The limits of learning to learn

Steven Phillips (steve@ni.aist.go.jp)
Mathematical Neuroscience Research Group, National Institute of Advanced Industrial Science and Technology (AIST),
Tsukuba, Ibaraki 305-8568 JAPAN

Abstract

Learning to learn is a reduction in the amount of training needed for task attainment across a series of similar tasks. Transfer differentiates (adult) humans from other species, portending a window into unique aspects of human learning. However, its unclear whether such differences are quantitative, or qualitative and what it means for the evolution/development of cognition. In this paper, learning is regarded as a (categorical) limit. A limit is a universal construction, and so transfer follows from a (generalized) optimization process. This result provides a formal basis for comparison/contrast of learning transfer in humans and other species—another step to bringing the empirical question into sharper relief.

Keywords: Learning to learn; learning set transfer; relational schema induction; category theory; category; functor; limit

Introduction

Learning to learn (also called *learning set transfer*) refers to an improvement in training rate across a series of similar learning tasks (Harlow, 1949). For example, suppose each task is to learn a configural association, where the target responses depend on interacting cues (colour and shape): in one task instance, subjects are trained to select square over triangle when presented with a square on a blue coloured background, and triangle over square when a square is presented on a green background; in another instance, subjects are trained to select circle over cross when a circle is presented on a red background, and cross over circle when the background for the presented circle is yellow. Evidence of learning to learn is observed when the number of training trials needed to reach some criterion for successful learning decreases on subsequent instances of the task.

Many species have a capacity for learning, yet its unclear whether differences in transfer are quantitative, or qualitative (Bitterman, 1975; Warren, 1965). For instance, some authors have argued for an association-based account of learning that is extendable to propositions, so providing a basis for higher cognition (Mitchell, Houwer, & Lovibond, 2009). An associative model could assume that learning rate changes with prior experience (Miller, Barnet, & Grahame, 1995), thereby providing a quantitative explanation for differences in learning transfer: greater transfer is linked to greater change in learning rate. Yet, other authors argue that the propositional (relational) aspects of cognition—inferring targets from relations between stimuli—are qualitatively unique to humans (Penn, Holyoak, & Povinelli, 2008), and most developed in adults (Halford, Wilson, & Phillips, 1998). Resolving such disputes over accounts of learning transfer should inform the nature, evolution and development of cognition.

To this end, the *relational schema induction* paradigm was developed to distinguish associative versus relational models of learning transfer (Halford, Bain, Maybery, & Andrews, 1998; Halford & Busby, 2007). These models make contrasting predictions following feedback on *information trials* that are necessary to determine the responses to the other stimuli. For example (above), having seen that circle is preferred over cross when presented on a red background (information trial), a relational model predicts that cross is preferred over circle when the background is yellow, because the second task involves the same relation(al schema). By contrast, the associative model makes no prediction for a novel stimulus, having not been paired with a target before. The empirical results support a relational model for learning transfer (Halford, Bain, et al., 1998; Halford & Busby, 2007).

Relational schema induction is ideal for comparing species and age groups, because it does not require language to administer. Yet, the empirical results raise a conundrum. On one hand, associative processes appear incapable of accounting for this form of transfer. On the other hand, this form of transfer may be beyond the capacities of non-humans and young children (Halford, Wilson, Andrews, & Phillips, 2014). How, then, does a systematic capacity for learning transfer develop?

A first step towards redressing this conundrum is to employ a more general theory incorporating both relational and associative aspects of cognition. So, the purpose of this paper is to provide a common theoretical grounding of relational schema induction and learning set transfer. That common ground is the category theory concept of *limit* (Mac Lane, 1998), which is motivated by the following points. Consistent induction of relational schemas (Halford, Bain, et al., 1998) is another form of *systematicity* (Fodor & Pylyshyn, 1988), which is explained by the category theory concept of *universal morphism* (Phillips & Wilson, 2010). And, universal morphisms are equivalently certain kinds of limits (Mac Lane, 1998). Accordingly, we show that these two forms of learning to learn obtain from a particular kind of limit process.

The paper proceeds with an outline of the basic theory of limits, in the next section. This categorical theory of limits is then applied to examples of relational schema induction and learning set transfer in the subsequent two sections, where it is shown that both forms of learning to learn are captured by a certain kind of limit, called the *end of a functor*. (Formal details appear in the Appendix.) These results are discussed in the broader context of learning transfer, comparative and developmental cognition, in the final section.

Limits

A (categorical) limit is a kind of optimal solution to a given problem. Such limits generalize the more familiar limit of a function. This section provides an intuitive guide to categorical limits for the purpose of modeling learning transfer.

A categorical limit depends on more basic concepts, which are introduced first. In brief, category theory starts with the concept of a *category* (definition 1), which consists of *objects*, relations between objects, called *morphisms*, and an operation for combining morphisms, called *composition*. Most models of cognition involve sets and functions, which are objects and morphisms in the category **Set** (example 1). A map between categories is a *functor* (definition 3), preserving categorical structure in the homomorphic sense—a functor is a category homomorphism. Two functors are compared by a *natural transformation* (definition 4), and an optimal comparison is a *universal morphism* (definition 5). A *limit* (definition 6) is a kind of universal morphism, and so an optimal construction.

A *product* (definition 2) is a limit, affording a basic form of compositonality (example 2) in models of cognition. Products are: constructed by functors (example 3), instances of universal morphisms (example 4), and derived by limit processes (example 5). So, categorical limits provide a framework for inducing (compositional) structure and learning transfer.

Learning transfer is modeled by a further generalization of limit, called the *end of a (bi)functor* (definition 8). A bifunctor is an analog of bivariate function. The intuition here is to regard each variable as pertaining to a task instance, whereby the end computes (reconstructs) the common relation between the structures underlying each instance (theorem 1). As we shall see in the following two sections, this reconstruction process is the formal foundation for our categorical treatment of learning to learn.

Relational schema induction

First we describe the relational schema induction paradigm and then we provide a category theory account of induction.

Relational schema induction consists of a series of tasks conforming to a group-like structure (Halford, Bain, et al., 1998; Halford & Busby, 2007). Suppose stimuli are drawn from the set of shapes $Sh = \{\heartsuit, \clubsuit\}$ and the set of trigrams $Tri = \{\text{BEH}, \text{FUT}, \text{PEJ}, \text{ROY}\}$, and the task is to learn a map from the set of shape-trigram pairs to the set of trigrams: $\tau_1 = Sh \times Tri \to Tri$, e.g., $(\heartsuit, \text{BEH}) \mapsto \text{FUT}$ and $(\clubsuit, \text{FUT}) \mapsto \text{PEJ}$. If we view trigrams as vertices of a square, the shapes correspond to horizontal and vertical reflections. Shapes and trigrams are unique across tasks and each task conforms to the same group-like structure: e.g., in another task, shapes are drawn from the set $\{\diamondsuit, \spadesuit\}$ and trigrams from the set $\{\text{HUQ}, \text{KES}, \text{NIZ}, \text{XAY}\}$ and the mappings follow accordingly (see figure 1). Participants consistently induce the relational schema after several task instances (Halford, Bain, et al., 1998): the limit of learning transfer as indicated by correct prediction on novel trials.

This form of induction is modeled as a kind of limit in category theory. Specifically, the relations between cues and targets are modeled as *monoid* (definition 9) *actions on a set* (definition 10): each shape is treated as an action that sends trigrams to trigrams (example 8), and the set of trigrams for each task instance is called an *M-set*. Each M-set corresponds

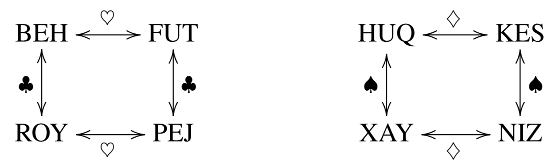

Figure 1: Two task instances for relational schema induction.

to a task, and a map between tasks that is compatible with (i.e. "preserves") the actions on trigrams is an *equivariant map* (definition 11). The collection of (possible) tasks and (action-preserving) maps between tasks, for a monoid M, forms a category, denoted **MSet** (remark 10). Participants are only given feedback on the target (trigram) that is associated with the given cue (shape-trigram pair), not the monoid generating those associations. This situation is modeled as the forgetful functor $U : \mathbf{MSet} \to \mathbf{Set}; (S, \sigma) \mapsto S$, which supplies the set of elements, S, forgetting the actions, σ (remark 11). Computing the end of a functor involving U recovers the monoid M, which contains the actions (theorem 1). (The unit of the monoid, e, corresponds to the "no-change" action where trigrams map to themselves, which is assumed but not part of the experiment.) The recovered monoid affords correct predictions on novel task instances: the two information trials determine the correspondence between shapes and elements of the monoid, hence how those shapes act on the other trigrams in novel trials of a new task instance.

In detail, suppose the task instances depicted in figure 1. The first task is represented as the pair (S, σ) consisting of the set of trigrams $S = \{\text{BEH}, \text{FUT}, \text{PEJ}, \text{ROY}\}$ and the set of shapes (actions) $\sigma = \{\heartsuit, \clubsuit\}$ where, e.g., $\heartsuit : \text{BEH} \mapsto \text{FUT}$ and $\clubsuit : \text{BEH} \mapsto \text{ROY}$. The second task is represented as (R, ρ), where $R = \{\text{HUQ}, \text{KES}, \text{NIZ}, \text{XAY}\}$ and $\rho = \{\diamondsuit, \spadesuit\}$, e.g., $\diamondsuit : \text{HUQ} \mapsto \text{KES}$ and $\spadesuit : \text{HUQ} \mapsto \text{XAY}$. Computing the end of the functor $\text{Hom}(U-, U-)$ reconstructs the monoid, i.e. $\int_{\mathbf{MSet}} \text{Hom}(U-, U-) \cong M$, consisting of the set of actions $\{h, v\}$ and monoid operation: e.g., $h \cdot h = e$ (see example 8). Ends are limits, which are *unique up to unique isomorphism* (Mac Lane, 1998), so this monoid is essentially the same as the monoid with the two actions relabeled as particular shapes.

Ends obtain as optimal constraint satisfaction (remark 12), where the constraints are implicitly specified by feedback on stimulus-response trials for the tasks. Sets $\text{Hom}(S, S)$ and $\text{Hom}(R, R)$ consist of all possible maps between trigrams within a task, and $\text{Hom}(S, R)$ consists of all possible maps of trigrams between tasks. Commutativity (diagram 4) constrains the candidate solution sets to only those sets whose elements pick out the trigram mappings for each task instance that conjointly satisfy *equivariance* (definition 11) between task instances. Universality (remark 4) further constrains the candidates to only those sets whose elements are necessary and sufficient for commutativity, i.e. the relational schema (monoid) common to all task instances. In this way, relational schema induction is a form of optimal constraint satisfaction.

Learning set transfer: configural association

Configural association can be seen as a kind of relational schema induction (Halford, Bain, et al., 1998), which has the structure of the logical operation, *exclusive-or* (remark 7). In this section, we show how learning set transfer for configural association is also a limit in the same manner as relational schema induction, as shown in the previous section.

In a configural association task, participants learn to associate cues to targets depending on context: e.g., in the context of a green display background, triangle is associated to triangle and square is associated to square; in the context of a blue background, triangle associates to square and square associates to triangle. After learning these associations a new instance of configural association is administered. This new instance consists of different shapes and colours: e.g., in the context of a yellow display background, circle associates to circle and cross associates to cross; in the context of a brown background, circle associates to cross and cross associates to circle. Learning set transfer is observed when the number of training trials to criterion for subsequent tasks decreases.

Configural association is also modeled as monoid actions on sets, whereby colours correspond to actions on shapes: e.g., $\mathrm{G} : \triangle \mapsto \triangle$, $\mathrm{G} : \square \mapsto \square$, $\mathrm{B} : \triangle \mapsto \square$, $\mathrm{B} : \square \mapsto \triangle$. In this case, the monoid corresponds to exclusive-or: $\mathbb{Z}_2$ (remark 7). Accordingly, the collection of such tasks and their equivariant maps forms a category, and the monoid is reconstructed by computing the end of the functor, $\mathrm{Hom}(U-, U-)$, as in the previous example. The monoid is then applied to a new task instance given a single information trial, affording target prediction for the other three cues. In this way, learning set transfer for configural association and relational schema induction are two instances of reconstruction.

Discussion

The approach presented here places relational schema induction and learning set transfer on a common footing: both forms of learning to learn obtain from the same (general) limit process—the end of a functor. Indeed, this approach clarifies the close connection between the two examples of relational schema induction and configural association: the common structure underlying task instances of the former is the (categorical) product of the common structure underlying task instances of the latter (remark 8). The rest of this section considers the implications of this theory for the nature, development and evolution of cognition.

Relational versus associative processes

The relational schema induction paradigm was introduced to assess whether learning transfer depends on relational or associational processes (Halford, Bain, et al., 1998). On one hand, some authors have argued that developmental differences depend on a capacity to process relational information (Halford, Wilson, & Phillips, 1998; Penn et al., 2008). However, other authors argue that associative processes are sufficient (Leech, Mareschal, & Cooper, 2008; Mitchell et al., 2009). Our approach shows how these disparate views are reconciled.

Specifically, to induce the common structure, subjects must first learn the basic cue-target associations that constitute a task instance. The trigram-trigram associations for a task S and a given action constitute a map in the set $\mathrm{Hom}(S, S)$. Thus, there is an associative (first-order) component to induction. However, to recover the structure, subjects must also compute a limit (end): a relation between the associations within task instances, i.e. $\mathrm{Hom}(S, S)$ and $\mathrm{Hom}(R, R)$, that is constrained by equivariance, i.e. the maps $\mathrm{Hom}(1_S, f)$ and $\mathrm{Hom}(f, 1_R)$. So, there is also a relational (second-order) component in both paradigms. These components are independently manipulated as the cardinalities of set X: number of elements acted on, and monoid M: number of relations acting on X, respectively. The arity of the (product) monoid can also be varied from unary to binary, and so on—i.e. as the n-ary product monoid, $\mathbb{Z}_2^n \cong \Pi_1^n \mathbb{Z}_2$ (remark 8). Hence, the empirical implications of associative and relational information on learning transfer can both be assessed.

Notice that although the relational schema induction and learning set paradigms (as considered here) involve one-to-one correspondence between the elements (e.g., shapes and trigrams) of different task instances, this principle is derived by our approach, not assumed. Relational schema induction was considered to involve the mapping of structure (Halford, Bain, et al., 1998), for example, as specified by *structure mapping theory* (Gentner, 1983). However, structure mapping theory and related models of analogy generally assume a one-to-one correspondence constraint (Gentner & Forbus, 2011). Here, this principle is derived from computing the end of a functor, which affords a further generalization of the induction/transfer paradigm whereby task instances are homomorphic but not necessarily isomorphic. This situation is automatically captured by the theory of ends.

Systematicity

The systematic consistency with which subjects induce the relational schema and transfer this knowledge across task instances (Halford, Bain, et al., 1998; Halford & Busby, 2007) raises a familiar systematicity challenge (Fodor & Pylyshyn, 1988). In the current context, this challenge is to explain *why* subjects who exhibit transfer on one task also exhibit transfer on another task. This property pertains to a systematicity of learning, or *second-order systematicity* (Aizawa, 2003; Phillips & Wilson, 2016a).

A category theory explanation is that systematicity obtains from a universal construction (Phillips & Wilson, 2010). Ends are a form of universal construction (Mac Lane, 1998). Hence, the explanation for the (second-order) systematicity of learning exhibited in the relational schema induction experiments (Halford, Bain, et al., 1998; Halford & Busby, 2007) follows from computing the end of the appropriate functor. Computing the end (or universal morphism) follows from another kind of universal process: *categorical (co)recursion* (Phillips & Wilson, 2016b). So, the current work affords both structural and computational explanations for learning to learn.

Continuous versus punctate development

An overarching question concerns the development of such learning to learn capacity, *Is development of a capacity for learning transfer continuous, or punctate (i.e. a change in quantity, or quality)?* Some authors have argued that a capacity for processing relational information is unique to humans (Penn et al., 2008). And, other authors have gone further to argue that relations of varying complexity also differentiate age groups within humans (Halford, Wilson, & Phillips, 1998; Halford et al., 2014). Our approach provides two intimately related perspectives on this issue of evolution/development. On one hand the relational induction task involves a (binary) product of the structure of the configural association task (unary), indicating discontinuity. On the other hand unary relations are equivalent to binary relations with one argument fixed, suggesting continuity from the lower to higher complexity situation. This equivalence is expressed as the isomorphism $A \times 1 \cong A$, where 1 is the *terminal object*—another kind of limit—which is any singleton set in **Set**. There is a quantitative difference in terms of product arity (i.e. one versus two) and a qualitative difference in terms of relations (i.e. unary versus binary). In this regard, the transition from unary to binary products is essentially a transition between binary products with one variable argument to two variable arguments, which concords with *relational complexity theory* as the number of *dimensions of variation* that must be considered conjointly (Halford, Wilson, & Phillips, 1998; Halford et al., 2014).

Further work

New approaches raise new questions and directions for further work. In this paper, we focused on the induction of common structure, but we have not provided a categorical account of how that structure affords generalization to novel cues given the information trials. One approach to generalization is to employ a *free functor* (Mac Lane, 1998), whereby partial knowledge of the task instance learned from the information trials is "freely" completed for prediction of the targets for the other (novel) cues. This functor constructs the *free M-set representation* on the given set, i.e. $F : S \mapsto (S, \sigma)$, which is the "pseudo-inverse" of the forgetful functor $U : (S, \sigma) \mapsto S$ (remark 11). Free and forgetful functors form an *adjoint pair*, which is another kind of universal construction (Mac Lane, 1998), hence related to our categorical explanation for induction. (Diagonal and product functors also form an adjoint pair.) The details are beyond the space available, but a closely related example was given in regard to systematicity with respect to this task (Phillips & Wilson, 2010).

Another important direction is to extend this approach to the probabilistic setting. Typically, transfer is multi-shot, not one-shot. We focused on one-shot transfer, because it is regarded as a hallmark of human-level transfer (Halford et al., 2014). Further work is needed to understand the link between one-shot and multi-shot learning transfer, as commonly exhibited in non-human studies (Bitterman, 1975; Harlow, 1949).

Our approach has been to consider a more general theory to incorporate apparently different forms of cognitive process: relational versus associative. Yet, more general theory seems more removed from the underlying neuroscience, which raises questions about the link to the neurocomputational system. As observed elsewhere (Phillips, 2019), some category theory constructions pertaining to constraint satisfaction, like the ones employed here, are reminiscent of a neural network model of analogy (Doumas, Hummel, & Sandhofer, 2008). So, another direction is to investigate the formal links between the current theory and such models.

The theorem used here is a simple case of a far more general theory that affords the reconstruction of other kinds of algebras and is not limited to just sets. Thus, we expect our approach also applies to other forms of induction beyond the examples presented here. The present work is an entrée to a course of cognitive theory, whereby induction is modeled as reconstruction and generalization as completion.

Despite the formal elegance of a category theory approach, this approach does not directly say why subjects fail to induce the relevant structure. Inducing structure depends on both associative and relational components, as discussed earlier, which are putatively linked to differences in working memory systems across different cohorts (Halford et al., 2014). By unifying various forms of learning transfer, the theory makes clearer how such questions can be addressed experimentally. A valuable aspect of good theory is to bring such questions into sharper relief.

References

Aizawa, K. (2003). *The systematicity arguments*. New York: Kluwer Academic.

Bitterman, M. E. (1975). The comparative analysis of learning. *Science*, *188*, 699–709.

Doumas, L. A. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review*, *115*(1), 1–43.

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

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., 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., Andrews, G., & Phillips, S. (2014). *Categorizing cognition: Toward conceptual coher-*

ence 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*, *42*, 51–65.

Leech, R., Mareschal, D., & Cooper, R. P. (2008). Analogy as relational priming: A developmental and computational perspective on the origins of a complex cognitive skill. *Behavioral and Brain Sciences*, *31*(4), 357–414.

Mac Lane, S. (1998). *Categories for the working mathematician* (2nd ed.). 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.

Mitchell, C. J., Houwer, J. D., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*, 183–246.

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. (2019). Sheaving—a universal construction for semantic compositionality. *Philosophical Transactions of the Royal Society B*, *375*(1791), 20190303.

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. (2016a). 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. (2016b). Systematicity and a categorical theory of cognitive architecture: universal construction in context. *Frontiers in Psychology*, *7*, 1139.

Warren, J. M. (1965). Primate learning in comparative perspective. In *Behavior of nonhuman primates* (Vol. 1, pp. 249–281). Academic Press.

Appendix A: Basic theory

Definition 1 (Category)**.** A *category* $\mathbf{C}$ consists of a collection of *objects*, $O(\mathbf{C}) = \{A, B, \dots\}$, 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 $A \in O(\mathbf{C})$ 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 1. The collection of morphisms in $\mathbf{C}$ with domain A and codomain B is called a *hom-set*, denoted $\mathrm{Hom}(A, B)$.

Example 1 (**Set**)**.** In **Set**, the objects are sets, the morphisms are functions, and composition is composition of functions. The identity morphisms are the identity functions.

Remark 2. The *opposite category*, denoted $\mathbf{C}^{\mathrm{op}}$, has all the objects and "reversed" morphisms of $\mathbf{C}$: a morphism $f : A \to B$ in $\mathbf{C}$ is the morphism $f^{\mathrm{op}} : B \to A$ in $\mathbf{C}^{\mathrm{op}}$.

Definition 2 (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 2 (Cartesian product)**.** In **Set**, the product of sets A and B is the *Cartesian product*: $A \times B = \{(a, b) | a \in A, b \in B\}$ and projections $\acute{\pi} : (a, b) \mapsto a$ and $\grave{\pi} : (a, b) \mapsto b$. The unique morphism, u, is the function $\langle f, g \rangle : z \mapsto (f(z), g(z))$.

Definition 3 (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 $A \in O(\mathbf{C})$, and
- *compositionality*: $F(g \circ_{\mathbf{C}} f) = F(g) \circ_{\mathbf{D}} F(f)$ for every pair of compatible morphisms $f, g \in \mathcal{M}(\mathbf{C})$.

Example 3 (Diagonal, product)**.** The *diagonal functor* $\Delta : \mathbf{C} \to \mathbf{C} \times \mathbf{C}$ sends each object and each morphism to their pairs, and the *product functor* $\Pi : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ sends each pair of objects and each pair of morphisms to their products:

- $\Delta : A \mapsto (A, A), f \mapsto (f, f)$, and
- $\Pi : (A, B) \mapsto A \times B, (f, g) \mapsto f \times g$.

Definition 4 (Natural transformation)**.** Let $F, G : \mathbf{C} \to \mathbf{D}$ be functors. A *natural transformation* $\eta : F \dot{\to} G$ is a family of $\mathbf{D}$-morphisms $\{\eta_A : F(A) \to G(A) | A \in O(\mathbf{C})\}$ such that $G(f) \circ \eta_A = \eta_B \circ F(f)$ for every morphism $f : A \to B$ in $\mathbf{C}$, as indicated by the following commutative diagram:

$$\begin{array}{ccc} F(A) & \xrightarrow{\eta_A} & G(A) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F(B) & \xrightarrow[\eta_B]{} & G(B) \end{array} \quad (1)$$

Remark 3. $\acute{\pi} : \Pi \dot{\to} \acute{\Pi}$ and $\grave{\pi} : \Pi \dot{\to} \grave{\Pi}$, where functors $\acute{\Pi} : (A, B) \mapsto A$ and $\grave{\Pi} : (A, B) \mapsto B$, are natural transformations.

Definition 5 (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 (B, ψ) consisting of an object B in $\mathbf{C}$ and a morphism

$\psi : F(B) \to Y$ in **D** such that for every object X in **C** and every morphism $g : F(X) \to Y$ in **D** there exists a unique morphism $u : X \to B$ in **C** such that $g = \psi \circ F(u)$.

Example 4 (Product as 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 $\pi = (\acute{\pi}, \grave{\pi})$.

Definition 6 (Limit)**.** A *limit* of a functor $D : \mathbf{C} \to \mathbf{C}^J$ is a universal morphism from D to an object in $\mathbf{C}^J$—the category of functors (from J to **C**) and natural transformations.

Example 5 (Product)**.** A product of sets A and B is the limit of the functor $\Delta : \mathbf{Set} \to \mathbf{Set}^2$ to (A, B). (NB. $\mathbf{C}^2 \cong \mathbf{C} \times \mathbf{C}$.)

Definition 7 (Wedge)**.** A *wedge* to a functor $F : \mathbf{C^{op}} \times \mathbf{C} \to \mathbf{D}$ is a *di*natural transformation $\omega : D \overset{..}{\to} F$ consisting of a family of **D**-morphisms $\{\omega_A : D \to F(A,A) | A \in \mathcal{O}(\mathbf{C})\}$ such that for each $f : A \to B$ in **C** the following diagram commutes:

$$\begin{array}{ccc} D & \xrightarrow{\omega_A} & F(A,A) \\ {\scriptstyle \omega_B}\downarrow & & \downarrow{\scriptstyle F(1_A,f)} \\ F(B,B) & \xrightarrow[F(f,1_B)]{} & F(A,B) \end{array} \qquad (2)$$

Definition 8 (End)**.** The *end* of a functor $F : \mathbf{C^{op}} \times \mathbf{C} \to \mathbf{D}$ is a pair (E, ω) consisting of an object E in **D** and a wedge $\omega : \mathcal{D} \overset{..}{\to} F$ such that for every wedge $\beta : Z \overset{..}{\to} F$ 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 4. An end is a universal wedge; equivalently, a limit in two variables.

Example 6 (Hom-set)**.** Hom-sets of natural transformations, $\mathrm{Nat}(F, G)$, are constructed from the ends of hom-functors.

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 5. For example 6(b), substitution yields

$$\begin{array}{ccc} E & \xrightarrow{\omega_A} & \mathrm{Hom}(FA, GA) \\ {\scriptstyle \omega_B}\downarrow & & \downarrow{\scriptstyle \mathrm{Hom}(1_{FA}, Gf)} \\ \mathrm{Hom}(FB, GB) & \xrightarrow[\mathrm{Hom}(Ff, 1_{GB})]{} & \mathrm{Hom}(FA, GB) \end{array} \qquad (3)$$

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).

Definition 9 (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:

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

Remark 6. A monoid is a one-object category—morphisms are the elements and composition is the monoid operation.

Example 7 (Integers)**.** Examples of monoids include:

- $(\mathbb{Z}, +, 0)$: the integers together with addition,
- $\mathbb{Z}_2$: $\{0, 1\}$ together with addition modulo-2, and
- $\mathbb{Z}_2^2 \cong \mathbb{Z}_2 \times \mathbb{Z}_2$: the product of monoid $\mathbb{Z}_2$ with itself.

Remark 7. $\mathbb{Z}_2$ is isomorphic to exclusive-or, as indicated by the following tables for the respective operations:

+	0	1
0	0	1
1	1	0

⊕	F	T
F	F	T
T	T	F

Remark 8. $\mathbb{Z}_2 \times \mathbb{Z}_2$ (with projections) is a categorical product in the category of monoids and monoid homomorphisms.

Definition 10 (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*.

Remark 9. A monoid action is a functor, $M \to \mathbf{Set}$, which identifies each action in M with an endomorphism $X \to X$.

Example 8 (Actions)**.** The actions of $\mathbb{Z}_2$ ($\mathbb{Z}_2^2$) on a set of shapes (trigrams) is shown in the following left (right) table:

$\mathbb{Z}_2$	△	□
0	△	□
1	□	△

$\mathbb{Z}_2^2$	BEH	FUT	PEJ	ROY
h	FUT	BEH	ROY	PEJ
v	ROY	PEJ	FUT	BEH

where h (horizontal) and v (vertical) in $\mathbb{Z}_2^2$ correspond to the elements $(0, 1)$ and $(1, 0)$ in $\mathbb{Z}_2 \times \mathbb{Z}_2$, respectively. Actions e (no change) and d (diagonal) in $\mathbb{Z}_2^2$, corresponding to elements $(0, 0)$ and $(1, 1)$ in $\mathbb{Z}_2 \times \mathbb{Z}_2$, are not shown.

Definition 11 (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 $x \in X$.

Remark 10. An M-set S is *represented* by the functor (S, σ) identifying set S and the homomorphism $\sigma : M \to (X \to X)$. M-set representations (functors) and equivariant maps (natural transformations) form a (functor) category, denoted **MSet**.

Remark 11. The forgetful functor $U : \mathbf{MSet} \to \mathbf{Set}$ forgets the actions, i.e. $U : (S, \sigma) \mapsto S$.

Theorem 1 (Reconstruction)**.** Let **MSet** be a category of M-set representations for a monoid M, and $U : \mathbf{MSet} \to \mathbf{Set}$ the forgetful functor. We have $\int_{\mathbf{MSet}} \mathrm{Hom}(U-, U-) \cong M$.

Remark 12. Substituting $\mathrm{Hom}(U-, U-)$ for functor F in definition 8 yields commutative diagram

$$\begin{array}{ccc} E & \xrightarrow{\omega_a} & \mathrm{Hom}(S, S) \\ {\scriptstyle \omega_b}\downarrow & & \downarrow{\scriptstyle \mathrm{Hom}(1_S, f)} \\ \mathrm{Hom}(R, R) & \xrightarrow[\mathrm{Hom}(f, 1_R)]{} & \mathrm{Hom}(S, R) \end{array} \qquad (4)$$

where E identifies with (set) M, and ω_a/ω_b identifies action a/b with the endomorphism. See Tannaka duality.