認知科学 第 1 巻 1 号 (9999) 1–14 doi:
Cognitive Studies: Bulletin of the Japanese Cognitive Science Society

Feature: Category theory **Invited Paper**

A category theory principle for cognitive science: cognition as universal construction

Steven Phillips

National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

What can category theory contribute to cognitive science? We argue that the category theory principle of construction via a *universal mapping property* affords a significant contribution. Such universal constructions explain why, not just how cognition is systematic/compositional, i.e. the "best" one can do in the given context. The significance of this principle is indicated by examples.

Keywords: category, functor, universal construction, adjunction, sheaf, analogy, systematicity, compositionality

Received ; Available online at J-STAGE

1. Introduction

What can *category theory* contribute to cognitive science? Presumably, this question begs a more basic one for many cognitive scientists: *What is category theory?* Category theory (Awodey, 2010; Mac Lane, 1998; Leinster, 2014) is a branch of (meta)mathematics invented as a formal language for comparing mathematical constructions.

> In a metamathematical sense our theory provides general concepts applicable to all branches of abstract mathematics, and so contributes to the current trend towards uniform treatment of different mathematical disciplines. In particular, it provides opportunities for the comparison of constructions and of the isomorphisms occurring in different branches of mathematics; in this way it may occasionally suggest new results by analogy. (Eilenberg & Mac Lane, 1945, p. 236)

In that sense, category theory can be regarded as a formal theory of analogy (Brown & Porter, 2006). Or, perhaps more aptly, *category theory is to mathematics as analogy theory is to cognitive science*.

Mathematical models have been a mainstay of psychology (Townsend, 2008), so the above suggests that category theory can contribute to cognitive science, given analogy as a core cognitive process (Hofstadter, 2001). Indeed, most cognitive models of analogy are driven by structural consistency (Gentner & Forbus, 2011)—cf. *structure mapping theory* (Gentner, 1983)—and a categorical approach to structural consistency was recognized early on (Halford & Wilson, 1980). Despite this parallel, contributions to cognitive models of analogy have been few and far between (see, e.g., Halford, Wilson, Guo, Gayler, Wiles, & Stewart, 1994; Navarrete & Dartnell, 2017).

The sparsity of category theory contributions for models of analogy may reflect different aspirations. Category theorists want formal accounts of structural relations between formal concepts. Cognitive scientists seek explanations for conceptual thinking more broadly, including commonsense concepts which are notoriously difficult to formally define (Laurence & Margolis, 1999). Formal concepts typically have succinct and precise definitions. Unsurprisingly, then, a significant contribution of category theory to a field outside mathematics has been computer science (see, e.g., Arbib & Manes, 1975; Barr & Wells, 1990; Walters, 1991). Still, category theory may have a deeper role to play in our understanding of cognition that has yet to be fully recognized—a role that goes beyond analogy as a core cognitive process.

There is more to category theory than arbitrarily identifying interconnections.

> What matters is the many *real* [emphasis added] interconnections, not the wholly artificial ones. (Mac Lane, 1997, p. 121)

The sense of artificial is exemplified by *anomalous cancellation*, e.g., $\frac{16}{64} = \frac{1}{4}$ (Weisstein, 2020). Anomalous cancellation is analogous to the *cancellation rule* for fractions: $\frac{a}{b} \times \frac{b}{c} = \frac{a}{c}$. But the analogy is artificial in that the equality only holds coincidentally (idiosyncratically) for a handful of cases, whereas the cancellation rule applies systematically to all non-zero numbers and derives from deeper properties pertaining to groups.[1] Category theory provides a formal sense in which an analogy is *natural* as opposed to artificial.

1.1 Natural and universal connections

A focus on the real connections resonates with the *systematicity* property of language (Chomsky, 1980) and thought (Fodor & Pylyshyn, 1988). Systematicity is a property of cognition whereby certain structurally-related capacities tend to coexist. The archetypal example is where having a capacity to understand the expression *John loves Mary* implies having a capacity to understand the (structurally related) expression *Mary loves John*. A basic observation to be explained is why you only find native (English) speakers who can understand one statement *if and only if* they can understand the other statement. A classical explanation is that these two capacities are intrinsically connected by the same syntactic process that respects their semantic relationship, i.e. the *loves* relations between agent and patient. Hence, consistent with the data, one can only have either both, or neither capacity, but nothing in between. By contrast, idioms are idiosyncratic: e.g., *John kicked the bucket* (i.e, *John died*) is not semantically related to *Mary kicked the ball*. The second expression can be understood independently of the first, despite the syntactic analogy. The expression *John bit the dust* (i.e, *John died*) is likewise idiomatic, despite the semantic analogy. A classical explanation also accounts for this situation, because two different syntactic processes are responsible for understanding these expressions. Hence, one can have both, neither, or just one of these capacities, which is again consistent with the data. A theory of cognition should account for the evident systematicity properties.

Note that counterexamples of systematicity are not just about idioms. For example, *John loves Mary* is also analogous to *loves Mary John* as three-word lists. Clearly, one can have a capacity for three-word lists without having a capacity to understand the meanings conveyed by *John loves Mary* and *Mary loves John*. In the latter two cases, *John* and *Mary* are "bound" to the *loves* relation. As lists, those words are just related by order. The "shape" of those structures differs, which is an important aspect of systematicity.

Accounts of systematicity generally assume some form of *compositionality* whereby representations of complex entities are composed from representations of constituents. Classical theory assumes symbolic representations and processes, i.e. a combinatorial syntax and semantics whereby the semantic relations between constituents are mirrored by syntactic relations between the corresponding symbols (Fodor & Pylyshyn, 1988). Non-classical (e.g., neural network) theory assumes representations are composed by other means, e.g., vector operations (e.g., Smolensky, 1990).

Whether such theories account for systematicity and to what extent cognition is systematic has been extensively debated (see Aizawa, 2003, for a review and analysis). The essential criticism has been that such theories generally fall short, because they also admit constructions that fail to exhibit the requisite systematicity property without some additional, so-called *ad hoc* assumptions that must be added to the theory to pick out just the systematically related capacities: auxiliary assumptions whose sole purpose is to fit the errant cases (Aizawa, 2003).

A category theory approach was introduced to account for systematicity, without such assumptions, in terms of *natural* and *universal constructions* (Phillips & Wilson, 2010, 2016c). The basic intuition, which we elaborate upon later, is that natural and universal constructions capture the generality and commonality alluded to in the *John loves Mary* example. Such constructions satisfy a *universal mapping property* (Awodey, 2010), hence are called *universal constructions*. Put briefly, the systematicity properties occur just where there are universal constructions, so no further (ad hoc) assumptions are needed. Moreover, universal constructions follow from a universal process (see Phillips & Wilson, 2016c), i.e. every situation in the given context (category) points to the universal construction, hence systematicity necessarily follows.

1.2 The *shape* of things to come

An empirical challenge to this categorical view is where systematicity is not evident despite the presence of a universal construction. This possibility was tested in a task where participants learned cue-target relations conforming to a particular universal construction, called a *product map*: specifically, a map from pairs of letters to coloured shapes that was the product of letter-colour and letter-shape maps (Phillips, Takeda, &Sugimoto, 2016). To illustrate, suppose cues and targets are composed from three-element sets of letters, colours and shapes. Hence, there are nine mappings from letter pairs to coloured shapes. Participants were trained on a subset of these mapping and tested on the other pairs. In the *product* condition, each letter at one position uniquely maps to a colour and each letter at the other position uniquely maps to a shape. If participants learned the training pairs as the product of letter-colour

1) That is the axioms for a group, which the set of non-zero real numbers (under multiplication) constitutes an instance.

and letter-shape maps, then training on four mappings (including all constituent letters, colours and shapes, but not all possible combinations) is sufficient for correct prediction on the test pairs. A product is a universal construction, so participants should exhibit systematic capacity to map letter pairs to coloured shapes.

In some product conditions, however, participants did not exhibit systematicity as evidenced by response errors on novel pairs. Two groups of participants were given a series of learning tasks in either *ascending* of *descending* order of set size, i.e. the number of cue-target mappings constituting the task. Participants in the ascending group exhibited systematicity on large, but not small sized tasks.[2] (Participants in the descending group exhibited systematicity at all sizes.) As a product map, the targets for novel pairs were predictable from the constituent (letter-colour and letter-shape) maps. By contrast, novel pairs were not predictable when a cue-target map was learned without regard to the constituent maps, i.e. without representing the universal construction. These errors revealed a situation where systematicity did not follow from the presence of a universal construction, challenging the categorical view.

An alternative categorical view is that systematicity and non-systematicity are two aspects of a more general universal construction. This situation is familiar in the form of (Type 1/Type 2) *dual-process* theories (see, e.g., Evans, 2003; Kahneman, 2011). Subsequent work showed that the two aspects are related by particular category theory constructions (Phillips, 2018a), which we elaborate later. The point of departure from other modeling approaches is to regard cues and targets as data attached to some (topological) space. The *shape* (topology) of the space determines systematicity/non-systematicity (idiosyncracy), as evidenced by the empirical data.

The challenge for a categorical approach appears to be problematical. How can category theory explain the (systematic) emergence of systematicity without simply appealing to a universal construction whenever systematicity emerges—the kind of ad hoc assumption that was problematic for connectionist (and classical) accounts of systematicity (Aizawa, 2003)? An answer lies with another kind of universal construction, called an *adjunction* (Mac Lane, 1998), that bridges the gap between systematicity and non-systematicity, i.e. as a universal construction relating other constructions (Phillips, 2018a). This situation can be seen as a form of *second-order* systematicity, i.e a systematic capacity to *learn* certain cognitive capacities (Aizawa, 2003; Chomsky, 1980; Phillips & Wilson, 2016b). In contrast, systematicity is typically introduced as a relation between base capacities, independently of learning, as exemplified in the previous section. Naturally, then, our explanation for this form of systematicity involves a second-order universal construction.

2) Failures were attributed to a cost/benefit trade-off between the relative difficulty of associative learning and induction of the product (relational) structure (Phillips et al., 2016).

The discovery of adjunctions, in general form, was a major contribution of category theory to mathematics (see Marquis, 2009, for a historical perspective).

> The slogan is "Adjoint functors [adjunctions] arise everywhere." (Mac Lane, 1998, p. vii)

In this paper, we expand upon the earlier suggestion (Phillips, 2018a) that these opposing properties are themselves systematically related by presenting an adjunction that pertains to a change in the shape of the underlying space. We further show how some other opposing properties of cognition—often the subject of dual-process theories—are too related by adjunctions (see also Phillips, 2018b, 2020b). The general point expounded here is that universal constructions (and, in particular, adjunctions) afford an important category theory contribution to cognitive science.

We proceed by presenting basic category theory concepts for our categorical approach (section 2). A specific example follows (section 3). The adjunction used in this example builds upon a suggestion from earlier work (Phillips, 2018a). Then we examine how adjoints arise in other aspects of cognition (section 4). We discuss this principle as a category theory contribution to cognitive science (section 5). No familiarity with category theory is assumed. Some theory is given in the appendices for convenience and specificity.

2. Categorical constructions

For readers unfamiliar with category theory, basic definitions are given first before presenting the concept of universal construction.

> As Eilenberg-Mac Lane first observed, "category" has been defined in order to be able to define "functor" and "functor" has been defined in order to be able to define "natural transformation". (Mac Lane, 1998, p. 18)

And "natural transformations" are defined in order to be able to define "universal morphisms" and "universal morphisms" are used to define "adjunctions". To help mitigate the seemingly endless abstraction, these concepts are introduced by comparison to concepts more familiar to cognitive scientists (Table 1).

2.1 Basics

A proportional analogy is used to bootstrap some intuitions about categories, functors and natural trans-

Table 1 A comparison of concepts

Concept	Category
entity	object
relation(al structure)	morphism
domain	category
source/target	functor (image)
analogy	natural transformation
optimal	universal morphism
correspondence	adjunction

formations: *Ebb is to flow as wax is to wane*. Analogy is generally recognized as a map of entities in a source domain to entities in a target domain that preserves their relationships (Gentner, 1983). For the current example, the entities (concepts) *ebb* and *flow* in the source are antonymous, and respectively map to antonyms *wax* and *wane* in the target (see fig. 1). The source and target are instances of a common relation(al schema), which we write as Opposes(Primal, Dual), i.e. Opposes(ebb, flow) and Opposes(wax, wane).

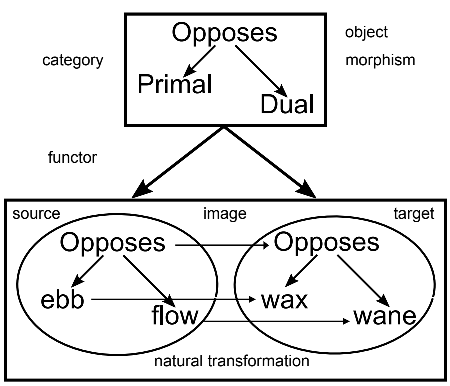

Fig. 1 *Ebb is to flow as wax is to wane* as a natural transformation from source to target functors.

2.1.1 Objects and morphisms

An interpretation of this analogy in terms of category theory concepts follows.

- The schema constitutes a *category* (definition 1) consisting of three *objects* (Opposes, Primal and Dual) and two (non-identity) *morphisms*, which capture the roles played by Primal and Dual in the relation Opposes.
- The image of each *functor* (definition 5) corresponds to the source and target instances of the relation.
- The proportional analogy is the *natural transformation* (definition 6) from source to target functor consisting of three maps (one for each object in the schema) satisfying a consistency condition.

Analogy is modeled as a natural transformation between functors that pick out the source and target relations constituting the proportion.[3)]

2.1.2 Generalized (shaped) "elements"

A quintessential difference between category and set theory views is the focus on maps versus elements. To wit, an element a in a set A is equivalently a map that points to a, i.e. $\overline{a} : * \mapsto a$, where $*$ is some element whose identity is unimportant. The map $\overline{a} : 1 \to A$ is called a *generalized element* of A, where object 1 is a *terminal* object (definition 3). In general, a morphism $f : X \to A$ is called a *generalized X-shaped element f of A*. Focusing on maps affords a convenient treatment of certain kinds of (universal) constructions.

2.2 Universals

In the formal language of category theory, a universal construction is a construction that satisfies a universal mapping property—a kind of optimal (unique existence) condition for the given situation, or context, pertaining to a special case of natural transformations. Examples are given to illustrate these connections.

2.2.1 Pairs and products

Pairing elements is one of the most basic forms of compositionality, and this form is an instance of a more general class of universal constructions. We use the notion of generalized element, introduced in the previous section, to derive this class of universal constructions, called a (categorical) *product* (definition 4), from the more familiar notion of sets of element pairs.

As sets, the *(Cartesian) product* of A and B, written $A \times B$, is the set of all pairs, (a,b), where a is drawn from A and b is drawn from B (example 5). As maps, each pair (a,b) is equivalently the generalized element $* \mapsto (a,b)$, written $\langle a,b \rangle : 1 \to A \times B$. The product of functions $f : X \to A$ and $g : Y \to B$, written $f \times g : X \times Y \to A \times B$, is defined elementwise, i.e. $f \times g : (x,y) \mapsto (f(x), g(y))$. Pair (a,b) obtains from product $\overline{a} \times \overline{b} : (*,*) \mapsto (a,b)$, and pair $(*,*)$ is equivalently generalized element $\langle 1,1 \rangle : * \mapsto (*,*)$. So, every generalized element $\langle a,b \rangle$ is the composite of $\langle 1,1 \rangle$ and $\overline{a} \times \overline{b}$. The map $\langle 1,1 \rangle$ is universal among such elements, constituting a *universal element*, which is an example of a universal construction.

The product construction for sets generalizes to objects in a category by regarding $\langle f,g \rangle : Z \to A \times B$ as a generalized Z-shaped element of the product of ob-

3) What counts as an object/morphism depends on the category: e.g., categories are also objects in a larger category whose morphisms are functors; a *functor category* has functors for objects and natural transformations for morphisms.

jects $A \times B$ in some category $\mathbf{C}$. Accordingly, $\langle f, g \rangle$ is the composite of morphisms $\langle 1, 1 \rangle : Z \to Z \times Z$ and $f \times g$, where $\langle 1, 1 \rangle$ constitutes the *universal morphism* among such elements, which is another example of a universal construction.

The Cartesian product comes with two functions that retrieve the first and second elements of each pair, i.e. $\pi_1 : (a, b) \mapsto a$ and $\pi_2 : (a, b) \mapsto b$. These functions are also universal, as all such functions $f : Z \to A$ and $g : Z \to B$ are composites of $\langle f, g \rangle$ and (π_1, π_2). Again, generalizing to objects in a category, the (canonical) categorical product is the object $A \times B$ and the pair of morphisms $\acute{\pi} : A \times B \to A$ and $\grave{\pi} : A \times B \to B$, satisfying the universal mapping property.

2.2.2 Universal morphisms and limits

All previous instances of universal constructions are *universal morphisms* (definition 7), which are optimal in a formal sense. A universal morphism consists of a *mediating morphism* (remark 4) that pertains to a natural transformation involving a *constant functor* (example 11). The mediating morphism is an "extreme" member among the family of morphisms constituting the natural transformation—every member factors through the mediating morphism (see example 14).

An important class of universal morphisms is called *limit* (definition 8), including products (example 15). All limits are determined by *shape*, i.e. a category used by a functor (called a *diagram*) to pick out a collection of objects and morphisms. For instance, a product is the limit of a diagram whose shape is a two-object category, one object for each constituent of the product. As we shall see in the next section, shape plays an important role in accounting for cognitive capacity.

2.2.3 Duals

The examples of universal construction involved two "opposite" forms. For products, the morphism $\langle f, g \rangle$ building up the product of objects $A \times B$ is a composite of the common morphism $\langle 1, 1 \rangle$ and the (unique) morphism $f \times g$, i.e. $\langle f, g \rangle = f \times g \circ \langle 1, 1 \rangle$. The pair of morphisms (f, g) retrieving constituent objects A and B is the composite of the (unique) morphism $\langle f, g \rangle$ and the common pair of morphisms $(\acute{\pi}, \grave{\pi})$, i.e. $f = \acute{\pi} \circ \langle f, g \rangle$ and $g = \grave{\pi} \circ \langle f, g \rangle$. In the first case, the common morphism is composed *before* the unique morphism, but is composed *after* the unique morphism in the second case.

This situation is typical in that constructions have two forms: *primal* and *dual* (remark 1). For instance, product is the primal construction and *coproduct* is the dual (example 6). A construction in regard to one category generally has a dual construction obtained by reversing the directions of the morphisms.

2.3 Adjunctions

In certain situations, primal and dual forms of universal constructions are closely related by a pair of opposing functors, and this relationship is called an *adjoint situation*, or simply an *adjunction* (definition 9). An adjunction between $F : \mathbf{C} \to \mathbf{D}$ and $G : \mathbf{D} \to \mathbf{C}$ involves universal constructions in both directions, i.e. from F to every object in $\mathbf{D}$ and from every object in $\mathbf{C}$ to G. Such situations afford an important category theory contribution to cognitive science.

The motivation for adjunctions is two-fold. Firstly, adjunctions express a weaker form of correspondence that is generally more useful than *isomorphism* (definition 2)—a kind of second-order isomorphism when a first-order isomorphism may not exist. Secondly, adjunctions embody the kinds of trade-offs that interest dual-process theorists, alluded to previously. Two examples follow that illustrate these motivations.

2.3.1 Approximation and precision

We consider *approximation* and *precision* as conceptual, but not actual inverses (isomorphisms), since approximation discards information—real numbers are approximated by integers, but their sets are not isomorphic. Yet, there is a "second-order isomorphism" affording comparisons with reals in terms of integers *without* loss of precision. We can interpret an adjunction as a relation between opposing functors: e.g., one functor sends each real to its *ceiling* (approximation: e.g., $2.3 \mapsto 3$) and the other functor sends each integer to its corresponding real (precision: e.g., $3 \mapsto 3.0$). The integers, likewise reals, with the usual order form a category (example 2). The two number systems are related by adjunctions (example 17). Effectively, comparisons of reals to upper and lower (integer) bounds can be computed in terms of integers, thus avoiding a need for infinite precision. In a cognitive context, an infinite world is represented by finite resources. This example also shows how adjunctions convey a second-order isomorphism even though a first-order isomorphism does *not* exist.

2.3.2 Parts and wholes

One commonly held distinction is whether cognitive processes operate on cognitive representations wholistically or componentially (whole/part). For example, *kicked the bucket* can be interpreted wholistically (idiomatically) as *died*, or componentially as an act of *kicking*. We have already seen a basic example of compositionality in the form of the categorical product. Products are constructed by the *product functor* (example 7), which can be seen as putting together parts into wholes. The conceptual inverse is to take wholes and regard them as parts of a larger construction. The *diagonal functor* (example 9) plays this role as the conceptual inverse of the product functor. The two functors are adjoints (example 18): the diagonal functor is left adjoint to the product functor.

As with the previous example, these functors are not actual inverses. The diagonal functor does not take a whole and return its original parts. However, as we shall see, adjointness is more useful. In the psychological sense, the part-whole relationship pertains to chunking and dechunking, often seen as a way of circumventing cognitive capacity limitations (see Halford, Wilson, Andrews, & Phillips, 2014).

2.3.3 A categorical principle: adjointness

Every adjunction induces a natural isomorphism between a pair of set-valued functors (remark 8). The natural isomorphism is a local connection between two categories that need not be globally isomorphic, as we saw with the integers and reals. This situation alludes to a general category theory principle for cognitive science: adjointness as the bridge between apparently disparate dual-processing forms of cognition.

Adjointness is a special case of the universal construction principle that affords a formal framework for thinking about cognition as a collection of dual-process systems. Dual-process situations typically realize a trade-off between resources and goals. The two situations just given are examples. The first exemplifies a common situation that is a speed-accuracy trade-off in that real numbers afford the benefit of precision at the expense of response time. The second is a kind of space-accuracy trade-off that we will explore in the next section, as an empirical implication of the theoretical principle.

3. Systematicity: an adjoint situation

Recall the cue-target learning task from the introduction (section 1.2), which was designed to explore an interaction between systematic and idiosyncratic properties of cognition, as a kind of trade-off characterized by an adjunction. We hypothesized that failures of systematicity in the presence of a universal construction were due to task demands. For relatively simple situations, participants would learn the task without the cost of representing the universal construction, thereby exhibiting idiosyncracy. For more complex situations, the benefit of universal construction would outweigh the cost, thereby exhibiting systematicity. Response errors on novel cues for various levels of learning difficulty provided support for this hypothesis (Phillips et al., 2016).

There are two aspects of the data, with regard to novel cues, that interest us here for the purpose of relating the universal construction (adjunction) principle to empirical implications. They pertain to the successes and failures to predict targets, regarded as instances of systematicity and idiosyncracy (respectively). The first aspect pertains to generalization from the training set to the testing set, which is all possible cue-target mappings for the given task (section 3.1). The second aspect pertains to one group of participants that showed a transition from no generalization on tasks involving a small number of cue-target mappings to generalization on tasks involving a larger number of mappings (section 3.2). We show how both aspects are related by adjunctions that depend on the "shape" of the representations for the cues and targets.

3.1 Representation as "shaped" data

A standard approach to cognitive modeling is to represent inputs and outputs as data encoded as vectors *in* some vector space. We depart from this approach by modeling inputs and outputs as data *on* (attached to) some *topological space* (definition 11). These representations are special kinds of functors, called *presheaves* (definition 12) and *sheaves* (definition 13). In this way, the shape (topology) of the underlying space is explicitly modelled. The data are shaped by the space, which has implications for cognitive capacity, as we shall see in this section.

The cue-target learning task is modeled using presheaves and sheaves, which can be conceptualized in terms of relational database tables (Abramsky & Brandenburger, 2011). For example, suppose pairs of letters GA, GE and KA. Each pair is regarded as data attached to a topological space that consists of the first and second letter positions as the points of the space. In the case that the space has a *discrete topology* (example 25) this arrangement is a presheaf corresponding to a two-column table with three rows—one column per position and one row per pair. There are two possible letters in the first position, G and K, and two possible letters in the second position, A and E. This presheaf is not a sheaf because it does not contain the pair KE. A sheaf is a "complete" presheaf. Targets are modeled likewise, where the attached data are colour-shape pairs. A cue-target map is a *presheaf morphism* (remark 13), i.e. a natural transformation.

Generalization from the training to the testing set is modeled by a sheaf theory construction, called *sheaving*, or *sheafification* (definition 14). Sheaving is a functor that completes each presheaf by sending it to the nearest sheaf. For example, applying the sheaving functor to the presheaf representing the pairs GA, GE and KA constructs the sheaf with the pair KE added. Likewise, sheaving sends presheaf morphisms to sheaf morphisms. Generalization is modeled by the sheaving functor which sends the training set (presheaf morphism) to the testing set (sheaf morphism). Applying the sheaving functor to a sheaf just returns that sheaf.

For a given topological space, the category of sheaves is a *subcategory* of the category of presheaves, hence are related by an *inclusion functor* (example 10). The sheaving functor sends presheaves to sheaves, and the inclusion functor going in the other direction identifies those presheaves that are sheaves. Presheaves model (possibly partial) knowledge afforded by the training set. Sheaves model complete knowledge of the task, in the form of correct responses for all testable cues. The sheaving functor is left adjoint to the inclusion functor. In other words, training and testing are related by a pair of adjoint functors.

3.2 Re-representation as change-of-shape

One group of participants showed no generalization to the novel cues on tasks involving a small number of cue-target mappings, but showed generalization on tasks involving more mappings (Phillips et al., 2016). These results were interpreted as participants having learned the tasks as two different kinds of maps. The tasks involving fewer mappings were learned as cue-target associations that disregarded the constituent maps, i.e. the cues and targets were regarded as single items. The larger tasks were learned as pairs of letter-colour and letter-shape maps. Participants re-represented the structure of the task from a single map to a pair of maps, as the task required learning more mappings. The suggestion was that these two situations (i.e. generalization and no generalization) were related by another kind of adjunction that pertains to a change in the shape of the underlying space (Phillips, 2018a). We expand upon this suggestion, here, by showing how the two situations relate by a *continuous function* (definition 15) that changes the topological space. This function induces a pair of adjoint functors that correspond to familiar psychological processes: chunking and dechunking.

We model this situation by considering the generalization and no generalization situations as data attached to discrete and *indiscrete topological spaces* (example 25), respectively. We have already seen the discrete case, where participants regard the cues/targets as data composed along two component dimensions. The indiscrete space corresponds to data attached to a single dimension, i.e. a psychological notion of chunking in the sense that each pair of letters is regarded as single item; likewise, each target is regarded as a single object, not a colour-shape pair. The corresponding relational tables have just a single column whose rows contain the cues/targets. In this case, all presheaves are sheaves, so the result of applying the sheaving functor to the training set is just the training set. No new mappings are added. So, sheaving with respect to the indiscrete space models the no generalization case. All presheaves are sheaves, in this situation, so the adjunction is trivial whereby the adjoints are the identity functor.

The indiscrete and discrete spaces are related by a continuous function, which induces two functors: the *direct image functor* and the *indirect image functor* (example 26). The direct image functor corresponds to chunking the data attached to the two-dimensional space to data attached to the one-dimensional space: e.g., treating a pair of letters as a single letter chunk. The inverse image functor corresponds to dechunking—the reverse operation that splits chunks into their constituents. Dechunking involves sheaving. So dechunking the training set generates the missing test mappings, hence obtains generalization. The direct and inverse image functors form an adjoint pair. Put another way, dechunking is left adjoint to chunking. Thus, the associative-relational form of dual-process is connected by an adjunction.

4. Cognitive adjoint situations

Adjoint situations in cognition were supposed to arise as trade-offs over cognitive resources, such as memory and attention. Adjunctions arise in many areas of mathematics. Their prevalence suggests that other resource trade-offs that motivate dual-process theories in cognition (Evans, 2003; Kahneman, 2011) are likewise adjoint situations. Dual-process theories are controversial, in part because they are vaguely defined (Evans & Stanovich, 2013). A strength of category theory is the precise formalization of relations between formal systems—a kind of meta-theory (see opening quote). In this regard, category theory can help make precise in what sense a cognitive system supposedly involves dual-processes. We sketch out two examples, in this section, showing how category theory can help in this way. (See Phillips, 2018b, for further discussion and examples.)

4.1 Serial versus parallel processes

The *relational complexity theory* of cognitive capacity says that the number of task dimensions (relational arity) is an important factor in determining task difficulty (Halford, Wilson, & Phillips, 1998; Halford et al., 2014). The theory also says that relational complexity can be mitigated by chunking (as we saw in the previous section), or *segmentation*, i.e. serialization of a single step into multiple sequential steps. This example views segmentation as another instance of an adjunction.

Segmentation, as a serial/parallel distinction, is expressed by the *product-exponential* adjunction (example 20). As applied to functions, this adjunction is called the *(un)curry* transform in computer science: e.g., $+(x,y) \Leftrightarrow \tilde{+}(x)(y)$. Here, the general advantage of parallel processing is speed, since multiple arguments are applied concurrently, but at the expense of attention to more than one input.

4.2 Automatic versus controlled processes

Another dual-process distinction pertains to automatic versus controlled processes.[4] For example, counting a list of items, one-by-one, is a controlled (serial) process. However, a small number of items, up to about four, can be determined automatically and effortlessly, by a putative (parallel) process called *subitizing* (Rensink, 2013). These two processes are viewed as another adjoint situation.

The free-forgetful adjunction (example 21) expresses this distinction. Counting is modelled as a *monoid*: the natural numbers with addition, and zero as the identity element—counting is a serial process starting at zero and adding one until all items are counted. The free functor sends a set to the *free monoid* on that set, which affords the counting process. The forgetful functor sends a monoid to its underlying set, forgetting the monoid operation. So, the free functor constructs the control process, whereas the forgetful functor constructs the corresponding automatic process, a map that obviates the control steps, expressing the distinction between automatic and controlled processes. The automatic process associates lists of items to their count, the controlled process steps through each list item, hence the automatic process is fast, effortless and effectively parallel since the intermediate counting steps are obviated.

5. Discussion

We return to the original question: What can category theory contribute to cognitive science? The emerging view from the examples presented here is that category theory contributes the principle of cognition as universal construction. We look at how this principle extends to other aspects of cognitive science.

5.1 Beyond analogies

A recurring theme throughout this paper is the use of analogies, as a topic of study and as an expository device. Analogies play a key role in mathematics.

> A mathematician is a person who can find analogies between theorems; a better mathematician is one who can see analogies between proofs and the best mathematician can notice analogies between theories. One can imagine that the ultimate mathematician is one who can see analogies between analogies. (Randrianantoanina & Randrianantoanina, 2007, p. v—quoting Stefan Banach)

Mathematics is a cognitive activity (Lakoff & Nunez, 2000), and analogies play a key role in cognition. So, where do analogies come from?

A straightforward answer to this question comes from our earlier interpretation of analogies in terms of natural transformations. Natural transformations obtain from another kind of universal construction, called the *end of a (bi)functor* (Mac Lane, 1998). So, analogies come from another kind of universal construction.

An example is *relational schema induction* (Halford et al., 1998), where subjects are trained on a series of tasks conforming to a common relational (analogical) structure, which is needed to predict targets for novel cues in novel task instances. The analogies obtain by the end of a functor modeling the task instances (Phillips, 2020a).

5.2 Categorical compositionality

All constructions are derived from a more general construction, called *Kan extension* (definition 10).

> The notion of Kan extensions subsumes all the other fundamental concepts of category theory. (Mac Lane, 1998, p. 248).

Duality is not a universal mapping property as such. However, duality is incorporated as the left and right Kan extensions (remark 10). For example, products obtain from a right Kan extension (example 22); dually, coproducts obtain from the left Kan extension (example 23). Other limits are constructed likewise (remark 9).

Products, and limits generally, are constructed in two closely related ways. The direct way is by application of the product (or, limit) functor that directly constructs the product (or, limit) from its constituents. The indirect way is via the relevant universal mapping property. The product (limit) functor is right adjoint to the diagonal functor. Adjoints obtain from Kan ex-

4) See Evans and Stanovich (2013) for a list of characteristic distinctions between Type 1 and Type 2 processes.

tensions (example 24), hence the product (limit) construction is uniquely determined by an extension (remark 11). Thus, category theory provides the how and why of compositionality, as a functorial construction that derives from a universal mapping property (i.e. as an optimally consistent construction).

5.3 Empirical/Methodological implications

Categories are abstractly defined. No specific claims are made about the nature of the objects, morphisms and composition operation, only that they cohere in a certain way. However, there is more to category theory than abstraction.

> Put differently, good general theory does not search for the maximum generality, but for the *right* [emphasis added] generality. (Mac Lane, 1998, p. 108)

Good theories in the natural sciences make testable predictions. What testable predictions follow from the category theory principle of universal construction?

All universal constructions share the same general form. There is a family of morphisms whereby every member is uniquely composed from a common mediating morphism. The empirical implication is that if one has the mediating morphism and each of the unique constituent morphisms then necessarily one has their compositions. This situation arises as the product of two sets of morphisms. So, a test of this empirical implication is analogous to the test of generalization for the product construction in the cue-target learning task (section 3). Here, the constituents are themselves morphisms.

The converse situation is that removing the mediating morphism implies removing the entire family of morphisms for the universal construction. This situation was detailed in an analysis (Phillips & Wilson, 2016a) of the claim of compositionality in birdsong (Suzuki, Wheatcroft, & Griesser, 2016). The methodological implication follows from our category theory account of systematicity. If birdsong is compositional in a way that is analogous to human language, then there needs to be some test for systematicity, i.e. interfering with the process corresponding to the mediating morphism implies interfering with all systematically related processes (universally) constructed that way. Without such tests, claims of human-like compositionality are premature.

Another example pertains to cognitive development. For instance, young children below five years of age typically find certain tasks difficult relative to older children, despite comparative performance in other areas, and this difference has been attributed to a capacity to process relations (Halford et al., 1998, 2014). An early example used category theory to make empirical predictions in this regard (Halford & Wilson, 1980). Later work related development of cognitive capacity to differences between certain universal constructions (Phillips, Wilson, & Halford, 2009).

The categorical explanation for compositionality from universal construction has another methodological implication. Universal constructions are only *unique up to a unique isomorphism* (theorem 1). So, for example, any object and pair of maps that together satisfy the unique-existence conditions qualifies as a product.[5] Thus, cognitive representations need not be (canonically) composed from constituent symbols or dimensions (Phillips, 2020b), whereby constituents are "tokened" (inscribed, or written out) as part of the tokening of their complex hosts—a characteristic feature of classical (symbolic) compositionality (Fodor & Pylyshyn, 1988). The "atoms of thought" need not be classically tokened, and may only have meaning in an established context. The methodological implication is that context is primal, whence compositionality necessarily follows by universal construction.

5.4 Other contributions

Category theory approaches to cognitive science are in their infancy. We have just scratched the surface with one particular approach, vis-a-vis, various instances of universal constructions. Applications of category theory to other areas of cognitive science, not yet mentioned, include language (Clark, Coecke, & Sadrzadeh, 2008), memory and neural systems (Ehresmann & Gomez-Ramirez, 2015; Healy, Olinger, Young, Taylor, Caudell, & Larson, 2009), including learning by backpropagation (Fong, Spivak, & Tuyeras, 2017; Fong & Johnson, 2019), consciousness (Tsuchiya, Taguchi, & Saigo, 2016; Tsuchiya & Saigo, 2020), and philosophical aspects of cognition (Ellerman, 2016a, 2016b). More generally, category theory provides a unifying framework for recursion and, dually, corecursion (Hinze & Wu, 2016), which serves as a basis for (co)recursive aspects of cognition (Phillips & Wilson, 2012), including methods for obtaining universal constructions (Phillips & Wilson, 2016c).

Looking further ahead, Quantum Probability theory was developed to address non-classical (quantum-like) aspects of cognition (Busemeyer & Bruza, 2012). Quantum systems are closely related to presheaves (Abramsky & Brandenburger, 2011), suggesting sheaf theory can be used to model quantum-like cognition (Phillips, 2020b). Category theory has revealed many hitherto unrecognized connections between formal

5) Just being isomorphic to the product object is not enough.

systems, foreshadowing other categorical approaches to cognitive science that are yet to be explored.

One promising direction is *enriched category theory* (Kelly, 2005), where the sets of morphisms between objects are replaced by more general structures. Enriched category theory affords methods for modeling resources (Fong & Spivak, 2018), which has an important role in our treatment of dual-process cognition.

5.5 Occam's razor and Chatton's dual

We have shown how universal construction affords a category theory principle for cognitive science. A universal construction appears to embody Occam's razor, i.e. the principle of simplicity over complexity, in that all solutions to a family of problems factor through a common component. Adjunctions appear to be the antithesis of this principle, in admitting distinct paths that are functionally identical. This preference for complexity over simplicity embodies the lesser known principle called *Chatton's anti-razor*, seen here as Occam's dual in the category theory sense. However, a universal construction is essentially a balance of these two principles: the uniqueness condition requires at most one morphism (Occam); at least one morphism is required for the existence condition (Chatton). Occam's razor is used as a heuristic to justify, e.g., biases in machine learning. Category theory affords a formal foundation for this heuristic as part of the universal construction principle.

5.6 Final remark

Note that the expository style of the current work alludes to our adjoint basis for dual-process cognition: an informal main text on one hand and a formal appendix on the other. The main text is an approximate, yet more readily accessible exposition with links to the latter more precise, yet densely written alternative. Bidirectional exploitation of such trade-offs, in general form, is seen as the quintessential advance of human cognition (Phillips, 2017).

References

Abramsky, S. & Brandenburger, A. (2011). The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, *13*, 113036.

Aizawa, K. (2003). *The systematicity arguments*. Studies in Mind and Brain. New York: Kluwer Academic.

Arbib, M. A. & Manes, E. G. (1975). *Arrows, structures, and functors: The categorical imperative*. London, UK: Academic Press.

Awodey, S. (2010). *Category theory* (2nd ed.). Oxford Logic Guides. New York, NY: Oxford University Press.

Barr, M. & Wells, C. (1990). *Category theory for computing science* (First ed.). Prentice Hall International Series in Computer Science. New York: Prentice Hall.

Brown, R. & Porter, T. (2006). *Category Theory: An abstract setting for analogy and comparison*, Vol. 3 of *Advanced Studies in Mathematics and Logic*, pp. 257–274. Monza, Italy: Polimetrica Publisher.

Busemeyer, J. R. & Bruza, P. D. (2012). *Quantum models of cognition and decision*. Cambridge, UK: Cambridge University Press.

Chomsky, N. (1980). *Rules and representations*. New York, NY: Columbia University Press.

Clark, S., Coecke, B., & Sadrzadeh, M. (2008). A compositional distributional model of meaning. In P. Bruza, W. Lawless, K. van Rijsbergen, D. Sofge, B. Coecke, & S.Clark (Eds.), *Proceedings of the Second Symposium on Quantum Interaction*, pp. 133–140. Oxford, UK: College Publications.

Ehresmann, A. C. & Gomez-Ramirez, J. (2015). Conciliating neuroscience and phenomenology via category theory. *Progress in Biophysics and Molecular Biology*, *119*, 347–359.

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

Ellerman, D. (2016a). Four ways from universal to particular: How Chomsky's language-acquisition faculty is not selectionist. *Journal of Applied Non-Classical Logics*, *26* (3), 193–207.

Ellerman, D. (2016b). On adjoint and brain functors. *Axiomathes*, *26* (1), 41–61.

Evans, J. S. B.T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7* (10), 454–459.

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

Fodor, J. A. & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28* (1–2), 3–71.

Fong, B. & Johnson, M. (2019). Lenses and learners. In J. Cheney & H.-S. Ko (Eds.), *Proceedings of the Eighth International Workshop on Bidirectional Transformations*, pp. 16–29.

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

Fong, B., Spivak, D. I., & Tuyeras, R. (2017). Backprop as functor: A compositional perspective on supervised learning. Preprint: arXiv:1711.10455v2.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7* (2), 47–59.

Gentner, D. & Forbus, K. D. (2011). Computational models of analogy. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2* (3), 266–276.

Halford, G. S. & Wilson, W. H. (1980). A category theory approach to cognitive development. *Cognitive Psychology*, *12*, 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., Guo, J., Gayler, R. W., Wiles, J., & Stewart, J. E. M. (1994). Connectionist implications for processing capacity limitations in analogies. In K. J. Holyoak & J. Barnden (Eds.), *Advances in Connectionist and Neural Computation Theory: Analogical Connections*, Vol. 2, chap. 7, pp. 363–415. Norwood, NJ: Ablex.

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.

Healy, M. J., Olinger, R. D., Young, R. J., Taylor, S. E., Caudell, T. P., & Larson, K. W. (2009). Applying category theory to improve the performance of a neural architecture. *Neurocomputing*, *72* (13–17), 3158–3173.

Hinze, R. & Wu, N. (2016). Unifying structured recursion schemes: an extended study. *Journal of Functional Programming*, *26*, e1.

Hofstadter, D. R. (2001). Analogy as the core of cognition. In D. Gentner, K. Holyoak, & B. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science*, pp. 499–538. Cambridge, MA: MIT Press.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kelly, G. M. (2005). Basic concepts of enriched category theory. *Reprints in Theory and Applications of Categories*, *10*, 1–136.

Lakoff, G. & Nunez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York, NY: Basic Books.

Laurence, S. & Margolis, E. (1999). Concepts and cognitive science. In E. Margolis & S. Laurence (Eds.), *Concepts: Core readings*, chap. 1, pp. 1–81. Cambridge, MA: MIT Press.

Lawvere, F. W. & Schanuel, S. H. (1997). *Conceptual mathematics: A first introduction to categories*. Foundations of Computing. Cambridge, UK: Cambridge University Press.

Leinster, T. (2014). *Basic category theory*, Vol. 143 of *Cambridge Studies in Advanced Mathematics*. Cambridge, UK: Cambridge University Press.

Mac Lane, S. (1997). Categorical foundations of the protean character of mathematics. In E. Agazzi & G. Darvas (Eds.), *Philosophy of mathematics today*, pp. 117–122. Dordrecht, The Netherlands: Kluwer Academic.

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

Mac Lane, S. & Moerdijk, I. (1992). *Sheaves in geometry and logic: A first introduction to topos theory*. New York, NY: Springer.

Marquis, J.-P. (2009). *From a geometrical point of view: A study of the history and philosophy of category theory* (Hardcover ed.). Logic, epistemology, and the unity of science. Cambridge, UK: Springer.

Navarrete, J. A. & Dartnell, P. (2017). Towards a category theory approach to analogy: Analyzing re-representation and acquisition of numerical knowledge. *PLoS Computational Biology*, *13* (8), e1005683.

Phillips, S. (2017). A general (category theory) principle for general intelligence: duality (adjointness). In T. Everitt, B. Goertzel, & A. Potapov (Eds.), *Artificial General Intelligence*, Vol. 10414 of *Lecture Notes in Artificial Intelligence*, chap. 16, pp. 57–66. Cham, Switzerland: Springer.

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

Phillips, S. (2018b). What underlies dual-process cognition? Adjoint and representable functors. In C. Kalish, M. Rau, J. Zhu, & T. T. Rogers (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, pp. 2250–2255. Austin, TX: Cognitive Science Society.

Phillips, S. (2020a). The limits of learning to learn. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*, pp. 2486–2491. Austin, TX: Cognitive Science Society.

Phillips, S. (2020b). Sheaving—a universal construction for semantic compositionality. *Philosophical Transactions of the Royal Society B*, *375* (1791), 20190303.

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.

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

Phillips, S. & Wilson, W. H. (2012). Categorial compositionality III: F-(co)algebras and the systematicity of recursive capacities in human cognition. *PLoS ONE*, *7* (4), e35028.

Phillips, S. & Wilson, W. H. (2016a). Commentary: Experimental evidence for compositional syntax in bird calls. *Frontiers in Psychology*, *7*, 1171.

Phillips, S. & Wilson, W.H. (2016b). Second-order systematicity of associative learning: a paradox for classical compositionality and a coalgebraic resolution. *PLoS ONE*, *11* (8), e0160619.

Phillips, S. & Wilson, W. H. (2016c). Systematicity and a categorical theory of cognitive architecture: universal construction in context. *Frontiers in Psychology*, *7*, 1139.

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.

Randrianantoanina, B. & Randrianantoanina, N. (Eds.) (2007). *Banach spaces and their applications in analysis: Proceedings of the International Conference at Miami University, May 22-27, 2006, in Honor of Nigel Kalton's 60th Birthday*. New York, NY: Walter de Gruyter.

Rensink, R. A. (2013). Perception and attention. In D. Reiserg (Ed.), *Oxford handbook of cognitive psychology*, chap. 7, pp. 97–116. New York, NY: Oxford University Press.

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

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence*, *46* (1-2), 159–216.

Spivak, D. I. (2014). *Category theory for scientists*. Cambridge, MA: MIT Press.

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications*, *7*, 10986.

Townsend, J. T. (2008). Mathematical psychology: Prospects for the 21st century: A guest editorial. *Journal of Mathematical Psychology*, *52* (5), 269–280.

Tsuchiya, N. & Saigo, H. (2020). Applying Yoneda's lemma to consciousness research: categories of level and contents of consciousness. Preprint: 10.31219/osf.io/68nhy.

Tsuchiya, N., Taguchi, S., & Saigo, H. (2016). Using category theory to assess the relationship between consciousness and integrated information theory. *Neuroscience Research*, *107*, 1–7.

Walters, R. F. C. (1991). *Categories and computer science*. Cambridge Computer Science Texts. Cambridge, UK: Cambridge University Press.

Weisstein, E. W. (2020). Anomalous cancellation. From MathWorld—A Wolfram Web Resource.

Appendix

A. Basic theory

This appendix provides the basic category theory used for our approach to cognition. Deeper and broader coverage of category theory, assuming different backgrounds, can be found in many textbooks on the topic (see e.g., Lawvere & Schanuel, 1997; Leinster, 2014; Simmons, 2011; Spivak, 2014).

Definition 1 (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

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

Example 1 (Sets)**.** The collection (class) of sets and functions between sets forms a category, denoted **Set**. Each set is an object and each function is a morphism. The identity morphisms are the identity functions, and composition is function composition.

Example 2 (Naturals, reals)**.** The natural numbers, N, with the usual order form a category, denoted $(N, \leq)$, that has numbers $n \in N$ for objects and order relations $m \leq n$ for morphisms. The identities are $n \leq n$, and composition is given by transitivity. Likewise, the set of reals, R, constitute a category, denoted $(R, \leq)$.

Definition 2 (Isomorphism)**.** A morphism $f : A \to B$ is called an *isomorphism* if there exists a morphism $g : B \to A$ such that $f \circ g = 1_B$ and $g \circ f = 1_A$. If g exists then it is called the *inverse* of f.

Definition 3 (Terminal)**.** In a category $\mathbf{C}$, a *terminal* object is an object, denoted 1, such that for every object Z in $\mathbf{C}$ there exists a unique morphism $u : Z \to 1$.

Example 3 (Singleton)**.** A terminal object in the category **Set** is any singleton set.

Example 4 (Infinity)**.** Infinity is the terminal object in the category $(N \cup \{\infty\}, \leq)$.

Definition 4 (Product)**.** In a category $\mathbf{C}$, a *product* of objects A and B is an object P, also denoted $A \times B$, together with morphisms $\acute{\pi} : P \to A$ and $\grave{\pi} : P \to B$, called *projections*, such that for every object Z and morphisms $f : Z \to A$ and $g : Z \to B$, all in $\mathbf{C}$, there exists a unique morphism $u : Z \to P$ such that $(f, g) = (\acute{\pi}, \grave{\pi}) \circ u$. Morphism u is also denoted $\langle f, g \rangle$ as it is determined by f and g.

Example 5 (Cartesian product)**.** In **Set**, the product of sets A and B is the *Cartesian product*—set of pairs drawn from A and B, i.e. $A \times B = \{(a, b) | a \in A, b \in B\}$.

Remark 1. Category theory constructions come in two forms: *primal* and *dual*. The dual construction is obtained by reversing the directions of the morphisms in the primal construction. The dual construction of a primal construction in the category $\mathbf{C}$ is the primal construction in the category, denoted $\mathbf{C}^{\mathrm{op}}$, which consists of the objects of $\mathbf{C}$ and the morphisms $f^{\mathrm{op}} : B \to A$ whenever $f : A \to B$ is a morphism of $\mathbf{C}$.

Example 6 (Coproduct)**.** The *coproduct* of objects A and B, in $\mathbf{C}$, is (Q, ι), where ι is the pair of morphisms $\grave{\iota} : A \to Q$ and $\acute{\iota} : B \to Q$, called *injections* (cf. product in $\mathbf{C}^{\mathrm{op}}$). In **Set**, the coproduct is *disjoint union*.

Definition 5 (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 all $A \in \mathcal{O}(\mathbf{C})$, and
- *compositionality*: $F(g \circ_{\mathbf{C}} f) = F(g) \circ_{\mathbf{D}} F(f)$ for all compatible morphisms $f, g \in \mathcal{M}(\mathbf{C})$.

Example 7 (Product)**.** The *product functor* constructs products, i.e. $\Pi : (A, B) \mapsto A \times B, (f, g) \mapsto f \times g$.

Example 8 (Projections)**.** *Projection functors* $\Pi_1 : (A, B) \mapsto A, (f, g) \mapsto f$ and $\Pi_2 : (A, B) \mapsto B, (f, g) \mapsto g$.

Example 9 (Diagonal)**.** The *diagonal functor* doubles objects and morphisms, i.e. $\Delta : A \mapsto (A, A), f \mapsto (f, f)$.

Example 10 (Inclusion)**.** The *inclusion functor* is the category theory analog of inclusion between sets: a functor identifying a *subcategory* of a category.

Example 11 (Constant)**.** A *constant functor*, written $K : \mathbf{C} \to \mathbf{D}$, sends every object and morphism in $\mathbf{C}$ to the object K and identity morphism 1_K in $\mathbf{D}$.

Definition 6 (Natural transformation)**.** Suppose functors $F, G : \mathbf{C} \to \mathbf{D}$. A *natural transformation* from F to G, written $\eta : F \dot{\to} G$, is a family of $\mathbf{D}$-morphisms $\{\eta_A : F(A) \to G(A) | A \in \mathcal{O}(\mathbf{C})\}$ such that $G(f) \circ \eta_A = \eta_B \circ F(f)$ for every morphism $f : A \to B$ in $\mathbf{C}$.

Remark 2. A natural transformation is called a *natural isomorphism* when every η_A is an isomorphism.

Example 12 (Projections)**.** $\acute{\pi} : \Pi \dot{\to} \Pi_1$ and $\grave{\pi} : \Pi \dot{\to} \Pi_2$.

Definition 7 (Universal morphism)**.** Let $F : \mathbf{C} \to \mathbf{D}$ be a functor and Y an object in $\mathbf{D}$. A *universal morphism* from F to Y is a pair (A, ψ) consisting of an object A in $\mathbf{C}$ and a morphism $\psi : F(A) \to Y$ in $\mathbf{D}$ such that for every object X in $\mathbf{C}$ and morphism $g : F(X) \to Y$ in $\mathbf{D}$ there exists a unique morphism $u : X \to A$ in $\mathbf{C}$ such that $g = \psi \circ F(u)$. ψ is called the *mediating morphism*.

Remark 3. The *dual* form of universal morphism is defined by arrow reversal: a universal morphism from an object X in $\mathbf{C}$ to a functor $G : \mathbf{D} \to \mathbf{C}$ is a pair (A, ϕ) consisting of an object A in $\mathbf{D}$ and a morphism $\phi : X \to G(A)$ in $\mathbf{D}$ such that for every object Y in $\mathbf{D}$ and morphism $f : X \to F(Y)$ in $\mathbf{C}$ there exists a unique morphism $u : A \to Y$ in $\mathbf{D}$ such that $f = G(u) \circ \phi$.

Example 13 (Product)**.** Pair $(A \times B, \pi)$ is a universal morphism from Δ to (A, B), where $\pi = (\acute{\pi}_A, \grave{\pi}_B)$, i.e. $\acute{\pi}_A : A \times B \to A$ and $\grave{\pi}_B : A \times B \to B$.

Remark 4. The mediating morphism (relabeled ψ_A) is the "optimal" member of natural transformation $\psi : F \dot{\to} Y$, where Y is the constant functor—every member of ψ factors through ψ_A.

Example 14 (Projection as mediating morphism)**.** The projections $(\acute{\pi}_A, \grave{\pi}_B)$ are the mediating morphisms of the product $A \times B$, hence the optimal members of the natural transformation $(\acute{\pi}, \grave{\pi}) : \Delta \dot{\to} (A, B)$.

Definition 8 (Limit)**.** A *limit* of a (J-shaped) functor $D : \mathbf{C} \to \mathbf{C}^J$ is a universal morphism from D to an object (functor) in $\mathbf{C}^J$—the category of functors (from J to $\mathbf{C}$) and natural transformations.

Example 15 (Product)**.** The product $(A \times B, \pi)$ is the limit of the diagonal functor, Δ, to the functor $(A, B) : 2 \to \mathbf{C}$, picking out the pair (A, B), where $J = 2$ is a two-object category with no non-identity morphisms.

Example 16 (Limits)**.** Other important limits (of shape J) are: *terminal* (0), *equalizer* ($* \rightrightarrows *$), and *pullback* ($* \to * \leftarrow *$). In **Set**, the terminal is any singleton set; the equalizer of $f, g : A \to C$ is the set of elements $a \in A$ such that f and g agree (i.e., $f(a) = g(a)$); and the pullback of $f : A \to C$ and $g : B \to C$ is the set of pairs $(a, b) \in A \times B$ such that $f(a) = g(b)$.

Remark 5. All (finite) limits are constructed from products and equalizers.

Theorem 1 (Uniqueness)**.** Let $F : \mathbf{C} \to \mathbf{D}$ be a functor, Y an object in $\mathbf{D}$, and (A, ψ) be a universal morphism from F to Y. If (A', ψ') is a universal morphism from F to Y then there exists a unique isomorphism $\upsilon : A \cong A'$ such that $\psi = \psi' \circ F(\upsilon)$.

Remark 6. In other words, universal morphisms are unique up to unique isomorphism.

Definition 9 (Adjunction)**.** An *adjunction* from a category $\mathbf{C}$ to a category $\mathbf{D}$ is a triple, $(F, G, \eta) : \mathbf{C} \rightharpoonup \mathbf{D}$, consisting of functors $F : \mathbf{C} \to \mathbf{D}$ and $G : \mathbf{D} \to \mathbf{C}$, and a natural transformation $\eta : 1 \dot{\to} G \circ F$ such that for every object X in $\mathbf{C}$ the pair $(F(X), \eta_X)$ is a universal morphism from X to G. Functor F is called the *left adjoint* of G (G is called the *right adjoint* of F), denoted $F \dashv G$, and η is called the *unit* of the adjunction.

Remark 7. Equivalently, an adjunction is a triple, $(F, G, \varepsilon) : \mathbf{C} \rightharpoonup \mathbf{D}$, consisting of a natural transformation $\varepsilon : F \circ G \dot{\to} 1$ such that for every object Y in $\mathbf{D}$ the pair $(G(Y), \varepsilon_Y)$ is a universal morphism from F to Y. (ε is called the *counit* of the adjunction.)

Example 17 (Approximation)**.** The functions *Ceil* and *Floor* are adjoints to inclusion.

- i *Ceil* $\dashv$ *Incl* with unit $x \leq \lceil x \rceil$ and counit $y \leq y$.
- ii *Incl* $\dashv$ *Floor* with unit $x \leq x$ and counit $\lfloor y \rfloor \leq y$.

Example 18 (Product)**.** The diagonal functor is left adjoint to the product functor: $\Delta \dashv \Pi$ with unit $\langle 1, 1 \rangle : Z \to Z \times Z$ and counit $(\acute{\pi}, \grave{\pi}) : (A \times B, A \times B) \to (A, B)$.

Remark 8. Every adjunction induces a natural isomorphism: $\phi : \mathrm{Hom}(F-, -) \cong \mathrm{Hom}(-, G-) : \psi$.

Example 19 (Bounds)**.** Instantiating $F \dashv G$ as:

- i *Ceil* $\dashv$ *Incl* yields $\lceil x \rceil \leq y \Leftrightarrow x \leq y$, and
- ii *Incl* $\dashv$ *Floor* yields $x \leq y \Leftrightarrow \lfloor x \rfloor \leq y$.

Example 20 (Serial/parallel)**.** The product functor $\Pi_B : A \mapsto A \times B$. The *exponential functor* $\Lambda_B : C \mapsto C^B$. The exponential functor is right adjoint to the product functor, i.e. $\mathrm{Hom}(A \times B, C) \cong \mathrm{Hom}(A, C^B)$. For functions, this situation is the familiar *curry-uncurry* pair of operators in computer science: $f(a, b) = \tilde{f}(a)(b)$.

Example 21 (Free/forgetful)**.** A general class of adjoint situations arise from the relationship between an algebra and its underlying set. For example, a *monoid*, $(M, \cdot, e)$, is a set, M, with a binary operation, $\cdot$, that is associative (i.e. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$) and unital (i.e. $a \cdot e = e = e \cdot a$), and a *unit* element (i.e. $e \in M$). The *forgetful functor* sends monoids to their underlying set, i.e. $U : (M, \cdot, e) \mapsto M$, and the free functor sends each set S to the *free monoid* on S, i.e. $F : S \mapsto S^*$. The free/forgetful functor is a left/right adjoint. A familiar example is the free functor that constructs "words" (strings) from an alphabet, $F : A \mapsto A^*$.

Definition 10 (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, η) consisting of a functor $R : \mathbf{B} \to \mathbf{C}$ and a natural transformation $\eta : RF \dot{\to} X$

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 $\mu = \eta \circ \delta_F$.

Example 22 (Product as right Kan extension)**.** A product is obtained as the right Kan extension of $(A,B) : 2 \to \mathbf{C}$ along $! : 2 \to 1$.

Remark 9. All limits obtain this way.

Remark 10. The *left Kan extension* of X along F is the dual of the right Kan extension.

Example 23 (Coproduct as left Kan extension)**.** A coproduct is obtained as the left Kan extension of $(A,B) : 2 \to \mathbf{C}$ along $! : 2 \to 1$.

Example 24 (Adjunction as Kan extension)**.** Suppose $F : \mathbf{C} \to \mathbf{D}$ and $G : \mathbf{D} \to \mathbf{C}$ are left and right adjoints. Left adjoint F obtains as the right Kan extension of the identity functor on $\mathbf{D}$ along G. Dually, G obtains as the left Kan extension of the identity on $\mathbf{C}$ along F.

Remark 11. Kan extensions are unique up to unique isomorphism. Hence, for a given functor, the (left, or right) adjoint (if it exists) is likewise unique.

B. Presheaves and sheaves

Some basic constructions are provided here (see Phillips, 2018a, 2020b, for further details). Deeper introductions to sheaf theory from a category theory perspective are also available (see e.g., Leinster, 2014; Mac Lane & Moerdijk, 1992).

Definition 11 (Topological space)**.** A *topological space* is a pair (X,T) consisting of a set X and a collection of subsets T of X, called the *open sets* of X, that consists of the empty set and X, arbitrary unions (if U and V are open sets of X, then so is $U \cup V$), and finite intersections (if U and V are open sets of X, then so is $U \cap V$). The set T is called the *topology* of X, and the space is sometimes simply denoted X.

Example 25 (Indiscrete, discrete)**.** The *indiscrete topology* consists of just the empty set and X. The *discrete topology* consists of all subsets of X.

Remark 12. A topological space is a category consisting of open sets (objects) and inclusions (morphisms).

Definition 12 (Presheaf)**.** A *presheaf* is a functor $F : X \to \mathbf{Set}$ that sends each inclusion $V \to U$ of X to the *restriction morphism* $f|_V : F(U) \to F(V)$, i.e. each function over U is restricted to V.

Definition 13 (Sheaf)**.** A *sheaf* is a universal presheaf.

Remark 13. The collection of (pre)sheaves on a topological space X forms a (functor) category, whence *(pre)sheaf morphisms* are natural transformations.

Definition 14 (Sheaving)**.** Let $\mathbf{PSh}(X)$/$\mathbf{Sh}(X)$ be the category of presheaves/sheaves on X. The *sheaving functor*, $F^+ : \mathbf{PSh}(X) \to \mathbf{Sh}(X)$, sends presheaves to their universal presheaves (sheaves).

Remark 14. Sheaving is left adjoint to inclusion.

Definition 15 (Continuous function)**.** A *continuous function* is a function between topological spaces, $f : X \to Y$, that *reflects* open sets: if U is an open set of Y, then the preimage $f^{-1}[U]$ of U is an open set of X.

Example 26 (Images)**.** Every continuous function $f : Y \to X$ induces two functors.

i The *direct image functor*, $f_* : \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$, sends sheaves on Y to sheaves on X, by assigning to the open sets of X the data on their preimages.

ii The *inverse image functor*, $f^* : \mathbf{Sh}(X) \to \mathbf{Sh}(Y)$, sends sheaves on X to sheaves on Y. This operation involves (co)limits, because the image of an open set in Y may not be an open set in X, and the corresponding assignment may not be a sheaf.

The direct image functor is right adjoint to the inverse image functor.

Steven Phillips

Steven Phillips is Chief Senior Researcher at the National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan. His research pursues category theory approaches to cognitive science. He is co-author of the book, *Categorizing Cognition: Toward Conceptual Coherence in the Foundations of Psychology* (MIT Press).