses, i.e. a map from (shape, trigram) pairs to trigrams. However, each task contains additional structure: the relationships between the shapes constitute a *monoid* (definition 1) that can be interpreted as *acting* (definition 8) on the trigrams. For example (τ_1, Table 1), $\square$ acts on BEH to produce FUT. All tasks conform to the same action (relational schema), albeit with different shapes and trigrams. These tasks are related by *equivariant maps* (definition 13), essentially preserving this relational structure.

Categories. Sets and functions are instances of categorical constructions. Specifically, a monoid is a *category* (definition 16), the learning tasks as actions are *functors* (definition 22) and the equivariant maps are *natural transformations* (definition 26). The collections of (possible) learning tasks and their equivariant maps constitute a category (remark 30).

Universal construction. The more general, category theory formulation affords a unified view of associative and relational learning as *universal constructions* (definition 36), i.e. the best one can do in the given context. In the current context, induction of relational schemas follows by reconstructing the underlying monoid (theorem 51) from the category of these learning tasks. Using this result, we show how the two forms of learning unify as instances of constrained optimization, i.e. *Kan extensions* (definition 54).

Results

Our results build on work showing that induction of relational schemas is an instance of reconstruction (Phillips, 2021b). The novel aspect of the current work derives from the observation that associative learning corresponds to a special (trivial) case of relational schema induction. The result is two-fold:

1. (in comparison) associative learning of task instances and induction of the common relational schema are both instances of reconstruction, but

2. (by contrast) they differ in terms of the "dimensionality" of the underlying schema, in a formal sense provided here.

We sketch out their comparison/contrast in the rest of this section, and formally characterize their dimensionality.

Comparison. We recall that the relational schema (monoid, in this case) underlying each task is recovered by computing the *end* (definition 47) of a *hom-functor* (examples 49). For the current example, that is the cyclic group that affords learning transfer to new instances of the task. In the associative case, the underlying schema is the *trivial* (one-element) monoid (remark 5), whence the equivariant maps reduce to ordinary functions (remark 28). So, associative learning derives from the same reconstruction process.

Contrast. The trivial monoid, however, does not afford learning transfer, since there are no (non-trivial, i.e. non-identity) relations to base transfer. Consequently, the two forms of learning differ in terms of learning transfer. This difference is formally characterized by the dimensionality of the maps involved (remark 29): non-trivial relational schemas involve two-dimensional maps (relational learning); trivial schemas involve one-dimensional maps (associative learning).

Discussion

The relational schema induction paradigm was introduced to contrast associative versus relational learning (Halford, Bain, et al., 1998). The paradigm is used here to address two central questions for cognitive science. (1) What essentially distinguishes associative and relational (learning) processes? (2) How are these two forms of learning related? Our theory says that (1) associative and relational learning are two instances of induction by reconstruction where (2) association involves induction of a trivial relation schema—the mappings are one-dimensional (associative learning) versus two-dimensional (relational learning).

This formal comparison/contrast of associative and relational learning raises two further questions. First, how are these two forms of learning (formally) related, beyond their dimensionality? Second, what drives one form of learning to predominate over the other?

In regard to the first question, category theory provides another kind of universal construction—*change of base*—that we expect to play an important role here. In the current context, the categories of learning tasks are constructed on a base monoid. Monoids are related by monoid homomorphisms (definition 6). So, the transition between associative and relational learning is predicated on a change of base (monoid homo)morphism inducing a change in the categories of learning task. An analogous situation arises in *sheaf theory* (Mac Lane & Moerdijk, 1998), where the base is a topological space. Continuous functions between topological spaces induce functors between categories. Sheaf theory was used to model changes in learning transfer (Phillips, 2018). An analogous situation is expected here based on morphisms between trivial and non-trivial monoids (examples 7) corresponding to the transition between the two forms of learning.

In regard to the second question, it remains to be explained why associative learning prevails at all. According to the reconstruction theorem, induction of the relational schema is necessitated by the data, i.e., the category of learning tasks. However, younger children in contrast to older children can fail at induction, as observed in an earlier version of the relational schema induction paradigm (Halford & Wilson, 1980). Such differences were characterized in terms of *relational complexity* (Halford, Wilson, & Phillips, 1998), i.e. the number related dimensions of task variation, which is similar to our characterization. However, our use of reconstruction theory does not say why participants (in particular, younger children) fail to induce the relevant relational schema given the same data—series of learning tasks. One approach is to work with an enriched form of category theory to model resources (see, e.g., Fong & Spivak, 2018).

Resources can be modeled by giving the theory a suitable *categorical semantics*: e.g., an *adjoint functor* (remark 44) to a category modeling resources, such as *Petri nets* whose neurons activate when they have a sufficient number of *tokens* (neuronal resources). There are adjoint relationships between certain kinds of formal categories and certain categories of Petri nets (Baez, Genovese, Master, & Shulman, 2021). In this way, a model of the theory could be developed to address capacity as an explanation for failures of induction.

An implicit-explicit distinction

One way to characterize our reconstruction approach to relational schema induction is in terms of an implicit-explicit distinction. The actions are implicitly given by input-output relations between trigrams. Computing the end, by comparing those within-task relations across task instances, essentially makes those relations explicit. In effect, this is the role of hom-functors, by treating morphisms in some category **C** as elements of a set (function space) in **Set**, in effect, objectifying an action (compare, e.g., "I run" with "I went for a run"). Such distinctions are commonplace in so-called Type 1 versus Type 2 characterizations of cognitive processes (see, e.g., Evans & Stanovich, 2013). Thus, we expect that a category theory approach like ours also has applications for dual-process theories.

Relationship to previous work

As mentioned earlier, the current result extends previous work (Phillips, 2021b) by observing that an associative account of learning in the relational schema induction paradigm is a special (trivial) case of relational learning, vis-a-vis, a one-element set is a trivial monoid. Accordingly, the two forms of learning unify within the more general notion of Kan extension. The point is not to show that associative and relational cognition are the same, but rather how they are connected, by placing them within a common formal framework.

In doing so, our formal approach sheds some light on the nature of the conundrum over comparative claims of relational processes, mentioned earlier (Introduction). The theory essentially points out that an associative schema is trivially a relational schema, vis-a-vis, the trivial monoid. However, the "extra" dimension, does not add to the complexity (variance) of the problem. The methodological implication is that a comparative test of relational schemas in non-humans and younger age groups necessitates variation over instances of the task: cf. a single task is trivially a one-task series of tasks. Without such variation, caution is warranted over comparisons of relational processing across cohorts.

The more general message is that while much of the debate has focussed on distinguishing associative from relational processes, less attention has been given to relating these two forms of cognition. Category theory provides a formal framework for both comparison and contrast in the form of universal constructions: cognitive processes as composition of the common (mediating) arrow shared by all constructions in the given context (comparison) and the unique arrow that is construction specific (contrast). In this way, category theory provides a general principle for cognitive science (Phillips, 2021a).

Acknowledgments

I thank the reviewers for comments and suggestions to help improve the presentation of this work.

Appendix A: Sets

Definition 1 (Monoid)**.** A *monoid* $(M, \cdot, e)$ consists of a set M, a (closed) binary operation $\cdot$, and an element $e \in M$, called the *unit*, such that for all elements $a, b, c \in M$ the operation is:

- *unital*: $a \cdot e = a = e \cdot a$, and
- *associative*: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

Remarks 2. If every element $a \in M$ has an *inverse*, i.e. an element $b \in M$ such that $a \cdot b = e = b \cdot a$, then the monoid is also a *group*. The inverse of a is also denoted a^{-1}.

Example 3 (Cyclic-3 group)**.** The set $\{0, 1, 2\}$ together with addition modulo-3 is a (cyclic) group/monoid, denoted $\mathbb{Z}/3\mathbb{Z}$.

Example 4 (Shapes)**.** The set of shapes $Sh = \{\triangle, \square, \heartsuit\}$ constitute a cyclic-3 group/monoid, $(Sh, \cdot, \triangle) \cong \mathbb{Z}/3\mathbb{Z}$.

Example 5 (Trivial)**.** The *trivial monoid*, denoted 1, has the unit as its only element, i.e. $1 = (\{e\}, \cdot, e)$.

Definition 6 (Monoid homomorphism)**.** Let M and M' be monoids. A *monoid homomorphism* is a map $h : M \to M'$ that for all elements $a, b \in M$ preserves the

- *unit*: $h(e) = e'$, and
- *operation*: $h(a \cdot b) = h(a) \cdot h(b)$.

Examples 7 (Monoid homomorphism)**.** Every monoid M is associated with two homomorphisms:

a. $1 \to M$, which picks out the unit of M, and

b. $M \to 1$, which sends every element of M to the unit.

Definition 8 (Monoid action)**.** Let $(M, \cdot, e)$ be a monoid and X a set. A *(left) monoid action* on X is a function $\phi : M \times X \to X$ that satisfies the following laws for all $a \in M$ and $x \in X$:

- *identity*: $\phi(e, x) = x$, and
- *compatibility*: $\phi(a \cdot b, x) = \phi(a, \phi(b, x))$.

The set X is called an *M-set*.

Example 9 (Task)**.** Each task instance $\tau : Sh \times Tri \to Tri$ is a monoid action, e.g., see Table 1.

Definition 10 (Transpose)**.** Let $\phi : M \times X \to X$ be a monoid action. The *transpose* of ϕ, denoted $\tilde{\phi}$ (or simply ϕ), is the function $\tilde{\phi} : M \to (X \to X)$ sending each element $a \in M$ to the function $\phi_a : X \to X$, called the *component* of ϕ at a.

Example 11 (Transpose)**.** The transpose $\tilde{\phi}$ sends each shape to a component action on trigrams.

Remark 12. When M is trivial the transpose picks out the identity function, $\phi : e \mapsto (1_X : x \mapsto x)$.

Definition 13 (Equivariant map)**.** Let X and Y be M-sets for a monoid M. An *equivariant map* is a function $f : X \to Y$ such that $f(a \cdot x) = a \cdot f(x)$ for all $a \in M$ and $x \in X$.

Example 14 (Task map)**.** The tasks $\tau_1 : Sh_1 \times Tri_1 \to Tri_1$ and $\tau_2 : Sh_2 \times Tri_2 \to Tri_2$ (Table 1) are related by equivariant maps via an isomorphism $Sh_1 \cong Sh_2$ aligning the shapes in τ_1 to the shapes in τ_2. Specifically, for the set of shapes Sh_1 as the monoid and monoid action $\tau_1' : Sh_1 \times Tri_2 \to Tri_2$ we have the equivariant map $f : Tri_1 \to Tri_2$, and for the set of shapes Sh_2 as the monoid and monoid action $\tau_2' : Sh_2 \times Tri_1 \to Tri_1$ we have the equivariant map $g : Tri_2 \to Tri_1$. The isomorphism is called a *change of base* (see Discussion).

Remark 15. When M is the trivial monoid, every function $f \in Y^X$ is an equivariant map.

Appendix B: Categories

Definition 16 (Category)**.** A *category* $\mathbf{C}$ consists of a collection of *objects*, $\mathcal{O}(\mathbf{C}) = \{A, B, \dots\}$, a collection of *morphisms*, $\mathcal{M}(\mathbf{C}) = \{f, g, \dots\}$—a morphism written in full as $f : A \to B$ indicates object A as the *domain* and object B as the *codomain* of f—including for each object $A \in \mathcal{O}(\mathbf{C})$ the *identity morphism* $1_A : A \to A$, and a *composition* operation, $\circ$, that sends each pair of *compatible* morphisms $f : A \to B$ and $g : B \to C$ (i.e. the codomain of f is the domain of g) to the *composite* morphism $g \circ f : A \to C$, that together satisfy the laws of:

- *identity*: $f \circ 1_A = f = 1_B \circ f$ for every $f \in \mathcal{M}(\mathbf{C})$, and
- *associativity*: $h \circ (g \circ f) = (h \circ g) \circ f$ for every triple of compatible morphisms $f, g, h \in \mathcal{M}(\mathbf{C})$.

Remark 17. The opposite category to $\mathbf{C}$, denoted $\mathbf{C}^{\mathrm{op}}$, has the objects of $\mathbf{C}$ and the morphisms of $\mathbf{C}$ "reversed", i.e. a morphism $A \to B$ in $\mathbf{C}$ is a morphism $B \to A$ in $\mathbf{C}^{\mathrm{op}}$.

Example 18 (**Set**)**.** The category **Set** has sets for objects and functions for morphisms. The identity morphisms are identity functions and composition is composition of functions.

Remark 19. A set A is construed as the category having the elements $a \in A$ as objects and no non-identity morphisms.

Example 20 (Monoid)**.** A monoid, M, is a one-object category, whose morphisms are the elements of M. The identity is the unit and composition is given by the binary operation.

Example 21 (**Mon**)**.** The category **Mon** has monoids for objects and monoid homomorphisms for morphisms.

Definition 22 (Functor)**.** A *functor* is a "structure-preserving" map from a category $\mathbf{C}$ to a category $\mathbf{D}$, written $F : \mathbf{C} \to \mathbf{D}$, sending each object A and morphism $f : A \to B$ in $\mathbf{C}$ to the object $F(A)$ and the morphism $F(f) : F(A) \to F(B)$ in $\mathbf{D}$ (respectively) that satisfies the laws of:

- *identity*: $F(1_A) = 1_{F(A)}$ for every object $A \in \mathcal{O}(\mathbf{C})$, and
- *compositionality*: $F(g \circ f) = F(g) \circ F(f)$ for every pair of compatible morphisms $f, g \in \mathcal{M}(\mathbf{C})$.

Remarks 23. Functors preserve identities and composition.

Example 24 (Monoid homomorphism)**.** A monoid homomorphism is a functor that preserves the unit and operation.

Example 25 (Monoid action)**.** Suppose a monoid action $\phi : M \times X \to X$. The transpose $\phi : M \to (X \to X)$ corresponds to the functor $M \to \mathbf{Set}$ that sends each element $a \in M$ to the function $\phi_a : X \to X$.

Definition 26 (Natural transformation)**.** Let $F, G : \mathbf{A} \to \mathbf{C}$ be functors. A *natural transformation* $\eta : F \dot{\to} G$ is a family of $\mathbf{C}$-morphisms $\{\eta_A : F(A) \to G(A) | A \in \mathcal{O}(\mathbf{A})\}$ such that $G(f) \circ \eta_A = \eta_B \circ F(f)$ for every morphism $f : A \to B$ in $\mathbf{A}$, as indicated by the following commutative diagram:

$$\begin{tikzcd} A \arrow[d, "f"'] & F(A) \arrow[r, "\eta_A"] \arrow[d, "F(f)"'] & G(A) \arrow[d, "G(f)"] \\ B & F(B) \arrow[r, "\eta_B"'] & G(B) \end{tikzcd} \tag{1}$$

The morphism η_A is called the *component* of η at A.

Example 27 (Equivariant map)**.** Suppose a monoid M and actions $\phi : M \times X \to X$ and $\psi : M \times Y \to Y$. An equivariant map $f : X \to Y$ is a natural transformation, $f : \phi \dot{\to} \psi$, as indicated by the following commutative diagram:

$$\begin{tikzcd} * \arrow[d, "a"'] & X \arrow[r, "f"] \arrow[d, "\phi_a"'] & Y \arrow[d, "\psi_a"] \\ * & X \arrow[r, "f"'] & Y \end{tikzcd} \tag{2}$$

Remarks 28. When $M = 1$, naturality is trivially satisfied:

$$\begin{tikzcd} * \arrow[d, "e"'] & X \arrow[r, "f"] \arrow[d, "1_X"'] & Y \arrow[d, "1_Y"] \\ * & X \arrow[r, "f"'] & Y \end{tikzcd} \tag{3}$$

so every function $f \in Y^X$ is an equivariant map, shown by the following diagram with identities collapsed as they convey no additional information (variation):

$$\begin{tikzcd} * & X \arrow[r, "f"] & Y \end{tikzcd} \tag{4}$$

Remark 29. Contrast the dimensionality of diagram 4 (one-dimensional) and diagram 2 (two-dimensional).

Remark 30. An M-set S is *represented* by the pair (S, σ) consisting of a set S and a functor $\sigma : M \to (X \to X)$. M-set representations and their equivariant maps form a category, denoted **MSet**. The forgetful functor $U : \mathbf{MSet} \to \mathbf{Set}$ forgets the actions, i.e. $U : (S, \sigma) \mapsto S$.

Remark 31. When $M = 1$ the category of M-sets, denoted **1Set**, is isomorphic to the category of sets, i.e. $\mathbf{1Set} \cong \mathbf{Set}$.

Definition 32 (Functor category)**.** The collection of functors from a category $\mathbf{C}$ to a category $\mathbf{D}$ and their natural transformations is called a *functor category*, denoted $\mathbf{D}^{\mathbf{C}}$.

Example 33 (Diagrams)**.** Let 2 denote the category with two objects and no non-identity morphisms The functor category $\mathbf{C}^2$ contains all functors picking out pairs of objects and natural transformations picking out pairs of morphisms from $\mathbf{C}$.

Remark 34. $\mathbf{C}^2 \cong \mathbf{C} \times \mathbf{C}$.

Example 35 (Diagonal, product)**.** The diagonal and product functors pertain to functor categories.

a. *Diagonal.* $\Delta : \mathbf{C} \to \mathbf{C}^2; A \mapsto (A, A), f \mapsto (f, f)$.

b. *Product.* $\Pi : \mathbf{C}^2 \to \mathbf{C}; (A, B) \mapsto A \times B, (f, g) \mapsto f \times g$.

Appendix C: Universal constructions

Definition 36 (Universal morphism)**.** Let $G : \mathbf{A} \to \mathbf{C}$ be a functor and X an object in $\mathbf{C}$. A *universal morphism* from X to G is a pair (A, ϕ) consisting of an object A in $\mathbf{A}$ and a morphism $\phi : X \to G(A)$ in $\mathbf{C}$ such that for every object Y in $\mathbf{A}$ and every morphism $g : X \to G(Y)$ in $\mathbf{C}$ there exists a unique morphism $u : A \to Y$ in $\mathbf{A}$ such that $g = G(u) \circ \phi$, as indicated by commutative diagram

$$\begin{tikzcd} X \arrow[r, "\phi"] \arrow[dr, "g"'] & G(A) \arrow[d, dashed, "G(u)"] & A \arrow[d, dashed, "u"] \\ & G(Y) & Y \end{tikzcd} \tag{5}$$

Remark 37. ϕ corresponds to component ϕ_A of the natural transformation $\phi : X \dot{\to} G$, where X denotes the constant functor picking out the object X in diagram 5, cf. diagram 1.

Remark 38. The dual version of a universal morphism has the directions of arrows reversed, i.e. from a functor $F : \mathbf{A} \to \mathbf{C}$ to an object Y in $\mathbf{C}$ such that the following diagram commutes:

$$\begin{tikzcd} X \arrow[d, dashed, "u"'] & F(X) \arrow[d, dashed, "F(u)"'] \arrow[dr, "f"] & \\ B & F(B) \arrow[r, "\psi"'] & Y \end{tikzcd} \tag{6}$$

Definition 39 (Initial object)**.** In a category $\mathbf{C}$, an *initial object* (if it exists) is an object, denoted 0, such that for every object Z there exists a unique morphism $u : 0 \to Z$ (all in $\mathbf{C}$).

Remarks 40. The dual notion is the *terminal (final) object*, denoted 1, i.e. $u : Z \to 1$ exists uniquely for all Z.

Example 41 (Initial/final—**Set**)**.** In **Set**, the initial object is the empty set and the final object is any one-element set.

Example 42 (Initial/final—**Mon**)**.** The initial and final object in **Mon** is the trivial monoid.

Example 43 (Product—universal morphism)**.** A product of A and B is the universal morphism $(A \times B, \pi)$ from the diagonal functor, Δ, to the pair of objects (A, B), where $\pi = (\acute{\pi}, \grave{\pi})$ is the pair of projections $\acute{\pi} : A \times B \to A$ and $\grave{\pi} : A \times B \to B$.

Remarks 44. A universal morphism from every object Z in $\mathbf{C}$ to the functor $G : \mathbf{D} \to \mathbf{C}$ implies an *adjoint situation*, i.e. a functor $F : \mathbf{C} \to \mathbf{D}$ forming a pair, denoted $F \dashv G$, e.g., $\Delta \dashv \Pi$:

$$\begin{tikzcd} Z \arrow[r, "\langle 1,1 \rangle"] \arrow[dr, "\langle f,g \rangle"'] & Z \times Z \arrow[d, dashed, "f \times g"] \\ & A \times B \end{tikzcd} \qquad \begin{tikzcd} (Z,Z) \arrow[d, dashed, "(\langle f,g \rangle, \langle f,g \rangle)"'] \arrow[dr, "(f,g)"] & \\ (A \times B, A \times B) \arrow[r, "(\acute{\pi}, \grave{\pi})"'] & (A,B) \end{tikzcd} \tag{7}$$

F (G) is called the *left (right) adjoint* of G (F).

Remark 45. Products are instances of limit functors, $\mathrm{Lim} : \mathbf{C}^J \to \mathbf{C}$ forming a general class of adjoints, $\Delta \dashv \mathrm{Lim}$. For products, $J = 2$ picks out pairs of objects and morphisms.

Definition 46 (Wedge)**.** A *wedge* to a functor $F : \mathbf{C}^{\mathbf{op}} \times \mathbf{C} \to \mathbf{D}$ is a *di*natural transformation $\omega : D \stackrel{..}{\to} F$ consisting of a family of $\mathbf{D}$-morphisms $\{\omega_A : D \to F(A, A) | A \in O(\mathbf{C})\}$ such that for each $f : A \to B$ in $\mathbf{C}$ the following diagram commutes:

$$\begin{tikzcd} D \arrow[r, "\omega_A"] \arrow[d, "\omega_B"'] & F(A,A) \arrow[d, "F(1_A, f)"] \\ F(B,B) \arrow[r, "F(f, 1_B)"'] & F(A,B) \end{tikzcd} \tag{8}$$

Definition 47 (End)**.** The *end* of a functor $F : \mathbf{C}^{\mathbf{op}} \times \mathbf{C} \to \mathbf{D}$ is a pair (E, ω) consisting of an object E in $\mathbf{D}$ and a wedge $\omega : D \stackrel{..}{\to} F$ such that for every wedge $\beta : Z \stackrel{..}{\to} F$ there exists a unique morphism $u : Z \to E$ such that $\beta = \omega \circ u$. Object E is also denoted $\int_{A \in \mathbf{C}} F(A, A)$, or $\int_{\mathbf{C}} F$.

Remark 48. An end is a universal wedge.

Examples 49 (Hom-set)**.** Object and morphism relations are determined by *hom-functors*: e.g., $\mathrm{Hom}(A, -) : \mathbf{C} \to \mathbf{Set}$ sends each object B to the set of morphisms $\mathrm{Hom}(A, B)$ and each morphism f to composition operation $(f \circ) : g \mapsto f \circ g$. A set of natural transformations between a pair of functors is (re)constructed from the end of a hom-functor.

a. $\int_{\mathbf{C}} \mathrm{Hom}(-, -) \cong \mathrm{Nat}(1_{\mathbf{C}}, 1_{\mathbf{C}})$.

b. $\int_{\mathbf{C}} \mathrm{Hom}(F-, G-) \cong \mathrm{Nat}(F, G)$.

Remark 50. For example 49(b), substitution yields

$$\begin{tikzcd} E \arrow[r, "\omega_A"] \arrow[d, "\omega_B"'] & \mathrm{Hom}(FA, GA) \arrow[d, "\mathrm{Hom}(1_{FA}, Gf)"] \\ \mathrm{Hom}(FB, GB) \arrow[r, "\mathrm{Hom}(Ff, 1_{GB})"'] & \mathrm{Hom}(FA, GB) \end{tikzcd} \tag{9}$$

where E identifies with the set of natural transformations, $\{\eta\}$, and ω_A with the component, $\eta_A \in \mathrm{Hom}(FA, GA)$, according to the naturality condition (see diagram 1).

Theorem 51 (Reconstruction)**.** Let M be a monoid and **MSet** the category of M-set representations for M. ($U : \mathbf{MSet} \to \mathbf{Set}$ is the forgetful functor.) $\int_{\mathbf{MSet}} \mathrm{Hom}(U-, U-) \cong M$.

Remark 52. This theorem is a category theory version of *Tannakian reconstruction* (see NLab, 2014, 2019).

Remark 53. All previous universal constructions are subsumed by a single pair of constructions, defined next, regarded as a form of constrained optimization.

Definition 54 (Kan extension)**.** Let $X : \mathbf{A} \to \mathbf{C}$ and $F : \mathbf{A} \to \mathbf{B}$ be functors. The *(right) Kan extension* of X along F is a pair (R, ϵ) consisting of a functor $R : \mathbf{B} \to \mathbf{C}$ and a natural transformation $\epsilon : RF \dot{\to} X$ such that for any functor $M : \mathbf{B} \to \mathbf{C}$ and natural transformation $\mu : MF \dot{\to} X$ there exists a unique natural transformation $\delta_F : M \dot{\to} R$ such that $\mu = \epsilon \circ \delta_F$, as indicated by the right commutative diagram:

$$\begin{tikzcd} X \arrow[r, "\eta"] \arrow[dr, "\nu"'] & LF \arrow[d, dashed, "\gamma_F"] \\ & NF \end{tikzcd} \qquad \begin{tikzcd} MF \arrow[d, dashed, "\delta_F"'] \arrow[dr, "\mu"] & \\ RF \arrow[r, "\epsilon"'] & X \end{tikzcd} \tag{10}$$

Dually, the *left Kan extension* of X along F is the pair (L, η) given by the left commutative diagram (above).

Example 55 (Limits)**.** Limits of type (shape) J obtain as the left Kan extension of the identity functor along the diagonal functor (Δ) yielding the limit functor (Lim):

$$\begin{tikzcd} & \mathbf{C}^J \arrow[dr, "\mathrm{Lim}(\Pi)"] & \\ \mathbf{C} \arrow[ur, "\Delta"] \arrow[rr, "1"'] & & \mathbf{C} \end{tikzcd} \tag{11}$$

For pairs ($J = 2$), the left Kan extension along the diagonal yields the product functor, $\Pi : (A, B) \mapsto A \times B$. Compare diagram 10 with diagram 7. All adjoints extend this way.

Remark 56. Kan extensions can be regarded as constrained optimization: e.g., the optimal extension (Π) that recovers the original functor (1) given the constraints (Δ).

Example 57 (Ends)**.** Ends obtain as the left Kan extension of the identity along the hom functor (in theorem 51):

$$\begin{tikzcd} & \mathbf{Set}^P \arrow[dr, "\mathrm{End}"] & \\ \mathbf{Set} \arrow[ur, "\mathrm{Hom}"] \arrow[rr, "1"'] & & \mathbf{Set}, \end{tikzcd} \tag{12}$$

$P = \mathbf{MSet}^{\mathrm{op}} \times \mathbf{MSet}$. For $M = 1$, replace **MSet** with **1Set**.

References

Baez, J. C., Genovese, F., Master, J., & Shulman, M. (2021, January). *Categories of nets.* (Preprint: arXiv:2101.04238v1)

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, *8*(3), 223–241.

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

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170.

Giurfa, M. (2021). Learning of sameness/difference relationships by honey bees: performance, strategies and ecological context. *Current Opinion in Behavioral Sciences*, *37*, 1–6.

Halford, G. S., Bain, J. D., Maybery, M. T., & Andrews, G. (1998). Induction of relational schemas: Common processes in reasoning and complex learning. *Cognitive Psychology*, *35*(3), 201–245.

Halford, G. S., & Busby, J. (2007). Acquisition of structured knowledge without instruction: The relational schema induction paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 586–603.

Halford, G. S., & Wilson, W. H. (1980). A category theory approach to cognitive development. *Cognitive Psychology*, *12*(3), 356–411.

Halford, G. S., Wilson, W. H., Andrews, G., & Phillips, S. (2014). *Categorizing cognition: Toward conceptual coherence in the foundations of psychology*. Cambridge, MA: MIT Press.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, *21*(6), 803–831.

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, *56*(1), 51–65.

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

Mac Lane, S. (1998). *Categories for the working mathematician* (2nd ed.). New York, NY: Springer.

Mac Lane, S., & Moerdijk, I. (1998). *Sheaves in geometry and logic: A first introduction to topos theory*. New York, NY: Springer.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological Bulletin*, *117*(3), 363–386.

NLab. (2014, June). *The catsters.*

NLab. (2019, October). *Tannaka duality.*

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31*(2), 109–178.

Phillips, S. (2018). Going beyond the data as the patching (sheaving) of local knowledge. *Frontiers in psychology*, *9*, 1926.

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.