Title
Transitive inference in non-humans? Not so fast!

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Author
Phillips, Steven

Publication Date
2022

Peer reviewed

Transitive inference in non-humans? Not so fast!

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

Abstract

A capacity for transitive inference (i.e. if aRb and bRc then aRc) was thought to be uniquely human. However, evidence of transitive inference in other species suggests that this capacity is ubiquitous throughout the animal kingdom. This apparent ubiquity raises two basic questions for cognitive science. (1) Why is transitive inference so prevalent? (2) What is special about transitive inference in (adult) humans? Formal (category theory) methods are used to address these questions. To the first question, different (implicit and explicit) forms of transitive inference follow from a common (universal) operation over the premises, aRb and bRc, i.e. a category theory version of transitive closure, hence the ubiquity of this capacity. To the second question, this construction involves rapid (one-shot) premise integration in older humans, but not other cohorts. This formal comparison points to rapid encoding and integration of relational data as underlying the evolution and development of higher cognitive capacities.

Keywords: transitive inference; transitive closure; implicit; explicit; category theory; category; functor; profunctor

Introduction

According to Piaget, Transitive Inference (TI)—an inference of the form, if aRb and bRc, then aRc—is attainable only by adolescent, or adult humans. Although early work suggested a capacity for TI in children as young as four years of age (Bryant & Trabasso, 1971), subsequent studies controlling for "short-cut" strategies[1] indicated that TI appears later, around the median age of five (Andrews & Halford, 1998; Pears & Bryant, 1990), see also Halford, Wilson, Andrews, and Phillips (2014) for a review. So, a study of the processes underlying TI should afford a window into some uniquely human aspects of intelligence and its development.

Yet, in recent years, more than a few studies have provided evidence of TI in other species, including monkeys (Gazes, Lazareva, Bergene, & Hampton, 2014; Jensen, Altschul, Danly, & Terrace, 2013), rats (Dusek & Eichenbaum, 1997), ravens (Massen, Pašukonis, Schmidt, & Bugnyar, 2014), corvids (Bond, Kamil, & Balda, 2003; Bond, Wei, & Kamil, 2010), crows (Lazareva et al., 2004), pigeons (Daniels, Laude, & Zentall, 2014), and fish (Grosenick, Clement, & Fernald, 2007).[2] The growing evidence suggests that a capacity for TI is more widespread than previously thought.

Such results are potentially exciting for understanding the evolution and development of (higher) cognition. However, the methods used to assess TI in other species raise questions over its comparability to a form of TI observed in humans, given that TI is also known to be difficult even for adults (Maybery, Bain, & Halford, 1986). TI in non-humans is assessed via repeated reward-based reinforcement learning on adjacent pairs of stimuli, whereby response to a non-adjacent (test) pair that is consistent with the order implicit in the reward schedule is taken as evidence for TI.[3] By contrast, a test for TI in humans proceeds via trial-by-trial (unrepeated) presentation of all adjacent (premise) and non-adjacent (test) pairs, whereby a response to the test pair that is consistent with the order (explicitly) given by a spatial, or verbal relation that is transitive is taken as evidence for TI.[4] This procedural difference has led some to conclude that TIs in non-humans and humans are really of two different kinds, referred to as *implicit*[5] versus *explicit* TI, respectively (Goel, 2007; Halford et al., 2014; Wright, 2012), wherein the explicit form is found only in older humans and the implicit form is found in both humans and non-humans alike (Halford et al., 2014).

A distinction between non-human (implicit) and human (explicit) forms of TI, however, does not address a primary concern for cognitive science, which is an explanation for how such seemingly advanced cognitive capacity arises in the first place. On one hand, if the two forms share essentially nothing in common, then there is little reason to attribute implicit TI as a foundation for reasoning. On the other hand, if they are related, then such claims beg the question of how this is so. The literature sheds little light on this relationship, because an equivalent test of explicit TI in non-humans has not been developed (Halford et al., 2014), though adult humans evidently resort to an implicit form of TI when they are unaware of the relational structure (Frank, Rudy, Levy, & O'Reilly, 2005). Moreover, there has been little in the way of theory explicating a link between the two forms.

The purpose here is to address this theoretical shortcoming

[1] E.g., in a four-term series consisting of the premises $a \leq b$, $b \leq c$, $c \leq d$, one can infer $b \leq d$ simply by observing that d only appeared once in the list of premises. This strategy is circumvented in a five-term series by adding the premise $d \leq e$ and testing on $b \leq d$, thus avoiding the *end terms*, a and e.

[2] TI has also been reported in infants (Mou, Province, & Luo, 2014), but for a 3-term series, so did not control for short-cuts.

[3] That is, e.g., adjacent pairs AB^+, BC^+, CD^+, and DE^+, where "+" indicates the reinforced choice, and non-adjacent (test) pair BD.

[4] That is, e.g., adjacent pairs A < B, B < C, C < D and D < E, where "<" indicates spatial relation A *below* B, and test pair B?D.

[5] Implicit TI has also been called *transitivity of choice* (Halford et al., 2014; Libben & Titone, 2008).

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

by presenting a mathematical theory that is used to assess the relationship between implicit and explicit TI, as the basis for further empirical studies. Background theory provided in the next section forms the framework used to compare/contrast implicit and explicit TI in the section that follows. The main result is that both forms obtain from a categorical form of *transitive closure*, hence the ubiquity of TI across cohorts. However, they differ in the way that relational information is constructed. This difference reflects slow (implicit TI) versus fast (explicit TI) integration of premise relations as the basis for inference. Empirical implications of this result are discussed in the final section. In particular, placing implicit and explicit TI on common theoretical ground affords testable predictions for the emergence of relational reasoning (explicit TI) from basic associative processes (implicit TI).

Methods: paradigms and theory

TI is an inference from relationships between pairs of stimuli to a relationship between a novel stimulus pair. An empirical test of TI has two parts: (1) ensuring participants understand that a is R-related to b and b is R-related to c, and (2) testing whether they understand that a is R-related to c without further intervention. Differences between tests of TI in non-human cohorts versus older humans set the challenge for a comprehensive theory relating both forms. So, in this section, we recall the essential aspects of both paradigms as motivation for the theoretical framework that follows.

Experimental paradigms: implicit and explicit TI

Implicit TI. In the absence of language, the stimuli can be ordered by associated reward, e.g. $a \leq b$ means less reward associated with a than b. Given five distinguishable stimuli, A, B, C, D and E, stimulus pairs ordered by relative reward: e.g., AB+ indicates that when presented with spatially ordered pair (A, B), or (B, A), a reward is given only for a response corresponding to B. Likewise, BC+, CD+ and DE+ indicates that C is preferred over B, D over C, and E over D. Consistently selecting the rewarded item is taken as evidence that the participant has learned the premise relations. The participant is then presented with the novel pair (B, D) whence a response that corresponds to D is taken as evidence of TI.

Explicit TI. Explicit TI also involves five terms, but their pairwise relations are indicated by spatial order: e.g. as two-block towers, where say a green block sitting atop a red block indicates the order relation red < green. There are four two-block towers, e.g., blue atop green, yellow atop blue and brown atop yellow. Participants are presented with a pair of blocks and asked for the higher block when building a tower. E.g., given pair (yellow, green), a yellow response is taken as evidence for TI. The procedure is repeated with novel premise towers and test pairs for each trial. For comparison and variation, we consider an analogous test where the participant is presented with a novel two-block tower (e.g., yellow atop green, or green atop yellow) and ask whether block order is consistent with the premise towers, which is also a two-choice response (cf. Maybery et al., 1986).

Comparison/Contrast. Both forms involve one-shot responses to novel stimulus pairs as evidence of TI. Furthermore, the stimuli have no intrinsic ordering; rather their order is encoded in terms of rewarded response (implicit), or relative spatial location (explicit). However, they differ in three ways. First, relative spatial location of the stimuli is irrelevant to implicit TI, e.g., pairs (A, B) and (B, A) map to the same rewarded response, B, but significant to explicit TI as (A, B) and (B, A) are two different orders. Second, they differ in trials needed to encode premise relations: multiple reward reinforcement (implicit) versus one-shot presentation (explicit). And third, no feedback is given on the test trial for the implicit form, but feedback on the correct response is given during practice trials for the explicit form. These commonalities and differences are taken up next in our theoretical approach to TI.

Theoretical framework: category theory

Our approach is motivated by a desire to compare implicit and explicit TI as a (possible) window on the evolution and development of higher cognition. Comparison necessitates a common point of view to make sense of their relationship, assuming that they are meaningfully comparable. Analysis of relations between formal systems is the *raison d'être* of a branch of mathematics called *category theory* (Lawvere & Schanuel, 2009; Mac Lane, 1998). An important approach in this regard, called "categorification" (Baez & Dolan, 1998), is to recast set-theoretic constructions in terms of categorical (analogical) abstractions for the purpose of revealing connections that are otherwise obscured. We basically categorify implicit and explicit TI to reveal their connection.

Relations and joins. TI is afforded by relations (definition 1) that are transitive (definition 2), such as the order relation on numbers (example 3). Some relations are not transitive (remark 4), but every relation can be extended to a transitive relation by *transitive closure* (definition 5). Transitive closure can be computed via *relational joins* (remark 6), or for graphs (example 7) via matrix multiplication of *incidence matrices* (remark 8). We require constructions for relations that have more structure than just sets of pairs. For this reason, we turn to category theory.

Basic correspondences. Entities, called *objects*, and relations between objects, called *morphisms*, satisfying certain rules constitute a *category* (definition 9), such as the category of numbers and their order (example 10)[6], and the categories of sets and their functions, or relations (examples 12). Graphs also constitute the objects of a category (example 13) and for certain cases are categories in their own right (remark 14). Thus, we have basic correspondences between familiar set-theoretic constructions and categorical analogs relevant to TI. Categorical constructions reside in a category of some kind, so categories frame our theoretical approach.

Profunctors. Morphisms between categories considered as objects in a larger category are called *functors* (definition 15),

[6] All arrows are directed, so when reversed they usually constitute a new category (remark 11), though e.g., $\mathbf{Rel}^{\mathrm{op}} = \mathbf{Rel}$.

such as *hom-functors* (example 16). As functions generalize to relations between sets, functors generalize to *profunctors* (definition 17)—relations between categories (remark 18). Functors induce profunctors like functions as special kinds of relations (example 19). The join of relations between sets, or matrix multiplication as the basis for transitive closure is abstracted to *composition of profunctors* (definition 20) between categories as transitive closure categorified (remark 21).

Enrichment. For our purposes, we require the concept of *monoidal category* (definition 22)—a categorified monoid (remark 23)—to model situations where relations are not just sets (example 24), but have additional structure (example 25) in the form of *enriched profunctors*[7], which can be seen as generalized matrix algebra (example 27).

Result: theoretical comparison/contrast

We proceed to compare and contrast implicit and explicit TI. The basis for comparison is a categorical version of transitive closure: implicit and explicit TI follow from composition of enriched profunctors modeling the premises, thus deriving the relations between non-adjacent stimuli. We consider a five-term series consisting of the set of terms $T = \{A, B, C, D, E\}$, corresponding to the five stimuli of a TI task. The theory does not depend on this number of terms. No assumption is made about the stimuli beyond being distinguishable from each other. There are three aspects to a TI task: (1) encoding—representing the premise relationships, e.g., aRb and bRc, (2) completion—adding the implied relations between other pairs, e.g., aRc and bRd, and (3) inference—correct response to the test pair, i.e. bRd, which are detailed for each case.

Implicit transitive inference

We consider the following situation for implicit TI. On each learning trial, participants are presented with a pair of stimuli, e.g., (A, B) and two response choices, i.e. a and b corresponding to the stimuli A and B, respectively. A reward is given when participants choose the response corresponding to the preferred stimulus, i.e. b, or no reward otherwise. There are five stimulus-specific responses, a, b, c, d and e. Responses are only defined for adjacent stimuli, obtained by reward reinforcement, responses to non-adjacent stimuli are undefined, represented by the element $*$. The set of possible responses is $R_I = \{*, a, b, c, d, e\}$. The tensor product is given by the following table:

$\times_I$	$*$	a	b	c	d	e
$*$	$*$	$*$	$*$	$*$	$*$	$*$
a	$*$	a	b	c	d	e
b	$*$	b	b	c	d	e
c	$*$	c	c	c	d	e
d	$*$	d	d	d	d	e
e	$*$	e	e	e	e	e

constituting the monoidal category $\mathcal{R}_I = (R_I, \times_I, 1)$.

[7]That is a profunctor between *enriched categories* (remark 26).

Encoding. Premises are encoded as the response-enriched profunctor $\mathcal{P} : T^{\mathrm{op}} \times T \to \mathcal{R}_I$, e.g., $\mathcal{P}(A, B) = b$ and $\mathcal{P}(A, C) = \mathcal{P}(C, A) = *$.

Completion. Inference follows from profunctor composition: $\mathcal{P} \circ \mathcal{P}$, or $\mathcal{P}^2$, which can be viewed as (generalized) matrix multiplication, as shown for three-term subset $\{B, C, D\}$:

$$\begin{array}{|c|c|c|}\hline b & * & * \\ \hline c & c & * \\ \hline * & d & d \\ \hline\end{array} \times \begin{array}{|c|c|c|}\hline b & * & * \\ \hline c & c & * \\ \hline * & d & d \\ \hline\end{array} = \begin{array}{|c|c|c|}\hline b & * & * \\ \hline c & c & * \\ \hline d & d & d \\ \hline\end{array}$$

Inference. The response indicating TI is obtained by applying novel (non-adjacent) stimulus pair (B, D), or (D, B) to the composite profunctor, i.e. $\mathcal{P}^2(B, D) = d = \mathcal{P}^2(D, B)$, which assumes *symmetry* (i.e. $A \otimes B \cong B \otimes A$).

Explicit transitive inference

There are two stimulus responses for explicit TI: inconsistent, 0, and consistent, 1. As for implicit TI, we adjoin element $*$ for undefined. The set of possible responses is $R_E = \{*, 0, 1\}$. The tensor product is given by the following table:

$\times_E$	$*$	0	1
$*$	$*$	$*$	$*$
0	$*$	0	0
1	$*$	0	1

constituting the monoidal category $\mathcal{R}_E = (R_E, \times_E, 1)$.

Encoding. Premises are encoded as the response-enriched profunctor $\mathcal{P} : T^{\mathrm{op}} \times T \to \mathcal{R}_E$, e.g., $\mathcal{P}(A, B) = 1$, $\mathcal{P}(B, A) = 0$, $\mathcal{P}(A, C) = *$, and $\mathcal{P}(C, A) = *$.

Completion. As for implicit TI, inference is obtained via profunctor composition, $\mathcal{P}^2$, exemplified as matrix multiplication for three-term subset $\{B, C, D\}$:

$$\begin{array}{|c|c|c|}\hline 1 & 0 & * \\ \hline 1 & 1 & 0 \\ \hline * & 1 & 1 \\ \hline\end{array} \times \begin{array}{|c|c|c|}\hline 1 & 0 & * \\ \hline 1 & 1 & 0 \\ \hline * & 1 & 1 \\ \hline\end{array} = \begin{array}{|c|c|c|}\hline 1 & 0 & 0 \\ \hline 1 & 1 & 0 \\ \hline 1 & 1 & 1 \\ \hline\end{array}$$

Inference. Responses indicating TI are obtained by applying novel pairs (B, D) and (D, B) to composite profunctor, $\mathcal{P}^2$, i.e. $\mathcal{P}^2(B, D) = 1$ and $\mathcal{P}^2(D, B) = 0$.

Long-distance inference

This approach extends to *long-distance inference*: when there is more than one intermediate term. For example, suppose a six-term series, $T + \{F\}$, in which we have bRc and cRd and dRe implies bRe; equivalently, bRc and cRd implies bRd and bRd and dRe implies bRe. This situation is also obtained by profunctor composition: $\mathcal{P} \circ \mathcal{P} \circ \mathcal{P} = (\mathcal{P} \circ \mathcal{P}) \circ \mathcal{P} = \mathcal{P}^3$. For implicit TI, we have $\mathcal{P}^3(B, E) = \mathcal{P}^3(E, B) = e$. For explicit TI, we have $\mathcal{P}^3(B, E) = 1$ and $\mathcal{P}^3(E, B) = 0$.

Comparison and contrast

In comparison, both forms of TI involve transitive closure: the relationships between non-adjacent pairs of stimuli are obtained as (generalized) multiplication of the relationships between adjacent stimuli given by the premise pairs. By contrast, stimuli are ordered by selection reward for implicit TI, but relative spatial position for explicit TI. The implications of these similarities and differences are discussed next.

Discussion

What does this comparison of implicit and explicit TI reveal? The categorical view seems to support the extensive work in comparative psychology examining TI in other species as a window into the evolution of reasoning, as both forms are instances of a *universal construction*[8], i.e. a (categorified) form of transitive closure. In this way, transitive closure is common ground—the seedbed of higher cognition. However, this unifying view belies a circumspect outlook for comparative psychology in understanding the emergence of reasoning. Reasons for caution are discussed in the rest of this section.

A comparison also highlights the remaining differences. Generalized (categorified) transitive closure works for any monoidal category, like ordinary transitive closure works for any relation. So, for example, a simple neural network that realizes multiplication of boolean matrices, as a monoidal category, suffices to realize a form of TI. Thus, the theory suggests that some form of TI can be coerced from even very simple organisms, assuming a nervous system. However, the trend of attempting to find evidence of TI in ever simpler species obscures another difference, i.e. the encoding of premises via multi-shot reinforcement learning (implicit) versus one-shot presentation (explicit). We assumed the encoding of premises, without which neither paradigm can proceed. For explicit TI, participants encode premises and make inferences on a trial-by-trial basis, whereas implicit TI only requires an inference from a single set of premises. Presumably, (some) non-human species can be retrained on a different set of premises to make another inference. The key question then is about how rapidly new premises can be encoded, i.e. a capacity for transfer that has been investigated via the *learning set* paradigms developed for non-humans (Kendler, 1995). An analogous paradigm is needed for non-humans to assess such capacity as crucial to the development of higher cognition.

The *relational schema induction* paradigm was developed for such purposes (Halford, Bain, Maybery, & Andrews, 1998; Halford & Busby, 2007). Each task instance conforms to a common relational structure (schema). Induction is tested by first time performance on certain test trials. The difficulty of each task is manipulated by the number of related variables, or dimensions of task variation, i.e. *relational complexity* as a factor differentiating species and age-groups (Halford, Wilson, & Phillips, 1998; Halford et al., 2014).[9] This paradigm could be adapted as a comparable form of explicit TI for testing non-human and prelinguistic cohorts.

One possibility is to incorporate the *simultaneous chaining paradigm* (Terrace, 2005), used to test serial learning, as a test of explicit TI in non-humans. For example, premise pairs of say pictures are presented simultaneously on a screen so that relative order is explicitly indicated by a visual cue, not their relative spatial location: e.g., a red dot in proximity to the b item for each (a, b) premise pair indicating $a \leq b$. The non-adjacent test pair (b, d) is presented without the visual cue. Participants are rewarded when the stimulus associated with the red dot is touched for each pair in order (i.e. b, c, d, e). Participants are also rewarded when they touch the d element, but not the b element in the (b, d) test pair. The prediction that follows a capacity for explicit TI is straightforward: if participants correctly touch the locations of the red dots in the order corresponding to the sequence b, c, d, e and they correctly touch the location of the d item in the test pair (b, d) for one five-term series, then they will correctly touch the corresponding locations and item on first-time presentation of a novel five-term series, i.e. the locations of red dots in the sequence b', c', d', e' and the location of the d' item in the test pair (b', d'). Consistently responding this way on the first presentation of each novel 5-term series is evidence supporting a capacity for explicit TI.

This approach to testing explicit TI in nonverbal cohorts can also be used to investigate development as a transition from implicit to explicit TI. A category theory approach to relational schema induction was employed to show how the underlying schema is induced (reconstructed) from a series of task instances (Phillips, 2021). A similar account applies here, where TI is regarded as a relational schema over sets of stimuli, rather than a property of a relation on a particular set.

The moral of this work is a familiar one (see Penn, Holyoak, & Povinelli, 2008): explicit TI does not necessarily follow directly from implicit TI despite a common root (transitive closure).[10] Explicit TI is an inference over instances, i.e. a general rule versus a specific case, which can be seen as a difference in *representational rank* (Halford et al., 2014).

If implicit and explicit TI are so different, why do the two forms cohabit in older humans? An answer to this question is also familiar in the form of *dual-process* theories (Evans, 2003; Wright, 2012). Slow, yet robust encoding of relations affords survival advantages to individuals and their cohorts in, for example, a stable social hierarchy (Grosenick et al., 2007). Fast, yet fragile encoding of relations affords complementary advantages in a dynamically changing world. A challenge then is to balance these trade-offs and explain their emergence.

This analysis shows that implicit TI should be regarded as a test of transitive closure, not transitive inference *per se*. The premise relations need not be transitive. Indeed, preference is not transitive, in general, as one can prefer A over B and B over C, but not A over C. However, the inference is not necessarily fallacious, as participants may be seen as reinterpreting the relation R containing the premises as the extended relation R^+, which is always transitive.

A similar situation may arise for explicit TI if blocks of the same colour in different two-block towers are interpreted as different blocks: e.g., *red above blue* and *blue' above green*. In this situation, the transitive inference *red above green* does not follow, since either *blue* and *blue'* are not comparable (being in different towers), or *blue'* is regarded as being above

[8]See Mac Lane (1998) for this technical sense of universal.

[9]A category theory view is the arity of the underlying categorical *(co)product* (Phillips, Wilson, & Halford, 2009).

[10]For a formal analogy, all monoids trace back to the one-element monoid, i.e. the *initial object* in the category of monoids, but two monoids need not be related by a monoid homomorphism.

blue relative to the common ground on which the towers stand. However, instructions and feedback during practice trials provide the information needed to interpret *blue* and *blue'* as the same block, which affords the transitive inference.

Acknowledgments

This work was supported by a Japan Society for the Promotion of Science grant, Grant-in-Aid for Transformative Research (20H05710).

Appendix A: Basic theory

Definition 1 (relation)**.** Let A be a set. A *(binary) relation* on A is a subset R of the *Cartesian product* of A with itself, i.e. $R \subseteq A \times A = \{(a,b)|a,b \in A\}$. We also write aRb for pair $(a,b) \in R$, i.e. the relationship a is R-related to b.

Definition 2 (reflexive, transitive)**.** A relation R on A is called *reflexive* if aRa for every element a in A; *transitive* if aRb and bRc implies aRc for every triple of elements a,b,c in A.

Example 3 (preorder)**.** A *preorder* is a relation that is reflexive and transitive. The usual order on numbers is a preorder.

Remark 4. The *precede* relation is not transitive in the sense of 1 precedes 2 and 2 precedes 3, but 1 does not (immediately) precede 3. Compare with the *is-parent-of* relation.

Definition 5 (transitive closure)**.** Let A be a set and $R \subseteq A \times A$ a binary relation on A. The *transitive closure* of R, denoted R^+, is the smallest transitive relation on A containing R.

Remark 6. The transitive closure of a relation $R \subseteq A \times A$ can be computed by iteratively adding the *join* of R with itself, i.e. $\bowtie^2 R = R + \{(a,c)|\exists b \in A, (a,b) \wedge (b,c) \in R\}$, to its *fixed point* where no change obtains by further iterations, i.e. the relation $\bowtie^\infty R = R^+$, which terminates at k when $\bowtie^k R = R^+$.

Example 7 (graph)**.** A *directed graph* $G = (V,E,s,t)$ consists of a set of *vertices* V, a set of *edges* E, and two functions $s,t : E \to V$ returning the *source* and *target* vertex of each edge, respectively. If a graph has at most one edge from a given source to a given target, then the edges correspond to the relation $G_E = \{(s(e),t(e))|e \in E\}$. The transitive closure of G_E adjoins an edge for each path in G, i.e. the *reachability graph* with edges given by the fixed point relation $\bowtie^\infty G_E = G_E^+$, which says that for vertices v,w in G if $(v,w) \in G_E^+$, then w is reachable from v.

Remark 8. Transitive closure of a graph G is computed via matrix multiplication of the corresponding *incidence matrix* M whose cell $M_{ji} = 1$ if (i,j) is an edge in G, otherwise $M_{ji} = 0$, i.e. the fixed point matrix M^∞.

Definition 9 (category)**.** A *category* $\mathbf{C}$ consists of a collection of *objects*, $\mathcal{O}(\mathbf{C}) = \{A,B,\dots\}$, a collection of *morphisms*, $\mathcal{M}(\mathbf{C}) = \{f,g,\dots\}$—a morphism written in full as $f : A \to B$ indicates object A as the *domain* and object B as the *codomain* of f—including for each object $A \in \mathcal{O}(\mathbf{C})$ the *identity morphism* $1_A : A \to A$, and a *composition* operation, $\circ$, that sends each pair of *compatible* morphisms $f : A \to B$ and $g : B \to C$ (i.e. the codomain of f is the domain of g) to the *composite* morphism $g \circ f : A \to C$, that together satisfy the laws of:

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

$\mathbf{C}[A,B]$ denotes the collection of morphisms from A to B.

Example 10 (proset)**.** A *preordered set* (or, *proset*) is a set P together with a preorder $\leq$ on P, written $(P,\leq)$. A proset is a category with objects $p \in P$ and morphisms $p \to q$ whenever $p \leq q$. The set of natural numbers together with the usual order, $(\mathbb{N},\leq)$, constitute a proset, hence a category.

Remark 11. $\mathbf{C}^{\text{op}}$ denotes the *opposite category*: the objects and "reversed" morphisms of $\mathbf{C}$, e.g., $(P,\leq)^{\text{op}} = (P,\geq)$.

Examples 12 (functions, relations)**.** Categories **Set** and **Rel** have sets for objects and (respectively) functions and relations for morphisms with function composition and relational join as the respective composition operations.

Example 13 (graphs)**.** The category **Grph** has (directed) graphs for objects and graph homomorphisms for morphisms.

Remark 14. Directed graphs generally lack an edge (loop) on every vertex and an edge for every path to be a category, but every directed graph extends to a category by transitive (and *reflexive*[11]) closure of edges as relations.

Definition 15 (functor)**.** Let $\mathbf{C}$ and $\mathbf{D}$ be categories. A *functor* is a map $F : \mathbf{C} \to \mathbf{D}$ preserving:

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

A functor is a category (cf. graph) homomorphism.

Examples 16 (hom-functors)**.** A *hom-functor* preserves the composition of morphisms in $\mathbf{C}$ as functions in **Set**.

a. $\mathbf{C}[A,-] : \mathbf{C} \to \mathbf{Set}; g \mapsto (\mathbf{C}[A,g] : h \mapsto g \circ h)$.

b. $\mathbf{C}[-,A] : \mathbf{C}^{\text{op}} \to \mathbf{Set}; f \mapsto (\mathbf{C}[f,A] : h \mapsto h \circ f)$.

c. $\mathbf{C}[-,-] : (f,g) \mapsto (\mathbf{C}[f,g] : h \mapsto g \circ h \circ f)$.

Definition 17 (profunctor)**.** Let $\mathbf{C}$ and $\mathbf{D}$ be categories. A *profunctor* from $\mathbf{C}$ to $\mathbf{D}$, written $\mathcal{P} : \mathbf{C} \nrightarrow \mathbf{D}$, is a (bi)functor $\mathcal{P} : \mathbf{D}^{\text{op}} \times \mathbf{C} \to \mathbf{Set}$ sending each pair of:

- objects (D,C) to the set $\mathcal{P}(D,C)$ and
- morphisms $(d : D \to D', c : C \to C')$ to the function $\mathcal{P}(d,c) : \mathcal{P}(D,C) \to \mathcal{P}(D',C')$.

Remark 18. Profunctors are to functors as relations are to functions. Compare the graph of a function $f : A \to B$, i.e. the relation $\Gamma(f) = \{(a,f(a))|a \in A\} \subseteq A \times B$, with example 19a.

Examples 19 (functors to profunctors)**.** A functor $F : \mathbf{C} \to \mathbf{D}$ determines two profunctors.

[11] Add relation (a,a) for each $a \in A$ (cf. transitive closure).

a. $\mathbf{D}[1,F] : \mathbf{C} \nrightarrow \mathbf{D}; (d,c) \mapsto \mathbf{D}[d, F(c)]$.

b. $\mathbf{D}[F,1] : \mathbf{D} \nrightarrow \mathbf{C}; (c,d) \mapsto \mathbf{D}[F(c), d]$.

Hom-functors $\mathbf{C}[A,-]$ and $\mathbf{C}[-,A]$ (examples 16a, b) obtain from (endo)profunctors $\mathbf{C}[-,1]$ and $\mathbf{C}[1,-]$, where 1 is the identity functor on $\mathbf{C}$, by setting the other argument $(-)$ to A.

Definition 20 (profunctor composition)**.** Let $\mathcal{P} : \mathbf{C} \nrightarrow \mathbf{D}$ and $Q : \mathbf{D} \nrightarrow \mathbf{E}$ be profunctors. The *profunctor composition* of $\mathcal{P}$ and Q is the profunctor $Q \circ \mathcal{P} : \mathbf{C} \nrightarrow \mathbf{E}$ defined as

$$Q\mathcal{P}(E,C) = \left(\coprod_{D \in \mathbf{D}} Q(E,D) \times \mathcal{P}(D,C) \right) \Big/ \sim \qquad (1)$$

where $(e', c') \sim (e, c)$ whenever there exists an arrow $d \in \mathbf{D}$ such that $e' = d \circ e \in Q(E, D')$ and $c' \circ d = c \in \mathcal{P}(D, C)$.

Remark 21. Profunctor composition generalizes composition of relations. Compare profunctors $\mathcal{P} : \mathbf{C} \nrightarrow \mathbf{D}$ and $Q : \mathbf{D} \nrightarrow \mathbf{E}$ and their composition, $Q \circ \mathcal{P} : \mathbf{C} \nrightarrow \mathbf{E}$ as (relation) diagrams

$$\begin{tikzcd}
E_1 & D_1 \arrow[l, "q_{11}"'] & C_1 \arrow[l, "p_{11}"'] \arrow[dl, "p_{21}"] \\
E_2 & D_2 \arrow[ul, "q_{12}"] \arrow[l, "q_{22}"'] & C_2 \\
E_3 & D_3 \arrow[ul, "q_{23}"] \arrow[u, "d_{23}"'] & C_3 \arrow[l, "p_{33}"] \arrow[ul, "p_{23}"']
\end{tikzcd}
\qquad
\begin{tikzcd}
E_1 & C_1 \arrow[l, shift right, "(q_{11},p_{11})"'] \arrow[l, shift left, "(q_{12},p_{21})"] \arrow[dl, "(q_{22},p_{21})"] \\
E_2 & C_2 \\
E_3 & C_3 \arrow[uul, "(q_{12},p_{23})"] \arrow[ul, "(q_{22},p_{23})"]
\end{tikzcd}$$

with join of corresponding relations, $Q \bowtie P$ (see remark 6).[12]

Definition 22 (monoidal category)**.** A *monoidal category* $(\mathbf{M}, \otimes, I)$ consists of a category $\mathbf{M}$, a functor $\otimes : \mathbf{M} \times \mathbf{M} \to \mathbf{M}$, called the *tensor product*, and an object I in $\mathbf{M}$ such that $\otimes$ is unital and associative up to isomorphism.[13]

Remark 23. A monoidal category is a "categorified" monoid. For comparison, a *monoid* $(M, \cdot, e)$ consists of a set M, an operation $\cdot$ and a designated element $e \in M$, called the *unit*, such that $\cdot$ is *unital*, i.e. $m \cdot e = m = e \cdot m$ for all elements $a \in M$, and *associative*, i.e. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all triples of elements $a, b, c \in M$. A monoid is a one-object category whose arrows correspond to the elements of M and composition to the operation, e.g., the integers together with addition and 0 as the unit constitute a monoid, $(\mathbb{Z}, +, 0)$, hence a category.

Example 24 (sets)**.** $(\mathbf{Set}, \times, 1)$ is a monoidal category.

Example 25 (truth)**.** The set of truth values with *logical and* is monoidal, $(\{\bot, \top\}, \wedge, \top)$—equivalently, $(\{0,1\}, \times, 1)$.

Remark 26. Monoidal categories are the basis of *enriched category theory* used to model situations where the relations between objects have more structure than sets of morphisms, i.e. the hom-sets $\mathbf{C}[A,B]$ are replaced with the objects of a monoidal category, $\mathbf{M}$. For instance, ordinary categories are enriched in **Set** (example 24) and prosets are enriched in the monoidal category of truth values (example 25).

[12]Note the analogy to matrix multiplication (see example 27).

[13]The *unitors* and *associator* are omitted (see Mac Lane, 1998).

Example 27 (enriched profunctor)**.** An *enriched profunctor* is a (pro)functor of the form $\mathcal{P} : \mathbf{D}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{M}$. Composition takes on the monoidal structure, seen as a form of generalized matrix algebra with products and sums replaced accordingly. E.g., **Set** (example 24) for profunctor composition (remark 21) is replaced with (i.e. enriched in) boolean values (example 25) yielding multiplication of matrices, $M_{ED} \times M_{DC} = M_{EC}$:

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 0 \\ \hline 0 & 1 & 1 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \times \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 1 & 0 & 1 \\ \hline 0 & 0 & 1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 1 & 0 & 1 \\ \hline 1 & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline \end{array}$$

with sum as *coproduct* (Mac Lane, 1998), e.g., $1 + 1 = 1$.

References

Andrews, G., & Halford, G. S. (1998). Children's ability to make transitive inferences: The importance of premise integration and structural complexity. *Cognitive Development*, *13*(4), 479–513.

Baez, J. C., & Dolan, J. (1998). Categorification. In E. Getzler & M. Kapranov (Eds.), *Higher category theory* (Vol. 230, pp. 1–36). Providence, RI: American Mathematical Society.

Bond, A. B., Kamil, A. C., & Balda, R. P. (2003). Social complexity and transitive inference in corvids. *Animal behaviour*, *65*(3), 479–487.

Bond, A. B., Wei, C. A., & Kamil, A. C. (2010). Cognitive representation in transitive inference: a comparison of four corvid species. *Behavioural Processes*, *85*(3), 283–292.

Bryant, P. E., & Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature*, *232*, 456–458.

Daniels, C. W., Laude, J. R., & Zentall, T. R. (2014). Transitive inference by pigeons: Does the geometric presentation of the stimuli make a difference? *Animal cognition*, *17*(4), 973–981.

Dusek, J. A., & Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences*, *94*, 7109–7114.

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

Frank, M. J., Rudy, J. W., Levy, W. B., & O'Reilly, R. C. (2005). When logic fails: Implicit transitive inference in humans. *Memory & Cognition*, *33*(4), 742–750.

Gazes, R. P., Lazareva, O. F., Bergene, C. N., & Hampton, R. R. (2014). Effects of spatial training on transitive inference performance in humans and rhesus monkeys. *Journal of Experimental Psychology: Animal Learning and Cognition*, *40*(4), 477.

Goel, V. (2007). Anatomy of deductive reasoning. *Trends in cognitive sciences*, *11*(10), 435–441.

Grosenick, L., Clement, T. S., & Fernald, R. D. (2007). Fish can infer social rank by observation alone. *Nature*, *445*(7126), 429–432.

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 coherence in the foundations of psychology*. Cambridge, MA: MIT Press.

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

Jensen, G., Altschul, D., Danly, E., & Terrace, H. (2013). Transfer of a serial representation between two distinct tasks by rhesus macaques. *PloS one*, *8*(7), e70285.

Kendler, T. S. (1995). *Levels of cognitive development*. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Lazareva, O. F., Smirnova, A. A., Bagozkaja, M. S., Zorina, Z. A., Rayevsky, V. V., & Wasserman, E. A. (2004). Transitive responding in hooded crows requires linearly ordered stimuli. *Journal of the experimental analysis of behavior*, *82*(1), 1–19.

Libben, M., & Titone, D. (2008). The role of awareness and working memory in human transitive inference. *Behavioural processes*, *77*(1), 43–54.

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

Massen, J. J., Pašukonis, A., Schmidt, J., & Bugnyar, T. (2014). Ravens notice dominance reversals among conspecifics within and outside their social group. *Nature communications*, *5*(1), 1–7.

Maybery, M. T., Bain, J. D., & Halford, G. S. (1986). Information-processing demands of transitive inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4), 600–613.

Mou, Y., Province, J. M., & Luo, Y. (2014). Can infants make transitive inferences? *Cognitive psychology*, *68*, 98–112.

Pears, R., & Bryant, P. (1990). Transitive inferences by young children about spatial position. *British Journal of Psychology*, *81*(4), 497–510.

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. (2021). A reconstruction theory of relational schema induction. *PLoS Computational Biology*, *17*(1), e1008641.

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.

Terrace, H. S. (2005). The simultaneous chain: A new approach to serial learning. *Trends in cognitive sciences*, *9*(4), 202–210.

Wright, B. C. (2012). The case for a dual-process theory of transitive reasoning. *Developmental Review*, *32*(2), 89–124.