Category Theory

Alan Hylton ∈ NASA

History

Functions...

- started with an implicit definition - the dependency of some quantity on others
- were coined by Leibniz in 1673
- were given the $f(x)$ notation by Euler in 1734
- were formalized in the language of set theory from around 1847 to 1939
- have since been further abstract
- often capture and represent *structure*

Structure

$$f : X \to Y$$
$$x \mapsto y$$

Functions can be...

- smooth
- continuous
- order-preserving
- arithmetic operators
- injective (1-1) or surjective (onto)
- composed with other functions (when it makes sense)

Algebra

The study of algebraic structures is roughly the study of computational frameworks

Examples:

- groups
- rings
- modules

Groups

Let G be a set
Then G is a *group* if it has a basic arithmetic-satisfying operation $\cdot$

Let $a, b, c \in G$; then G is a group if it

- is associative : $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
- has an identity: $e \in G$ such that $a \cdot e = e \cdot a = a$
- has inverses : for each a, there is a unique a^{-1} such that $aa^{-1} = a^{-1}a = e$

If $\cdot$ is commutative – $a \cdot b = b \cdot a$ for all $a, b \in G$ – then G is abelian (write $+$ instead of $\cdot$)

Examples:

$G = \{0\}$ - the trivial group
$\mathbb{Z}$ with addition
$\mathbb{Z}_2 = \{0, 1\}$ with addition mod 2
$\mathbb{Z} \times \mathbb{Z}$ with pairwise addition
$n \times n$ invertible matrices
Free group over a set

Fundamental functions: Homomorphisms

Let $(G,\cdot)$ and $(H,*)$ be groups
A *homomorphism* f maps $G \to H$ in a way that preserves group structure

Specifically $f(a \cdot b) = f(a) * f(b)$

Example 1:
$$f : \mathbb{Z} \to \mathbb{Z}_2 \text{ by } f(x) = \begin{cases} 0 & \text{even} \\ 1 & \text{odd} \end{cases}$$

Example 2:
$g : \mathbb{Z} \to \mathbb{Z}_2$ by $g(x) = 0$

Example 3:
$h : \mathbb{Z}_2 \to \mathbb{Z}$ by $h(x) = 0$

$\mathbb{Z} \times \mathbb{Z}$

A typical element of $\mathbb{Z} \times \mathbb{Z}$ is (a, b) where $a, b \in \mathbb{Z}$

The group operator is pairwise addition: $(a, b) + (c, d) = (a + c,\ b + d)$

The identity is $(0, 0)$

The inverse of (a, b) is $(-a, -b)$

Associativity is naturally inherited from the integers

Composition of homomorphisms

Let G, H, K be groups, and let $f : G \to H$ and $g : H \to K$ be homomorphisms

Then $g \circ f : G \to K$ is a homomorphism

Example:
$f : \mathbb{Z} \to \mathbb{Z} \times \mathbb{Z}$ by $f(x) = (x, 0)$ $\leftarrow$ an embedding of $\mathbb{Z}$ into $\mathbb{Z} \times \mathbb{Z}$
$g : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ by $g((a, b)) = a \pmod 2$ $\leftarrow$ the even parity of a

Then $g \circ f : \mathbb{Z} \to \mathbb{Z}_2$ is given by $x \mapsto x \pmod 2$

Free group

Let S be a set whose elements are called generators

The *free group G over S*, written $\langle S \rangle$ is defined by:

- G is all combinations of generators, called *words*
- the operation is concatenation, written by juxtaposition
- the identity is the empty word
- inverses are simply added (for a generator a, we add a^{-1})

Example:
Let $S = \{a, b\}$; then
$G = \langle a, b \rangle = \{a, a^{-1}, b, b^{-1}, aa, ab, ab^{-1}, bb, aab, aab^{-1}, ab^{-1}, abb, \cdots\}$

Composition of homomorphisms

Let G, H, K be groups, and let $f : G \to H$ and $g : H \to K$ be homomorphisms

Then $g \circ f : G \to K$ is a homomorphism

Example:

$f : \mathbb{Z} \to \langle a, b \rangle$ by $f(x) = \overbrace{a \cdot \cdots \cdot a}^{x \text{ times}}$ ← x-fold concatenation of a

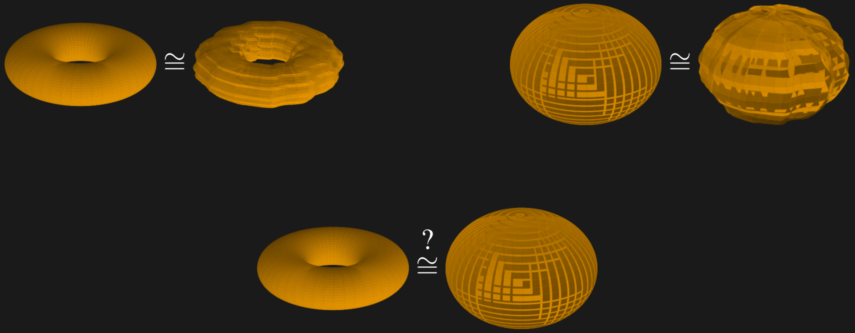
$g : \langle a, b \rangle \to \mathbb{Z}$ by $g(a) = 1$ and $g(b) = 0$ ← the even parity of as in a word

Then $g \circ f : \mathbb{Z} \to \mathbb{Z}_2$ is given by $x \mapsto x(\text{mod } 2)$

Topology

Topology is the generalization of geometry where we define continuity and convergence

Angles and distance are not fundamental to this geometric setting

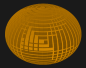


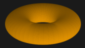
Fundamental functions: Continuous maps

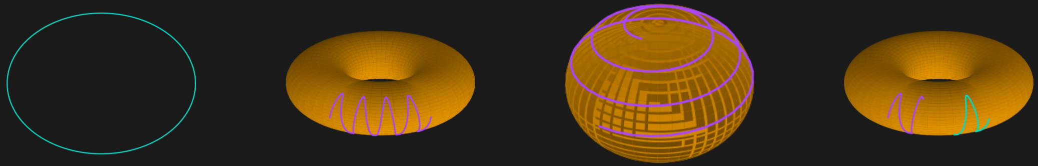
Continuous functions are fundamental to topology

Examples:

$I = [0, 1] =$

$S^2 =$ 

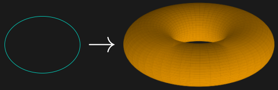
$T^2 =$ 

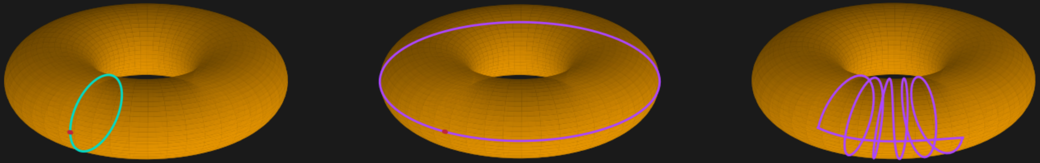


Loops

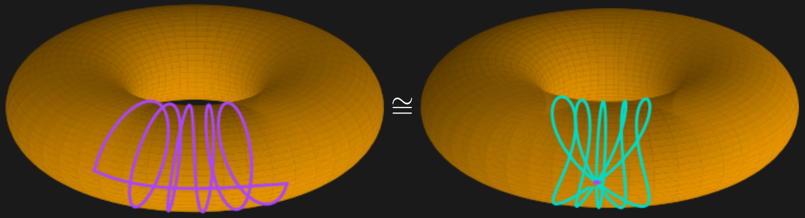
A *loop* is a mapping of S^1 into a topological space

Examples:

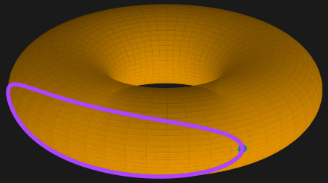


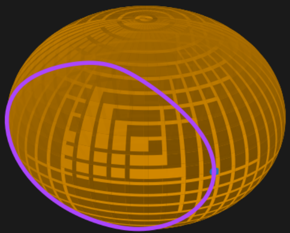


Equivalence of loops

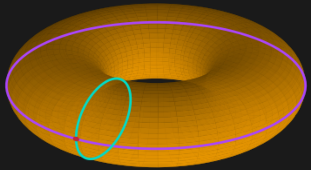


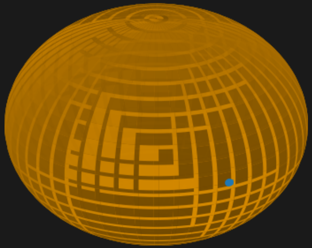
Trivial loops



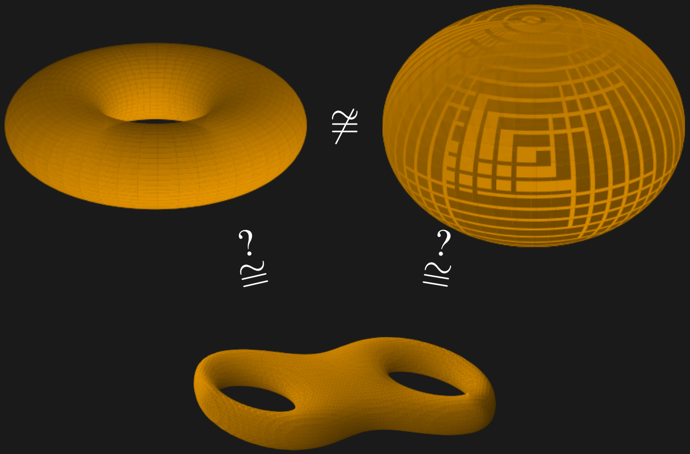


Basic Loops

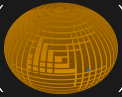


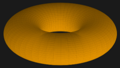
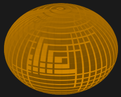


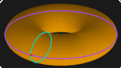
Equivalences



An algebra of geometry: The fundamental group

$\pi_1\Big($  $\Big) = \{0\}$

$\not\cong \quad \longrightarrow$  $\not\cong$ 

$\pi_1\Big($  $\Big) = \mathbb{Z} \times \mathbb{Z}$

Category Theory

A *category* is a collection of related objects...

Category Theory

A *category* is a collection of related objects...

- e.g. the category of all sets
- e.g. the category of all (real) vector spaces
- e.g. the category of all groups
- e.g. the category of all topological spaces

Category Theory

A *category* is a collection of related objects...

- e.g. the category of all sets
- e.g. the category of all (real) vector spaces
- e.g. the category of all groups
- e.g. the category of all topological spaces

...along with a collection of arrows between objects

- e.g. for sets X and Y, all functions from $X \to Y$
- e.g. for vector spaces W and V, all linear transformations from $W \to V$
- e.g. for any groups G and H, all homomorphisms $G \to H$
- e.g. for any spaces X and Y, all continuous maps $X \to Y$

Definition

A *category* $\mathcal{C}$ has

- Objects $\mathrm{Ob}(\mathcal{C})$
- For $A, B \in \mathrm{Ob}(\mathcal{C})$, all morphisms $A \to B$; denoted $\hom(A, B)$
- For each $A \in \mathrm{Ob}(\mathcal{C})$, an identity arrow $1_A : A \to A$

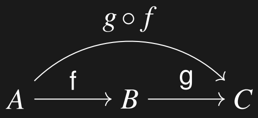
Definition

A *category* $\mathcal{C}$ has

- Objects $\mathrm{Ob}(\mathcal{C})$
- For $A, B \in \mathrm{Ob}(\mathcal{C})$, all morphisms $A \to B$; denoted $\hom(A, B)$
- For each $A \in \mathrm{Ob}(\mathcal{C})$, an identity arrow $1_A : A \to A$

Composition is *always* allowed: if there is a relation $f : A \to B$ and $g : B \to C$, then

Then $g \circ f \in \hom(A, C)$



Identities

Let $\mathcal{C}$ be a category, $A, B \in \mathrm{Ob}(\mathcal{C})$, and $f \in \hom(A, B)$

Then $1_B \circ f = f$, $f \circ 1_A = A$

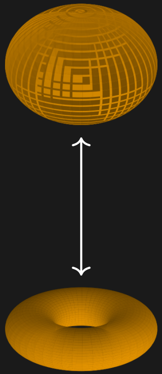
Example:

$$(1_B \circ f) \circ 1_A = 1_B \circ (f \circ 1_A) = 1_B \circ f \circ 1_A = f$$

Usage I

Categories form a system out of related objects and their morphisms

Geometric



Algebraic

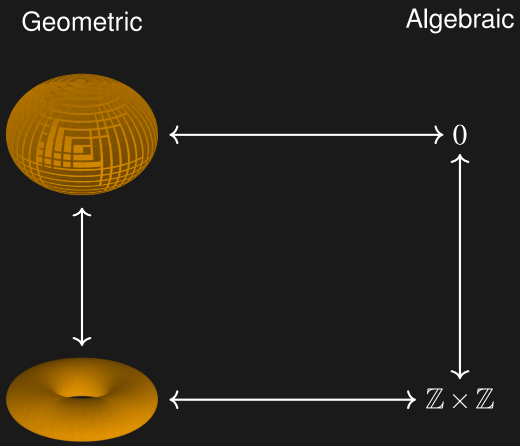
$$0$$

$$\updownarrow$$

$$\mathbb{Z} \times \mathbb{Z}$$

Usage II

Categories can be mapped to other categories using *functors*

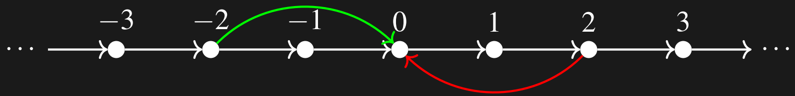


Example 1

Let $\mathcal{C}$ have integers as objects - $\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$

For any two integers x and y, define

$$\hom(x,y) = \begin{cases} \phi_{xy} & x \leq y \\ \varnothing & \text{else} \end{cases}$$



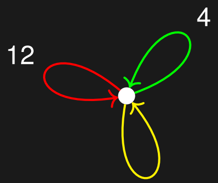
Example 2

Let $\mathcal{C}$ have one object $\bullet$

The morphisms from $\bullet$ to itself are

$$\hom(\bullet, \bullet) = \mathbb{Z}$$

Composition is given by addition



$$12 \circ 4 = 16$$

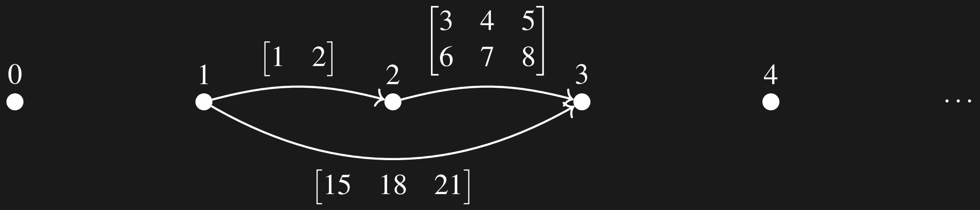
Example 3

Let $\mathcal{C}$ have natural numbers as objects - $\{0, 1, 2, 3, \ldots\}$

The morphisms from m to n are

$$\text{hom}(m, n) = \text{all } m \times n \text{ matrices}$$

Arrow composition is given by matrix multiplication

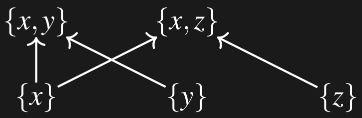


Example 4

Consider the power set of $\{x, y, z\}$: $\{\varnothing, \{x\}, \{y\}, \{z\}, \{x, y\}, \{x, z\}, \{y, z\}, \{x, y, z\}\}$

Order by set inclusion to make a category: for any elements x, y define

$$\hom(x, y) = \begin{cases} \phi_{xy} & x \subseteq y \\ \varnothing & \text{else} \end{cases}$$

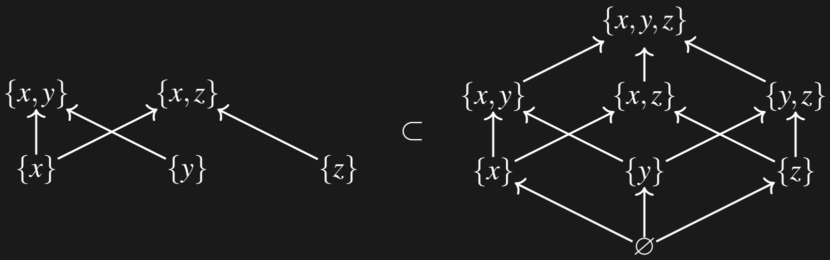


Example 4

Consider the power set of $\{x, y, z\}$: $\{\varnothing, \{x\}, \{y\}, \{z\}, \{x,y\}, \{x,z\}, \{y,z\}, \{x,y,z\}\}$

Order by set inclusion to make a category: for any elements x, y define

$$\hom(x,y) = \begin{cases} \phi_{xy} & x \subseteq y \\ \varnothing & \text{else} \end{cases}$$

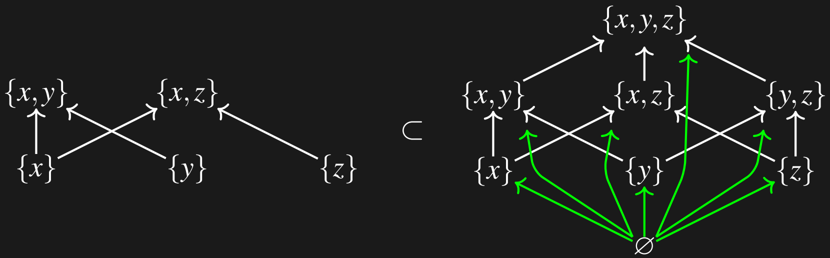


Example 4

Consider the power set of $\{x,y,z\}$: $\{\varnothing, \{x\}, \{y\}, \{z\}, \{x,y\}, \{x,z\}, \{y,z\}, \{x,y,z\}\}$

Order by set inclusion to make a category: for any elements x,y define

$$\hom(x,y) = \begin{cases} \phi_{xy} & x \subseteq y \\ \varnothing & \text{else} \end{cases}$$



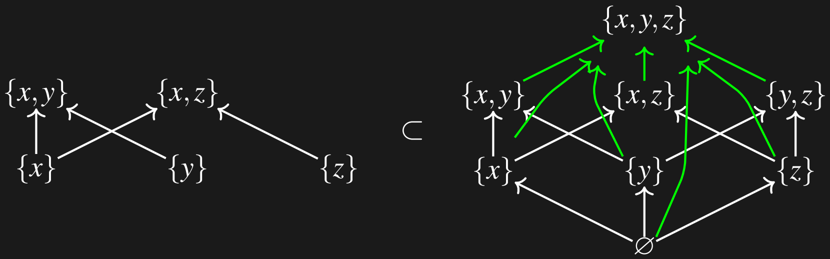
$\varnothing$ is the *initial object*

Example 4

Consider the power set of $\{x,y,z\}$: $\{\varnothing, \{x\}, \{y\}, \{z\}, \{x,y\}, \{x,z\}, \{y,z\}, \{x,y,z\}\}$

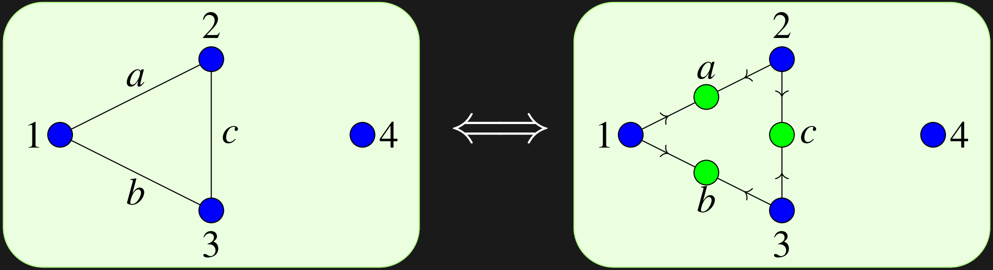
Order by set inclusion to make a category: for any elements x, y define

$$\hom(x,y) = \begin{cases} \phi_{xy} & x \subseteq y \\ \varnothing & \text{else} \end{cases}$$



$\{x,y,z\}$ is the *terminal object*

Example 5: Graphs



$G = (V, E)$

- Objects: $V \cup E$
- Arrows: unique morphism $v \hookrightarrow e$ if v is incident to e
- Name: $\mathcal{G}$

First cut

We can define a *category of neural networks* **NNet**

- Objects: natural numbers
- Morphisms: $\hom(m, n)$ is all neural networks with m inputs and n outputs
- Composition is concatenation where it makes sense

NNet has enough structure to define back propagation categorically!

(but I want more)

A new approach

A *neural network of length l* is a sequence of functions

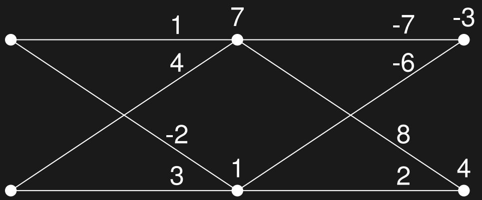
$$\left(\mathbb{R}^{n_0} \xrightarrow{N_0} \mathbb{R}^{n_1} \xrightarrow{N_1} \dots \xrightarrow{N_{l-1}} \mathbb{R}^{n_l}\right)$$

The functions N_i will be referred to as *layer functions of N*.

$$N_i : \mathbb{R}^{n_i} \to \mathbb{R}^{n_{i+1}} \text{ by } x \mapsto \sigma(Ax + b)$$

Notation: we use σ for activation functions

Example



$$\mathbb{R}^2 \xrightarrow{N_0} \mathbb{R}^2 \xrightarrow{N_1} \mathbb{R}^2$$

$$N_0(x) = \sigma\left(\begin{pmatrix} 1 & 4 \\ -2 & 3 \end{pmatrix} x + \begin{pmatrix} 7 \\ 1 \end{pmatrix}\right)$$

$$N_1(x) = \sigma\left(\begin{pmatrix} -7 & -6 \\ 8 & 2 \end{pmatrix} x + \begin{pmatrix} -3 \\ 4 \end{pmatrix}\right)$$

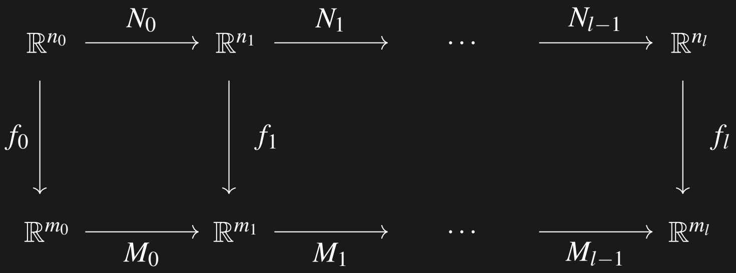
Note: in **NNet** this neural network is an arrow $2 \to 2$

Morphisms

$N = (N_0, N_1, \ldots, N_{l-1})$ and $M = (M_0, M_1, \ldots, M_{l-1})$ are neural networks of length l

A *morphism* $f : N \to M$ is a sequence of functions $(f_0, f_1, \ldots, f_l)$ such that

$$f_k \circ N_{k-1} \circ \cdots \circ N_1 \circ N_0 = M_{k-1} \circ M_{k-2} \circ \cdots \circ M_0 \circ f_0 \text{ for all } 1 \leq k \leq l$$

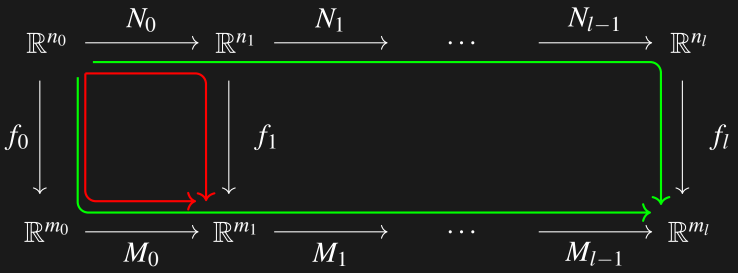


Morphisms

$N = (N_0, N_1, \dots, N_{l-1})$ and $M = (M_0, M_1, \dots, M_{l-1})$ are neural networks of length l

A *morphism* $f : N \to M$ is a sequence of functions $(f_0, f_1, \dots, f_l)$ such that

$$f_k \circ N_{k-1} \circ \cdots \circ N_1 \circ N_0 = M_{k-1} \circ M_{k-2} \circ \cdots \circ M_0 \circ f_0 \text{ for all } 1 \leq k \leq l$$



A new hope

We can form a new category whose

- objects are neural networks of length l
- morphisms are appropriate sequences $(f_0, \dots, f_l)$

What's left? Composition

This is done layer by layer:

$$(f_0, f_1, \dots, f_l) \circ (g_0, g_1, \dots, g_l) = (f_0 \circ g_0, f_1 \circ g_1, \dots, f_l \circ g_l).$$

Categories of neural nets

Recall layer functions:

$$N_i : \mathbb{R}^{n_i} \to \mathbb{R}^{n_{i+1}} \text{ by } x \mapsto \sigma(Ax+b)$$

We can make several categories by picking the activation function σ:

AffineNet$_l$	N_i required to be an affine function followed by any activation function
ReluAffineNet$_l$	N_i required to be an affine function followed by ReLU activation function

Category note: **ReluAffineNet**$_l$ is a subcategory of **AffineNet**$_l$

Recap

A *category* is a collection of objects and arrows between them

Examples:
The category of all sets and functions between them
The category of all categories and functors between them

The objects may or may not be sets, the arrows may or may not be functions
$\rightarrow$ a categorical construction should not require “looking inside” the objects

We want to transfer/utilize structure from one category to another
$\rightarrow$ *functors* embed one category into another

Arrows instead of elements: Injectivity

A function $f : X \to Y$ is *injective* if $f(a) = f(b)$ implies $a = b$

What's a good way to pick a and b? Mappings $\phi, \psi : * \to X$

We can write $* \underset{\psi}{\overset{\phi}{\rightrightarrows}} X \xrightarrow{f} Y$

Of course it doesn't matter if we pick one element out at a time to test f with...

$$W \rightrightarrows X \to Y$$

f is a *monomorphism* if for all W we always have $f \circ \phi = f \circ \psi$ implies $\phi = \psi$

Opposites: Surjectivity

A function $f : X \to Y$ is *surjective* if for all $y \in Y$, there is some $x \in X$ with $f(x) = y$

We can write $X \xrightarrow{f} Y \underset{\psi}{\overset{\phi}{\rightrightarrows}} W$

f is an *epimorphism* if for all W we always have $\phi \circ f = \psi \circ f$ implies $\phi = \psi$

This is just the definition of monomorphism with the arrows turned around!

Subobject

Let $G = \langle a \rangle$ and $H = \langle a, b \rangle$

Then G is a *subgroup* of H

As sets: $G \subset H$
As algebras: $xy \in G = xy \in H$

$$G \hookrightarrow H$$

Let $f : W \to X$ and $f' : W' \to X$ be monomorphisms
Define a preorder

$$(W', f') \leq (W, f) \text{ iff } \exists g : W' \to W \text{ st } f' \circ g = f$$

A *subobject* of X is an equivalence class of such (W, f)

In all cases, the image of a function or arrow can be defined in terms of subobjects

Functors

Let $\mathcal{C}$ and $\mathcal{D}$ be categories

A *functor* $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ transforms $\mathcal{C}$ into $\mathcal{D}$, keeping $\mathcal{C}$'s structure intact

- $A \in \mathrm{Ob}(\mathcal{C})$ becomes $\mathcal{F}(A) \in \mathrm{Ob}(\mathcal{D})$
- $A, B \in \mathrm{Ob}(\mathcal{C})$ with $A \xrightarrow{f} B \in \hom(A,B)$ becomes $\mathcal{F}(A) \xrightarrow{\mathcal{F}(f)} \mathcal{F}(B) \in \hom(\mathcal{F}(A), \mathcal{F}(B))$
- 1_A becomes $1_{\mathcal{F}(A)}$

The structure part: composition remains intact

$$\mathcal{F}(g \circ f) = \mathcal{F}(g) \circ \mathcal{F}(f)$$

Example I

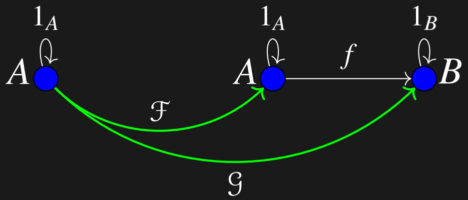
Let $\mathbf{1}$ be the category with one object A and the identity arrow

There is a functor $\mathcal{F} : \mathbf{1} \to \mathbf{1}$

Example II

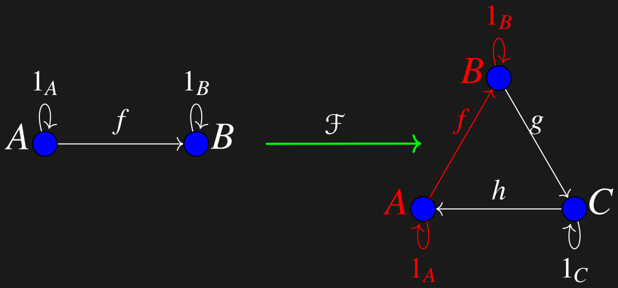
Let $\mathbf{2}$ be the category with two objects A and B and an arrow $A \to B$

There are two functors $\mathbf{1} \to \mathbf{2}$, $\mathcal{F}$ and $\mathcal{G}$



Example III

Let $\mathbf{3}$ be the category with three objects and arrows making a triangle
There are several functors from $\mathbf{2} \to \mathbf{3}$; we show one of them

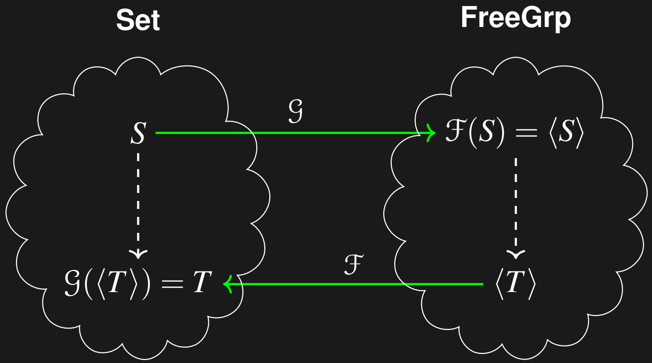


Are there any functors from $\mathbf{3} \to \mathbf{2}$?

Example IV: Adjunctions

There is a functor $\mathcal{G}$: **Set** $\to$ **FreeGrp** by taking S to $\langle S \rangle$

We can also forget the group structure of $\langle S \rangle$, getting a functor $\mathcal{F}$: **FreeGrp** $\to$ **Set**



Functors: objects and arrows interchanged

Adjointness: Any arrow from $S \to \mathcal{G}(\langle T \rangle)$ uniquely matched by arrow $\mathcal{F}(S) \to \langle T \rangle$

Questions?