ASPECTS OF CATEGORY THEORY

Esfandiar Haghverdi

School of Informatics and Computing
Indiana University
150 S. Woodlawn Ave.,
Bloomington, IN 47405 USA
ehaghver@indiana.edu

ABSTRACT

I give a very brief introduction to category theory. The paper covers the basic definitions, some important basic constructions and spends some time discussing the important notion of categorical trace that has recently found many applications in computer science and the proof theory of Linear Logic among other things. The goal is to introduce categorical ideas and notions to Information Theory community with the hopes that it might find some applications.

1. INTRODUCTION

Category theory was invented by S. Eilenberg and S. Mac Lane in 1945 in their study of algebraic topology. The first invented notion was that of a natural transformation which led to the formulation of a category and that of a functor. Since then category theory has found many applications in diverse areas from proof theory and logic to the foundations of quantum mechanics, semantics of programming languages, to the seminal work of Grothendieck on redefining the foundations of algebraic geometry. The aim of this paper is to introduce the basic notions of category theory to Information Theory community. I have emphasized the important notion of the categorical trace invented by Joyal, Street, and Verity [1] as I believe that this notion in particular, can find many applications in the field of Information Theory.

2. BASIC DEFINITIONS

A *category* $\mathbb{C}$ is a collection *ob*($\mathbb{C}$) of objects and for each pair A, B of objects a set $\mathbb{C}(A, B)$ called the homset of morphisms from A to B. Such categories are called *locally small* because $\mathbb{C}(A, B)$ is a set. For a general category it can be a class. In this paper, we shall only consider locally small categories. For each triple of objects A, B, and C there is a composition operation

$$\circ_{A,B,C} : \mathbb{C}(A, B) \times \mathbb{C}(B, C) \to \mathbb{C}(A, C)$$

and for each object A there is an identity morphism $1_A : A \to A$ subject to the following conditions:

- For any morphisms $f : A \to B, g : B \to C$, and $h : C \to D$,

$$h \circ (g \circ f) = (h \circ g) \circ f$$

- For any morphism $f : A \to B$, $1_B \circ f = f = f \circ 1_A$.

For readability we shall write gf instead of $g \circ f$. Here are some examples of categories:

1. The category **Sets** of sets and mappings, where objects are sets and morphisms are mappings. Composition is the usual functional composition which is associative and has the identity function as its neutral element.

2. The category **Rel** of sets and relations, where objects are sets and a morphism $R : A \to B$ is a binary relation from A to B. Given relations $R : A \to B$ and $S : B \to C$, $SR : A \to C$ is defined as

   $(a, c) \in SR$ iff $\exists b \in B$, $(a, b) \in R$ and $(b, c) \in S$.

   It can be easily checked that this composition operation is associative with the identity relation as its neutral element.

3. The category **Pfn** of sets and partial functions, where objects are sets and a morphism $f : A \to B$ is a partial function from the set A to B.

4. The category **SRel**, where objects are measurable spaces and a morphism $f : (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ is map $f : X \times \mathcal{F}_Y \to [0, 1]$ such that for a fixed $x \in X$, $f(x, .) : \mathcal{F}_Y \to [0, 1]$ is a probability measure, and for a fixed $B \in \mathcal{F}_Y$, $f(., B) : X \to [0, 1]$ is a bounded measurable function. Given $f : X \to Y, g : Y \to Z$,

$$gf(x, C) = \int_Y g(y, C) f(x, dy).$$

   The identity morphism $1_X : X \to X$ is given by $1_X(x, A) = 1$ if $x \in A$, and 0 otherwise.

5. The category **FDVec**$_k$ has finite-dimensional vector spaces over a field k as objects and a morphism $f : V \to W$ is a linear transformation from the vector space V to W.

6. The category **Graphs** of graphs and graph homomorphisms where objects are graphs and a morphism $f : G \to H$ is a graph homomorphism, that is an edge preserving map from the vertex set of G to that of H.

The basic idea and the main message of category theory is the emphasis on relationships, over that of objects themselves. For example, from a category theory perspective, set theory is about functions, linear algebra is about the study of linear transformations and graph theory is about graph homomorphisms, etc. Thus we shall try to relate categories together, we use the notion of a functor for this purpose.

A *functor* F from a category $\mathbb{C}$ to a category $\mathbb{D}$, $F : \mathbb{C} \to \mathbb{D}$ consists of a map from $ob(\mathbb{C}) \to ob(\mathbb{D})$ and a map $\mathbb{C}(A, B) \to \mathbb{D}(FA, FB)$ for each pair A, B of objects such that for any $f : A \to B, g : B \to C$ in $\mathbb{C}$, $F(gf) = F(g)F(f)$, and $F(1_A) = 1_{FA}$. For example, the mappings sending a set A to its power set 2^A and a function $f : A \to B$ to a function $\mathcal{P}(f) : 2^A \to 2^B$ defined as $\mathcal{P}(f)(S) = f(S)$ defines a functor $\mathcal{P} : \mathbf{Sets} \to \mathbf{Sets}$ called the *power set* functor.

In the same spirit we shall try to relate functors. Given functors $F, G : \mathbb{C} \to \mathbb{D}$, a *natural transformation* $\alpha : F \Rightarrow G$ is a family $\{\alpha_A\}_A : FA \to GA$ of $\mathbb{D}$-morphisms indexed over the objects of $\mathbb{C}$ such that for any morphism $f : A \to B$ in $\mathbb{C}$ we have $G(f)\alpha_A = \alpha_B F(f)$.

2.1. Some categorical constructions

Many notions and constructions that we use in everyday mathematics can be unified in a single categorical definition. One such notion (construction) is that of the categorical product. Let $\mathbb{C}$ be a category and A and B be two objects in $\mathbb{C}$. The product of A and B denoted $A \times B$ is an object together with morphisms $\pi_1 : A \times B \to A$ and $\pi_2 : A \times B \to B$ called projections such that the following universal property is satisfied: for any object C and pair of morphisms $f : C \to A$ and $g : C \to B$, there is a unique morphism $\langle f, g\rangle : C \to A \times B$ such that $\pi_1\langle f, g\rangle = f$ and $\pi_2\langle f, g\rangle = g$. For example, it is easy to check that in the category of sets and functions this yields the Cartesian product of sets. However, note that in the category of sets and relations the product is given by the disjoint union operation.

3. CATEGORICAL TRACE

In recent years the notion of categorical trace, introduced by Joyal, Street, and Verity [1] has gained a lot of attention ranging in applications from lambda calculus [2], to semantics of linear logic [3]. Before we introduce this notion we shall briefly mention the notion of a monoidal (tensor) category. A *monoidal category* $\mathbb{C}$ is a category together with a bifunctor $\otimes : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ (monoidal or tensor product), an object I in $\mathbb{C}$ called the *tensor unit* and structure isomorphisms $\alpha_{A,B,C} : A \otimes (B \otimes C) \to (A \otimes B) \otimes C$ (associativity), $\rho_A : A \otimes I \to A$ (right unit), and $\lambda_A : I \otimes A \to A$ (left unit) such that certain diagrams commute, for details see the canonical reference [4]. A *symmetric* monoidal category is a monoidal category with a natural isomorphism $s_{A,B} : A \otimes B \to B \otimes A$ such that $s_{B,A}s_{A,B} = 1_{A\otimes B}$ plus some more conditions. The canonical example to keep in mind is that of vector spaces and linear transformations, where the monoidal product is the familiar tensor product of vector spaces and the tensor unit is the ground field.

A *traced symmetric monoidal category* is a symmetric monoidal category $(\mathbb{C}, \otimes, I, s)$ with a family of functions $Tr^U_{X,Y} : \mathbb{C}(X \otimes U, Y \otimes U) \to \mathbb{C}(X, Y)$ called a *trace*, subject to the following axioms:

- **Natural** in X,
$$Tr^U_{X,Y}(f)g = Tr^U_{X',Y}(f(g \otimes 1_U))$$
where $f : X \otimes U \to Y \otimes U$, $g : X' \to X$,

- **Natural** in Y,
$$gTr^U_{X,Y}(f) = Tr^U_{X,Y'}((g \otimes 1_U)f)$$
where $f : X \otimes U \to Y \otimes U$, $g : Y \to Y'$,

- **Dinatural** in U,
$$Tr^U_{X,Y}((1_Y \otimes g)f) = Tr^{U'}_{X,Y}(f(1_X \otimes g))$$
where $f : X \otimes U \to Y \otimes U'$, $g : U' \to U$,

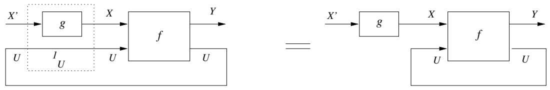
- **Vanishing (I,II)**,
$$Tr^I_{X,Y}(f) = f$$
and
$$Tr^{U\otimes V}_{X,Y}(g) = Tr^U_{X,Y}(Tr^V_{X\otimes U,Y\otimes U}(g))$$
for $f : X \otimes I \to Y \otimes I$ and
$g : X \otimes U \otimes V \to Y \otimes U \otimes V$,

- **Superposing**,
$$g \otimes Tr^U_{X,Y}(f) = Tr^U_{W\otimes X,Z\otimes Y}(g \otimes f)$$
for $f : X \otimes U \to Y \otimes U$ and $g : W \to Z$,

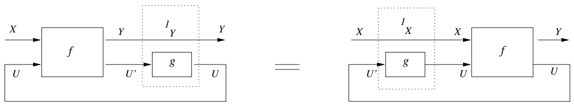
- **Yanking**,
$$Tr^U_{U,U}(s_{U,U}) = 1_U.$$

3.1. Graphical Representation

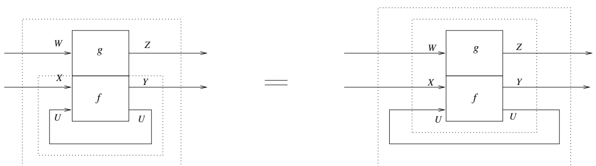
The axioms above admit a pictorial representation, for example we have the following:



Naturality in X



Dinaturality in U



Superposing

3.2. Examples

Consider the category **FDVec**$_k$ of finite dimensional vector spaces and linear transformations. Given $f : V \otimes U \to W \otimes U$, and $\{v_i\}, \{u_j\}, \{w_k\}$ bases for V, U, W respectively, write f as $f(v_i \otimes u_j) = \sum_{k,m} a_{ij}^{km}\, w_k \otimes u_m$, then

$$Tr_{V,W}^{U}(f)(v_i) = \sum_{j,k} a_{ij}^{kj}\, w_k.$$

Note that this is a simple generalization of the notion of trace for a matrix: sum of diagonal entries. Here the diagonal entries are block matrices of size $dim(W) \times dim(V)$ and one is adding $dim(U)$ many such diagonal blocks. Clearly if we choose $dim(V) = dim(W) = 1$, that is when $V \cong W \cong k$, we get back the usual trace.

As a second example consider the category **Rel** of sets and binary relations with $X \otimes Y = X \times Y$. Note that this is a tensor product with unit $I = \{*\}$. Given $R : X \otimes U \to Y \otimes U$, $Tr_{X,Y}^{U}(R) : X \to Y$ is defined by

$$(x, y) \in Tr(R) \text{ iff } \exists u.(x, u, y, u) \in R.$$

3.3. Traced Unique Decomposition Categories

Finally we shall consider an important class of traced categories called *traced Unique Decomposition Categories* (traced UDCs) without getting into too much detail. These are categories with special additive structures on their homsets called Σ-monoids.

A Σ*-monoid* is a pair (M, Σ), where M is a nonempty set and Σ is a partial operation on countable families in M where $\{x_i\}_{i\in I}$ is *summable* if $\Sigma_{i\in I} x_i$ is defined subject to:

- *Partition-Associativity:* $\{x_i\}_{i\in I}$ and $\{I_j\}_{j\in J}$ a countable partition of I

$$\Sigma_{i\in I} x_i = \Sigma_{j\in J}(\Sigma_{i\in I_j} x_i).$$

- *Unary sum:* $\Sigma_{i\in\{j\}} x_i = x_j$.

Here are some quick facts about Σ-monoids:

- $\Sigma_{i\in\emptyset} x_i$ exists and is denoted by 0. It is a countable additive identity.
- Sum is commutative and associative whenever defined.
- $\Sigma_{i\in I} x_{\phi(i)}$ is defined for any permutation ϕ of I, whenever $\Sigma_{i\in I} x_i$ exits.
- There are **no** additive inverses: $x + y = 0$ implies $x = y = 0$.

Let us look at some examples. The set $M = \mathbf{PInj}(X, Y)$ of partial injective functions where a family $\{f_i\}$ is summbale if f_i and f_j have disjoint domains and codomains for all $i \neq j$. The sum is then defined as $(\Sigma_I f_i)(x) = \begin{cases} f_j(x), & \text{if } x \in Dom(f_j) \text{ for some } j \in I; \\ undefined, & \text{otherwise.} \end{cases}$

The set $M = \mathbf{Pfn}(X, Y)$ of partial functions from a set X to a set Y where a family $\{f_i\}$ is summable if f_i and f_j have disjoint domains for all $i \neq j$, and $(\Sigma_I f_i)(x)$ is defined as above.

The set $M = \mathbf{Rel}(X, Y)$ of binary relations from X to Y. Here all families are summable, and $\Sigma_i R_i = \bigcup_i R_i$.

Finally here is a non-example, let M be an ω-complete poset, that is a partially ordered set where all countable chains have suprema. Define a family $\{x_i\}$ to be summable if it is a countable chain, and let the sum be $\Sigma_{i\in I} x_i = sup_{i\in I} x_i$. Suppose x, y, z are in this family, with $x \leq z, y \leq z$ and x, y incomparable, then $x + (y + z)$ is defined but $(x + y) + z$ is not defined.

A *unique decomposition category* $\mathbb{C}$ is a symmetric monoidal category where (1) Every homset is a Σ-monoid, (2) Composition distributes over sum whenever defined, satisfying the axiom:

(A) For all $j \in I$ (I finite) there are morphisms

- *quasi injection*: $\iota_j : X_j \to \otimes_I X_i$, and
- *quasi projection*: $\rho_j : \otimes_I X_i \to X_j$,

such that

- $\rho_k \iota_j = 1_{X_j}$ if $j = k$ and $0_{X_j X_k}$ otherwise,
- $\sum_{i\in I} \iota_i \rho_i = 1_{\otimes_I X_i}$.

Proposition 1 [Matricial Representation]
For $f : \otimes_J X_j \to \otimes_I Y_i$, there exists a unique family $\{f_{ij}\}_{i\in I, j\in J} : X_j \to Y_i$ with $f = \sum_{i\in I, j\in J} \iota_i f_{ij} \rho_j$, namely, $f_{ij} = \rho_i f \iota_j$.

In particular, for $|I| = m, |J| = n$

$$f = \begin{bmatrix} f_{11} & \dots & f_{1n} \\ \vdots & \vdots & \vdots \\ f_{m1} & \dots & f_{mn} \end{bmatrix}$$

Here are a few examples, the category **PInj** of sets and partial injective functions with disjoint union as the monoidal product. Here, $\rho_j : \otimes_{i\in I} X_i \to X_j$, $\rho_j(x, i)$ is undefined for $i \neq j$ and $\rho_j(x,j) = x$, and $\iota_j : X_j \to \otimes_{i\in I} X_i$ by $\iota_j(x) = (x,j)$.

The category **Rel** with disjoint union as the monoidal product, $\rho_j : \otimes_{i\in I} X_i \to X_j$, $\rho_j = \{((x,j),x) \,|\, x \in X_j\}$. And $\iota_j : X_j \to \otimes_{i\in I} X_i$, $\iota_j = \{(x,(x,j)) \,|\, x \in X_j\} = \rho_j^{op}$.

Proposition 2 [Standard Trace Formula]
Let $\mathbb{C}$ be a unique decomposition category such that for every X, Y, U and $f : X \otimes U \to Y \otimes U$, the sum $f_{11} + \sum_{n=0}^{\infty} f_{12} f_{22}^n f_{21}$ exists, where f_{ij} are the components of f. Then, $\mathbb{C}$ is traced and

$$Tr^U_{X,Y}(f) = f_{11} + \sum_{n=0}^{\infty} f_{12} f_{22}^n f_{21}.$$

Let us calculate some traces. Let $\mathbb{C}$ be a traced UDC. Then given any $f : X \otimes U \to Y \otimes U$, $Tr^U_{X,Y}(f)$ exists.

- Let $f : X \otimes U \to Y \otimes U$ be given by $\begin{bmatrix} g & 0 \\ h & 0 \end{bmatrix}$.
  Then
  $Tr^U_{X,Y}(f) = Tr^U_{X,Y}\left(\begin{bmatrix} g & 0 \\ h & 0 \end{bmatrix}\right) = g + \sum_n 00^n h$
  $= g + 0h = g + 0 = g.$

- Let $f : X \otimes U \to Y \otimes U$ be given by $\begin{bmatrix} g & 0 \\ 0 & h \end{bmatrix}$.
  Then
  $Tr^U_{X,Y}(f) = Tr^U_{X,Y}\left(\begin{bmatrix} g & 0 \\ 0 & h \end{bmatrix}\right) = g + \sum_n 0h^n 0$
  $= g + 0 = g.$

4. CONCLUSION

In this paper, I have tried to give a glimpse of category theory with special attention to the notion of trace hoping that such ideas might find some applications in Information Theory and/or Statistical Inference. Even though the treatment here is very cursory, I hope I was able to convey some of the general philosophy and the spirit of thinking in terms of categories. At this point the reader will naturally wonder what she or he might be able to do with categories. Let me try to point out some possibilities below.

- *Unification*
  Category theory, above all offers a unifying language whereof one can define once and for all a single notion or construction which upon instantiation yields the extant concrete definitions. We saw one such example here, namely that of a categorical product which yields the Cartesian product in **Sets** and graph product in **Graphs** and direct product of vector spaces in $\mathbf{FDVec}_k$, etc.

- *Technology transfer*
  Once you have organized your objects of study into a category you can start trying different constructions on them. For example you might wonder what the coproduct for two graphs could be or what kind of tensor (monoidal) products you could define on graphs, etc. Thus general categorical constructions (for example limits and colimits) might yield new, unknown constructions on the objects that you study. On the other hand, there might be some constructions that are possible in categories with additional structure. Once you prove that your category has this additional structure, you can transfer the existing technology and apply it to objects of your study.

- *Mathematical analysis*
  Quite often it is much more productive to study certain mathematical objects via their representations in terms of more familiar objects which are easier to understand or study, or it is the case that there are a lot of techniques developed for their study. For example think of group representation theory where one represents groups as linear transformations on vector spaces, or think of homology theory where one studies topological spaces using groups (homology groups). Such relations are often *functorial*, for example the functor that associates an abelian group to a topological space, etc. The lesson here, thus is to find functorial relations (representations) from the category of objects of study into one which is much better understood, in order to obtain new results about the former category.

5. REFERENCES

[1] A. Joyal, R. Street, and D. Verity, "Traced monoidal categories," *Math. Proc. Cambridge Philos. Soc.*, vol. 119, pp. 447–468, 1996.

[2] M. Hasegawa, "Recursion from cyclic sharing: traced monoidal categories and models of cyclic lambda calculi," in *Proc. 3rd International Conference on Typed Lambda Calculi and Applications (TLCA'97)*, 1997, pp. 196–213.

[3] E. Haghverdi and P.J. Scott, "A categorical model for the geometry of interaction," *Theoretical Computer Science*, vol. 350, pp. 252–274, 2006.

[4] S. Mac Lane, *Categories for The Working Mathematician*, Springer-Verlag, New York, 1998.