Hippocampal Categories: a mathematical foundation for navigation and memory

Jaime Gómez and Ricardo Sanz

Abstract It goes without saying that in science, experiments are essential; hypothesis need to be contrasted against empirical results, in order to build scientific theories. In a system of overwhelming complexity like the brain, it is very likely that hidden variables, unknown by the experimentalist are interacting with those few elements of which the values are expected and can be validated or rejected in the laboratory. Thus, at the end of the day, the experimentalist is refuting or validating models that are somehow prisoner of the lack of knowledge about the structure of the system. The global picture is missing, the key is to reveal the structure which must be found not in the objects, but in the relationships between the objects -morphisms. How components at the same level interact (objects in the category of neurons) and how superior levels constrain those levels below and emerge from those above is tackled here with a mathematical outlook. The mathematical theory of categories is proposed as a valid foundational framework for modeling in brain sciences.

1 The hippocampus as a representational device

How does the mind represent physical space?. This is a question that has kept philosophers busy for centuries. In 1975, the philosophical discussions about space representation acquired a extremely powerful and fresh insight when O’Keefe and Nadel, discovered the place cells in the hippocampus of the rat.

The experimental study of spatial representation has since then exploded. The 70’s was the decade of the place cells, neurons that discharge when the
rat is in a particular position. In the 80’s head direction cells, neurons that discharge significantly whenever the rat’s head changes direction, acquired the attention of scholars. Since 2005 we have been in the grid cell “era”.

These discoveries are of major importance in different research fields. Indeed the theory of the cognitive map [16] is rooted in the discovery of place cells in the hippocampus. One derivative of this theory is the map-based navigation capability, that some animals have, and that engineers have been able to replicate in robots[13].

The debate about whether the brain generates a map like or not, seems to have shifted in favour of those who back the cognitive map theory. Indeed the discovery of place cells, head cells and recently, grid cells suggest so.

Yet the underlying nature of the cognitive map remains elusive. Is the representation purely metrical or is topological? or are the maps constructed in the hippocampus without paying attention to the features of the environment i.e:metrical maps, or do they reflect the relationships between the environmental features i.e:topological maps?

The role of the hippocampus is to associate internal and external coordinates systems and to accommodate cue conflict situations (reinstantiate the context when there is a mismatch between internal and external relationships). Rather than debating whether the hippocampus is the depositary of the declarative memory or the index access of a collection of maps, it may be more productive to ask What is the role of hippocampus in navigation and memory? With this mind, in “The hippocampal debate: Are we asking the right questions?” [1], Redish suggests that there are multiple memory systems in the brain and multiple navigation systems.

1.1 Place cells

Place cells are neurons located in the brain hippocampus that fire in complex burst whenever an animal, for example a rat, moves through a specific location in an environment.

The striking thing about place cells is that they code the Cartesian position, irrespective of either the direction from which the position is reached or the behavior of the rat at any precise instant. Thus, there is a direct link between the neural activity of a single cell to a Cartesian position.

How does the animal know that it is in a particular position?. Apparently this could be done by computing the allocentric space, landmark or visual cues. The most important property of these cells is the omnidirectionality property and it can be observed in their conical shape (the firing rate increases when the rat approaches the location, independently of the direction is heading when it does it). Thus the place cells are necessarily, coding explicit (no contextual) locations in the environment and not particular sensorial cues.
The region in which a place cell fires the most is called place field. Thus, there is a correspondence place field/place cell. One cell has a place field when the firing rate within the field is much higher than outside - e.g: from 20 Hz to 0.1 Hz. For a given environment, we can determine a collection of place cells whose associated place fields cover the whole environment. Nobody denies that under certain circumstances, the hippocampal pyramidal cells show place fields - location-based firing.

Notwithstanding, it is less clear what they really represent; there are those who argue that place cells can be an epiphenomenon, produced by the spatial nature of the experiments where these cells are discovered. Granted that place cells are correlated to space, the question that arises is: Are the place cells the only neurons correlated to space? The other possible representational content of these cells and of the assemblies they form, can serve to further question how the hippocampus contributes to spatial representation, navigation and episodic memory.

1.1.1 Place cells as representational entities

The interest in these cells is rooted in the fact that they are candidates to be the direct representation of the external space. A place cell, fires maximally when the animal is in a specific position or place field, so the firing rate of a cell can be used to decode the position of the animal within the environment with striking accuracy.

The existence of place cells was not accepted until Muller [15] came out with the numerical method that allows to quantify the place fields.

At this point, we need to formally define the term place field. A place field \( F \), for a place cell, is an open ball of radius \( r \) and center \( x \) in a normed vector space \( V \) such that \( \text{fr}(F) > k \), where \( k \) is a constant that represents the threshold or firing rate, and \( f_r \) a function that returns the minimum firing rate for all the pixels (vectors) that fall into the vector space \( F \).

1.2 Grid cells

Grid cells, likewise place cells, are place-modulated neurons. The firing location of a grid cell is multiple, contrary to the place cells which are mono field. The multiple firing location of a grid cell is indeed a grid with a most striking property, it is an array of equilateral triangles [8].

It might be noted that grid cells were discovered, while researchers investigated whether place cells activity was endogenous to the hippocampus.

---

1 In the initial notation of Muller, the function \( f_r \) uses 6 colors to represent the firing rate, the darker the colour is, the higher is the firing rate of the cell at that pixel.
Fig. 1  The picture shows the 36 place fields one of each of 36 rat hippocampal CA1 place cells. Thus, as it is shown in the figure the position of the rat is contained in the firing of these cells. The place fields are cone shapes, this means that the firing rates increases irrespective of the direction from which the rat arrives.

Fig. 2  Grid map of a rat running on a linear track after 10 min [6].

The hypothesis was that CA3 and DG are the intra-hippocampal inputs to CA1, which is the area in the hippocampus where one can find most of the place fields. This idea was proven wrong, after removing CA3 and DB, the CA1 inputs, the spatial firing in CA1 persisted. So place signals did not exclusively arise within the hippocampus, spatial signal was brought to the CA1 from outside the hippocampus.

In 2004 Fyhn et al.[18] discovered a group of neurons in the medial entorhinal cortex (mEC) that shows spatial related firing. The mEC cells have sharply tuned spatial firing, much like the hippocampal place cells do, but with one difference, each of the mEC cells, has multiple firing fields rather than one as is the case in the place cells. One year later, Hafting et al.[8] discovered that the many firing fields of each neuron, generate a grid. Thus as the animal moves, the grid cells tile the environment with periodic triangles that reflect the changing position.

The grid cells have been found in the mEC, from the six layers of this cortical structure, it is in layer II where we find the highest density of this kind of cells.

The neurons in the layer II of the medial entorhinal cortex (mEC-II) are the main input of the place cells, but in contrast the entorhinal cells are activated throughout the environmental terrain, whenever the animal is at the vertex of some equilateral triangle, forming a tessellation or grid.
In short, both place cells and grid cells are neurons with spatial located firing, in other words, they have spatial representational power, allowing the animal to know its position and to navigate in an environment, for example find the way back home after eating. The difference, apart from the fact that place cells are hippocampal neurons and grid cells are in the mEC, is that whereas a place cell has a single firing location, a grid cell has multiple firing fields with a striking geometric regularity: the firing fields form periodic triangular arrays, tiling the entire environment available to the animal.

Fig. 3 Basic circuit of the hippocampus, as drawn by Ramón y Cajal[21]
1.2.1 Grid field

Three parameters totally describe the grid of a cell (grid cell): *spacing* is the distance between contiguous fields, *orientation* the rotation angle of the grid referred to the reference axis, *spatial phase* is how much the grid is translated relative to an external reference point.

A grid field for a grid cell is a set of open balls \( G_i : 1...n \), where for every ball \( G_i \), \( fr(G_i) > k \), that is to say, it has a significative firing rate. Thus, so far, the definition of \( G_i \) is identical to place field seen in 1.1.1.

Additionally, every ball \( G_j \) of a grid field, form a equilateral triangle with its two other closest balls \( G_i \) and \( G_k \) in \( \cup_{1...n} G \) or \( G \) to abbreviate. The grid field \( G \) is identified by the three parameters, spacing, orientation and phase that can be trivially obtained from the metric of the space defined above.

The processing of the place signal is therefore not an exclusive privilege of the hippocampus, the mEC plays a major role in the spatial representational. The majority of cells in mEC-II and mEC-III have grid properties, this means

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Place cells and Grid cells, similarities and differences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Brain area</strong></td>
<td><strong>Type of map</strong></td>
</tr>
<tr>
<td>Place cells</td>
<td>Hippocampus</td>
</tr>
<tr>
<td>Grid cells</td>
<td>mEC</td>
</tr>
</tbody>
</table>

that most of the cortical input to the hippocampal place cells that have to do with spatial representation, come from grid cells in the mEC. Grid cells can be found just one synapse upstream of the place cells[8]. Therefore, acknowledging that grid cells and place cells are intrinsically connected, to claim that place field may be extracted from grid fields[17] deserves to be appropriately explored.

Mathematically, using Fourier’s analysis, several grid fields with different spacing can combine linearly to yield a place field. In [11] is presented a computational model showing that place field can arise by the sum of 10-50 grid cells. When the spatial phase variation in the grid-cell input was higher, multiple, and irregularly spaced firing fields were formed. This idea has been very appealing in the hippocampus community, and it has helped to produce a large number of models with a common tenet: place cells in the hippocampus compete to receive the summed firing pattern activity of the cortical grid cells.

The problem with these kind of models that transform grid patterns into place pattern is that they do not tell us that much about the mechanisms that underlie the spatial firing pattern of grid cells and place cells. Besides, it is debatable that a linear sum of grid cell pattern which has a metric is the correct way to model a place cell pattern which represents topologically the environment without metrical relations.
It might be remarked here that the models of grid field formation, deal with timing rather than with structure or connectivity, and this is because they assume that the structure is already known. In these models the structure is a single cell whose firing activity needs to be understood.

There is a number of computational models that aim to simulate grid field, however they do not tell us much about the causes that originates that phenomenon, let alone a mechanistic explanation that unveils the real causes of the emergence of grid cells in the hippocampus.

As Zilli points out in [19], we must be prudent, “the study of grid cells is still in its infancy” and the mechanisms that underlie the spatial firing pattern are still waiting to be discovered.

2 A theory of brain’s spatial representation based on category theory

One of the highlights of this work is that it exemplifies the theory of categories in strong non-algebraic categories. Indeed, the crucial aspect and novelty in this work needs to be met in the categorical construction of biological (non algebraic) categories.

2.1 The category of neurons

Let us define a category of neurons $\text{CAT-Neur}$ as a category whose objects are either neurons or sets of neurons. $\text{CAT-Neur}$ as any other category, consists of three things, i. a set of objects $O$, ii. a set of morphisms $\text{Mor}(A, B)$ for any two objects $A, B$ of $O$, and iii. a rule of composition that fulfills the properties of associativity and identity.

We identify three possible categories for the category $\text{CAT-Neur}$ that are useful for the development of the theory introduced in this paper.

i the category $\text{Neur}$.
ii the category $\text{Neur}^+$.
iii the category $\text{Neur}^*$.

The category $\text{Neur}$, whose objects are neurons and the morphisms are the synaptic paths between them, with the convolution of paths as composition.

The category $\text{Neur}^*$ which is the category of neurons where the objects are topological spaces of neurons $(N, \theta)$ and the morphisms are continuous maps.

A function from two topological spaces $f : (N, \theta) \rightarrow (M, \upsilon)$ is continuous if $f^{-1}(B) \in \theta$ whenever $B \in \upsilon$. 
And the category \textbf{Neur$^+$}, which has as objects, metric spaces, and as morphisms, Lipschitz maps for $\lambda = 1$ that preserve distances.

Note that a Lipschitz map is always continuous but the contrary is not true. The morphisms in \textbf{Neur$^+$} preserve distances between metric spaces which exemplify neural assemblies.

### 2.2 The category of metric fields

Now we will define the category of place fields, that is the physical locations that produce the spike firing in the grid cells and place cells.

The category \textbf{Field$^+$} has as objects metric spaces (including hexagonal grids) and as morphisms contractions. And the category \textbf{Field$^*$} is composed of topological spaces and continuous functions.

![Diagram](image)

**Fig. 4** $(X, d)$ is a metric space where $X$ is the set of place cells in the hippocampus and $d$ the euclidean metric distance. $(Y, e)$ is a metric space in the bidimensional plane with identical distance $e = d$. The mapping between the metric spaces $f : X \rightarrow Y$ preserves the distances if $e(f(x_1), f(x_2)) = d(x_1, x_2)$. $f$ is said to be an isometry and is immediately a monomorphism (Demo: $x_1 \neq x_2, e(f(x_1), f(x_2)) = d(x_1, x_2) \neq 0 \Rightarrow f(x_1) \neq f(x_2)$). Then an isometry that is an epimorphism is an isomorphism.

Empirically speaking, it makes no sense to take the space $(Y, e)$ in figure 4 as a metric space, for example closed points in the space are not necessarily mapped onto closed place cells in the hippocampus.

Thus, objects in the category of neurons can not be mapped directly onto metric spaces.

The category of metric spaces is of course defined by objects and morphisms. An object is a metric space $(X, d)$ and a morphism is a mapping...
between two metric spaces \((X, d) \rightarrow (X', d')\). As in any other category, the composition of morphisms must satisfy associativity and identity.

Let us define first a metric space. The structure \((X, d)\) for \(X\) a set and the function \(d : X \times X \rightarrow \mathbb{R}^+\) satisfies:

1. \(d(x, y) = 0\) when \(x = y\)
2. \(d(x, y) = d(y, x)\) and
3. \(d(x, z) \leq d(x, y) + d(y, z)\)

The Euclidean distance is a map \(d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^+\). For \(n = 2\) the distance is \(d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}\).

Typically the function \(d\) is assumed to be the Euclidean distance, but of course, other distances are possible. One example of a metric that satisfies the three axioms above is the “Manhattan distance” or \(d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^+\), so for a two dimension space, \(d((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2|\).

**Definition 1.** A mapping \(f : (X, d) \rightarrow (X', d')\) preserves distances if for all pair of points, \(x_1, x_2 \in X\), it holds \(d(x_1, x_2) = e(f(x_1)f(x_2))\).

**Definition 2.** A function \(f : (X, d) \rightarrow (Y, e)\) between two metric spaces is continuous at \(x_0 \in X\) if for all \(\epsilon > 0\) there exists \(\delta > 0\) such that if \(d(x, x_0) < \delta\) then \(e(f(x_0), f(x)) < \epsilon\).

A contraction is a Lipschitz map with \(\lambda < 1\), while a map between two metric spaces \(f : (X, d) \rightarrow (X', e)\), is such that \(d(x_1, x_2) = e(f(x_1)f(x_2))\), is a distance preserving map. Note that every Lipschitz function is continuous and as a contraction is a Lipschitz map with \(\lambda < 1\), contractions are continuous [7].

Now we are able to define the category **Met** of metric spaces and Lipschitz maps that are structure preserving maps. The composition of Lipschitz maps, \(gf\), is a Lipschitz map and the properties associativity of composition and identity \(id_x : (X, d) \rightarrow (X, d)\), are trivially demonstrated.

The topological spaces are useful when we are interested in closeness and continuity rather than in distance as it is the case in metric spaces. The category of topological spaces **Top** is one that has topological spaces as objects and continuous maps as morphisms.

### 2.3 Functor between Neur and Field

At this point we wish to define the functor between the categories that have been defined.

Let us suppose that **Neur+** is a category whose objects are sets of neurons and the arrows all the functions between them, so in the case that there is only one place cell, the category **Neur+** is a set of a single element or singleton. For an object of category **C**, there is an unique functor \(F : C \rightarrow 1\). Thus,
there is a unique functor from the category of metric spaces and Lipschitz-distance perserving maps, \textbf{Field}_+, and the category of one place cell 1.

Functors preserve isomorphisms, so given the functor $F : C \to D$, the isomorphisms in category $C$ are preserved in category $D$.

An interesting feature of functors is that they may preserve properties. For example, since functors preserve composition of morphisms $\circ$, and identities, $id$, they preserve every property that can be positively expressed in the language of $\circ$ and $id$. In particular they preserve commutativity of diagrams [3]. So given a functor $F : C \to D$, for certain objects, arrows or composition of arrows in category $C$, that have the property $p$, the functor $F$ brings such property to the $F$-image.

\textbf{Definition 3.} Let $C$ and $C'$ two categories, a covariant functor $F$ from $C$ to $C'$ is defined as i. a rule which associates for every object $A$ in $C$ an object $F(A)$ in the category $C'$ and ii. a rule that associates for every morphism $\alpha : A \to B$ in $C$ a morphism $F(\alpha) : F(A) \to F(B)$ in the category $C'$. Then $F$ must satisfy the following two conditions:

ii.a The composition is preserved: for the diagram (Diagram is formally defined in 6) $A \xrightarrow{\alpha} B \xrightarrow{\beta} C$ in $C$, $F(\alpha \circ \beta) = F(\alpha) \circ F(\beta)$
ii.b Identities are preserved: for any object $A$ in the category $C$, $F(id_A) = id(F_A)$

Now, the functor (more precisely a covariant functor) from a category of neurons $\text{CAT-Neur}$ to the category $\text{Met}$ of metric spaces, $F : \text{CAT-Neur} \to \text{Met}$ is such that:

i every object $N$ in the category of neurons $\text{CAT-Neur}$ is mapped onto an object $F(N)$ in the category $\text{Met}$.
ii every morphism $\alpha : N \to N'$ in $\text{CAT-Neur}$ is mapped onto a morphism $F(\alpha) : F(N) \to F(N')$ in the category $\text{Met}$. $F$ preserves composition and identity.

ii.a The composition is preserved, so $A \xrightarrow{\alpha} B \xrightarrow{\beta} C$ in $\text{CAT-Neur}$, $F(\alpha \circ N \beta) = F(\alpha) \circ F(\beta)$ (both sides of the equation are morphisms in $\text{Met}$)
ii.b Identities are preserved, so for any object $A$ in the category $\text{CAT-Neur}$, $F(id_A) = id(F_A)$ (both sides of the equation are morphisms in $\text{Met}$)

The physiological interpretation of the functor is as follows. i means that it is possible for any object $N$ in the category of neurons $\text{CAT-Neur}$ to have associated a metric space $(X, d)$. As it was stated in 2.1, the objects in the sets of category $\text{CAT-Neur}$ are sets of neurons.

Note that this is different to assign a location to a set of neurons, rather we are associating a set of neurons with a metric space $N \to (X, d)$, where $d : X \times X \to \mathbb{R}^+$. 

For example, let $\text{Met}1$ be the category of metric planar space of diameter 1, $(M, \upsilon)$, that is, $d(m, m') \leq 1$ for all $m, m' \in \mathcal{M}$. Then $F(N), F : N \rightarrow (M, \upsilon)$, represents that the object $N$, a set of neurons, falls into a circumference of diameter 1 in the two-dimensional space $M$.

On the other hand, if we take for the category $\text{CAT-Neur}$ the category $\text{Neur}$, then condition $ii$ can be interpreted as follows, whenever there is a synapse between two neurons $n, n', \alpha : n \rightarrow n'$, there is a relationship between the metric spaces associated to each of the synaptic neurons, $F(\alpha) : F(N) \rightarrow F(N')$, such that $F$ is a map that preserves composition and identity.

Let $A \xrightarrow{\alpha} B \xrightarrow{\beta} C$, then $F(\alpha \circ \beta) = F(\alpha) \circ F(\beta)$ simply means that the map associated to a synaptic path is equivalent to the map associated to the synapses.

The other requirement, identity is preserved, can be interpreted as there is always a metric space for any neuron.

It might be remarked that the functor $F$ defined here, does not preserve the metric space defined in the category $\text{Met}$. This is in accordance with the empirical fact that the brain has no metric or at least not a Euclidean-like metric based on the concept of distance.

Indeed, what $F$ does is to bring the structure of the category of neurons over to the category of metric spaces $\text{Met}$. The very different nature of the two categories that are being mapped by $F$, makes difficult to see how $F$ works.

With an example we will try to make this point more clear. Let the objects of $\text{Neur}$ be place cells, that is, neurons that fire when the brain occupies a position in a plane surface like for example a maze or a box. The metric space for the environment is given by the category $\text{Met}$. For every synapse $\alpha$ coupling two place place cells, $N$ and $N'$ in $\text{Neur}$. $F(N)$ and $F(N')$ are called the place fields of $N$ and $N'$ respectively in the category $\text{Met}$.

Thus, the mapping $F$, in order to be a functor needs to be a structure preserving map between $\text{Neur}$ and $\text{Met}$, the two categories being mapped by $F$. In the case that $\text{CAT-Neur}$ is $\text{Neur}$ whose objects are neurons, the relationship between the place field of the postsynaptic cell $F(N')$ and the place field of the presynaptic cell $F(N)$ may be exemplified by $d(F(N_i), F(N_j)) \leq d(N'_i, N'_j)$, where $N_i, N_j$ are in category $\text{Neur}$, and $N'_i, N'_j$ in category $\text{Met}$.

3 A new framework for place and grid cells

Here we propose a radically new theoretical framework for the formation of place cells from grid cells. The computational models of the hippocampus [2], [12], [20] state that the sum of a set of elements, grid cells, directly produce another element, a place cell. In doing so, these models take for granted that the properties of the sum are directly reducible to those of its
components. This strict form of reductionism is at odds with the nature of complex systems. It is necessary to tackle the system as a whole, and bring to light the way in which the components interact, producing higher levels of complexity, exemplified in new systemic properties that are not present in the single components.

It might be remarked here, that this is not a criticism of the reductionist approach. Indeed the reductionist analysis is arguably the best plan of action that one may follow in order to understand how a system works. But this is just the half of the work, the synthetic endeavor must follow after the analysis.

In what follows, we describe the effect in injecting the concepts of co-product and colimit from category theory into the problem of place cell formation in the hippocampus.

The classical reductionism credo states that the whole is no more than the sum of its parts. Therefore the properties of the sum are reduced to those of its components, without introducing new properties. This is what the categorical concept coproduct exemplifies.

In a given category, all one needs to know is about the components $A_i$ in the coproduct $\coprod_i A_i$, this is possible because all the components play a symmetrical role in the construction coproduct.

**Definition 4.** A coproduct of two objects $A$ and $B$ is a an object $A + B$ together with the arrows $A \xrightarrow{\iota_1} A + B$ and $B \xrightarrow{\iota_2} A + B$, such that for any object $C$ and the pair of arrows $A \xrightarrow{\alpha} C$, $B \xrightarrow{\beta} C$, there exists an unique morphism $\pi$ that makes the diagram commutes.

\[ \begin{array}{ccc}
A & \xrightarrow{\iota_1} & A + B \\
\downarrow{\alpha} & & \downarrow{\pi} \\
A + B & \xrightarrow{\iota_2} & B
\end{array} \]

Coproduct generalizes to the direct sum shown in the next diagram[4]

\[ \begin{array}{ccc}
\coprod_i A_i & \xrightarrow{\pi_i} & A_i \\
\downarrow{\iota_i} & & \downarrow{\alpha_i} \\
C & \xrightarrow{\beta_i} & A_i
\end{array} \]
On the other hand, the colimit $cP$ embodies the collective operations made by the family of components $A_i$ which are made possible because the components cooperate by means of the links that connect them [14].

The colimit in a category of a family of components $A_i$ without any arrow between them is the coproduct.

The colimit, contrary to the coproduct, entails a non symmetric relationship with its components. As the figure 5 depicts, the coproduct can be compared to the colimit $cP$. This symmetry breaking process may be somehow quantified by the arrow $s$.

![Diagram](image)

**Fig. 5** The family of objects $A_1, A_2, A_3, A_4$ has both a colimit $cP$ and a coproduct $\coprod_i A_i$. The coproduct is linked by $s$ to the colimit. The link $s$ express the transit from the coproduct to the colimit and embodies the symmetry breaking in the relationship between the family of objects $A_i$ and the colimit.

### 3.1 Place field as colimit of grid fields

The cooperation of the grid fields gives rise to the colimit which is a place field. Thus the colimit of the metric system depicted in the figure 6 can be seen as an integrator of the information contained in the metric system components. It might be remarked that the colimit is an object of the category **Field**, a sort of complex object that actualizes the internal organisation of the objects that it is binding. Colimits and limits do not exist for all diagrams in all categories, but if they exist, they are unique up to isomorphism. The mathematical definition of colimit needs a prior definition, that of diagram.
Definition 5. A diagram $D$ in a category $C$ is a collection of vertices and directed edges consistently labeled with objects and arrows of the category $C$. Thus, if an edge in the diagram $D$ is labeled with an arrow $f$ such that $f : A \rightarrow B$ in $C$, then the vertices this edge in the Diagram $D$, must be labeled $A$ and $B$ [4].

Definition 6. Let $D$ be a diagram in a category $C$ with objects labeled $D_i$ and morphisms labeled $f_k : D_i \rightarrow D_j$. We call cocone $K$ for diagram $D$ to the apical object $B$, together with the set of morphisms $g_i : D_i \rightarrow B$ forming a commutative diagram, that is, $g_j \circ f_k = g_i$.

Given the cocones $K'$ and $K''$ for $D$, a cocone morphism $h : B' \rightarrow B''$ is a morphism in $C$ such that $g''_i = h \circ g'_i$. To simplify the notation we denote the cocone morphism determined by $h$ as $h : K \rightarrow K'$. Directly, the cocones form a category, the category of cocones $\text{coc}_D$.

Definition 7. A colimit for the diagram $D$ is an initial object $K$ in the category $\text{coc}_D$, that is, for any other cocone $K'$ for diagram $D$, there exists a unique cocone morphism $h : K \rightarrow K'$.

It follows from the definition that all colimits are isomorphic because all initial objects are isomorphic.

Fig. 6 A colimit $K$ for the base diagram $D$. For the sake of clarity in the figure the diagram $D$ has three objects $A_i, i = 1, 2, 3$.

The figure 7 shows that grid fields and grid cells in the medial entorhinal cortex (mEC), are linked by a map, likewise there is a map between place cells and place fields. Therefore for each grid cell, there is grid field, which is a metric space with the form of a regular hexagon, and for each place cell, there is one place field which is also an object of the category of metric spaces, Field, but in this case, its geometry is a simple point rather than a hexagon.

We can assume that the neurons -place cells and grid cells- depicted in the bottom of the figure, are in the category Neur having as objects neurons and as morphisms synaptic connections.
Fig. 7 The figure depicts a colimit where (4) acts as the place field of a place cell (6) in the hippocampus. The colimit is produced by several grid fields (one grid field (1) is produced by one grid cell (5)).

However, this is not always the case. For example, a category of neurons whose objects contain several neurons connected between them forming populations of neurons, rather than single neurons.

In this line, it is particularly valuable to shed light on how populations of grid cells contribute to the formation of one place cell. The colimit is the mathematical structure that allow us to encode the emergence of place field and the relationship between grid fields.

Now let us focus on the grid fields depicted as hexagons in figure 7 and their morphisms. It has been said above that regular hexagons are objects in the category Field, now we need to investigate the morphisms between the grid-field object in this category.

A contraction between two grid-field objects $G_1, d, o, \psi, G_2, d, o, \psi$ is a continuous function $f : (G_1, d, o, \psi) \rightarrow (G_2, d, o, \psi)$, satisfying $d(f(x, y)) \leq d(x, y)$ and $o(f(x, y)) \leq o(x, y)$. 
This restriction is in accordance with the experimental finding that shows that spacing in grid fields, increases along the dorsoventral axis in the medial entorhinal cortex (mEC). This fact appears to be correlated with the increase in size of place fields along the dorsoventral axis of the hippocampus [5], [9].

Neighbor cells in the mEC have similar spacing and orientation. However, there is no evidence that anatomical cell clusters, correspond to functionally segregated grid maps with their own spacing and orientation [17].

On the other hand, the phase of the grid, does not follow the restriction of continuity that spacing and orientation have. Indeed, firing vertices of colocalized grid cells are shifted randomly, that is to say, the mapping between vertices in the grid field and the external reference grid is not continuous. This is in fact how fields of neighboring hippocampal place cells behave.

The colimit is a universal property which is a remarkable fact that deserves to be explained. When a mathematical construction, in our case a colimit, satisfies an universal property, one can forget the details of the structure and focus on the universal property because all that has to be known about the colimit, is exemplified in the universal property.

One important point that needs emphasis is that the existence of a colimit, which imposes constraints, not only on the diagram of grid cells that determine the colimit, but also on all the objects of the category. Besides, the colimit, if it exists, is uniquely determined (up to isomorphism) but the reverse is not true, one colimit can have several decompositions. Put it in the context of figure 7, this means that when the coordinated activity of a group of grid cells produce a place cell, this is a colimit and it is unique. But given a place cell, its place field cannot be uniquely determined by a group of grid cells, as a matter of fact, several grid fields are possible for that place field.

4 A theory of declarative memory (episodic and semantic) based on category theory

The dual role of the hippocampus in formation and retrieval of concepts is not surprising, especially considering that the formation of new memory (knowledge) requires the retrieval of the old one. Thus, memory is knowledge, and perception is a condition of possibility of memory and therefore of knowledge.

Just as any other higher cognitive function, to try to give a definition of memory seems hopeless. The definition in the MIT encyclopedia of cognitive sciences [10] is open enough to satisfy everyone: “the term memory implies the capacity to encode, store, and retrieve information”. However, it is also too unrestricted to provide a clear idea about what memory is and how it works.

Certainly, memory is not an univocal term, it has several forms that depend on different brain mechanisms. So a well-founded strategy to get an
understanding of how memory works is to investigate how such cognitive process is implemented in the brain.

The idea behind this is that the layman’s view of memory, which is still commonly used, which will become irrelevant once the biological mechanisms of memory have been uncovered and, if possible, described in mathematical terms.

The main point that is being explored in this heading is that despite the diverse nature of episodic and semantic memory, it is possible to connect them via categorical objects like product, pullback or colimit.

Let us begin by the introduction of the categorical product and its application in a navigational task in one dimension, after which the results will be expanded to navigation in a two-dimensional arena and the use of the categorical concept of pullback.

### 4.1 Categorical product in acquisition of middle point concept in 1D navigation

Suppose a rat is placed in a track (one dimensional environment), the animal immediately starts moving back and forth in order to get and idea of the dimensions of the environment. As the rat moves from one point to the other, episodic memories are created. Thus the animal is able to make the association of self-centered information with the temporal order in which the different positions are reached.

Episodic memories are not explicit, contrary to the episodic memory, the explicit ones may be retrievable independent of the internal state of the rat.

Suppose there is no particular visual or smell stimulus that can make the rat remember any particular position. One may think that after a while, the rat will acquire an explicit memory, for example the concept of middle point which exemplifies the position in the track, from where it needs the same amount of time to get any of the extremes.

A cognitive behavior involves integration of information. The categorical concept of product is a formalisation of integration. Moreover, as it will be shown later, a product in a category that admits a final object, is an instance of a more general categorical form, pullback.

**Definition 8.** In a given category $\mathcal{C}$, a product of two objects $A$ and $B$, is another object $P$ equipped with two morphisms, $P \xrightarrow{p_1} A$ and $P \xrightarrow{q_1} B$, such that for any pair of morphisms, $X \xrightarrow{f} A$ and $X \xrightarrow{g} B$ there is an unique morphism $h$ making the following diagram commutes.

Note that the broken arrow $h$ means that is unique, and the morphisms $p_1, p_2$ are usually called projection morphisms. The main characteristic of a product is that the constituents are retrievable via the projection morphism.
The following diagram indicates the categorical product for the acquisition of the middle point.

For our purpose, the categorical product given by the object $P$ and the morphisms $p_1, p_2$ is a statement about a cognitive behavior of the rat, whereas $X$ and $x_1, x_2$ is a constraint on what constitutes a valid product, rather than a specific claim about cognition.

Note that there is not any particular commitment in the morphisms $p_1, p_2$. In fact, $p_1$ can mean the travel time to reach the wall $A$, $W_A$, but also the number of steps needed.

In figure 8, is represented one possible experimental setting, that could be used to explore the biological plausibility of our theory in the acquisition of the middle point concept in a rat moving in a single track maze. Thus $P$, $W_A$ and $W_B$ are objects in the category $C$ of memories or mental objects that will be described with more detail in future works.

4.2 Categorical pullback in acquisition of middle point concept in 2D navigation

Now suppose the rat is removed from the one dimensional track depicted in figure 8 and put upon a plane. The rat’s capacity to build the explicit memory for the middle point of the arena can be seen as analogous to the generalised product, i.e: a pullback.

**Definition 9.** In a category $C$, a pullback of two morphisms with common codomain $A \xrightarrow{f} C \xleftarrow{g} B$ is an object $P$ together with a pair of morphisms $P \xrightarrow{p_1} A$ and $P \xrightarrow{p_2} B$ that form a commutative diagram $f \circ p_1 = g \circ p_2$.

Moreover, the morphisms are universal among such squares because for any pair of morphisms $Z \xrightarrow{z_1} A$ and $Z \xrightarrow{z_2} B$ such that $f \circ z_1 = g \circ z_2$, there is an unique morphism $h$ such that the following diagram commutes.
A pullback may be seen as a constrained product, being the constrain given by \( f \) and \( g \), \( f \circ p_1 = g \circ p_2 \).

**4.3 Pullback and grid cell formation**

The concept of pullback may be useful in dealing with grid cells. The way in which grid cells are calculated in literature is tricky. One of the three
parameters refers to how accurate the representation outside the cell is. In
doing so you are putting the outcome of the system in the input.

\[
\begin{array}{c}
P \xrightarrow{p_1} A \\
\downarrow \quad \downarrow \\
B \xrightarrow{p_2} C
\end{array}
\]

In the diagram above, \( P \) can be seen as a grid cell where the projection
morphisms \( p_1 \) and \( p_2 \) refer to the orienting and the spacing respectively. The
morphisms \( f \) and \( g \) impose additional restrictions in the grid cell, like for
example the constant value of those parameters all over the arena.

## 5 Discussion

A theory that fully explains and predicts the highly complex cognitive abil-
ities like perception, memory or learning has not been produced yet. Our
society needs to deal with diseases like for example Alzheimer’s disease that
is ravaging a big sector of the population. It goes without saying that to shed
light on the role played by the hippocampal system in cognitive functions
like memory and learning can be of extraordinary value for the future of our
own species.

We must exploit the important fact that from the point of view of neurobi-
ological knowledge, memory and perception share the same neural substrate.
The time is ripe for a mature and rigorous approach to brain structure
and function that sets the basis for a shareable scientific framework, able to
carry out knowledge, commonly understandable among the different actors
in the brain sciences.

## References

1. Redish A. The hippocampal debate: Are we asking the right questions? Behavioral
2. Touretzky D.S. Redish A.D. Theory of rodent navigation based on interacting repre-
3. H. P. Adámek J. Gumm and Trnková V. Presentation of Set Functors: A Coalgebraic
5. Jung MW. Wiener SI. McNaughton BL. Comparison of spatial firing characteristics
6. Fuhs MC. Touretzky DS. A spin glass model of path integration in rat medial entorhi