

# Explicit Categorical Constructions Used in Modeling Sentences

*Monoidal Product Category*

Mia Goldstein and Emily Herbert  
SUNY New Paltz

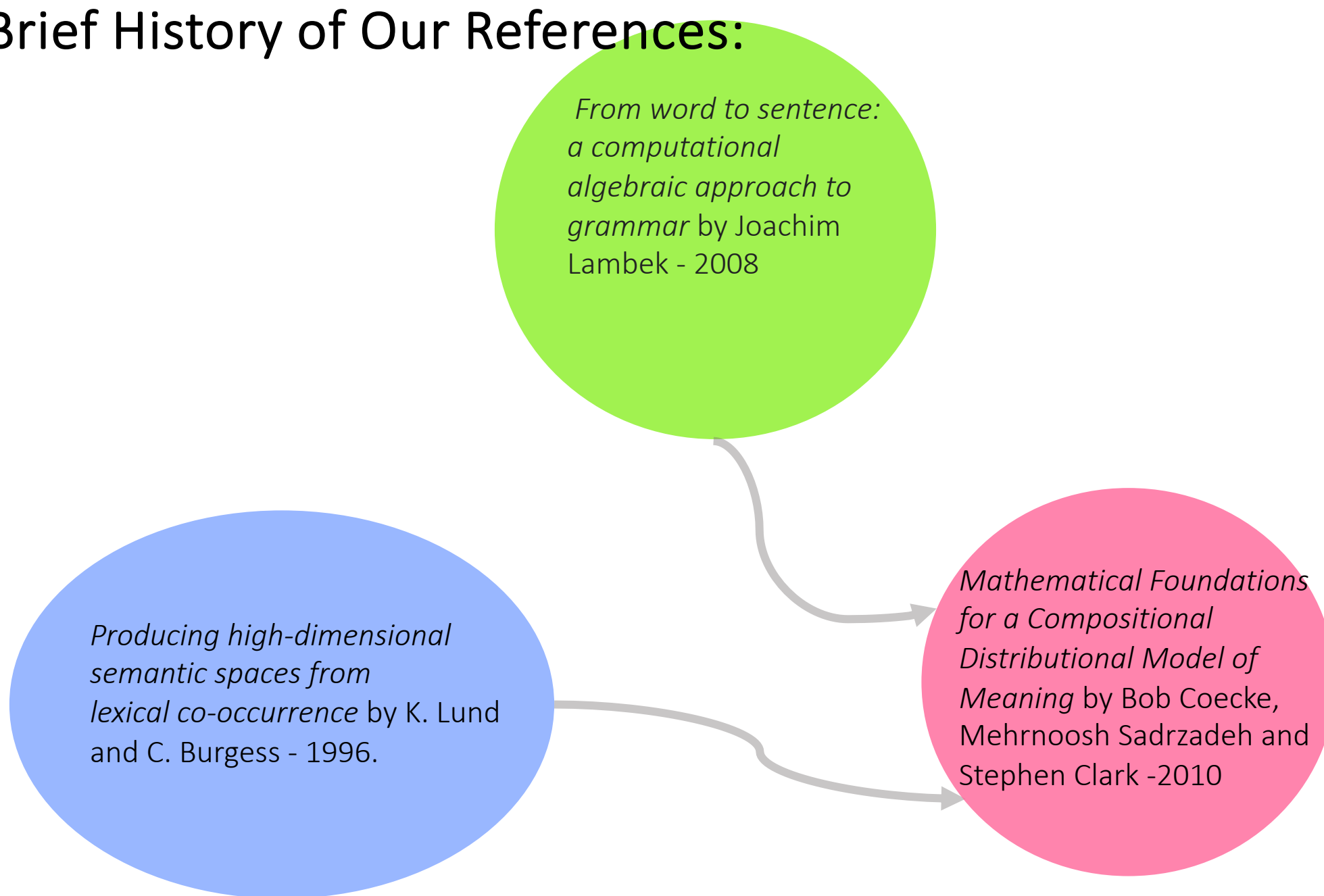
Thank you to RSCA and NCUWM for getting us here!

# Brief History of Our References:

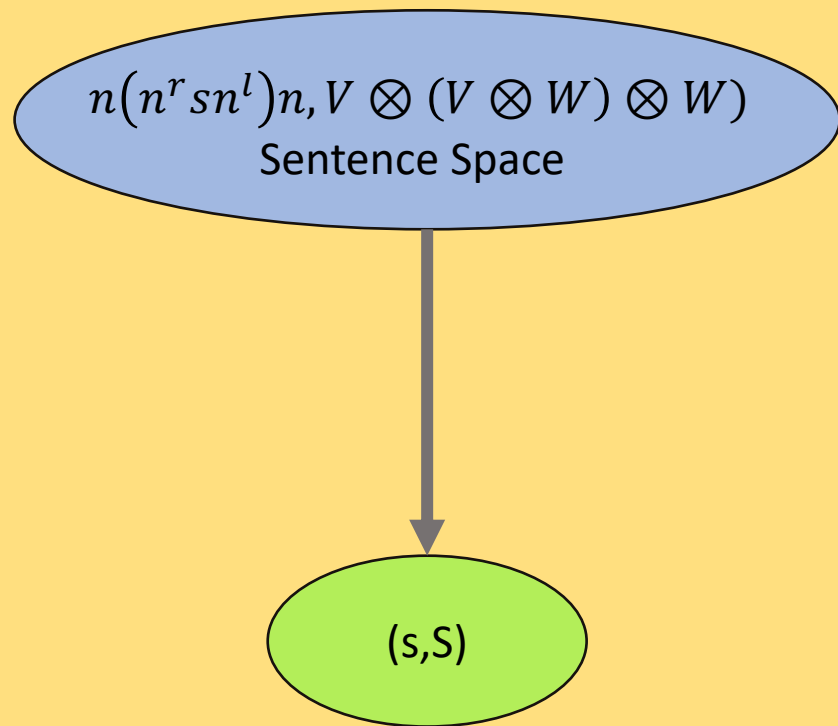
*From word to sentence:  
a computational  
algebraic approach to  
grammar* by Joachim  
Lambek - 2008

*Producing high-dimensional  
semantic spaces from  
lexical co-occurrence* by K. Lund  
and C. Burgess - 1996.

*Mathematical Foundations  
for a Compositional  
Distributional Model of  
Meaning* by Bob Coecke,  
Mehrnoosh Sadrzadeh and  
Stephen Clark -2010



# Our Motivation:



Original Example of sentence processing

arXiv:1003.4394v1 [cs.CL] 23 Mar 2010

## Mathematical Foundations for a Compositional Distributional Model of Meaning

Bob Coecke\*, Mehmoosh Sadrzadeh\*, Stephen Clark†

coecke, mehms@comlab.ox.ac.uk – stephen.clark@cl.cam.ac.uk

\*Oxford University Computing Laboratory

†University of Cambridge Computer Laboratory

### Abstract

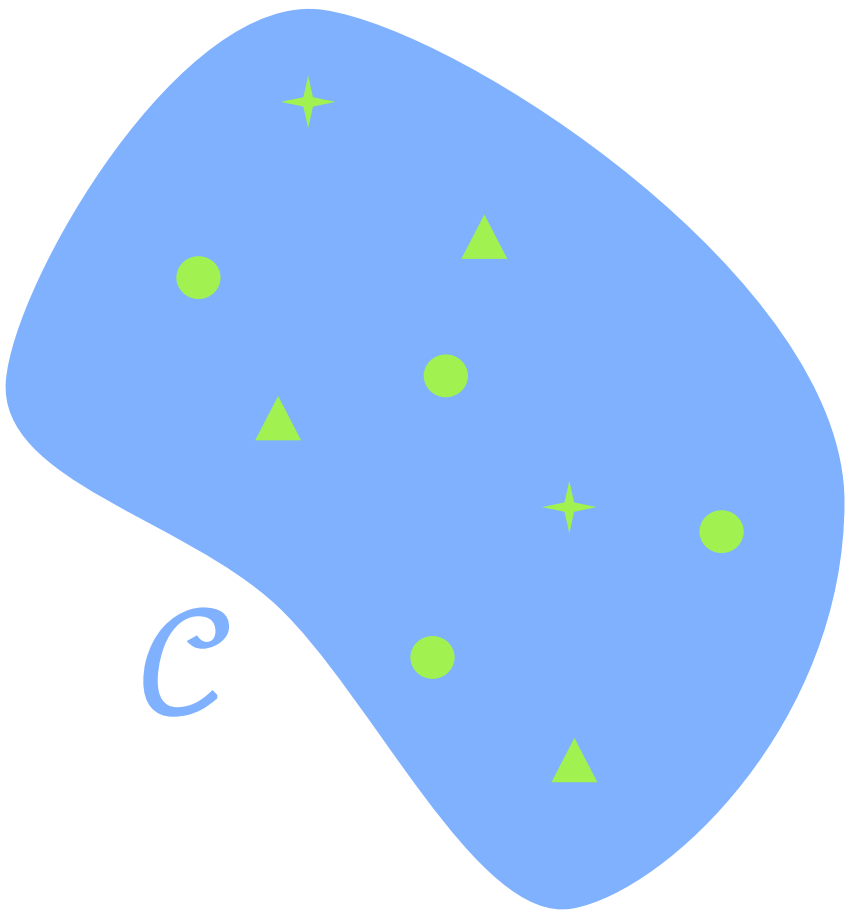
We propose a mathematical framework for a unification of the distributional theory of meaning in terms of vector space models, and a compositional theory for grammatical types, for which we rely on the algebra of Pregroups, introduced by Lambek. This mathematical framework enables us to compute the meaning of a well-typed sentence from the meanings of its constituents. Concretely, the type reductions of Pregroups are ‘lifted’ to morphisms in a category, a procedure that transforms meanings of constituents into a meaning of the (well-typed) whole. Importantly, meanings of whole sentences live in a single space, independent of the grammatical structure of the sentence. Hence the inner-product can be used to compare meanings of arbitrary sentences, as it is for comparing the meanings of words in the distributional model. The mathematical structure we employ admits a purely diagrammatic calculus which exposes how the information flows between the words in a sentence in order to make up the meaning of the whole sentence. A variation of our ‘categorical model’ which involves constraining the scalars of the vector spaces to the semiring of Booleans results in a Montague-style Boolean-valued semantics.

### 1 Introduction

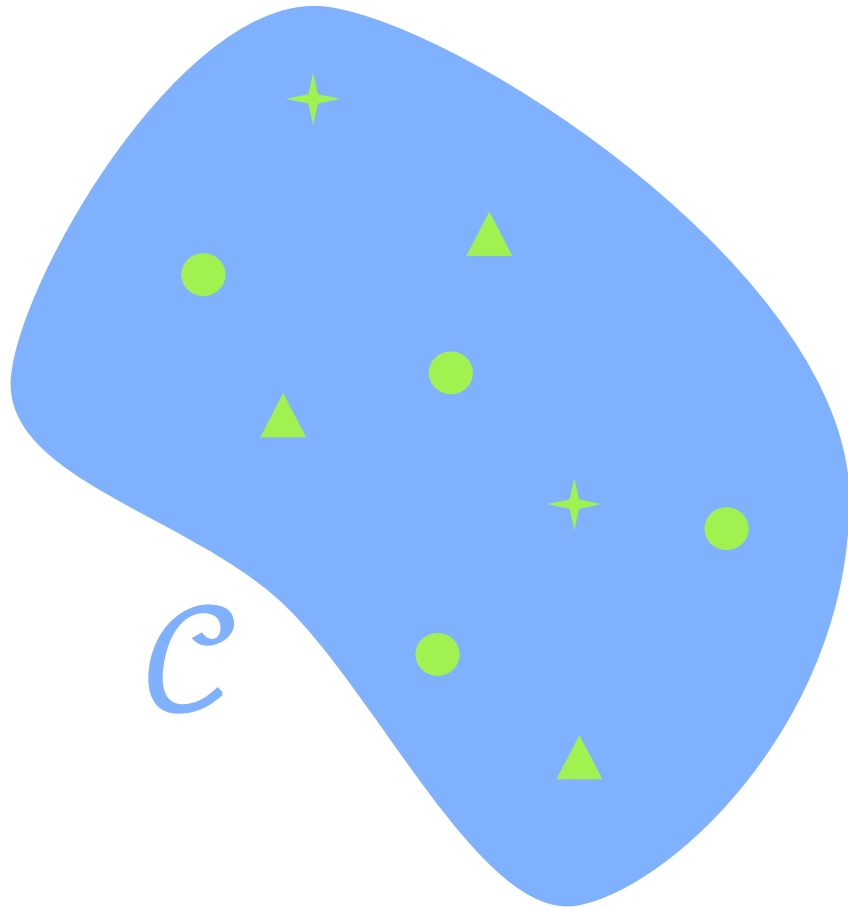
The symbolic [13] and distributional [36] theories of meaning are somewhat orthogonal with competing pros and cons: the former is compositional but only qualitative, the latter is non-compositional but quantitative. For a discussion of these two competing paradigms in Natural Language Processing see [15]. Following [39] in the context of Cognitive Science, where a similar problem exists between the

# Categories

Category:

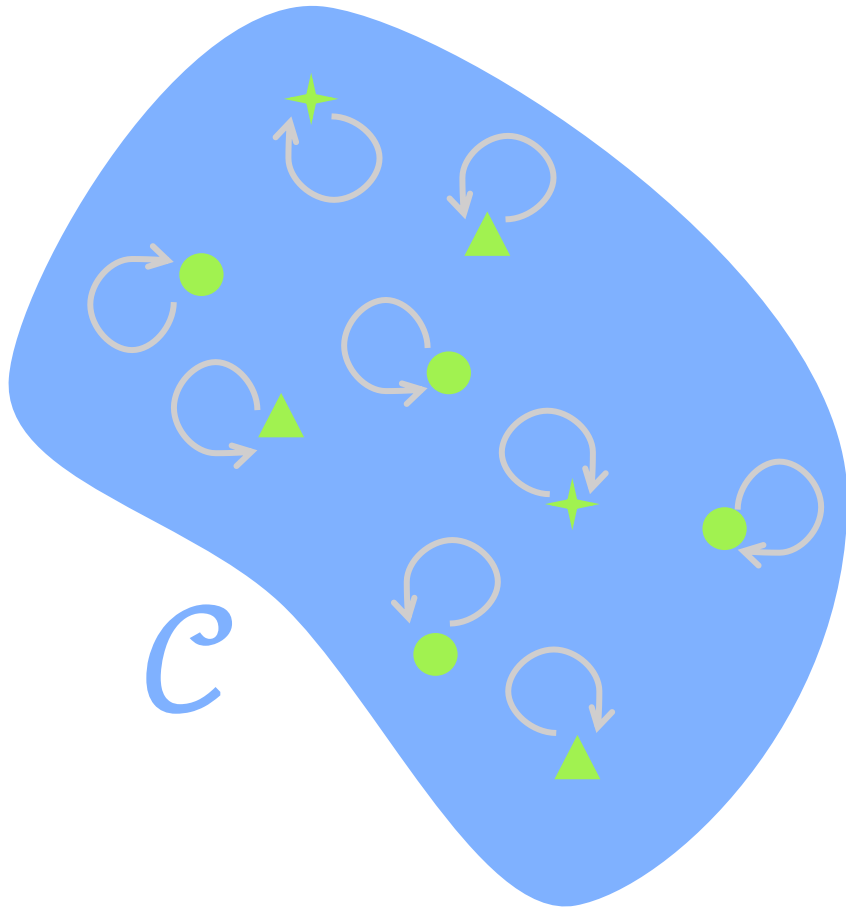


Category:



Objects: ✦, ▲, ●

Category:

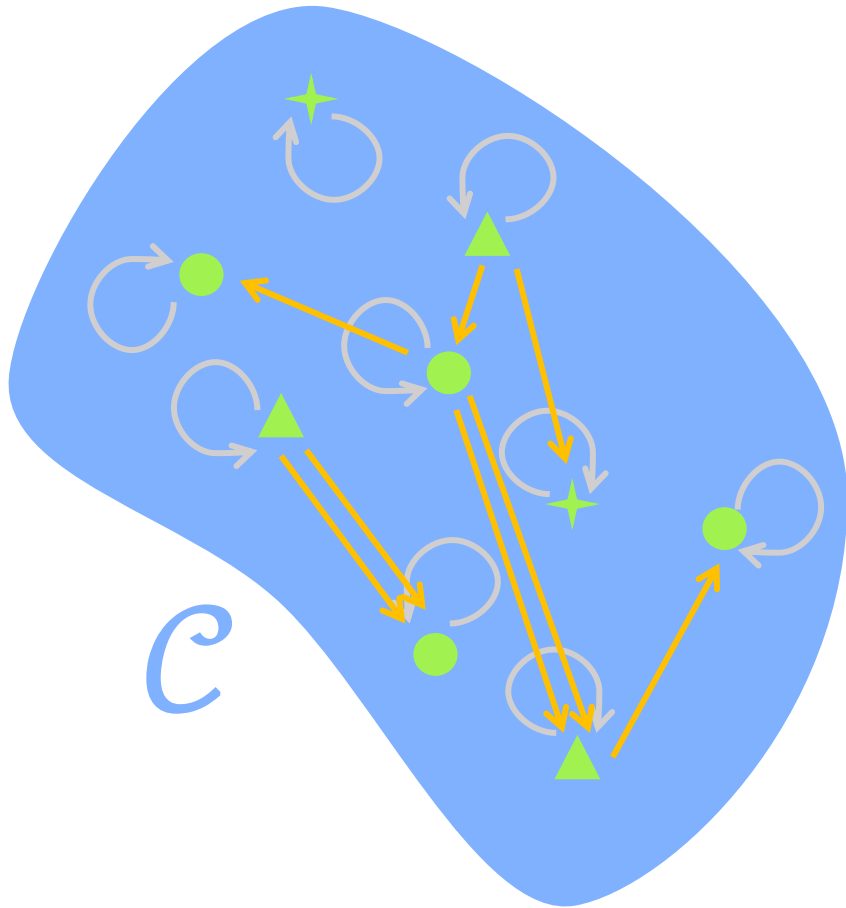


Objects:  $\star, \triangle, \bullet$

Morphisms:

- Identity Morphism  
 $id: c \mapsto c \in \mathcal{C}$

Category:



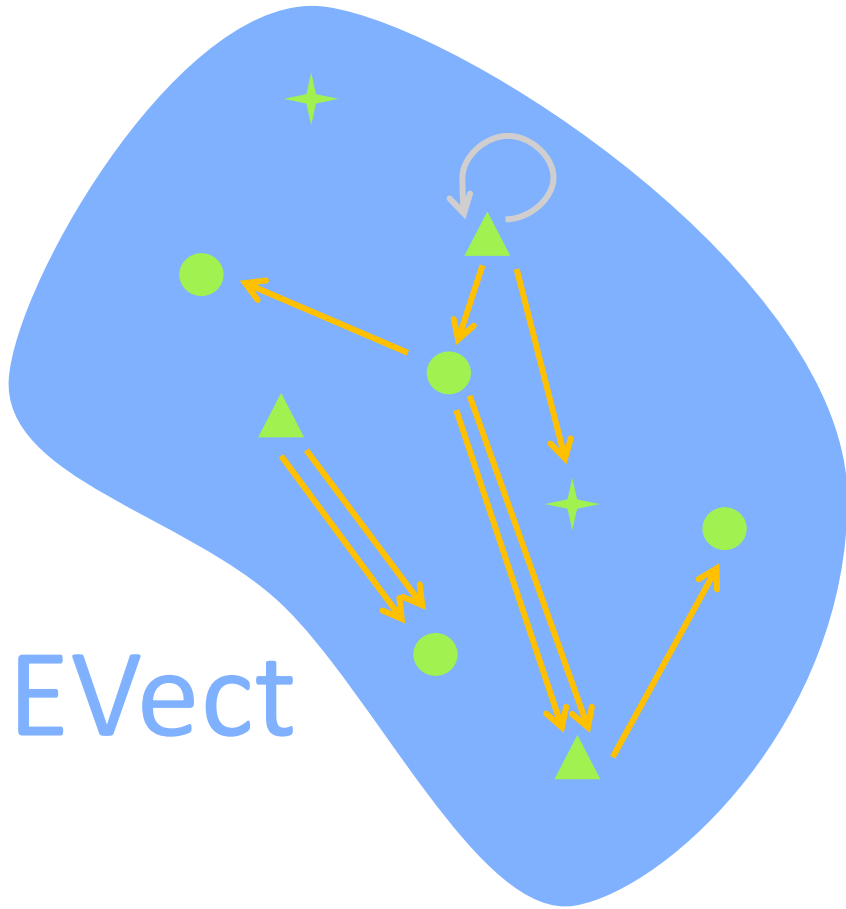
Objects:  $\star, \triangle, \bullet$

Morphisms:

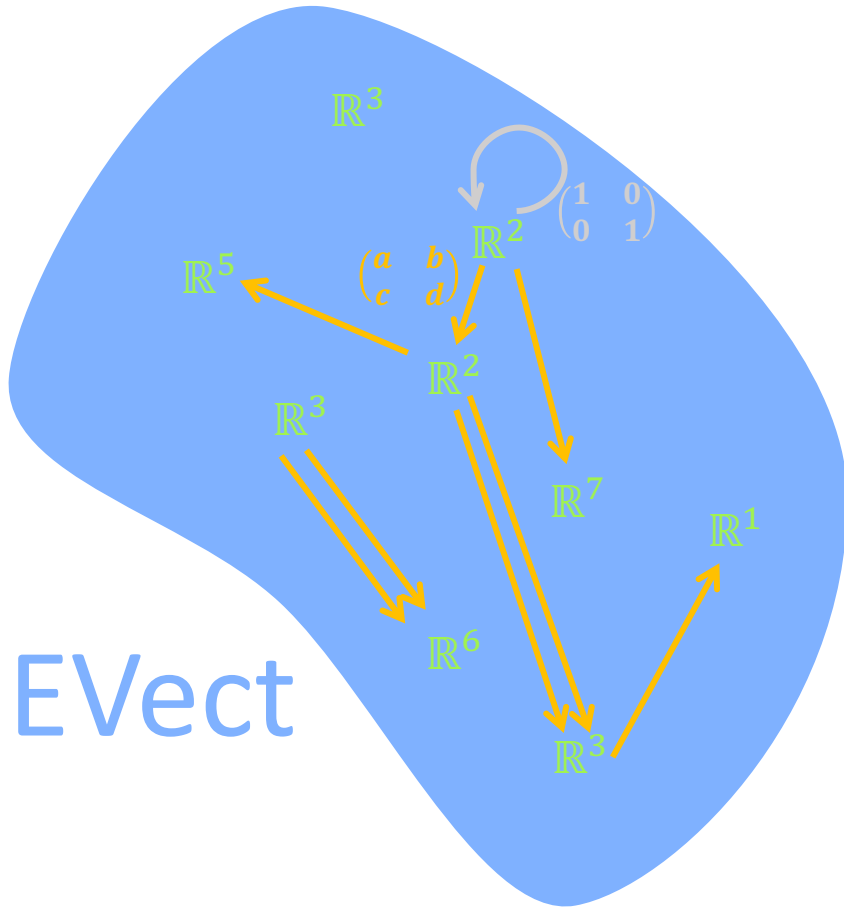
- Identity Morphism  
 $id: c \mapsto c \in \mathcal{C}$
- Morphisms between pairs of objects



# Example: EVect



# Example: EVect

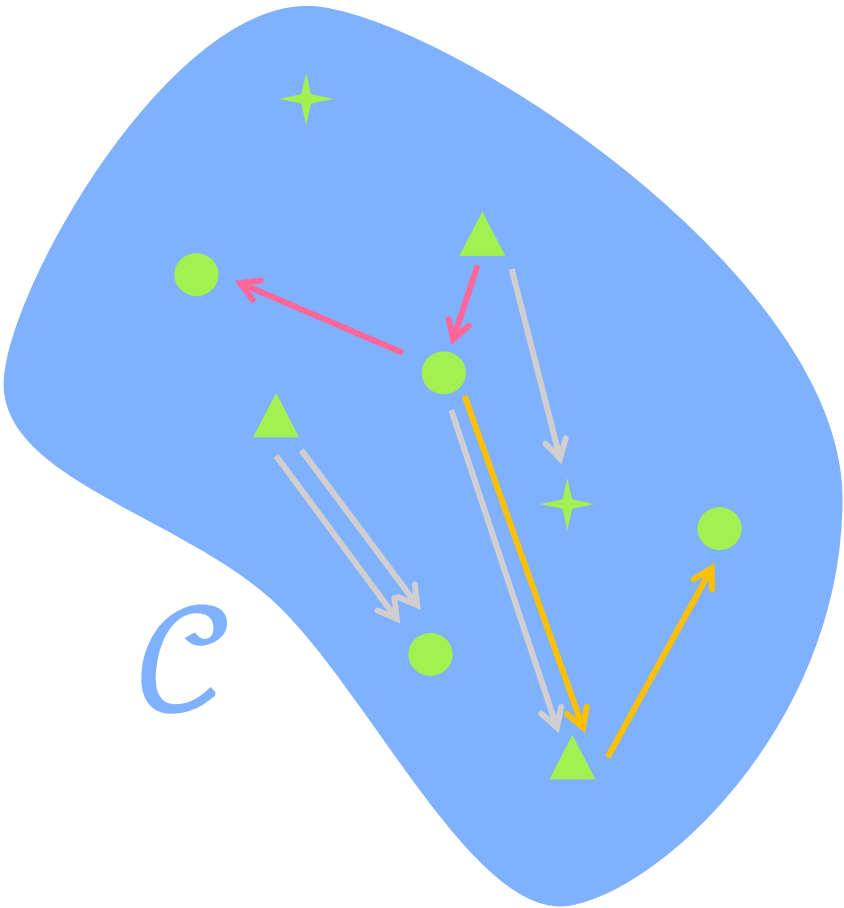


- Objects: Euclidian Vector Spaces ( $\mathbb{R}^n$ )
- Morphisms: matrices (Linear maps)

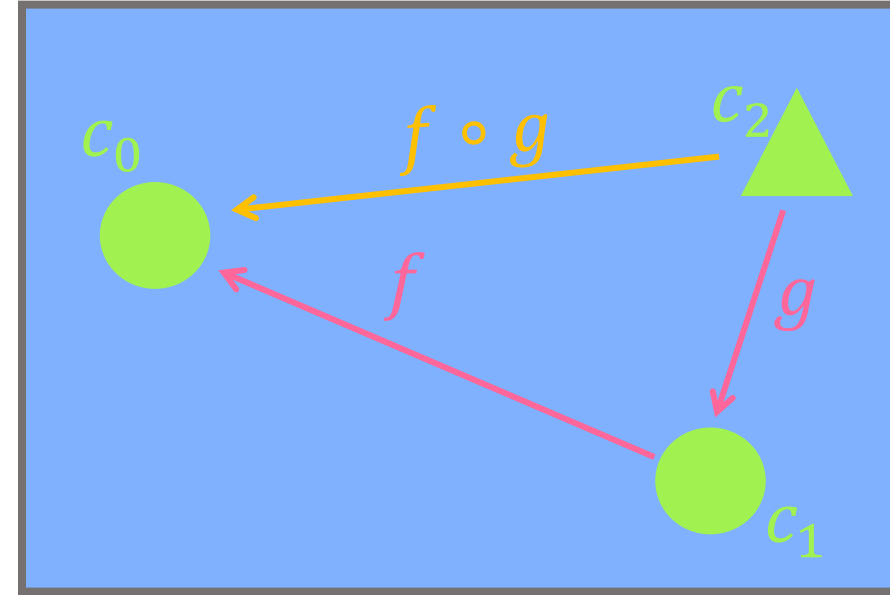
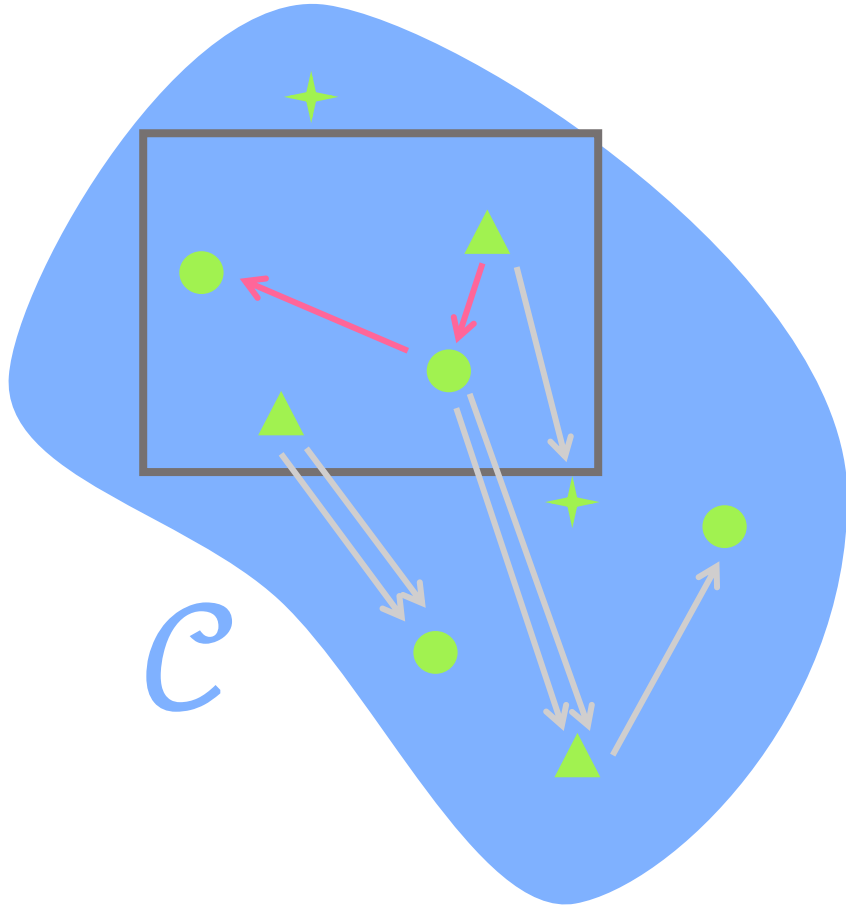
$$A = \begin{bmatrix} 4 & 6 \\ -3 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\mathbb{R}^2 \xrightarrow{A} \mathbb{R}^2$$

Category: Composition of Morphisms



# Category: Composition of Morphisms

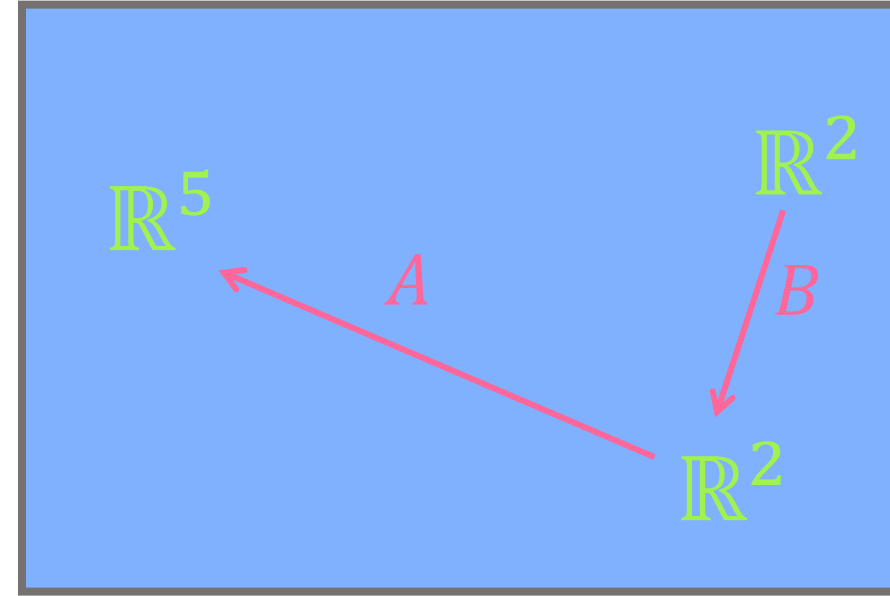
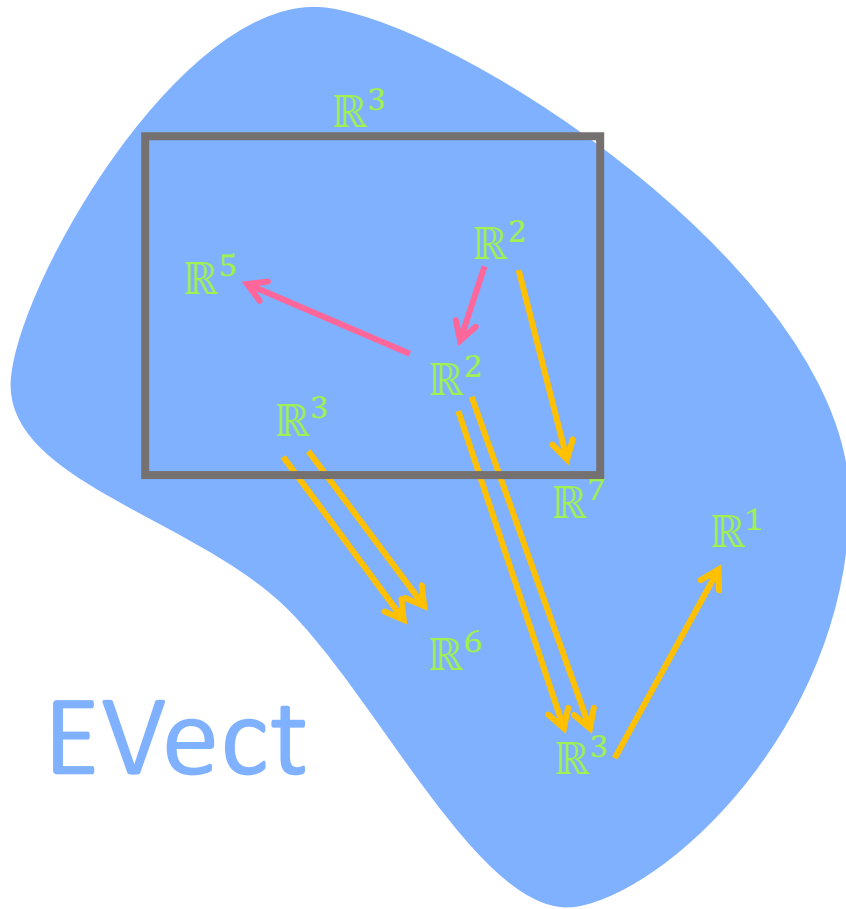


$$f: c_1 \mapsto c_0 \in \mathcal{C}$$

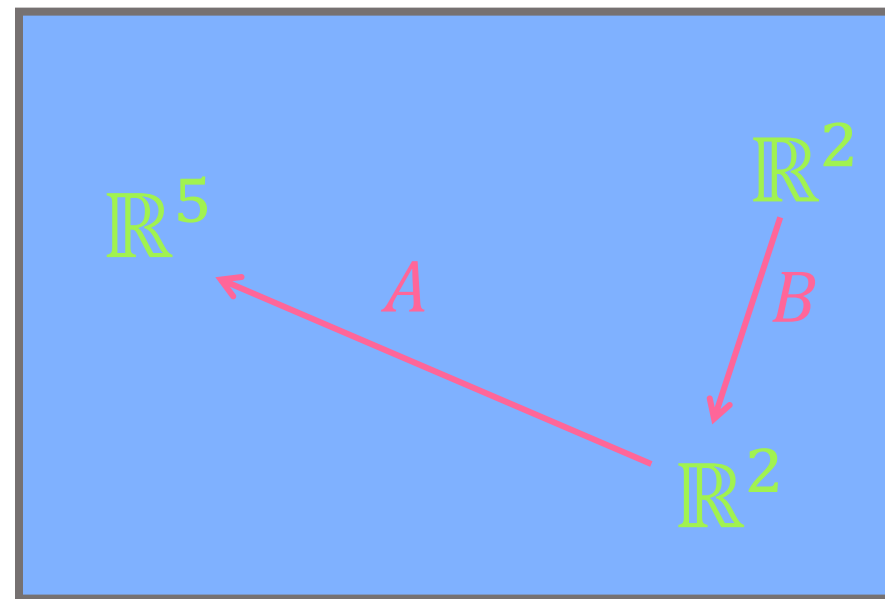
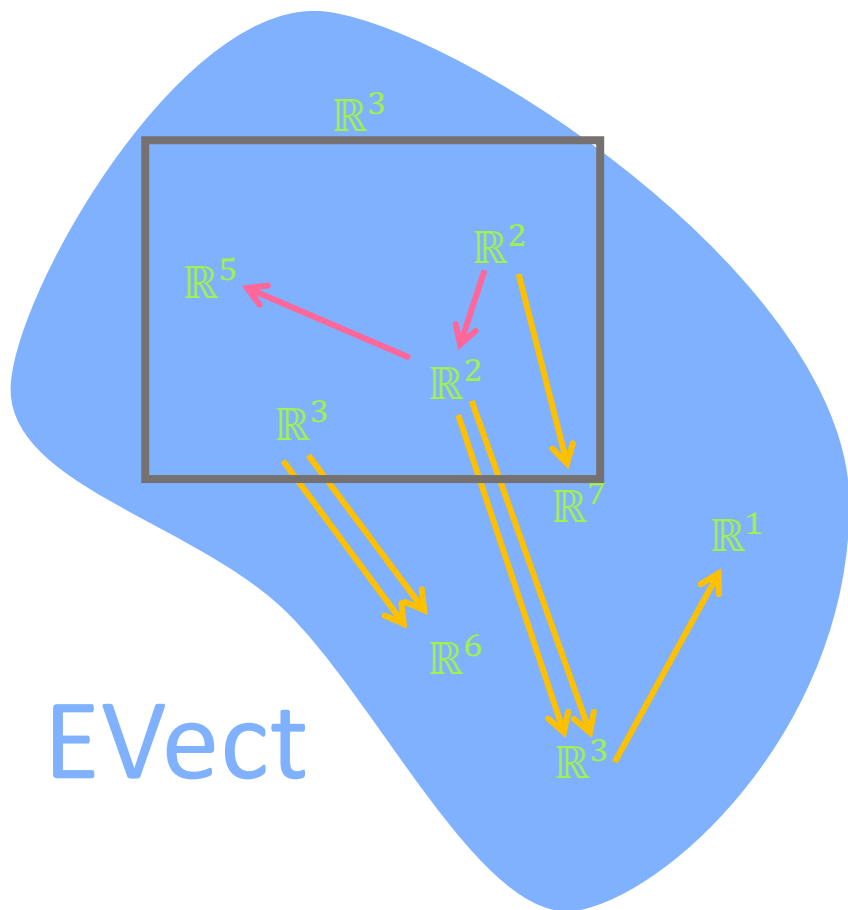
$$g: c_2 \mapsto c_1 \in \mathcal{C}$$

$$f \circ g: c_2 \mapsto c_0 \in \mathcal{C}$$

# Example: Matrix Multiplication

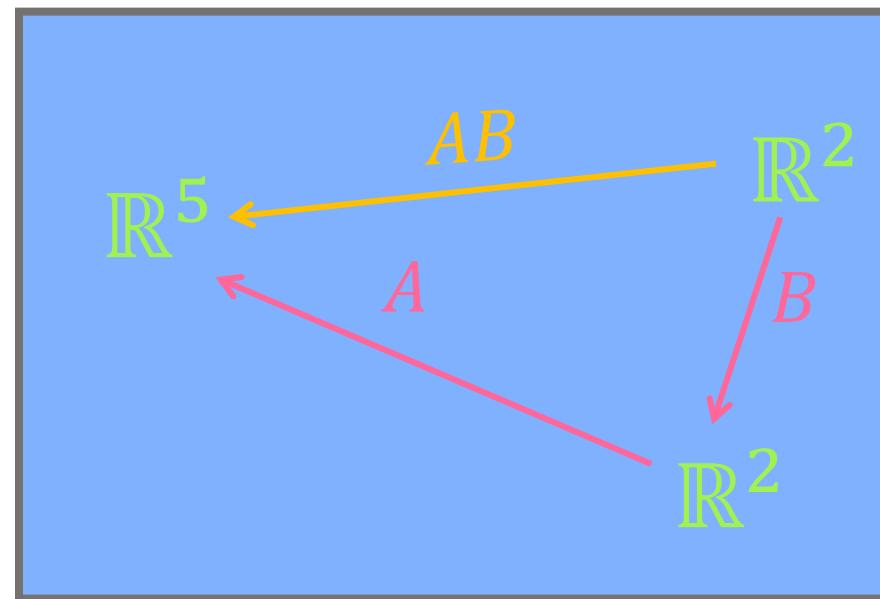
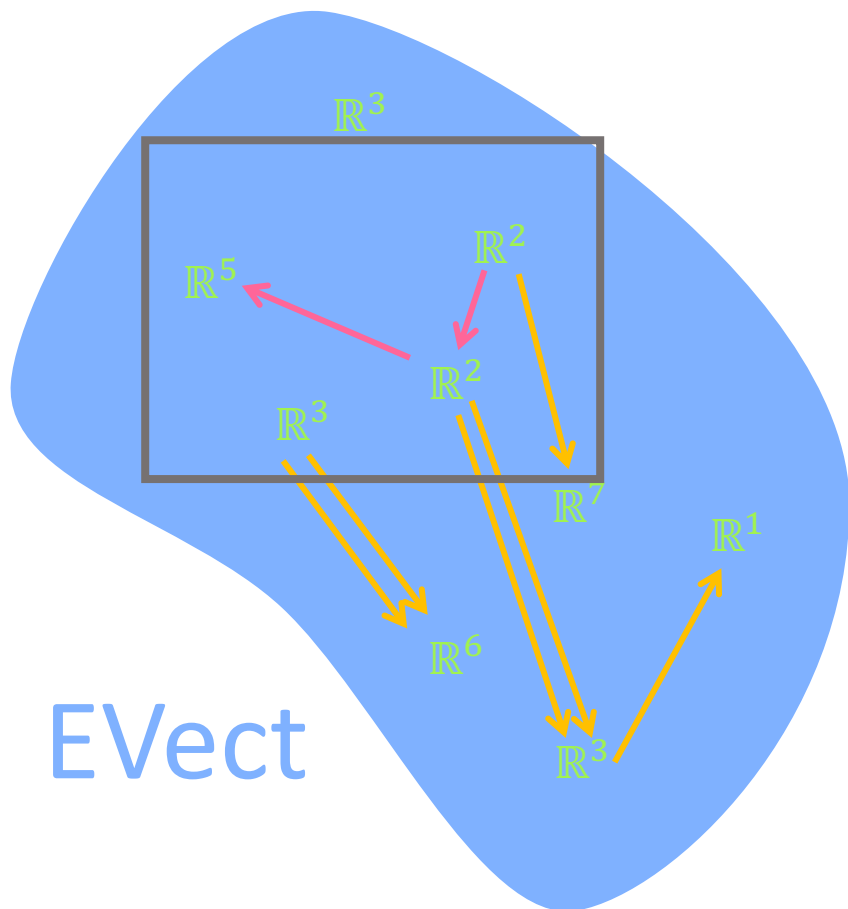


# Example: Matrix Multiplication



$$A \in \text{Mat}_{5 \times 2}$$
$$B \in \text{Mat}_{2 \times 2}$$

# Example: Matrix Multiplication



$$A \in \text{Mat}_{5 \times 2}$$

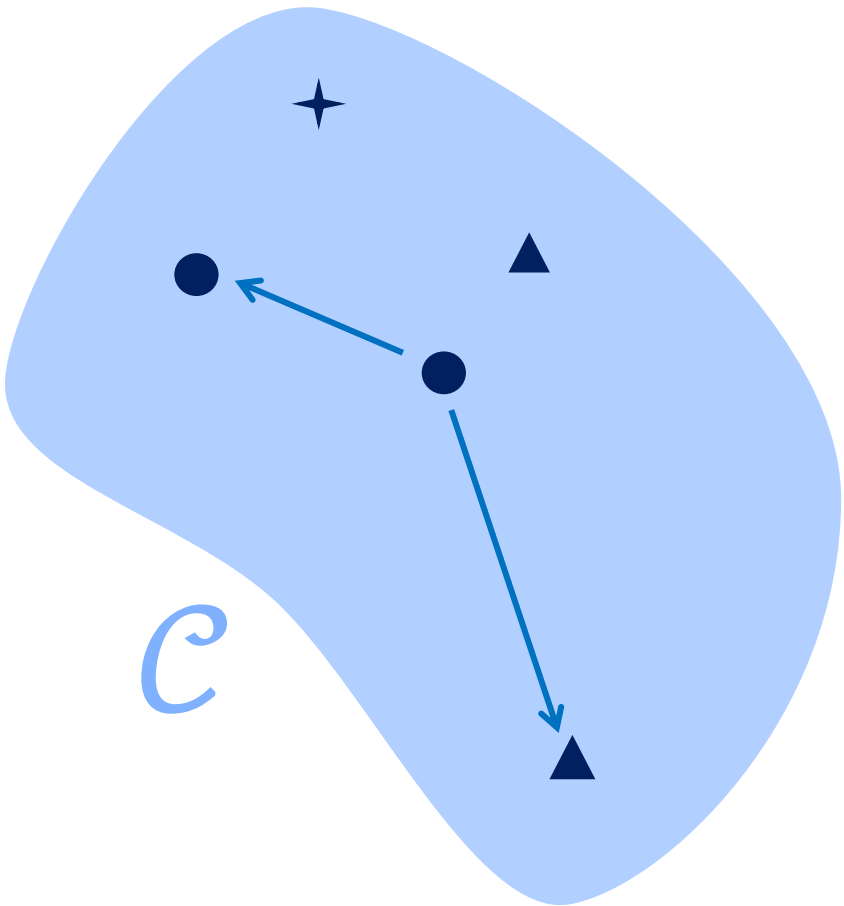
$$B \in \text{Mat}_{2 \times 2}$$

$$A \circ B = AB \in \text{Mat}_{5 \times 2}$$

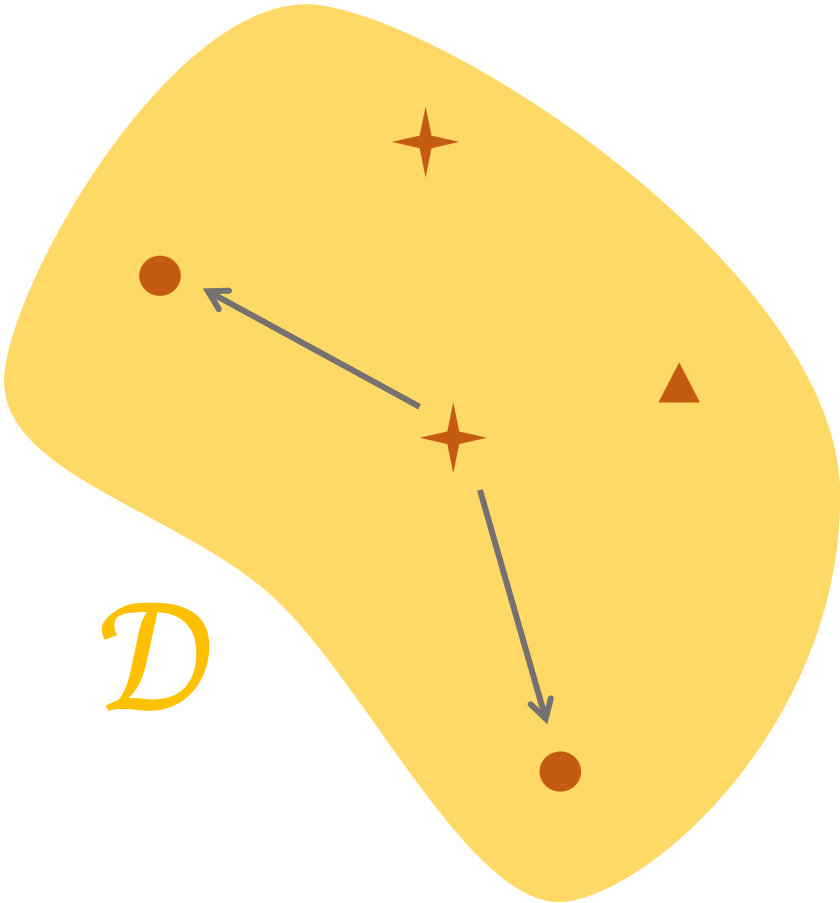
# Product Categories



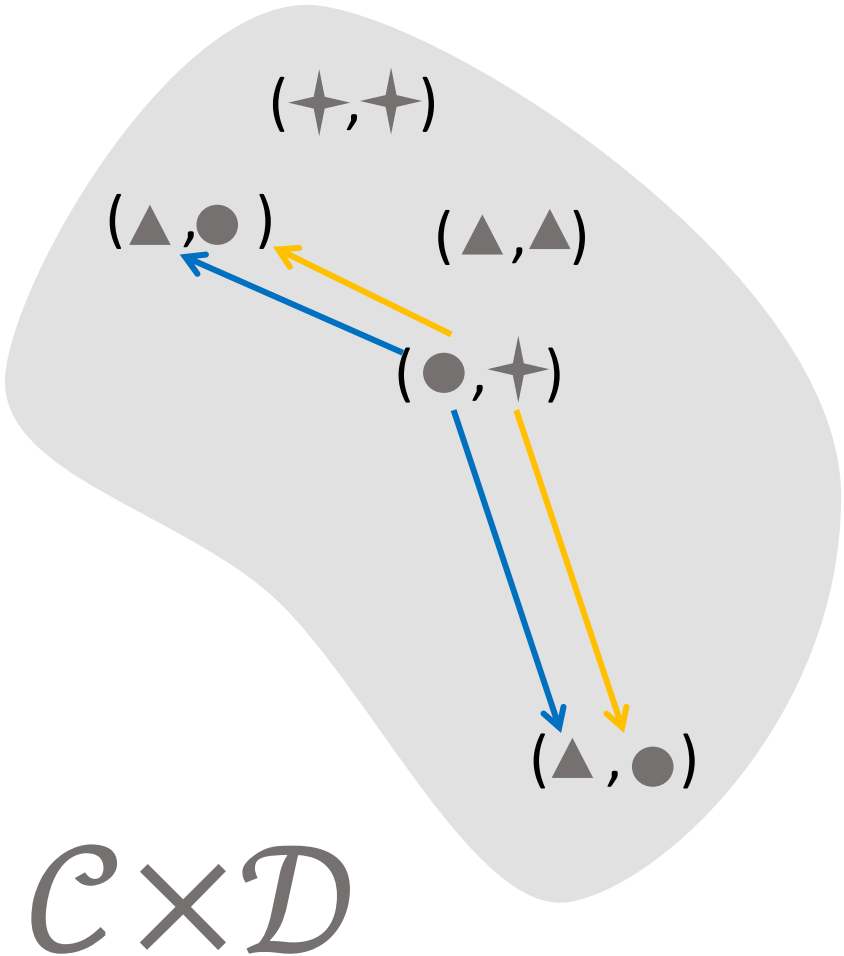
Product Category:



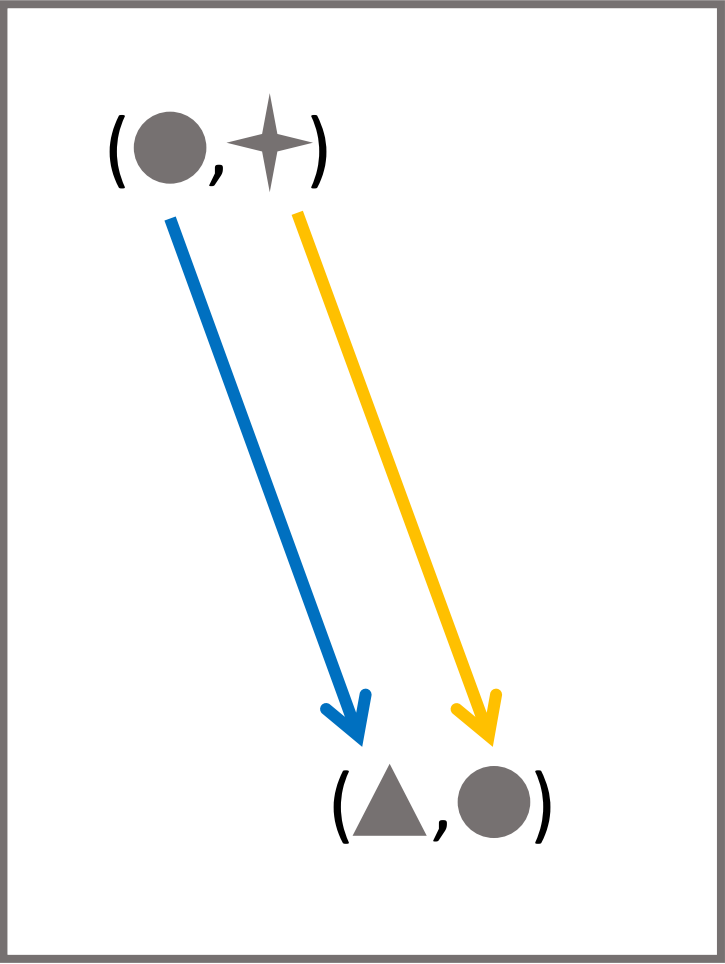
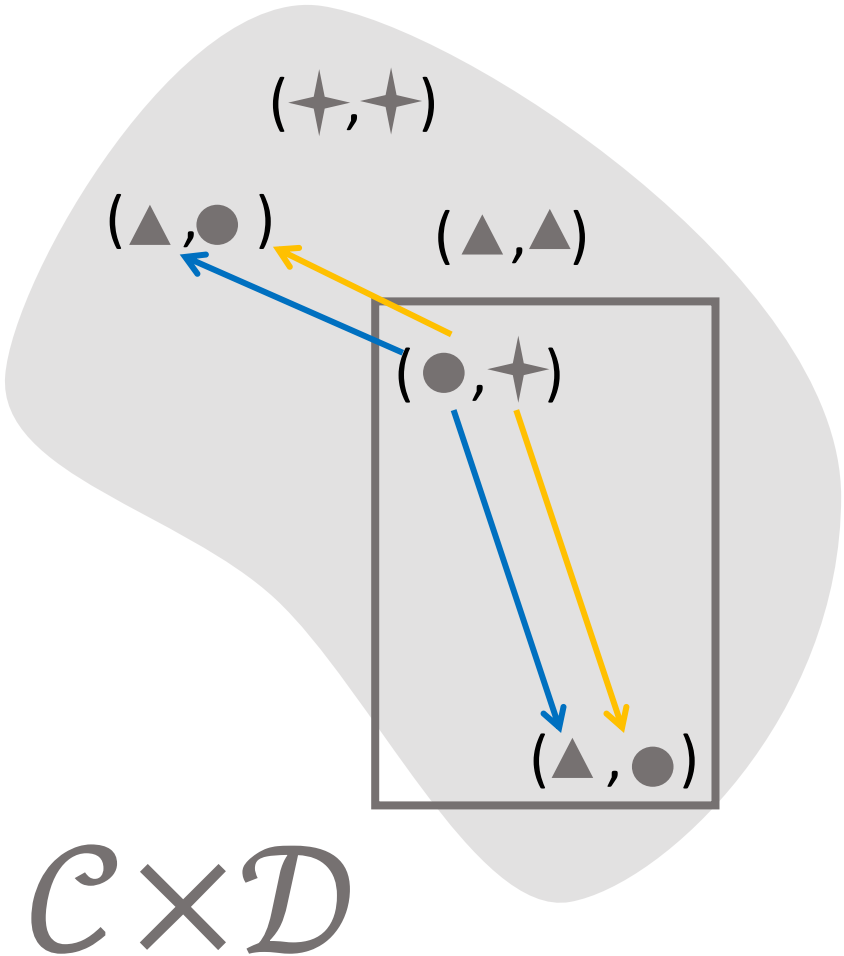
$\times$



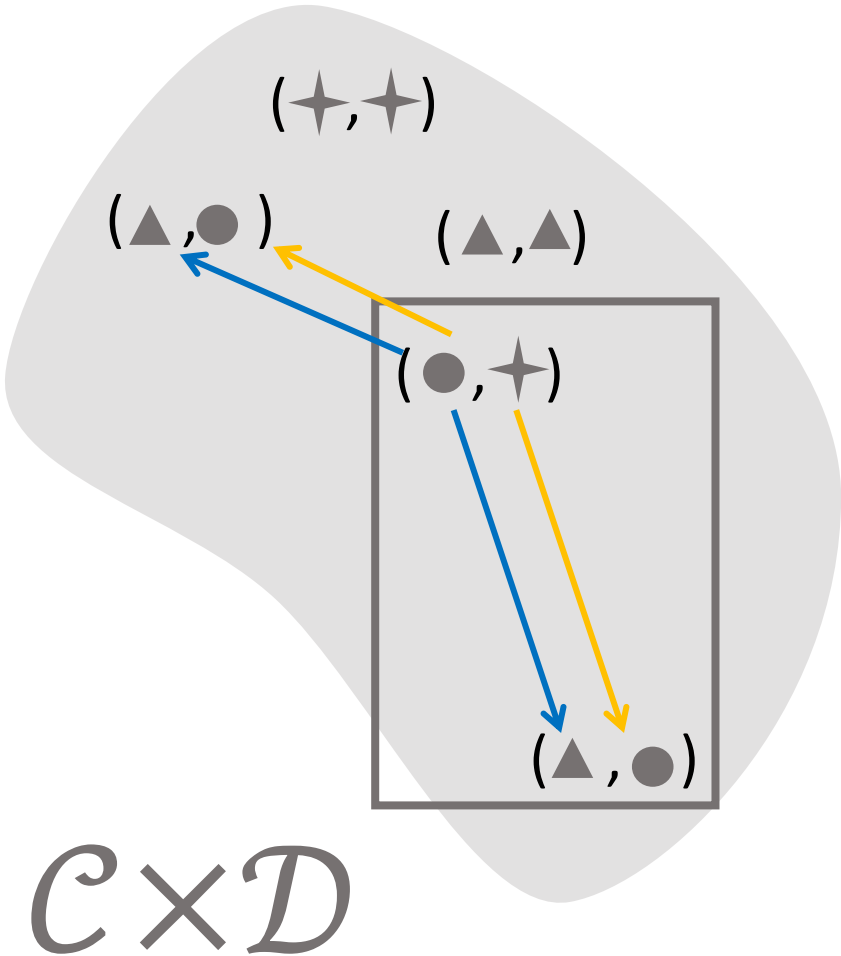
Product Category:



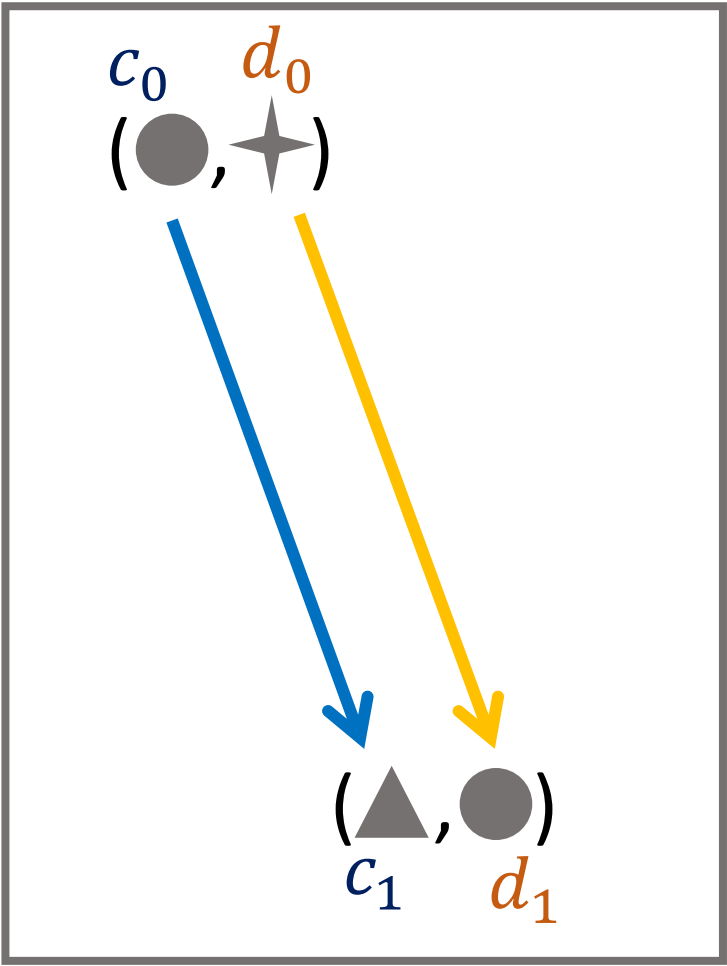
Product Category:



Product Category:



$\mathcal{C} \times \mathcal{D}$

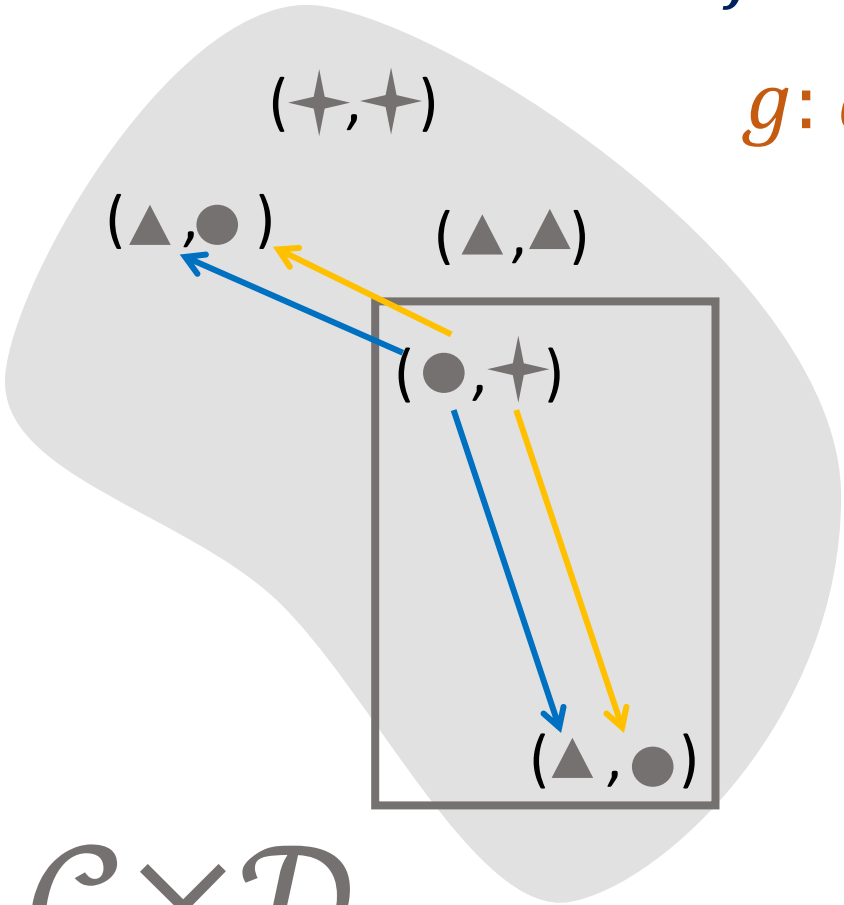


$(c_0, d_0) \in \mathcal{C} \times \mathcal{D}$  where  $c_0 \in \mathcal{C}$  and  $d_0 \in \mathcal{D}$

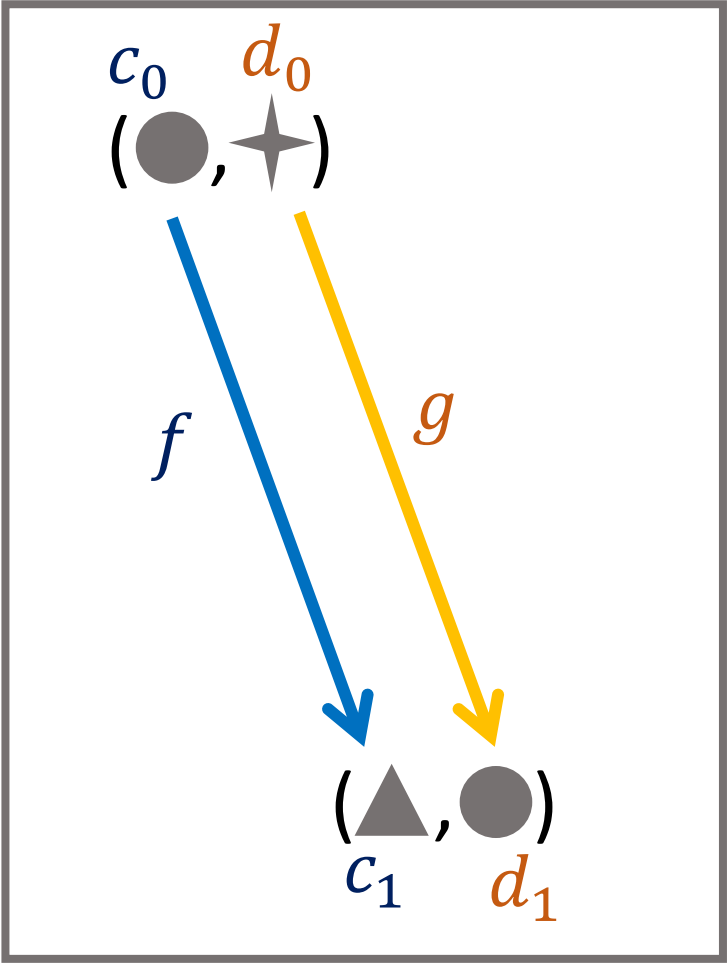
Product Category:

$$f: c_0 \mapsto c_1 \in \mathcal{C}$$

$$g: d_0 \mapsto d_1 \in \mathcal{D}$$



$\mathcal{C} \times \mathcal{D}$

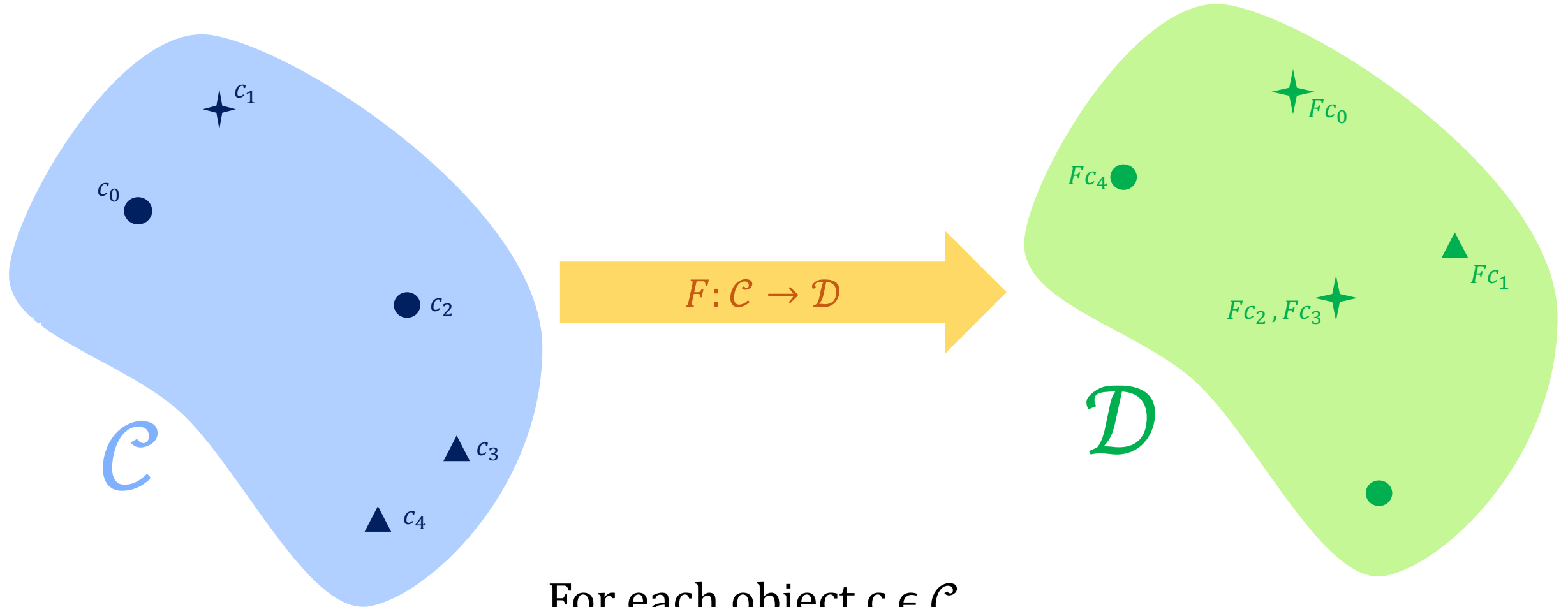


$$f \times g: (c_0, d_0) \mapsto (c_1, d_1) \in \mathcal{C} \times \mathcal{D}$$

# Functor

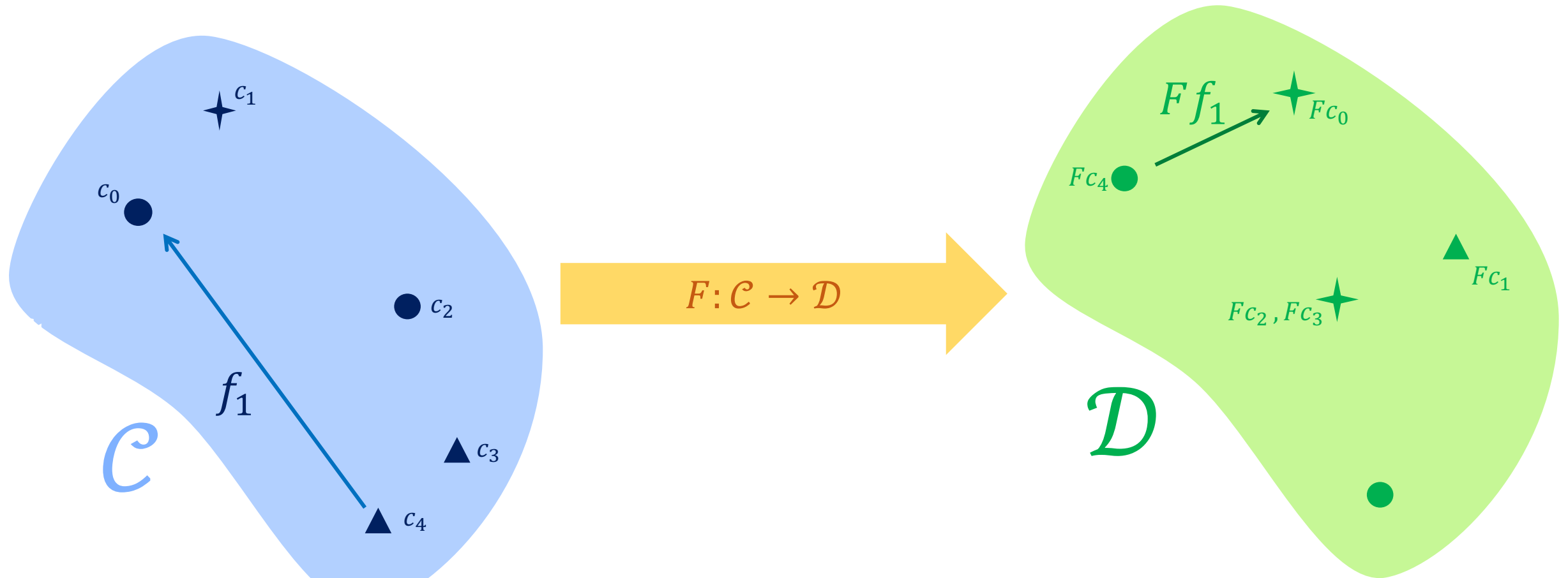
Assignments/Map between categories

Functor : Category  $\rightarrow$  Category



For each object  $c \in \mathcal{C}$ ,  
there is a corresponding  $Fc \in \mathcal{D}$

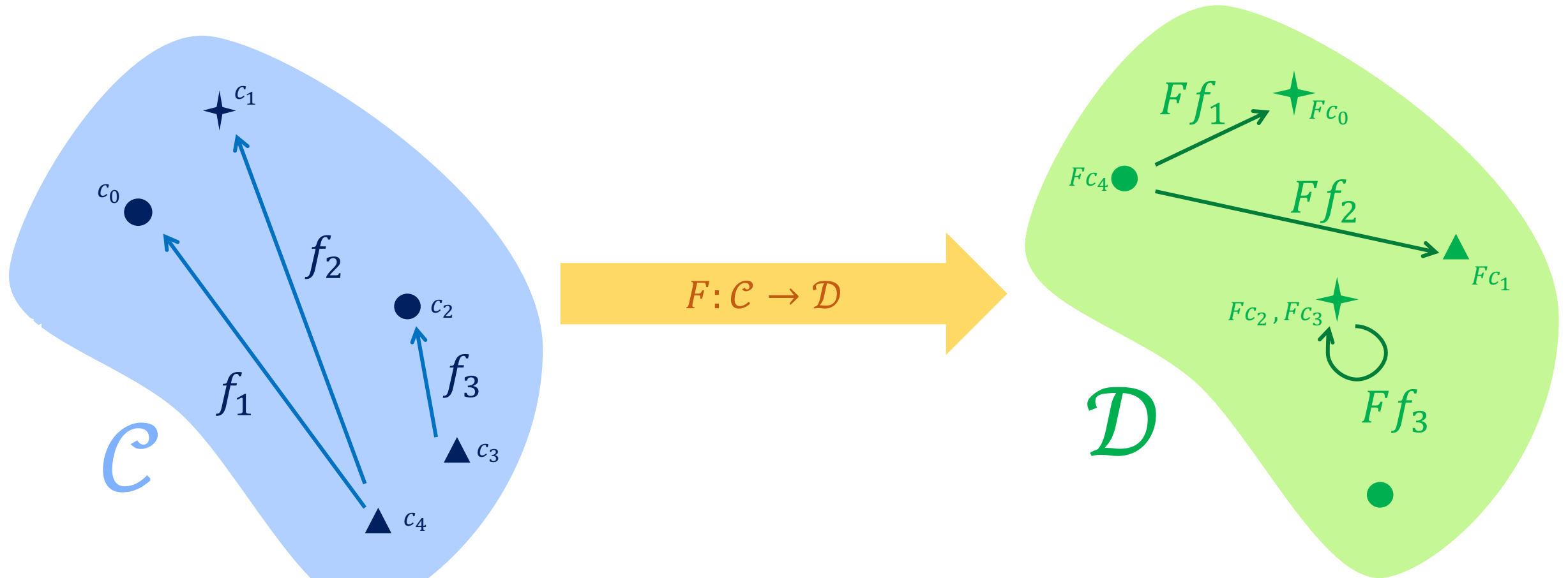
Functor : Category  $\rightarrow$  Category



For each morphism  $f: c \rightarrow c' \in \mathcal{C}$ ,  
there is a corresponding morphism  
 $Ff: Fc \rightarrow Fc' \in \mathcal{D}$



Functor : Category  $\rightarrow$  Category

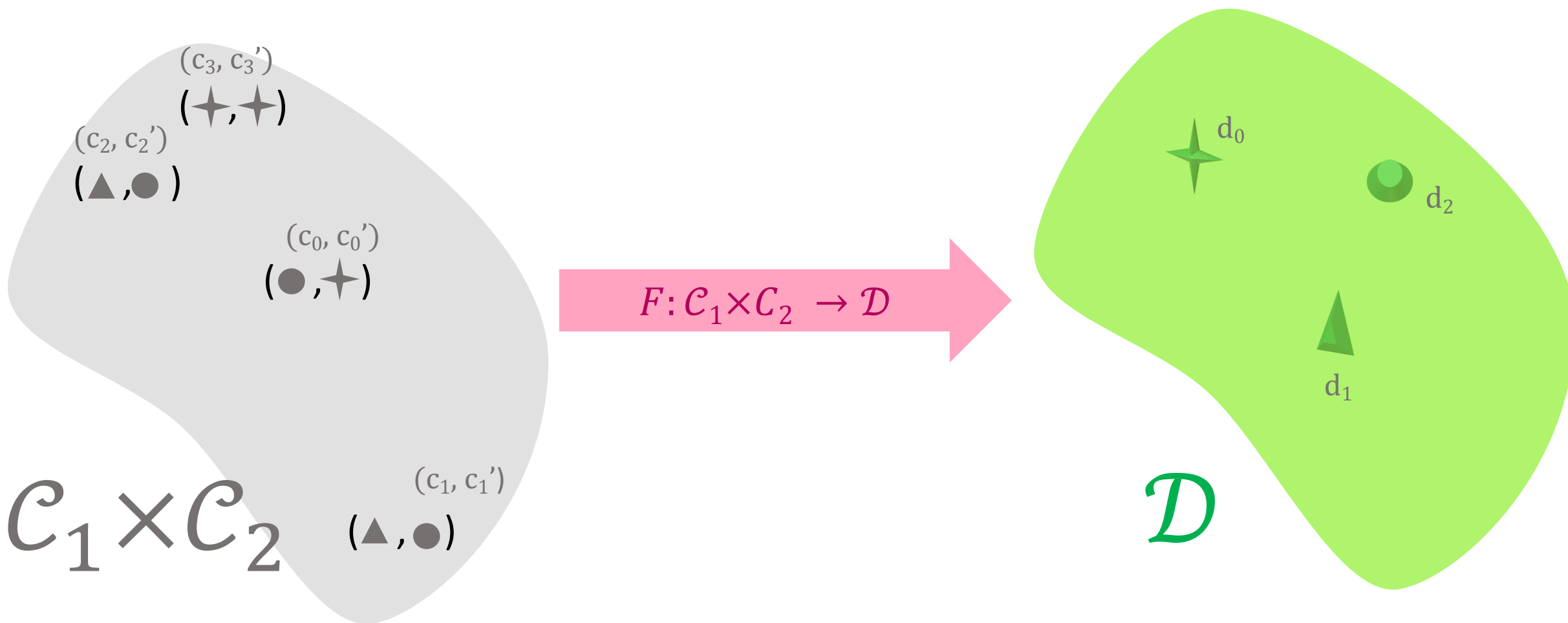


For each morphism  $f: c \rightarrow c' \in \mathcal{C}$ ,  
there is a corresponding morphism  
 $Ff: Fc \rightarrow Fc' \in \mathcal{D}$

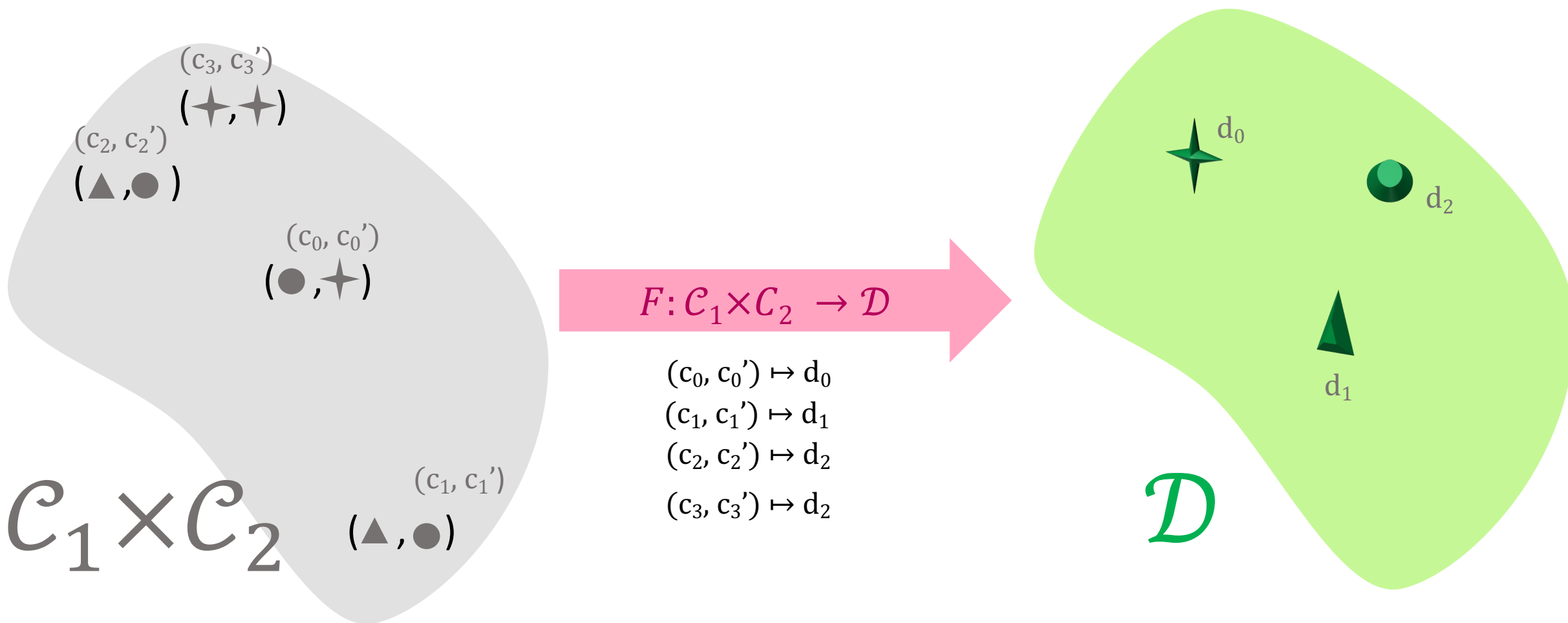
The most important functor in our story is the tensor product ( $\otimes$ ), which is a

bifunctor.

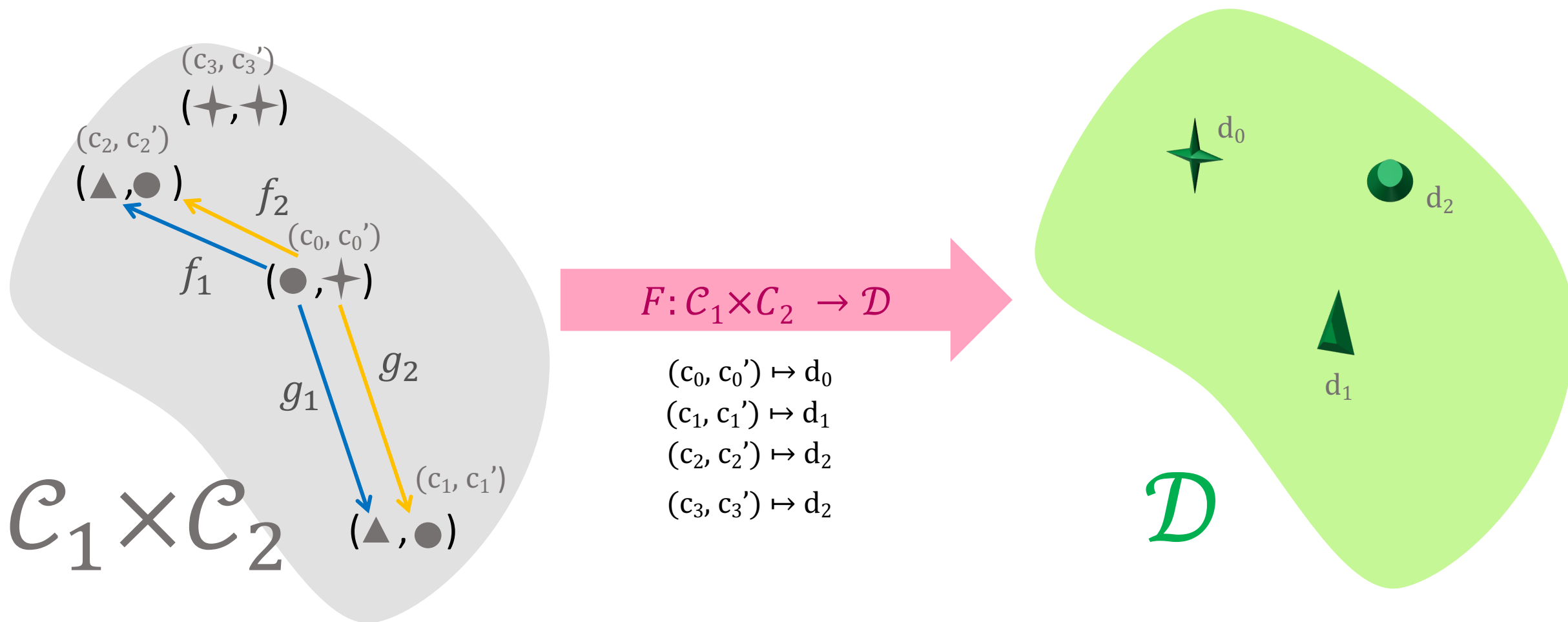
Bifunctor : Product Category  $\rightarrow$  Category



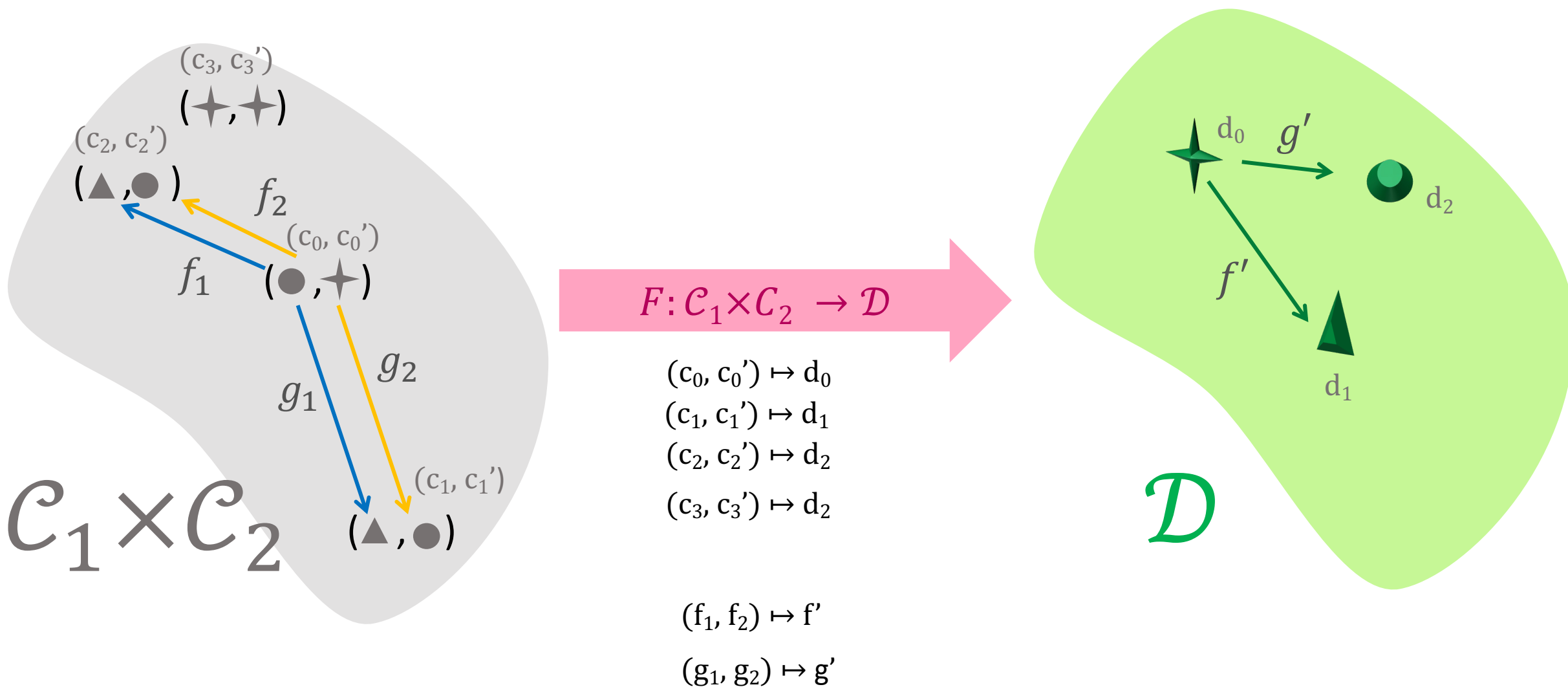
Bifunctor : Product Category  $\rightarrow$  Category



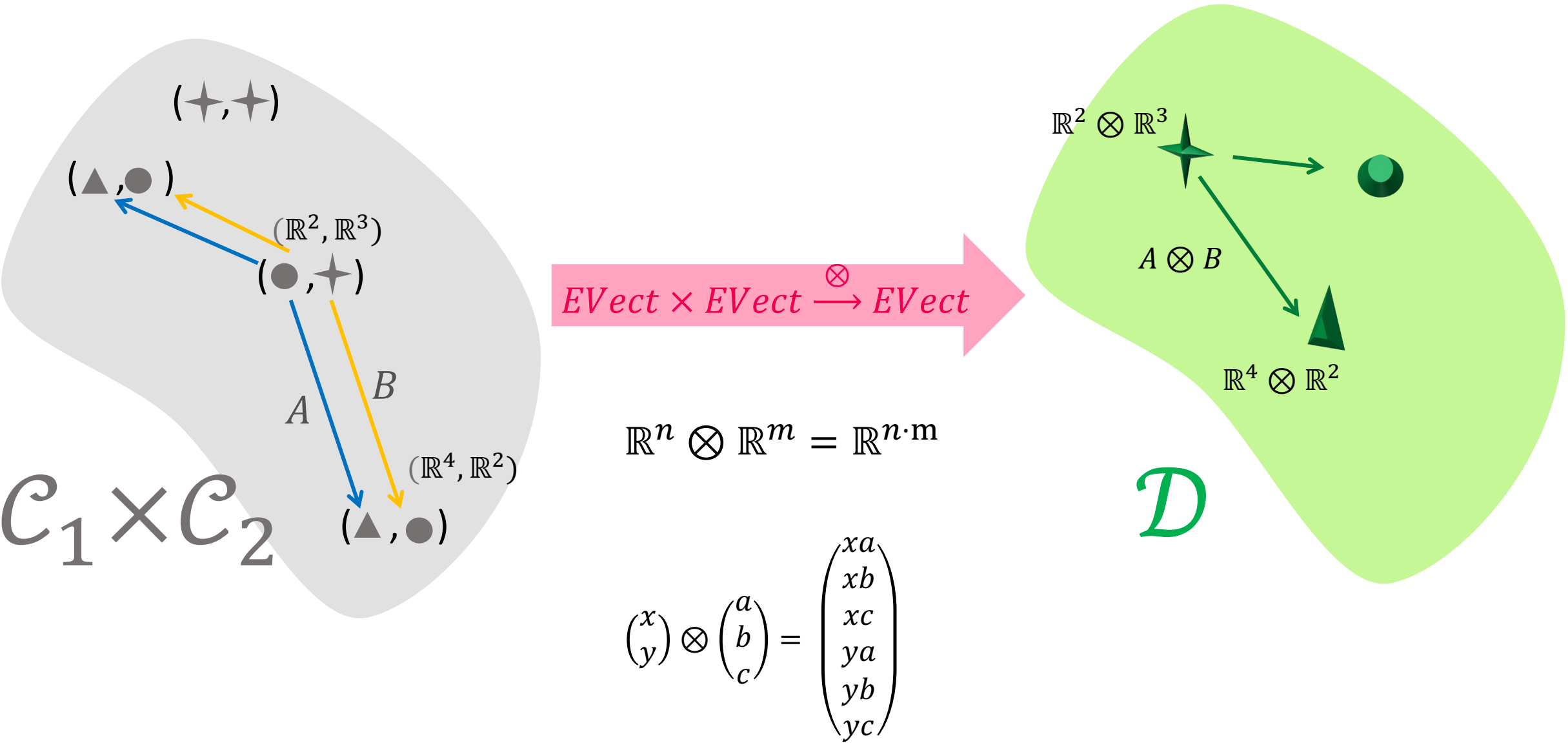
Bifunctor : Product Category  $\rightarrow$  Category



# Bifunctor : Product Category $\rightarrow$ Category



# Tensor Product in EVect



# Structures on Categories

Monoidal structures



A monoidal category  $(\mathcal{C}, \otimes, 1_{\mathcal{C}})$  consist of a category  $\mathcal{C}$  with the following structures:

- Bifunctor  $-\otimes -: \mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$
- Unit object  $1_{\mathcal{C}} \in \mathcal{C}$

For  $a, b, c \in \mathcal{C}$ , there exists natural isomorphisms

- Associator  $\alpha_c: (a \otimes b) \otimes c \xrightarrow{\sim} a \otimes (b \otimes c)$
- Unitor  $u_c: a \xrightarrow{\sim} a \otimes 1_{\mathcal{C}}$

A monoidal category  $(\mathcal{C}, \otimes, 1_{\mathcal{C}})$  consist of a category  $\mathcal{C}$  with the following structures:

- Bifunctor  $- \otimes -: \mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$
- Unit object  $1_{\mathcal{C}} \in \mathcal{C}$

For  $a, b, c \in \mathcal{C}$ , there exists natural isomorphisms

- Associator  $\alpha_c: (a \otimes b) \otimes c \xrightarrow{\sim} a \otimes (b \otimes c)$
- Unitor  $u_c: a \xrightarrow{\sim} a \otimes 1_{\mathcal{C}}$

(example) in  $(FVect, \otimes, \mathbb{R}^1)$

$$V \xrightarrow{\sim} V \otimes \mathbb{R}^1$$

### FVect Tensor Product

$$\begin{array}{ccc} FVect \times FVect & \xrightarrow{\otimes} & FVect \\ V \times W & \mapsto & (V \otimes W) \end{array}$$

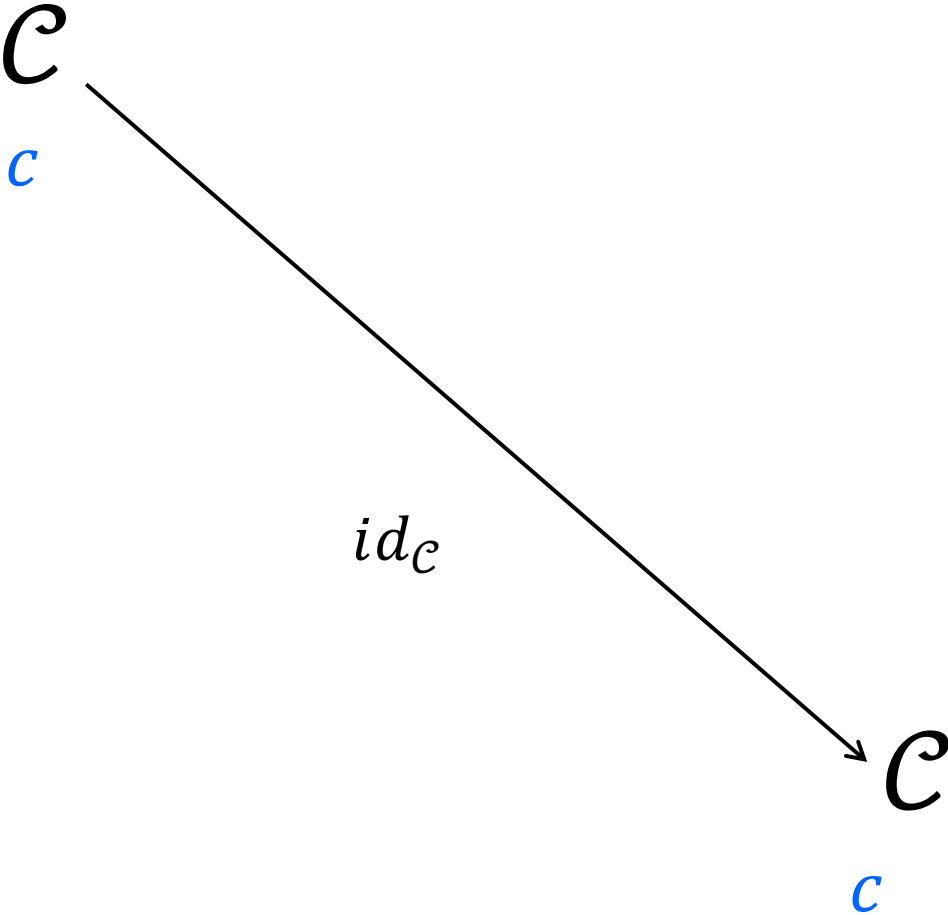
$\begin{array}{l} V, W \in FVect \\ V \simeq \mathbb{R}^n \text{ and } W \simeq \mathbb{R}^m \\ V \otimes W \simeq \mathbb{R}^{n \cdot m} \end{array}$
--

Unitor in a single Category

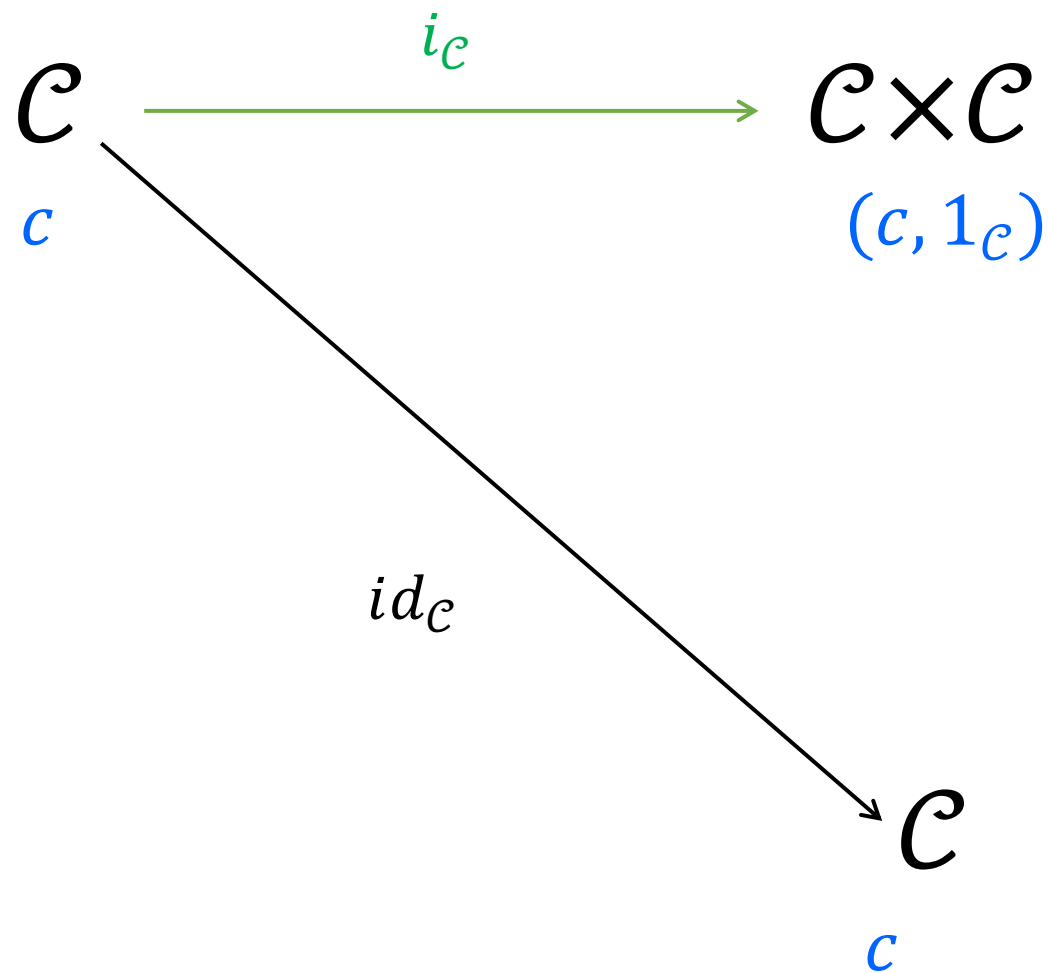
$\mathcal{C}$

$c$

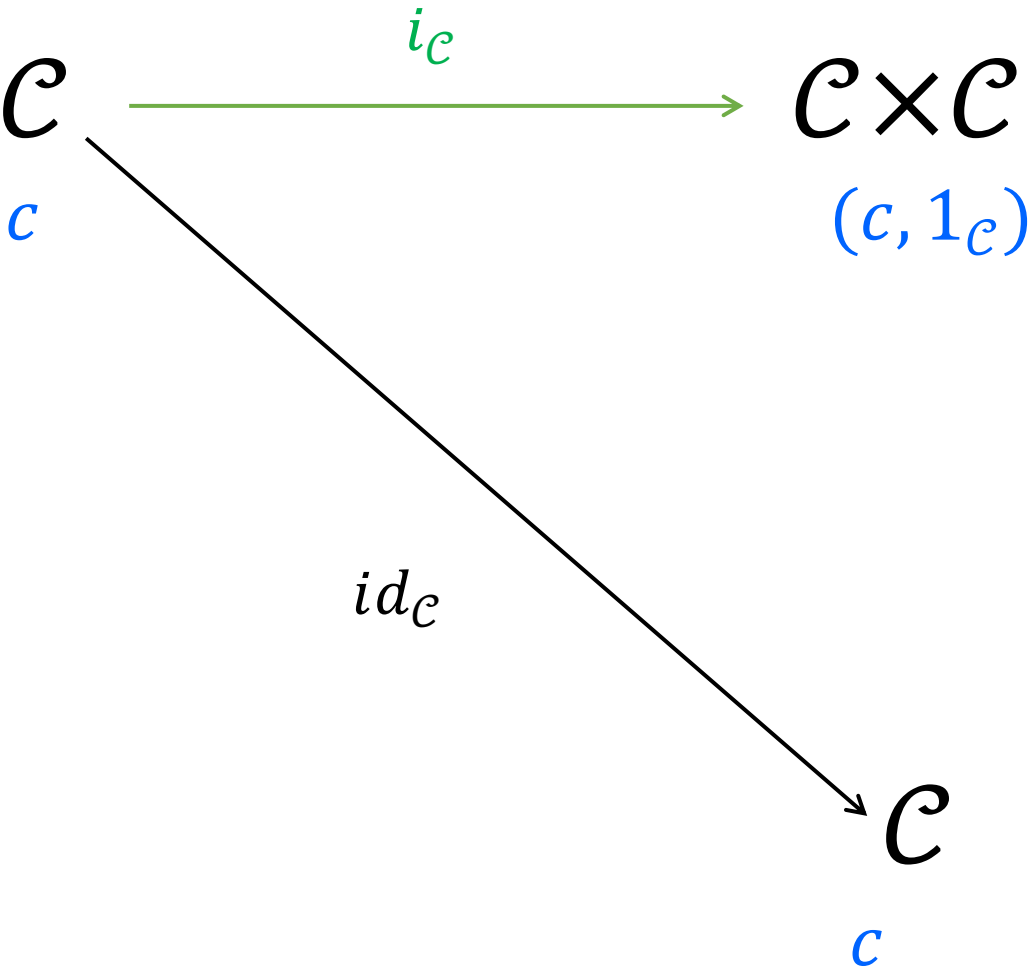
Unitor in a single Category



Unitor in a single Category



Unitor in a single Category



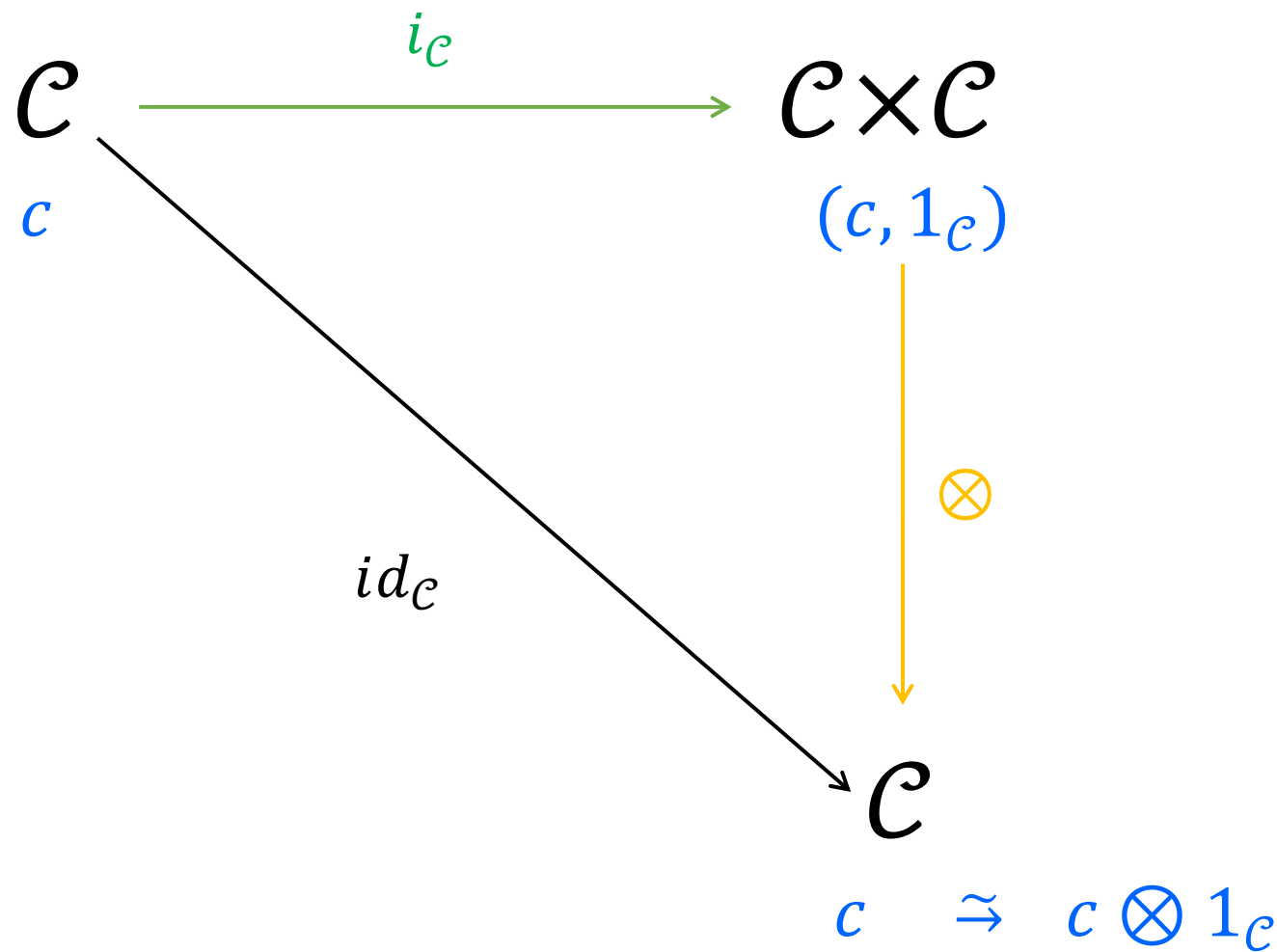
$i = (id \times cons) \circ \Delta$

Where

$\Delta: \mathcal{C} \rightarrow \mathcal{C} \times \mathcal{C}$   
 $c \mapsto (c, c)$

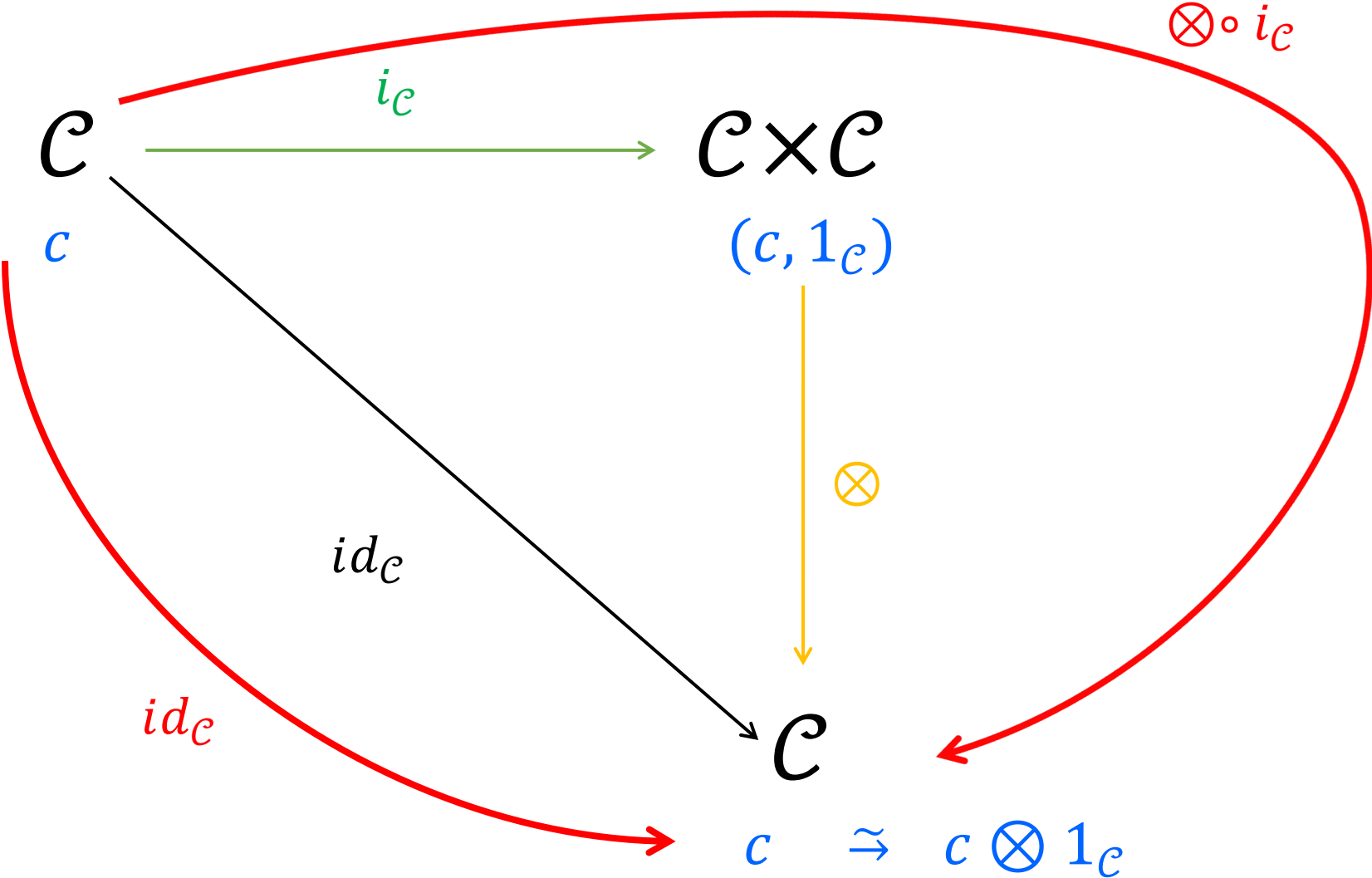
$cons: \mathcal{C} \rightarrow \mathcal{C}$   
 $c \mapsto 1_c$

Unitor in a single Category



$i_c := (id \times cons) \circ \Delta$

Unit in a single Category

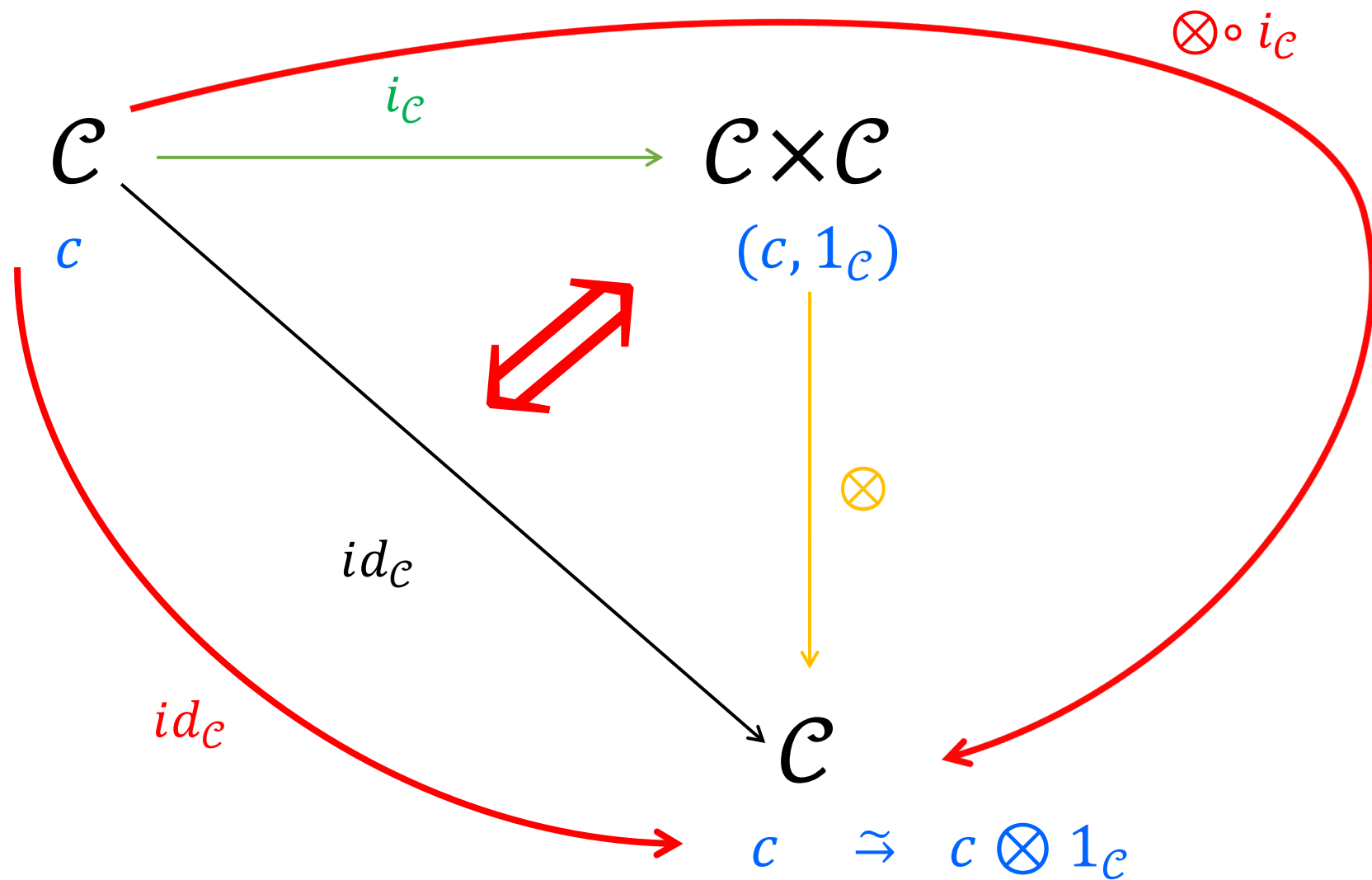


$i_c := (id \times cons) \circ \Delta$



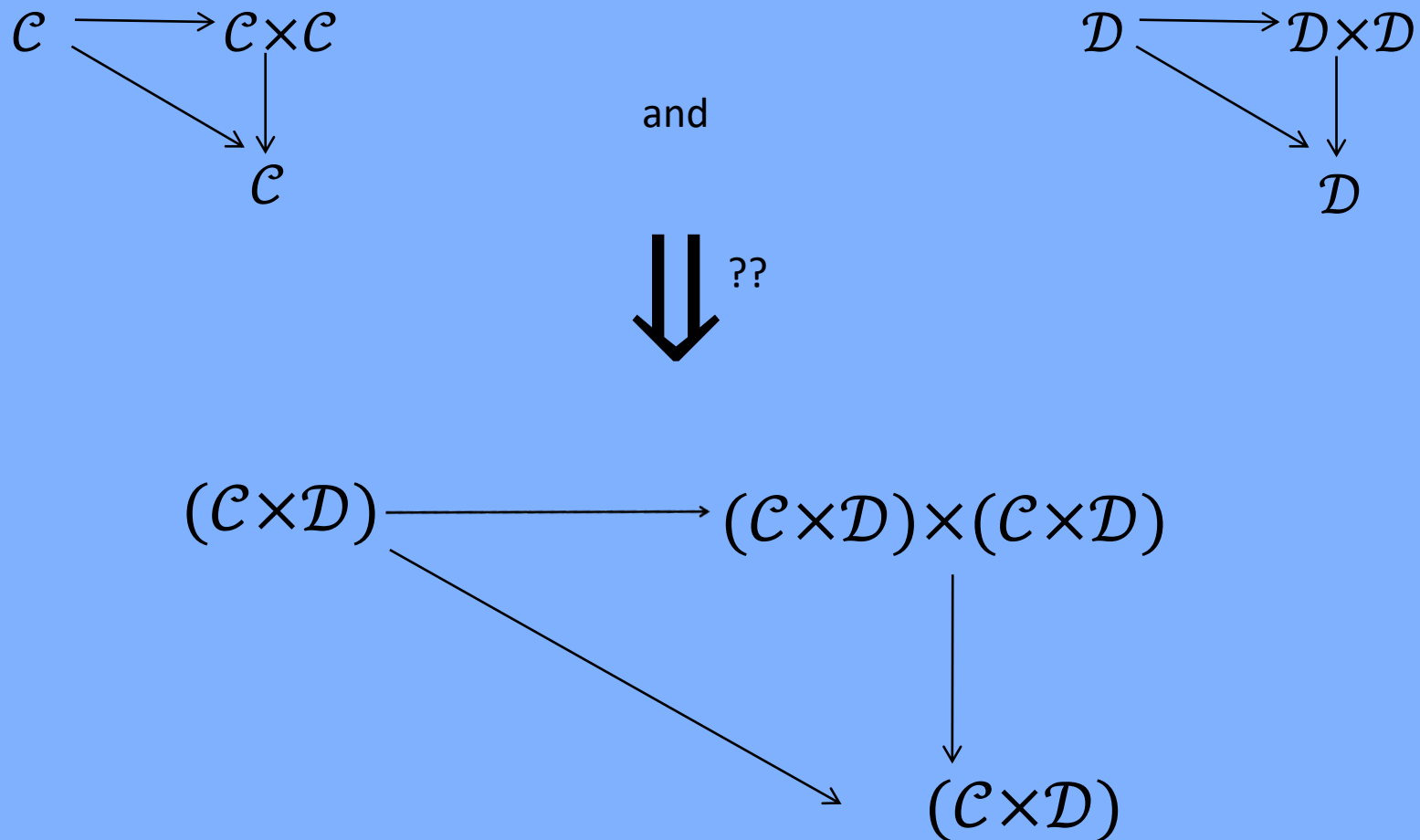
Unitor in a single Category

$u_c: \otimes \circ i_c \Rightarrow id_c$

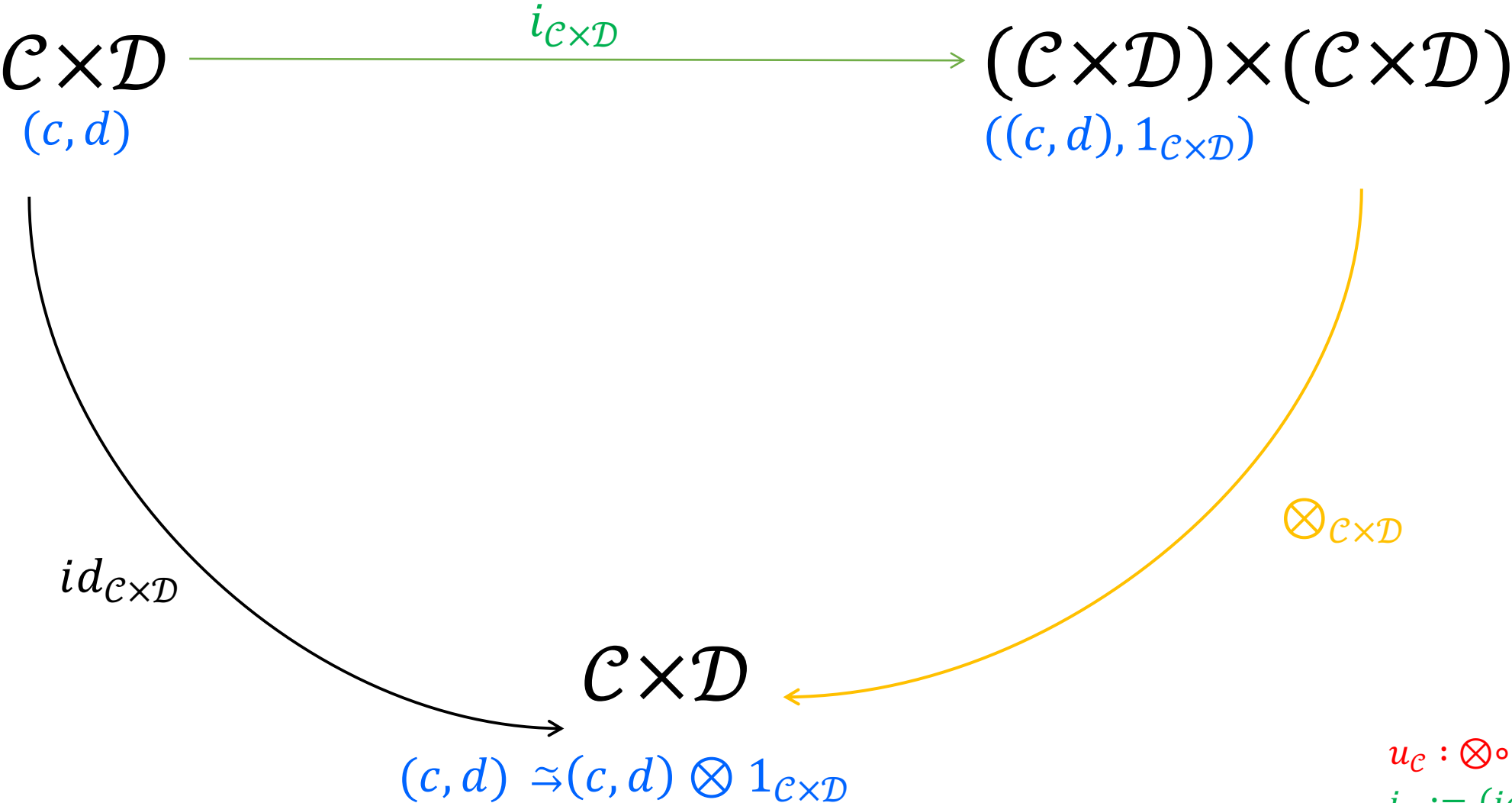


$i_c := (id \times cons) \circ \Delta$

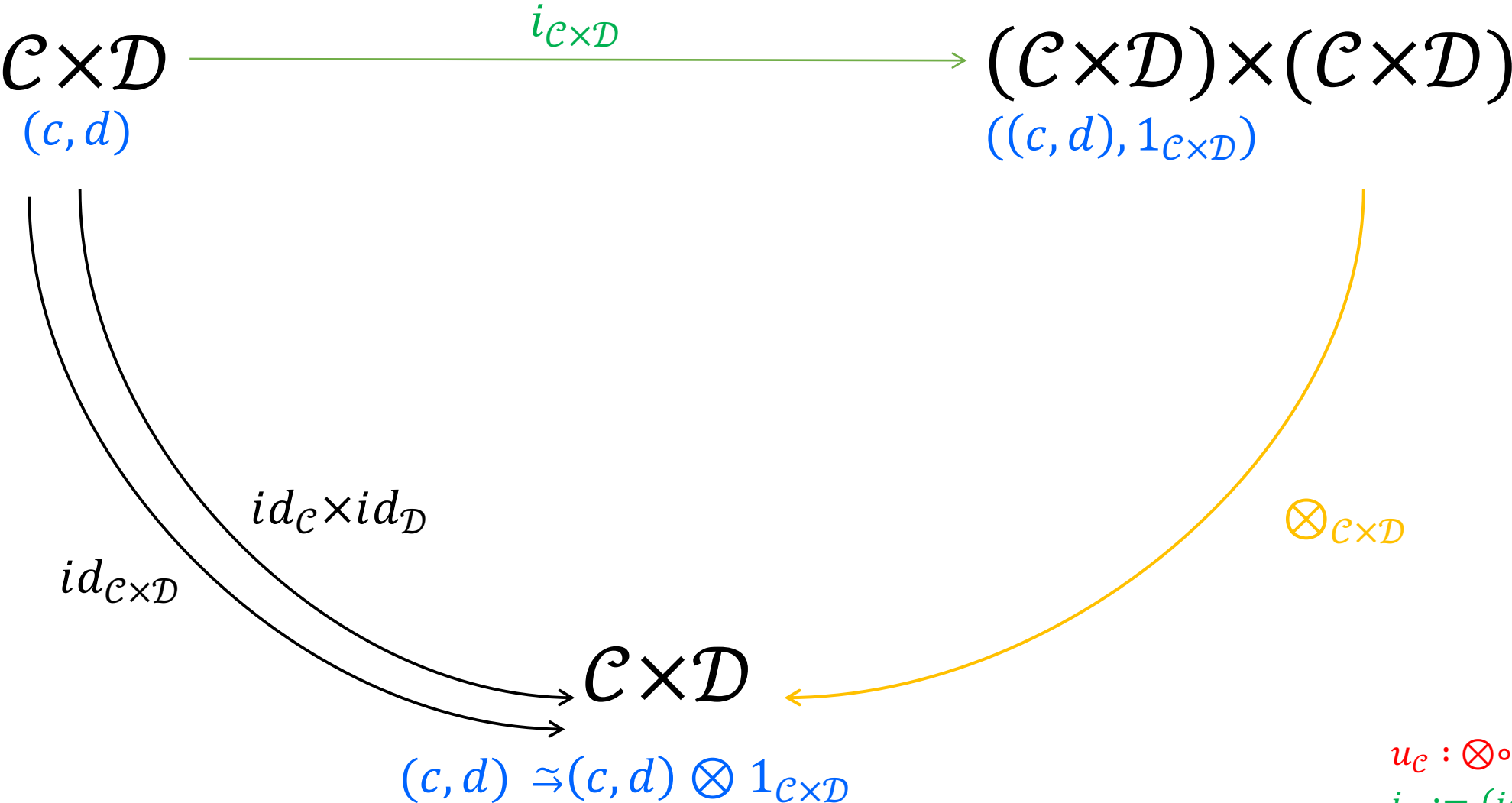
Claim: If  $\mathcal{C}$  and  $\mathcal{D}$  have a unitor, their product category  $\mathcal{C} \times \mathcal{D}$  has a unitor



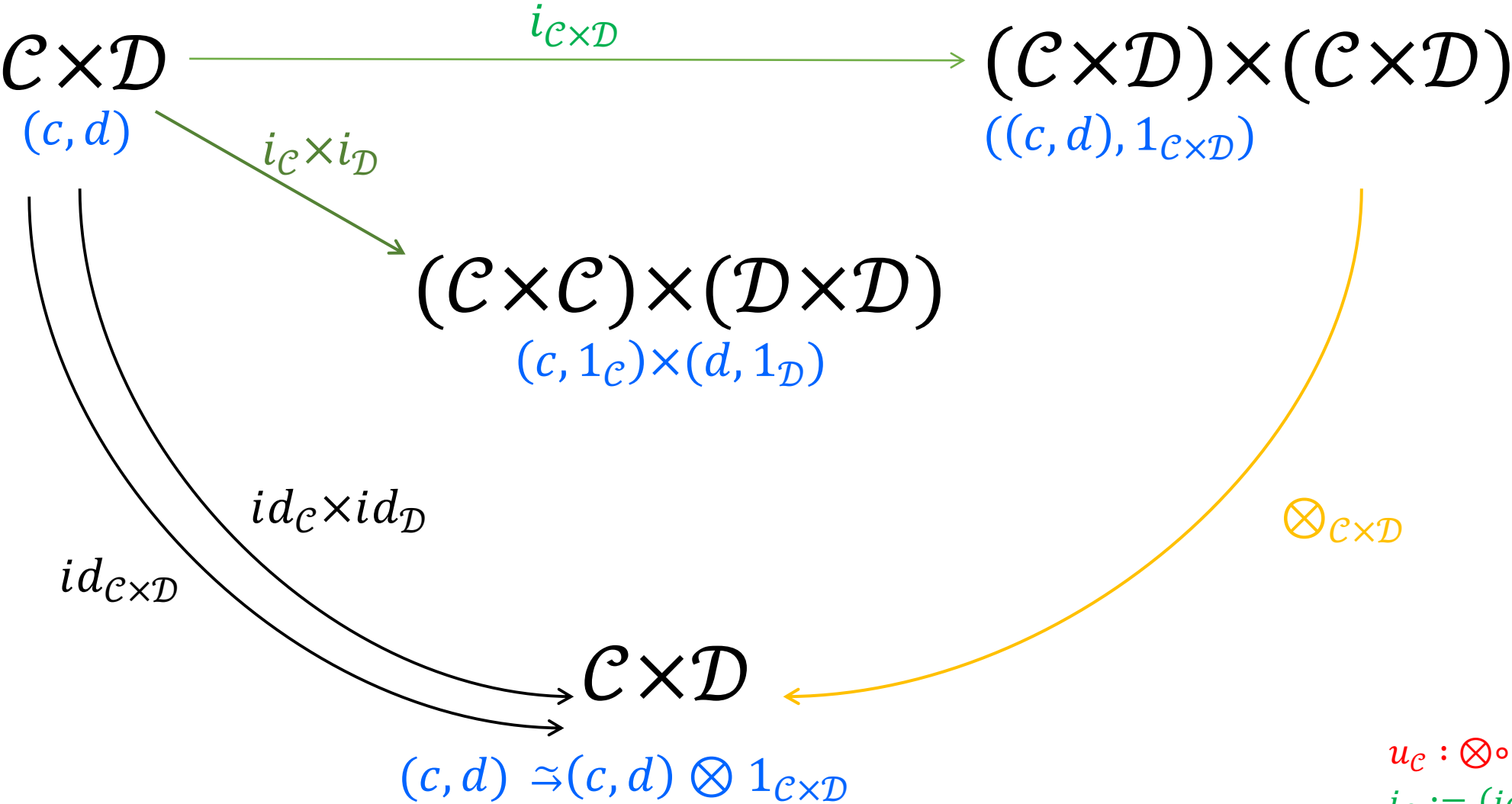
Unitors in a product category



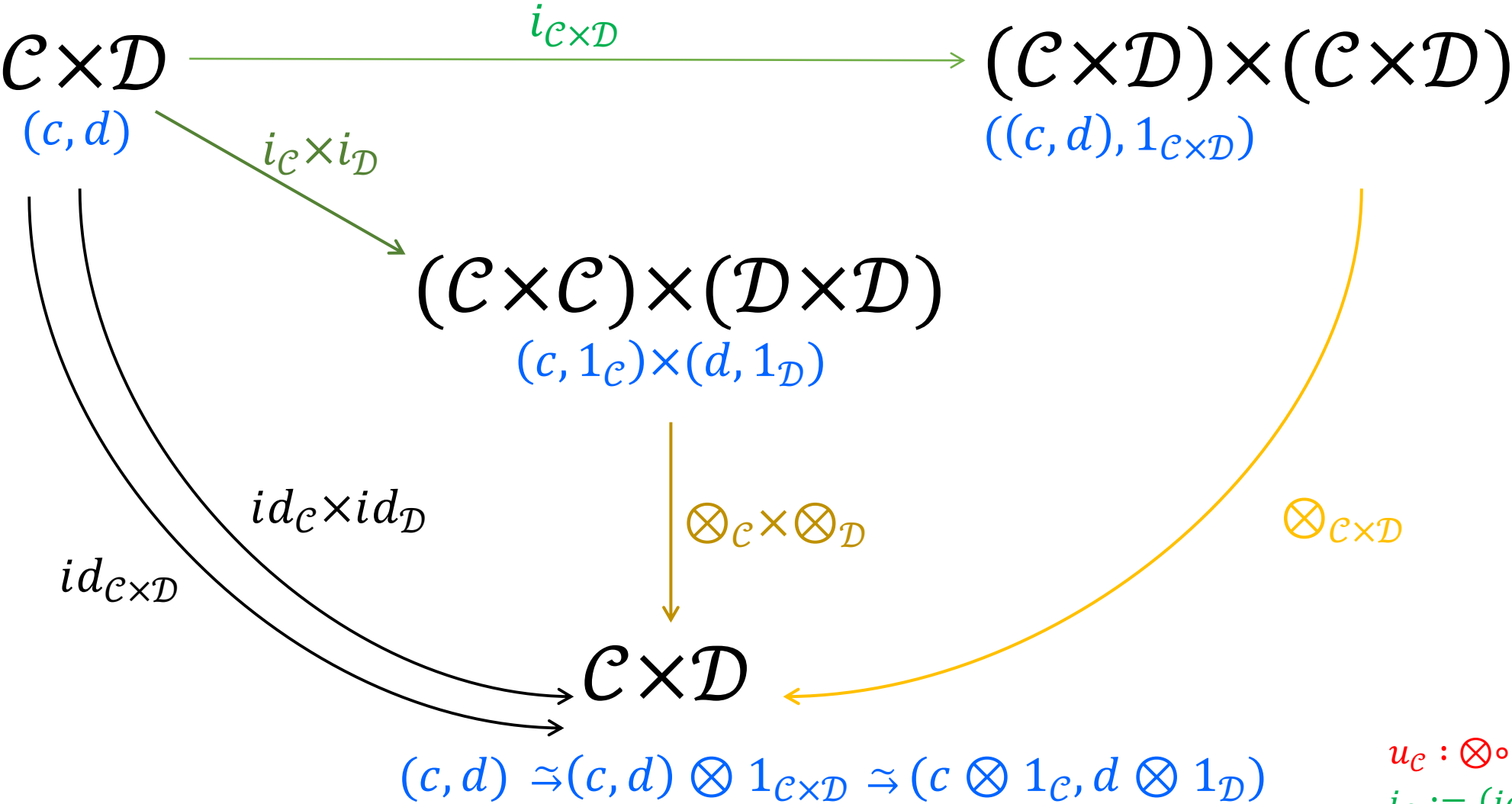
Unitors in a product category



Unitors in a product category

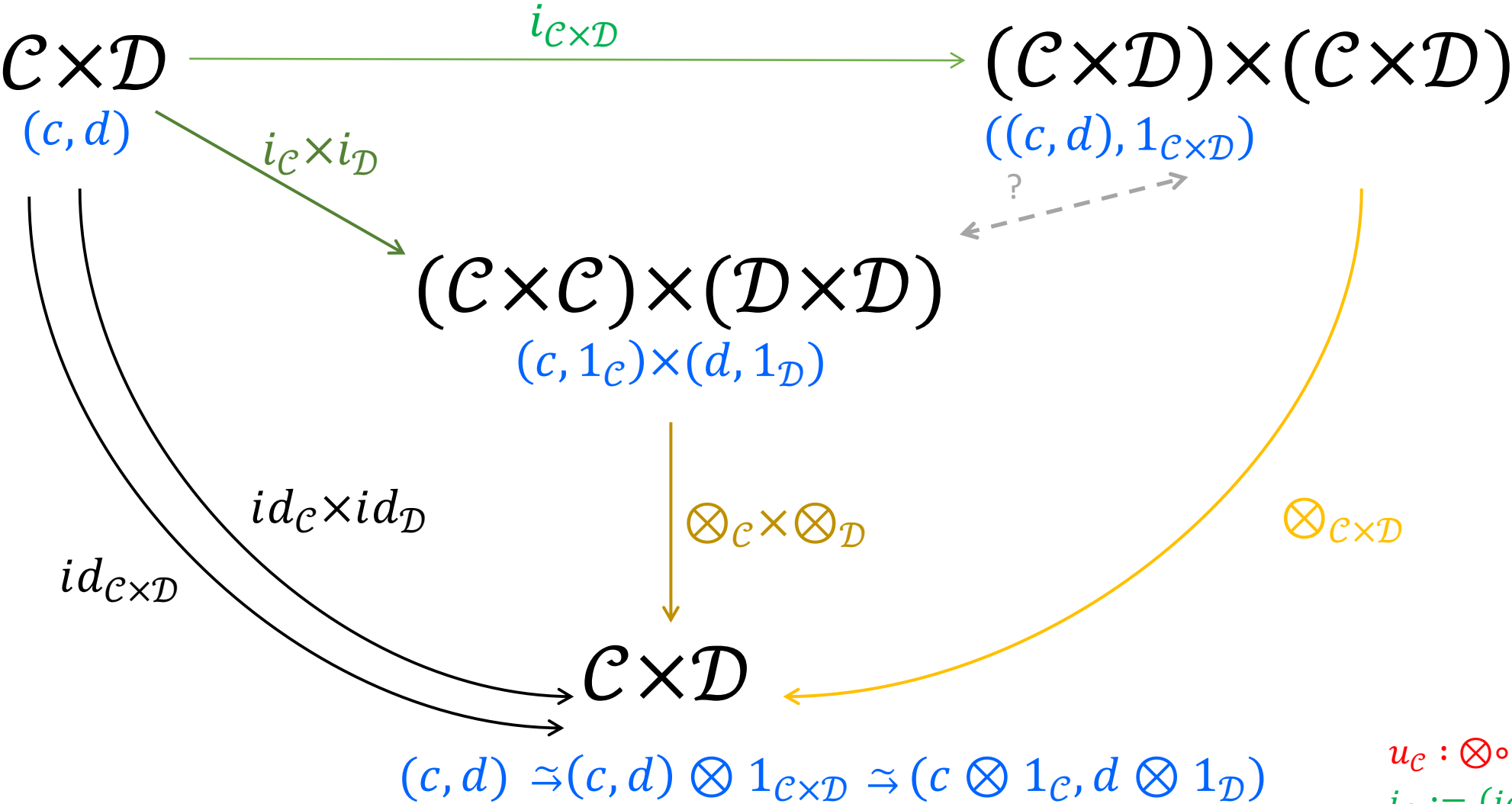


Unitors in a product category

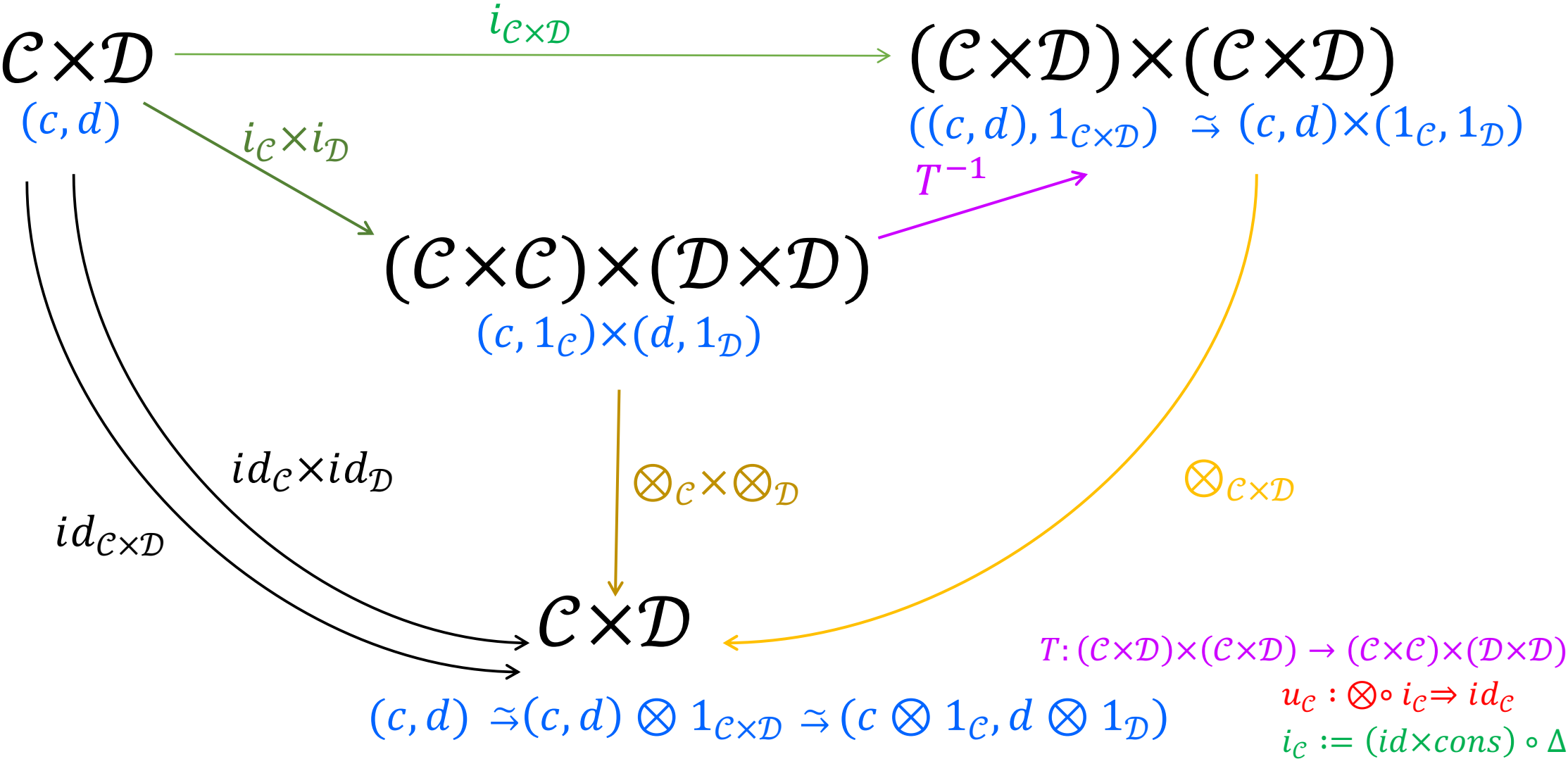


$u_{\mathcal{C}} : \otimes \circ i_{\mathcal{C}} \Rightarrow id_{\mathcal{C}}$   
 $i_{\mathcal{C}} := (id \times cons) \circ \Delta$

Unitors in a product category

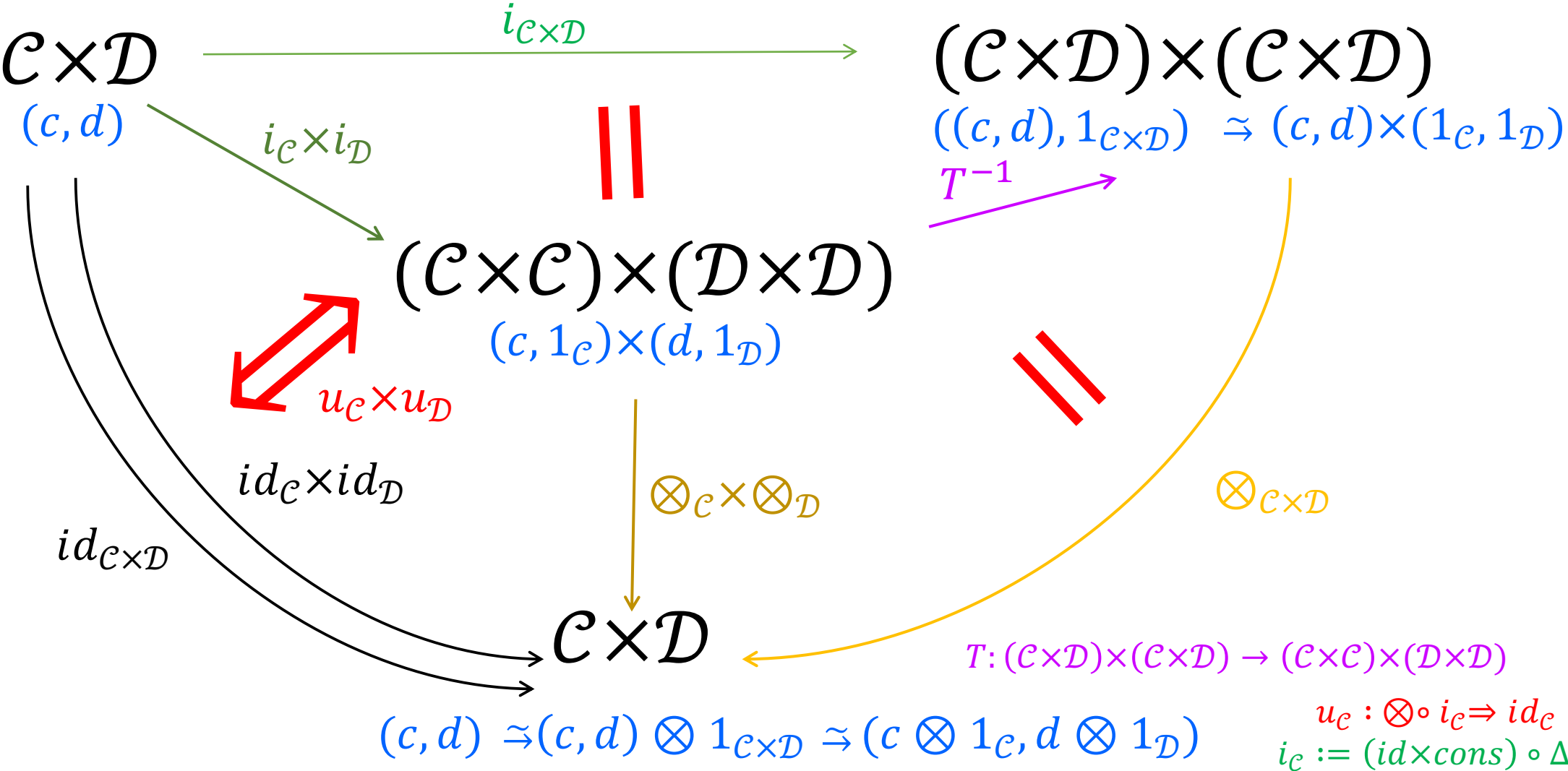


Unitors in a product category

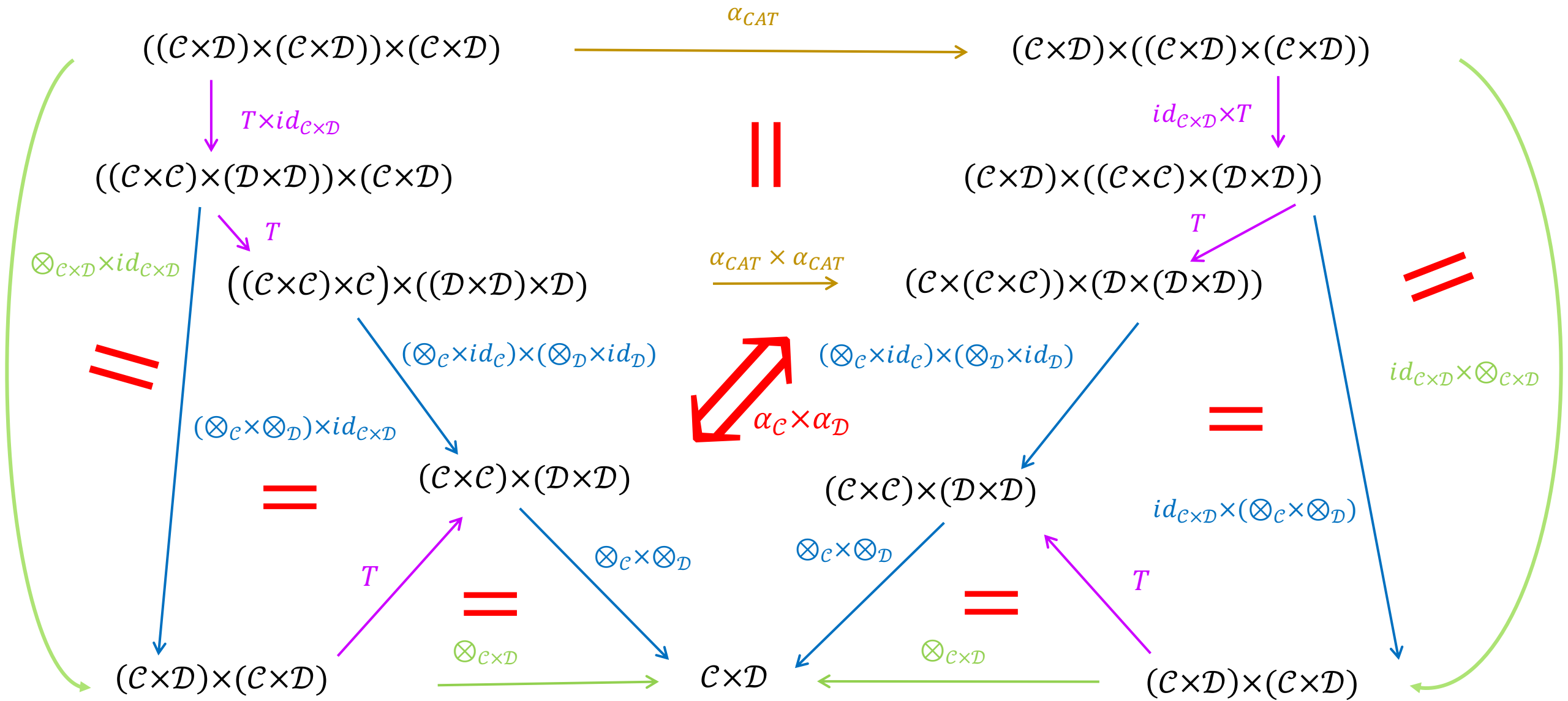




Unitors in a product category

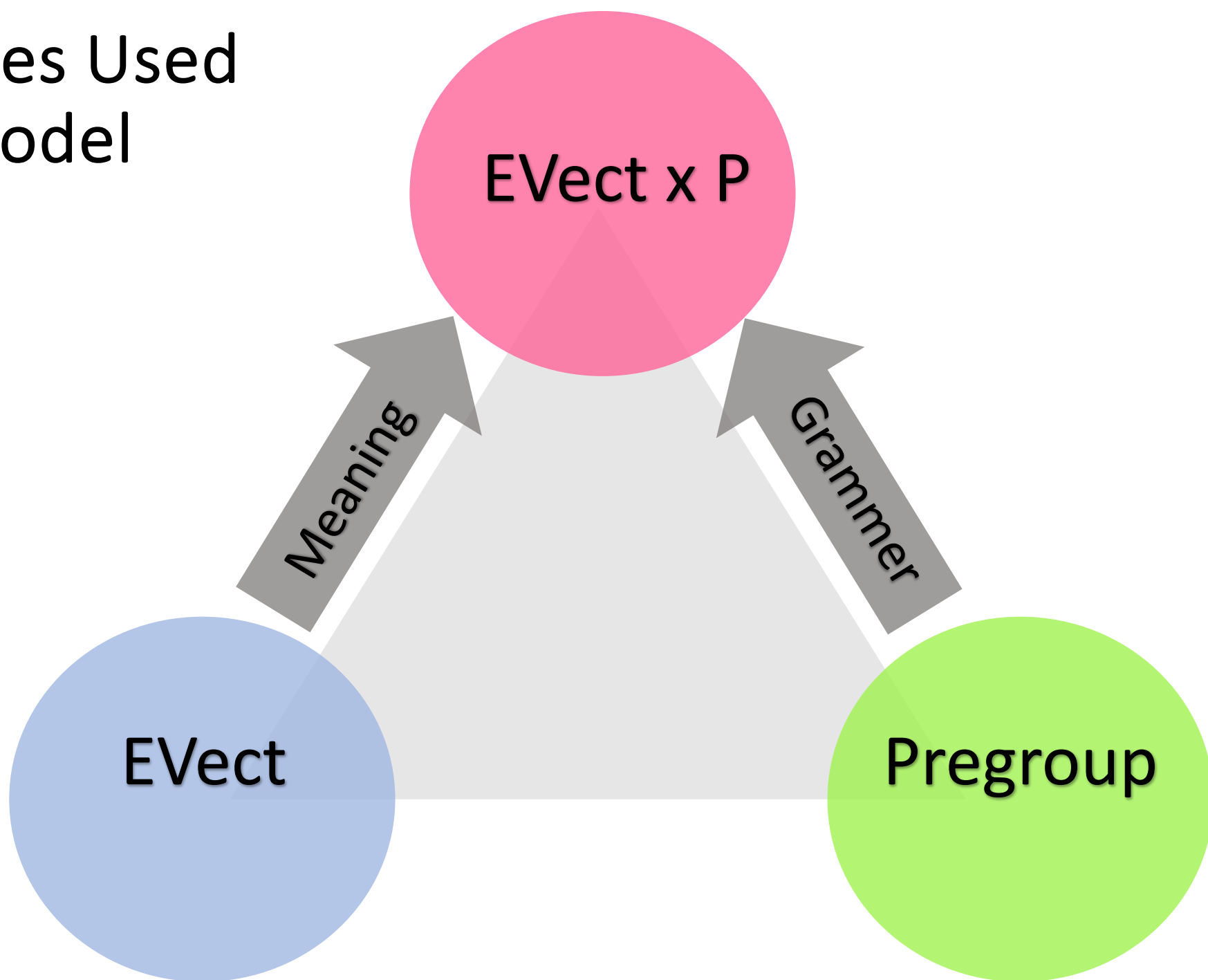


Associator in a product category



# Model in Sentence Structure

# Categories Used in the Model



# Application

The Sentence

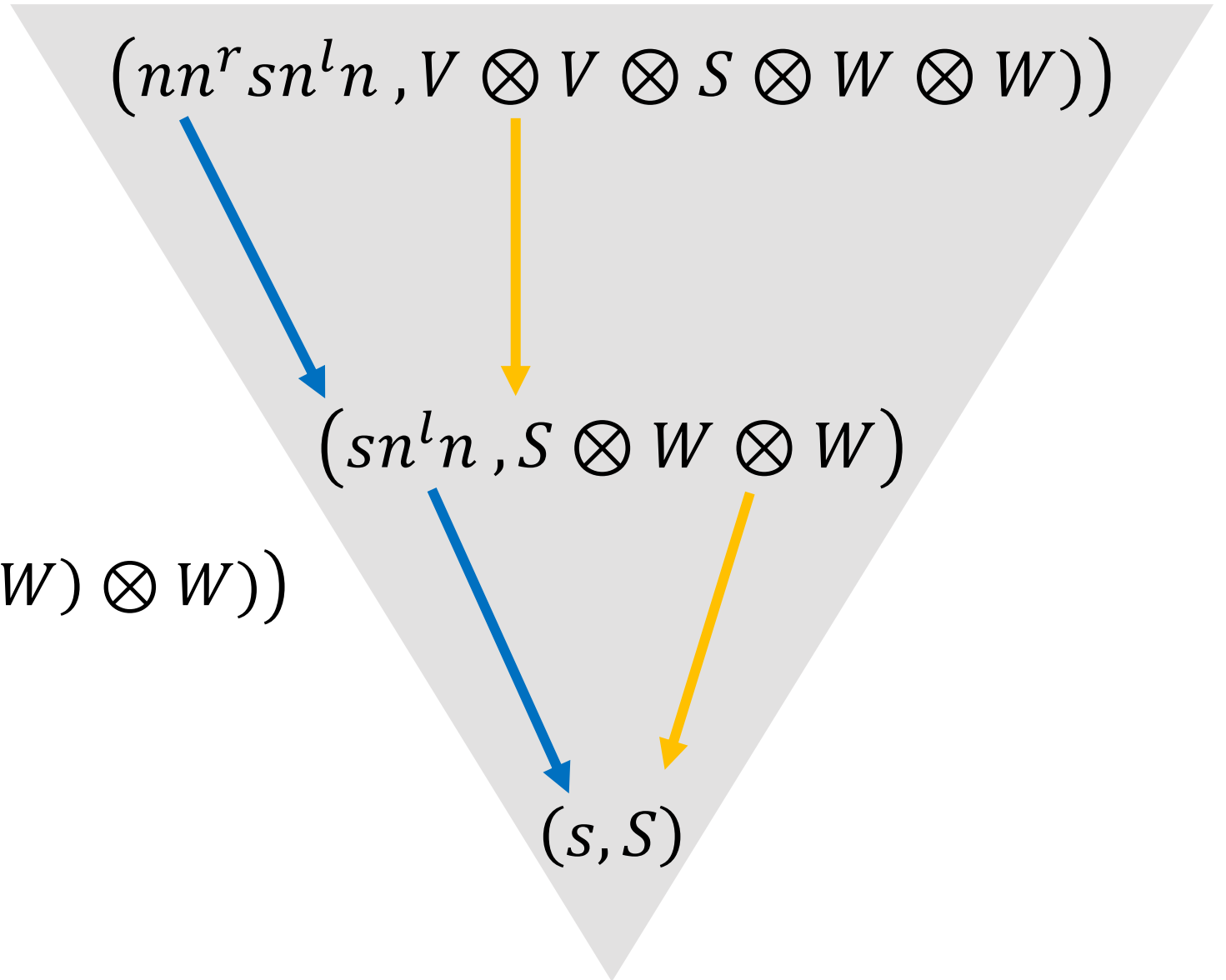
ex: *John likes Marie*

- Noun One:  $(n, V)$
- Verb:  $(n^r sn^l, V \otimes S \otimes W)$
- Noun Two:  $(n, W)$

$$(n(n^r sn^l)n, V \otimes (V \otimes S \otimes W) \otimes W)$$

Morphisms:

- $(\epsilon^r, \epsilon): (nn^r \rightarrow 1, V \otimes V \rightarrow \mathbb{R})$
- $(\epsilon^l, \epsilon): (n^l n \rightarrow 1, W \otimes W \rightarrow \mathbb{R})$



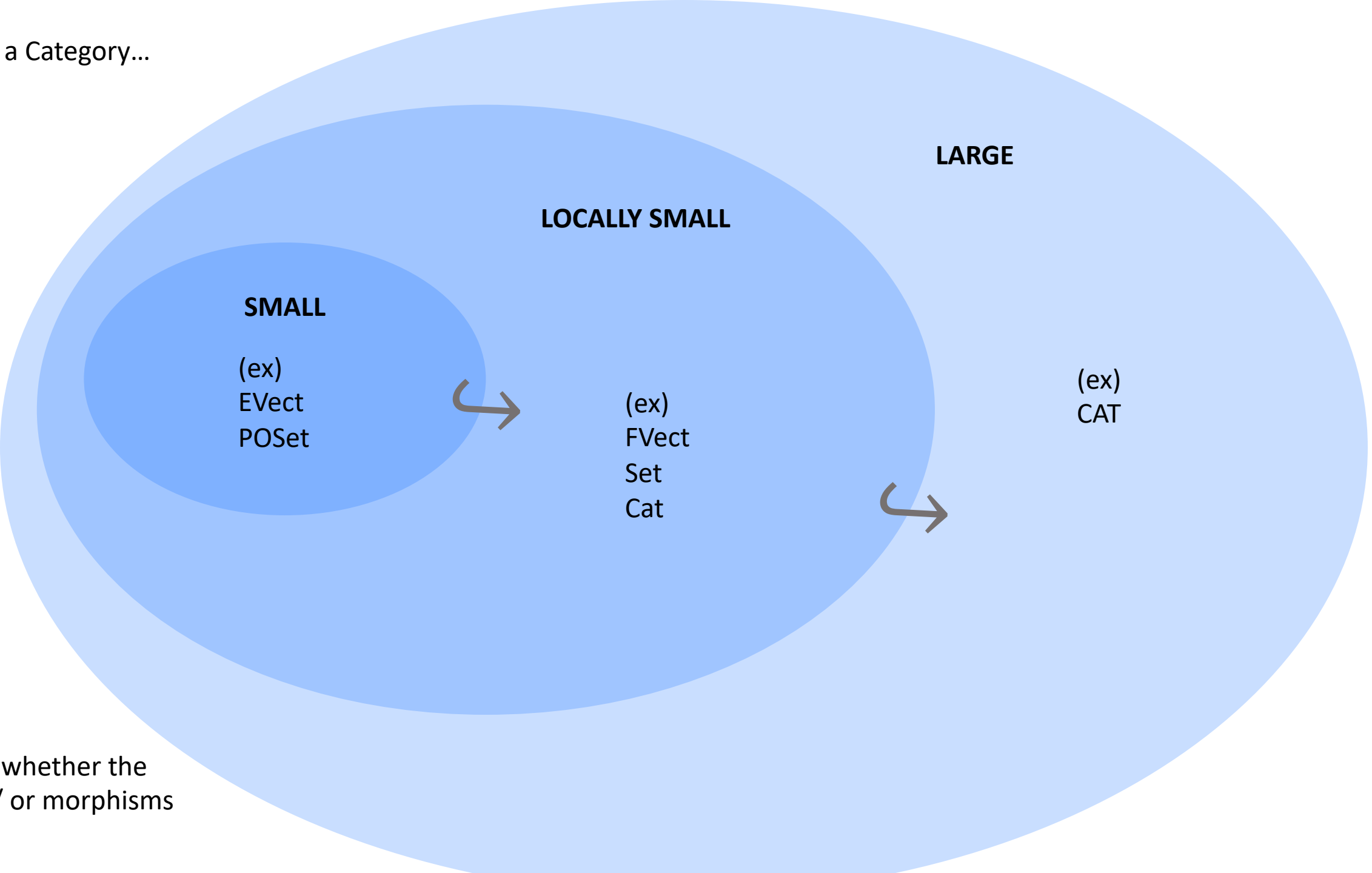
# Thank You! Any Questions?

Mia Goldstein and Emily Herbert  
SUNY New Paltz

Thanks again to RSCA, NCUWM, and Dr. Glass

# Additional Information

The size of a Category...



depends on whether the  
objects and/ or morphisms  
form sets



## Cat and CAT

A category of categories has:

- Categories as objects
- Functors as morphisms
- Natural transformations as functors

Objects in Cat are small while  
Cat is locally small

Objects in CAT are locally  
small while CAT is large

# Tensor Product

Universal Property:

$$\begin{array}{ccc} \mathbb{V} \times \mathbb{W} & \xrightarrow{\eta} & \mathbb{V} \otimes \mathbb{W} \\ & \searrow b & \downarrow f \\ & & \mathbb{Z} \end{array}$$

Concrete Model:

$$\begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^m & \xrightarrow{\phi} & \mathbb{R}^n \otimes \mathbb{R}^m \\ & \searrow b(A \times B) & \downarrow A \otimes B \\ & & \mathbb{R}^{nm} \end{array}$$

The vector space  $V \otimes W$  whose elements satisfy three properties:

$$(\alpha \cdot v, w) = (v, \alpha \cdot w) = \alpha \cdot (v, w)$$

$$(v_1 + v_2, w) = (v_1, w) + (v_2, w)$$

$$(v, w_1 + w_2) = (v, w_1) + (v, w_2)$$

If  $\mathcal{B}_1 = \{v_1, \dots, v_n\}$  is the basis for  $V$  and  $\mathcal{B}_2 = \{w_1, \dots, w_n\}$  for  $W$ , then  $\mathcal{B}_1 \otimes \mathcal{B}_2 = \{v_i \otimes w_j \mid v_i \in \mathcal{B}_1, w_j \in \mathcal{B}_2\}$  is the basis for  $V \otimes W$ .

Given linear maps  $V \xrightarrow{A} X$  and  $W \xrightarrow{B} Y$  there exists a unique linear map:  
 $V \otimes W \xrightarrow{A \otimes B} X \otimes Y$ .

# Pregroup

Pregroup:  $(P, \leq, \cdot, 1, p^r, p^l) \forall p \in P$

Pregroups are partially ordered monoids.

They hold the following properties:

$$p \leq q \Rightarrow p \cdot r \leq q \cdot r \quad \text{and} \quad r \cdot p \leq r \cdot q$$

$$p \cdot p^r \leq 1 \leq p^r \cdot p \quad \text{and} \quad p^l \cdot p \leq 1 \leq p \cdot p^l$$

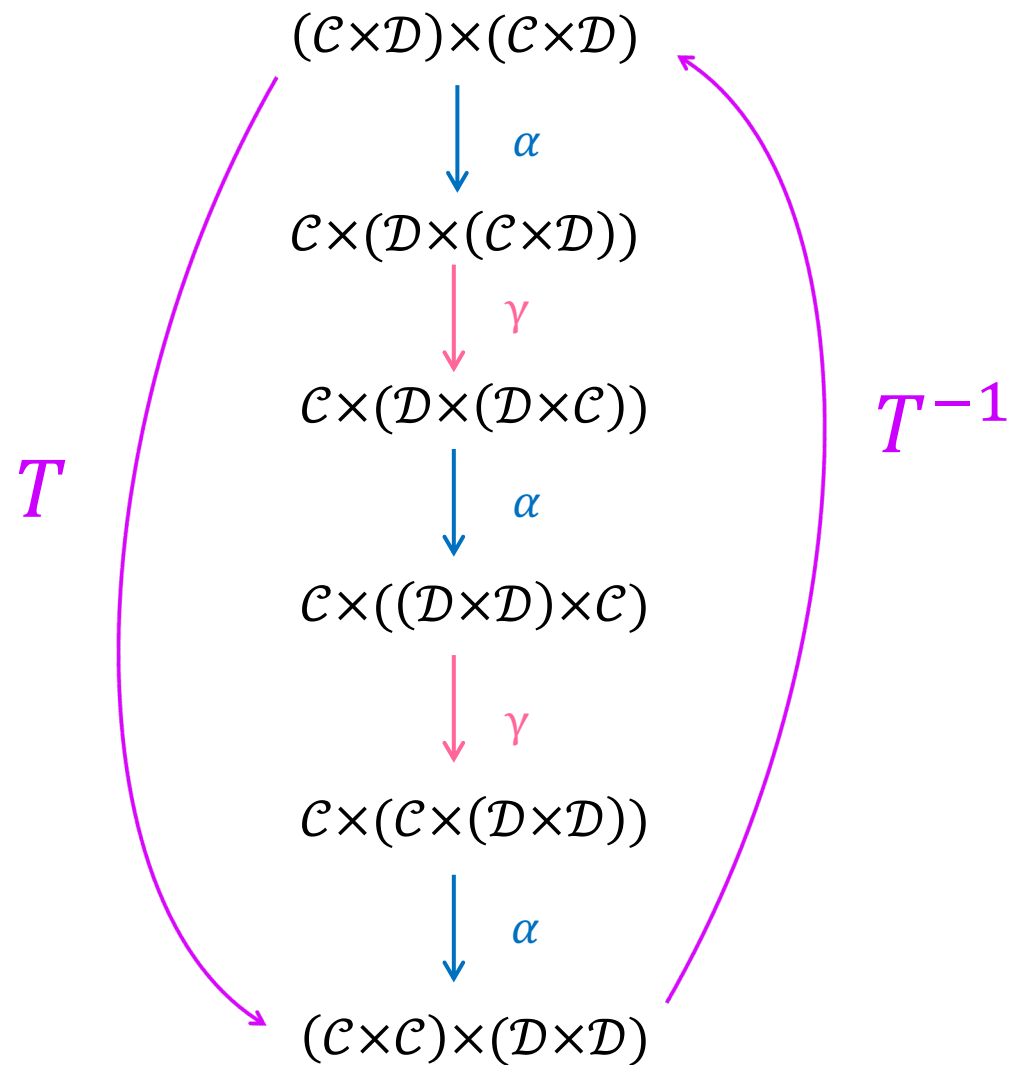
There can be no more than 1 morphism between any 2 objects.

The  $P$  used to model grammar is defined by Lambek grammar.

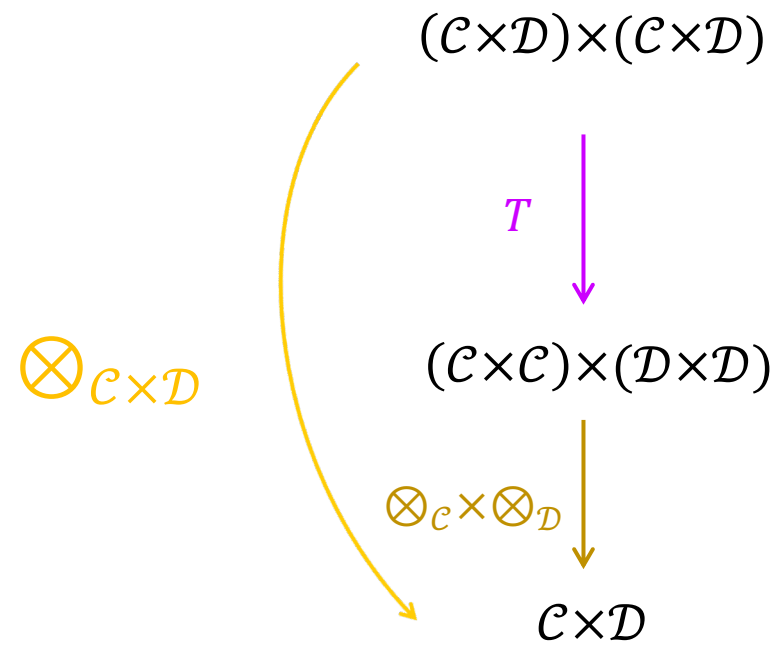
$$(\mathcal{C} \times \mathcal{D}) \times (\mathcal{C} \times \mathcal{D}) \longrightarrow (\mathcal{C} \times \mathcal{C}) \times (\mathcal{D} \times \mathcal{D})$$

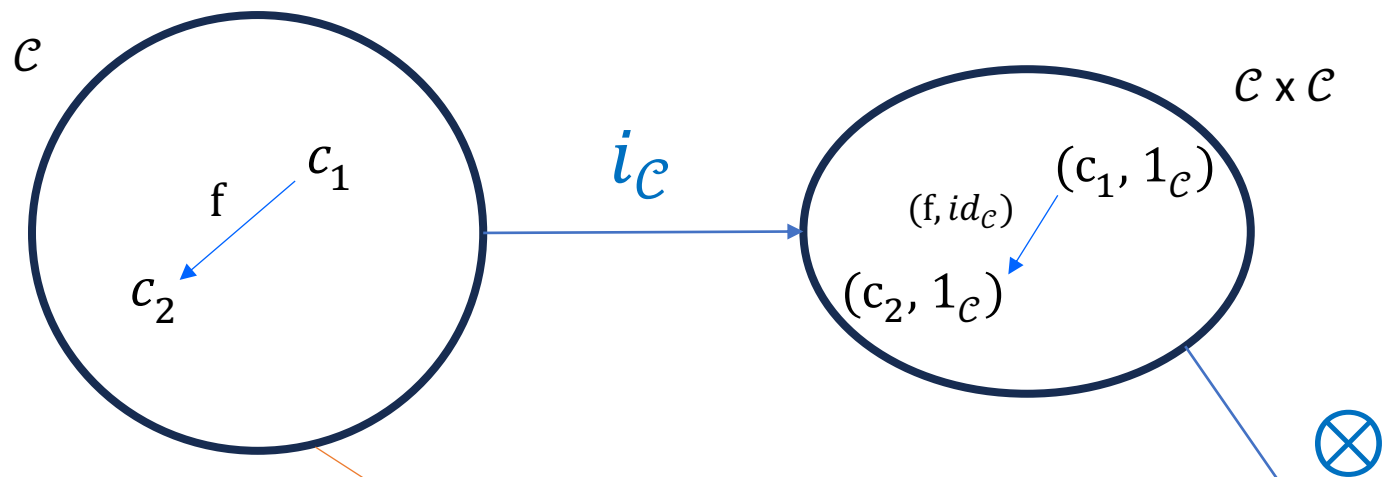
$$\begin{aligned} \alpha: (\mathcal{C} \times \mathcal{C}) \times \mathcal{C} &\rightarrow \mathcal{C} \times (\mathcal{C} \times \mathcal{C}) \\ ((c_0, c_1), c_2) &\mapsto (c_0, (c_1, c_2)) \end{aligned}$$

$$\begin{aligned} \gamma: \mathcal{C} \times \mathcal{C} &\rightarrow \mathcal{C} \times \mathcal{C} \\ (c_0, c_1) &\mapsto (c_1, c_0) \end{aligned}$$



$$(\mathcal{C} \times \mathcal{D}) \times (\mathcal{C} \times \mathcal{D}) \longrightarrow (\mathcal{C} \times \mathcal{D})$$





$id_{\mathcal{C}}$

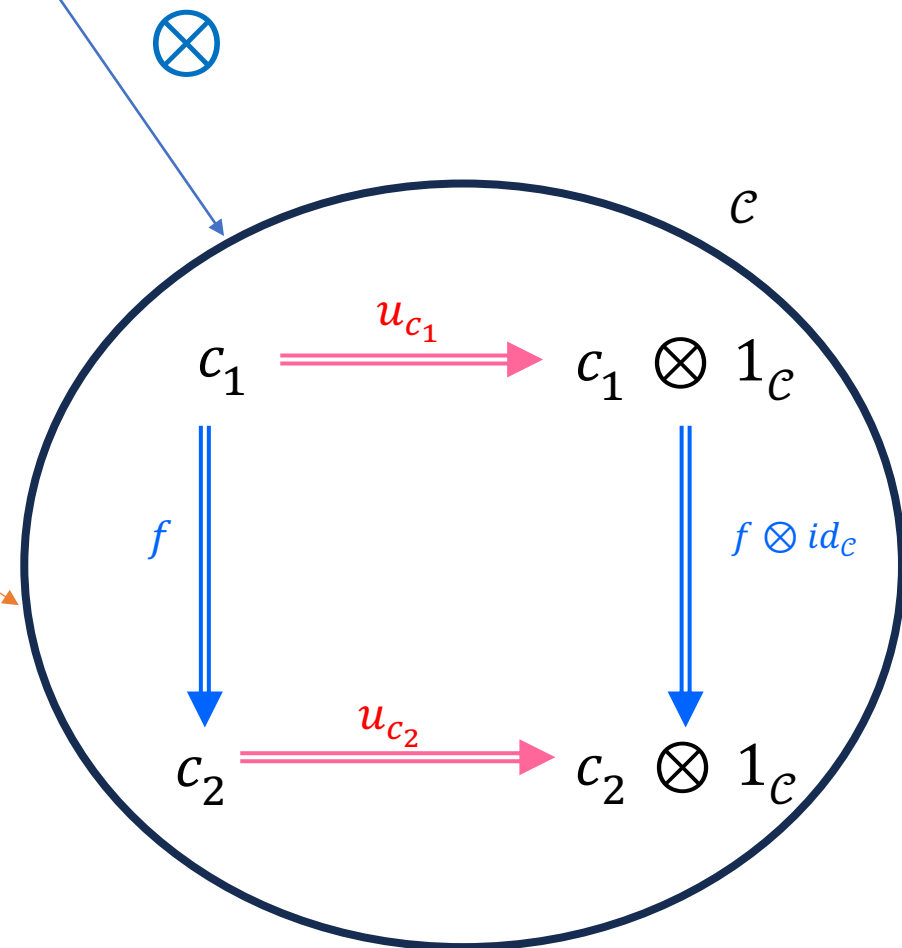
### Natural Isomorphism – Unitor of $\mathcal{C}$

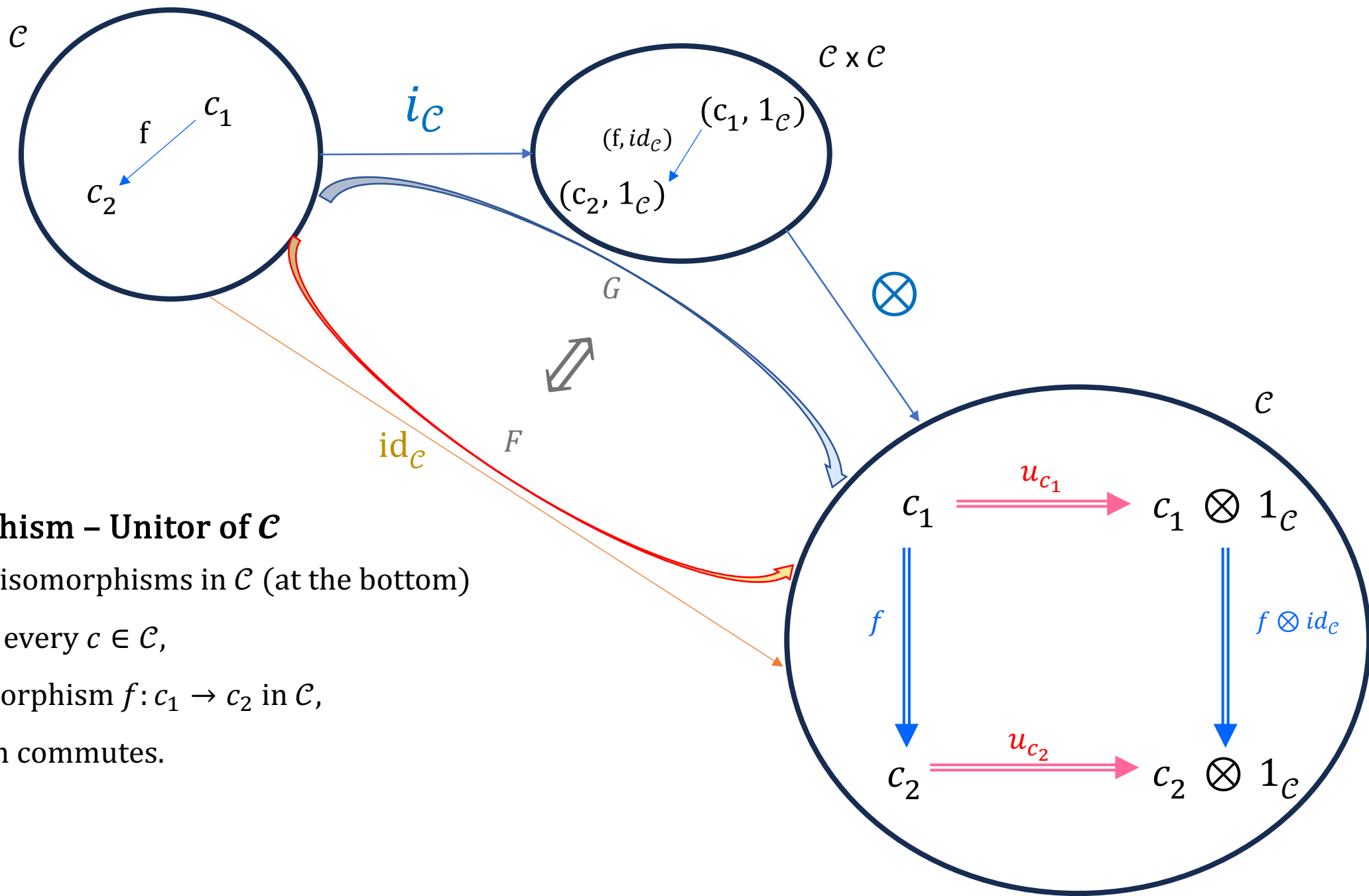
is the collection of isomorphisms in  $\mathcal{C}$  (at the bottom)

$u_c: c \rightarrow c \otimes 1_{\mathcal{C}}$  for every  $c \in \mathcal{C}$ ,

such that for any morphism  $f: c_1 \rightarrow c_2$  in  $\mathcal{C}$ ,

the square diagram commutes.





## Natural Isomorphism – Unitor of $\mathcal{C}$

is the collection of isomorphisms in  $\mathcal{C}$  (at the bottom)

$u_c: c \rightarrow c \otimes 1_{\mathcal{C}}$  for every  $c \in \mathcal{C}$ ,

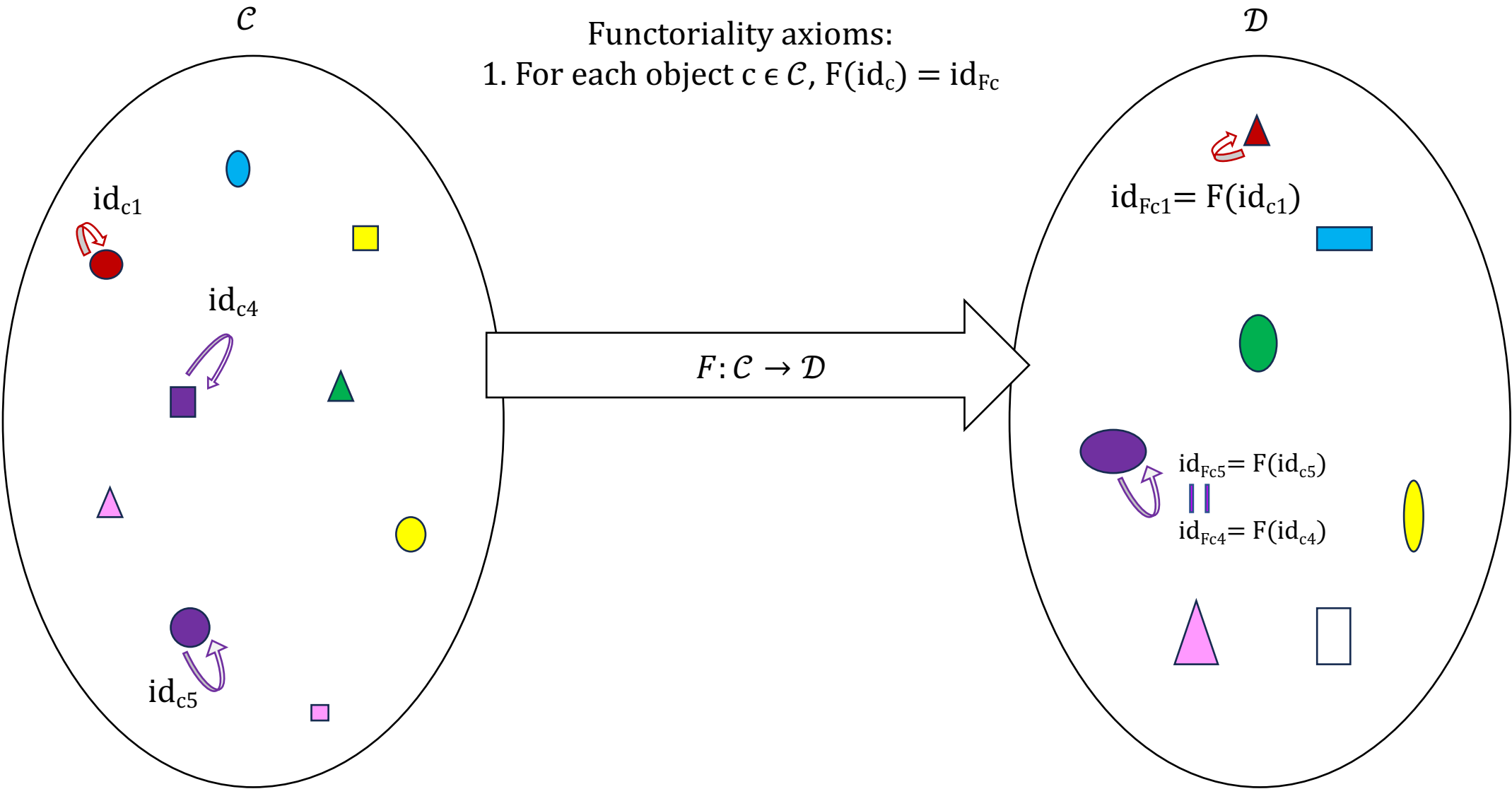
such that for any morphism  $f: c_1 \rightarrow c_2$  in  $\mathcal{C}$ ,

the square diagram commutes.

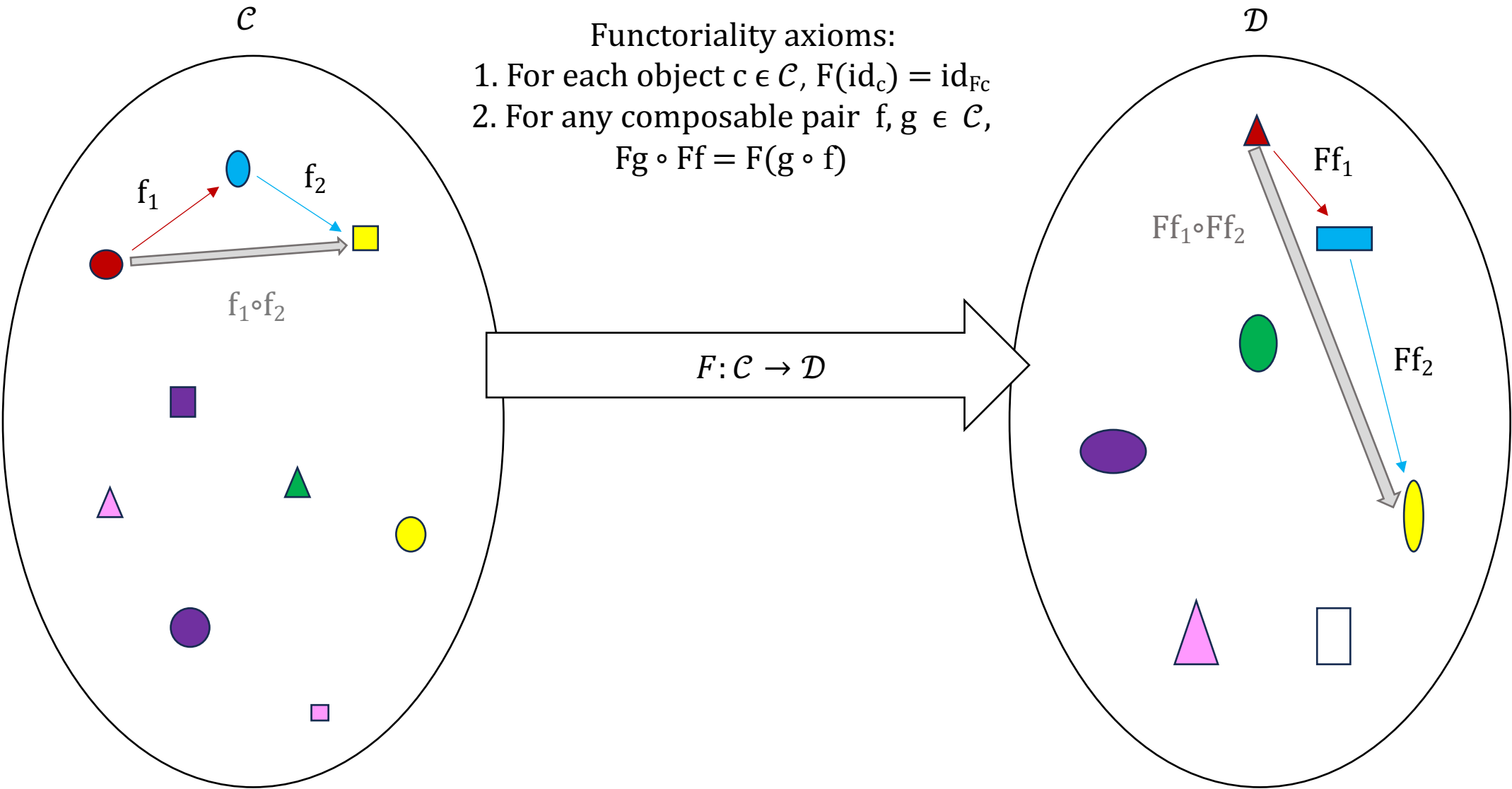
# Functoriality Axioms



Functor : Category  $\rightarrow$  Category



Functor : Category  $\rightarrow$  Category



Functor : Category  $\rightarrow$  Category

