



# Redes Neuronales

Ismael Acosta  
([iacosta@fcien.edu.uy](mailto:iacosta@fcien.edu.uy))  
Facultad de Ciencias, UdelaR

# Contenido de la clase

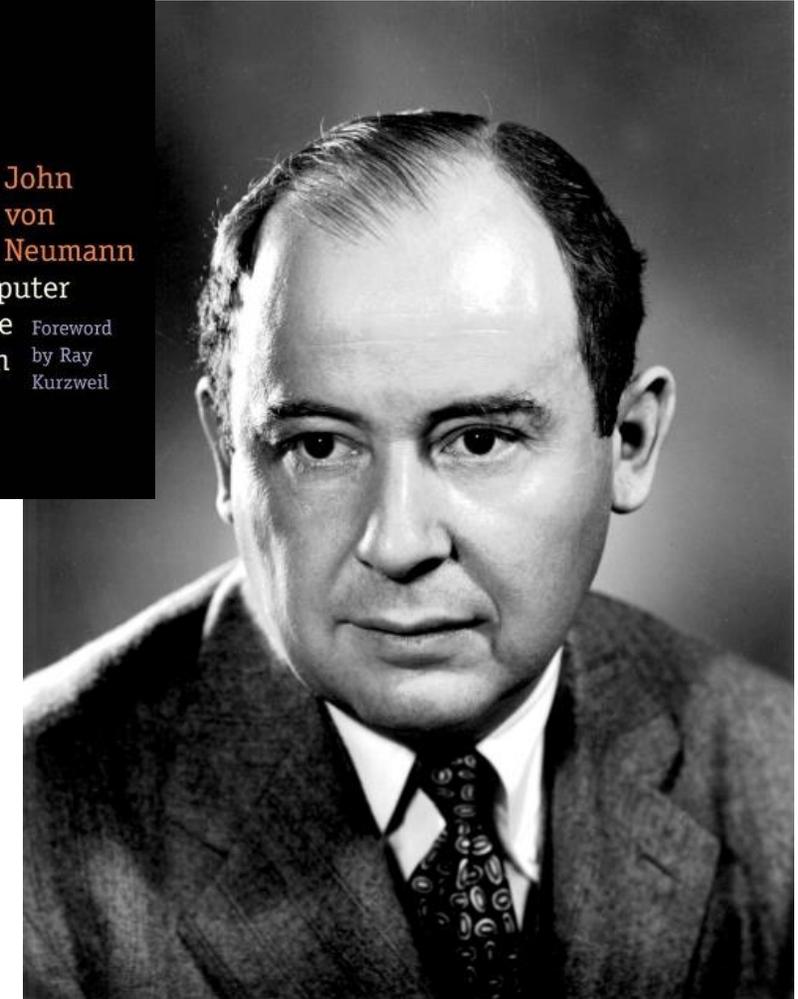
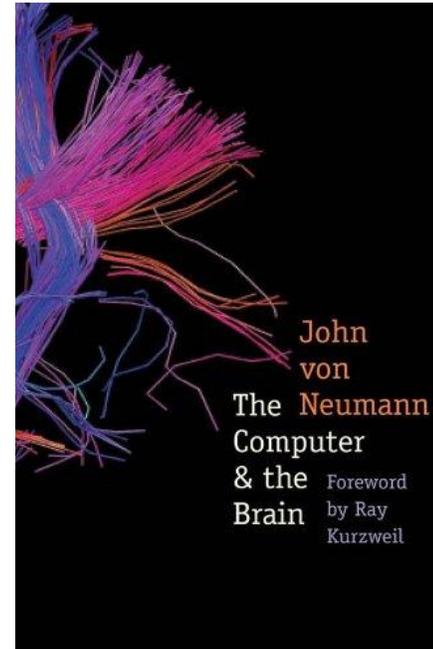
- Introducción: Teoría Computacional del Cerebro
- Modelo de McCulloch-Pitts
- Modelo de Memorias Asociativas

# Contenido de la clase

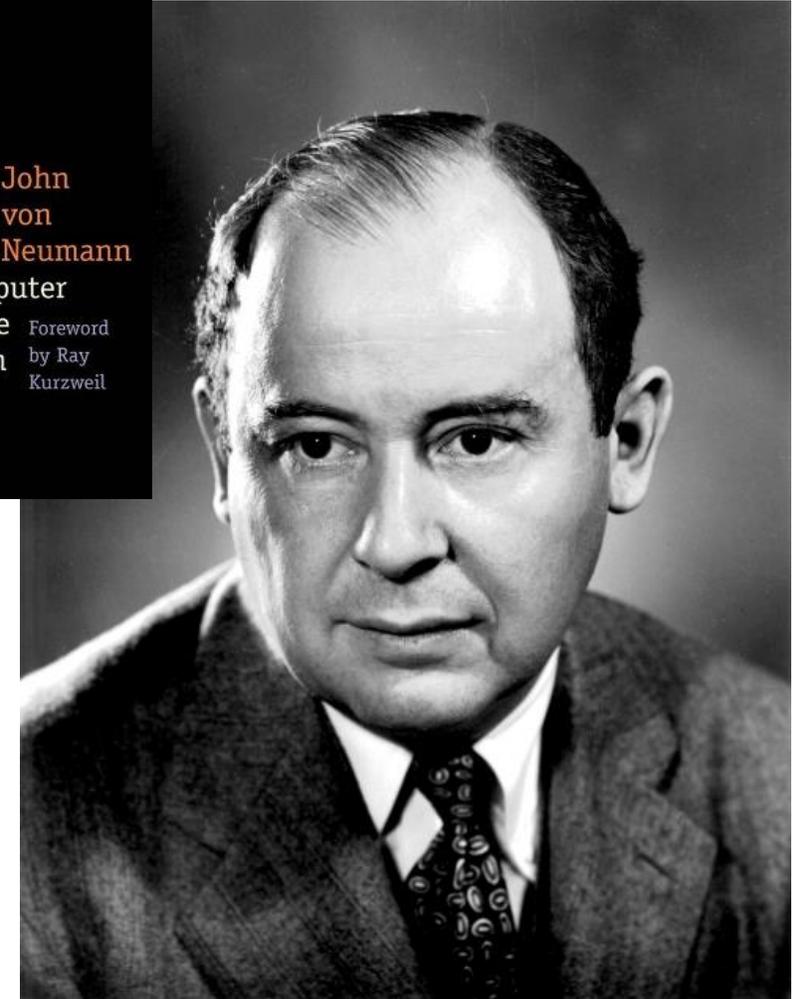
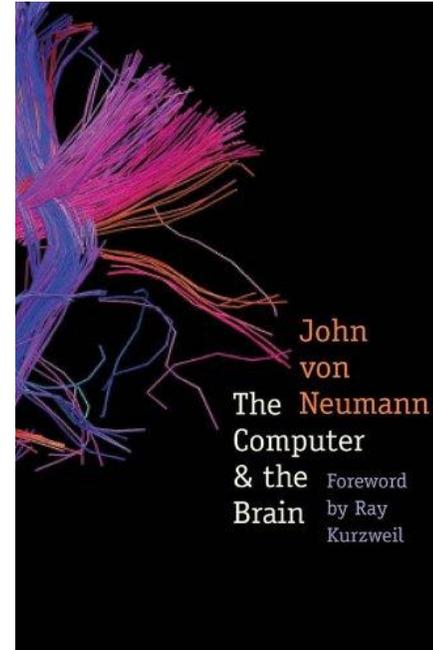
- **Introducción: Teoría Computacional del Cerebro**
- Modelo de McCulloch-Pitts
- Modelo de Memorias Asociativas

Desde los comienzos de la Era de la Computación existió una fuerte conexión con la Biología del Cerebro, al punto de desarrollarse una **Teoría Computacional del Cerebro**, que involucra aspectos biológicos, psicológicos, matemáticos, computacionales y filosóficos<sup>[1]</sup>.

[1] [Von Neumann, J. \(1958\). \*The Computer and the Brain\*. Yale University Press.](#)

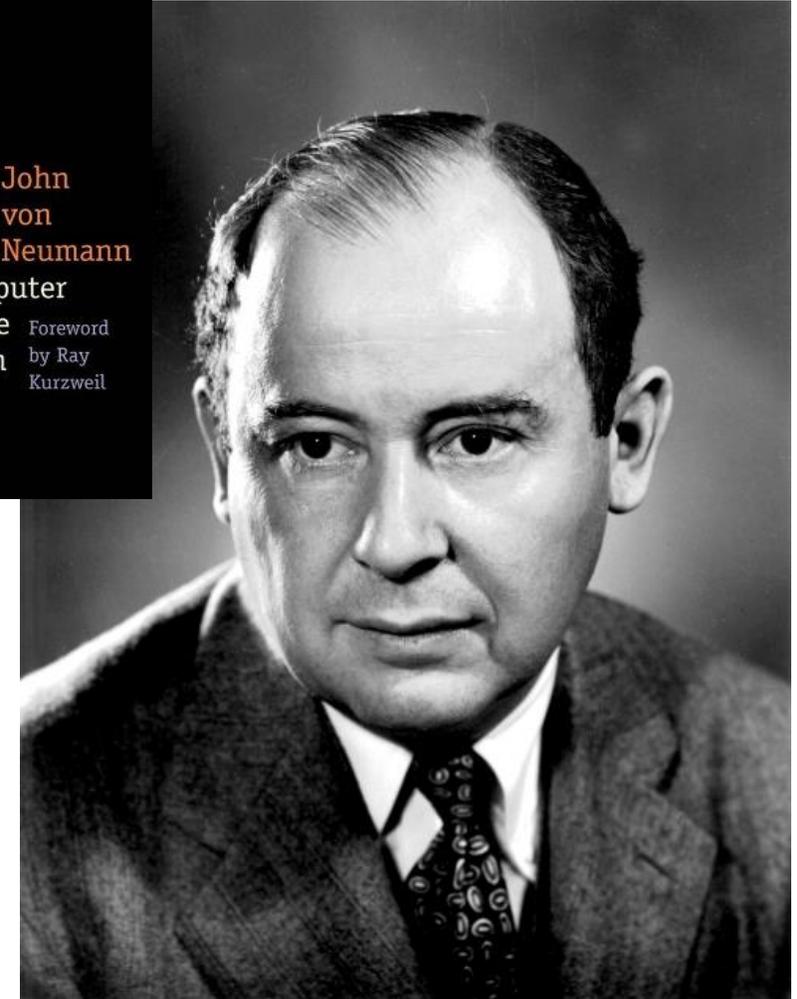
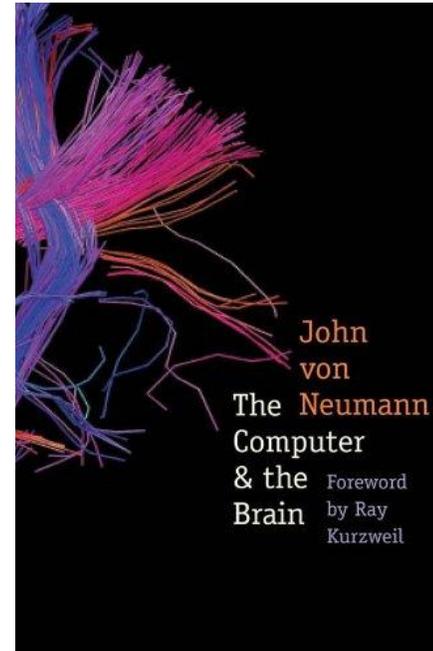


Esta teoría sostiene que el cerebro puede concebirse como un sistema de cálculo y procesamiento de información, con un funcionamiento (a través de la sinapsis y los potenciales de acción) y una arquitectura (formada por neuronas) *comparables* a de un computador.



Diferentes modelos biofísicos intentan explicar el funcionamiento de las redes neuronales utilizando ideas desarrolladas por la matemática y la lógica computacional.

Estos modelos deben estar en sintonía con el conocimiento adquirido por la Biología (biología celular, bioquímica, fisiología, etc.).



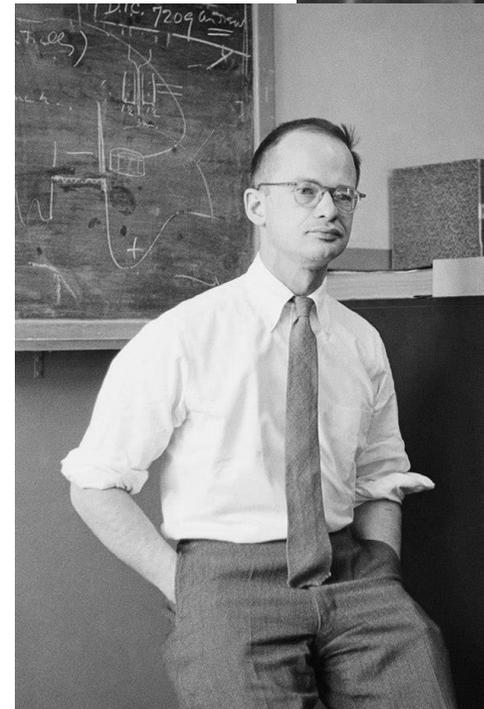
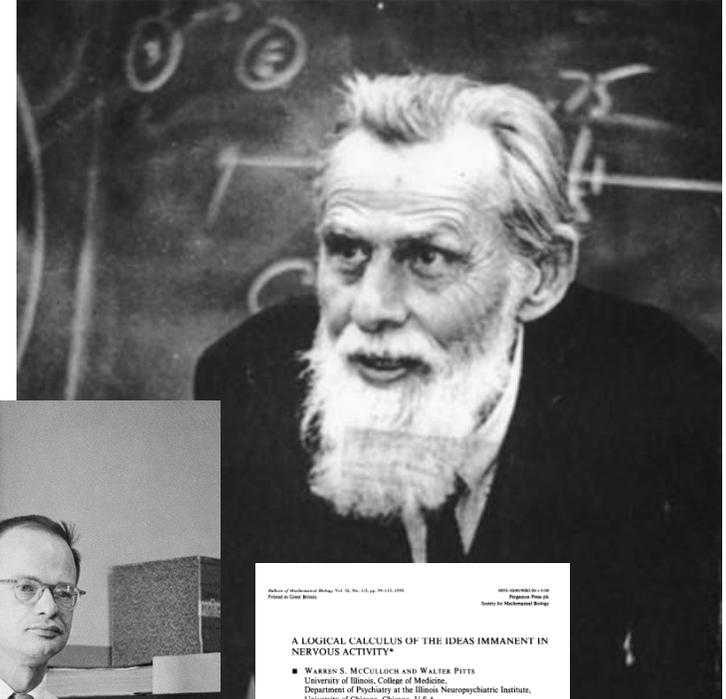
# Contenido de la clase

- Introducción: Teoría Computacional del Cerebro
- Modelo de McCulloch-Pitts
- Modelo de Memorias Asociativas

# Contenido de la clase

- Introducción: Teoría Computacional del Cerebro
- **Modelo de McCulloch-Pitts**
- Modelo de Memorias Asociativas

En 1943 McCulloch y Pitts publican un modelo matemático<sup>[2]</sup> que intentó describir la actividad de una red neuronal, basándose en la respuesta binaria del **Principio de “Todo o Nada”** que caracteriza a los **Potenciales de Acción**, en donde solo ocurrirán si el voltaje inducido supera cierto *umbral* activador.



**A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY\***

WARREN S. MCCULLOCH AND WALTER PITTS  
University of Illinois, College of Medicine,  
Department of Psychiatry at the Illinois Neuropsychiatric Institute,  
University of Chicago, Chicago, U.S.A.

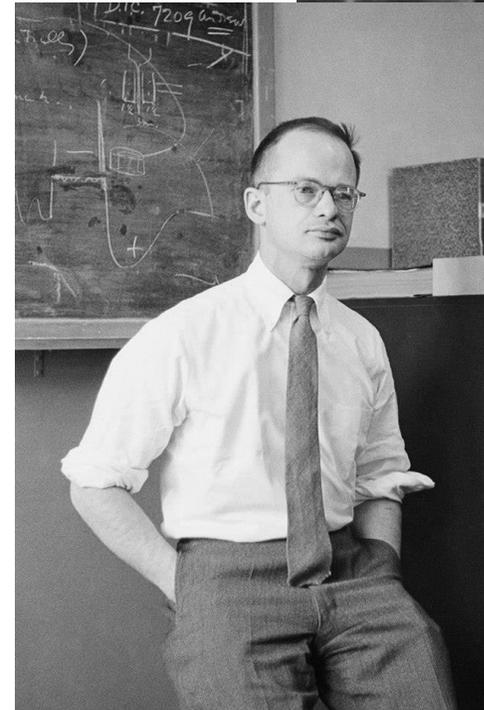
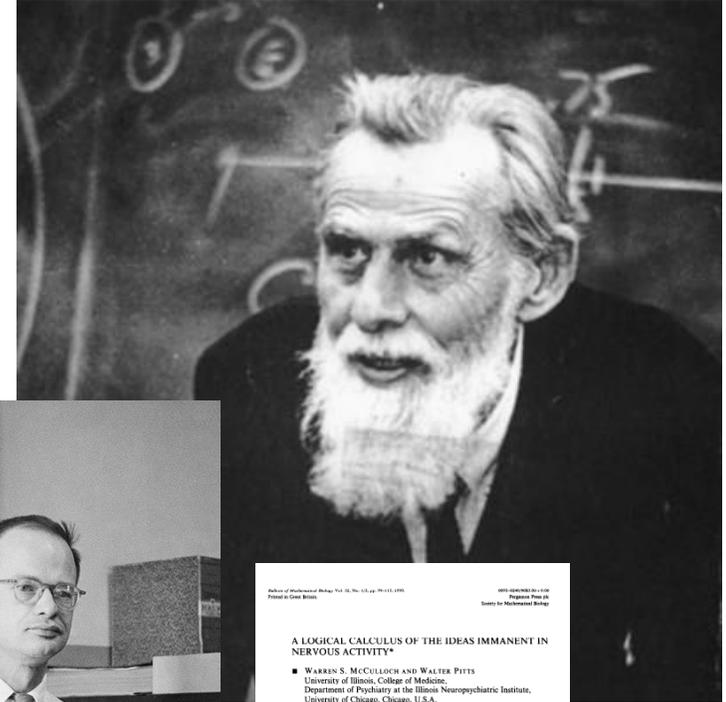
Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for excitation circles, and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular devices among possible neurophysiological arrangements are equivalent, in the sense that for every set following under one assumption, there exists another set which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.

**1. Introduction.** Theoretical neurophysiology rests on certain cardinal assumptions. The nervous system is a set of neurons, each having a soma and an axon. Their adjunctions, or synapses, are always between the axon of one neuron and the soma of another. At any instant a neuron has some threshold, which excitation must exceed to initiate an impulse. This, except for the fact and the time of its occurrence, is determined by the neuron, not by the excitation. From the point of excitation the impulse is propagated to all parts of the neuron. The velocity along the axon varies directly with its diameter, from  $< 1 \text{ ms}^{-1}$  in thin axons, which are usually short, to  $> 150 \text{ ms}^{-1}$  in thick axons, which are usually long. The time for axonal conduction is consequently of little importance in determining the time of arrival of impulses at points unequally remote from the same source. Excitation across synapses occurs predominantly from axonal terminations to somata. It is still a moot point whether this depends upon irreversibility of individual synapses or merely upon prevalent anatomical configurations. To suppose the latter requires no hypothesis *ad hoc* and explains known exceptions, but any assumption as to cause is compatible with the calculus to come. No case is known in which excitation through a single synapse has elicited a nervous impulse in any neuron, whereas any neuron may be excited by impulses arriving at a sufficient number of neighboring synapses within the period of latent addition, which lasts  $< 0.25 \text{ ms}$ . Observed temporal summation of impulses at greater intervals

\* Reprinted from the *Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115-133 (1943).

[2] [McCulloch, W. S., & Pitts, W. \(1943\). A logical calculus of the ideas immanent in nervous activity. \*The bulletin of mathematical biophysics\*, 5\(4\), 115-133.](#)

Bajo este modelo, las redes neuronales pueden ser construidas de tal forma que son capaces de computar **funciones de la lógica proposicional** (*funciones booleanas*), y como tal constituirían una forma análoga a un computador. La *complejidad* de las funciones computables sería proporcional a la cantidad de neuronas involucradas y el número de conexiones entre ellas.



**A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY\***

WARREN S. MCCULLOCH AND WALTER PITTS  
University of Illinois, College of Medicine,  
Department of Psychiatry at the Illinois Neuropsychiatric Institute,  
University of Chicago, Chicago, U.S.A.

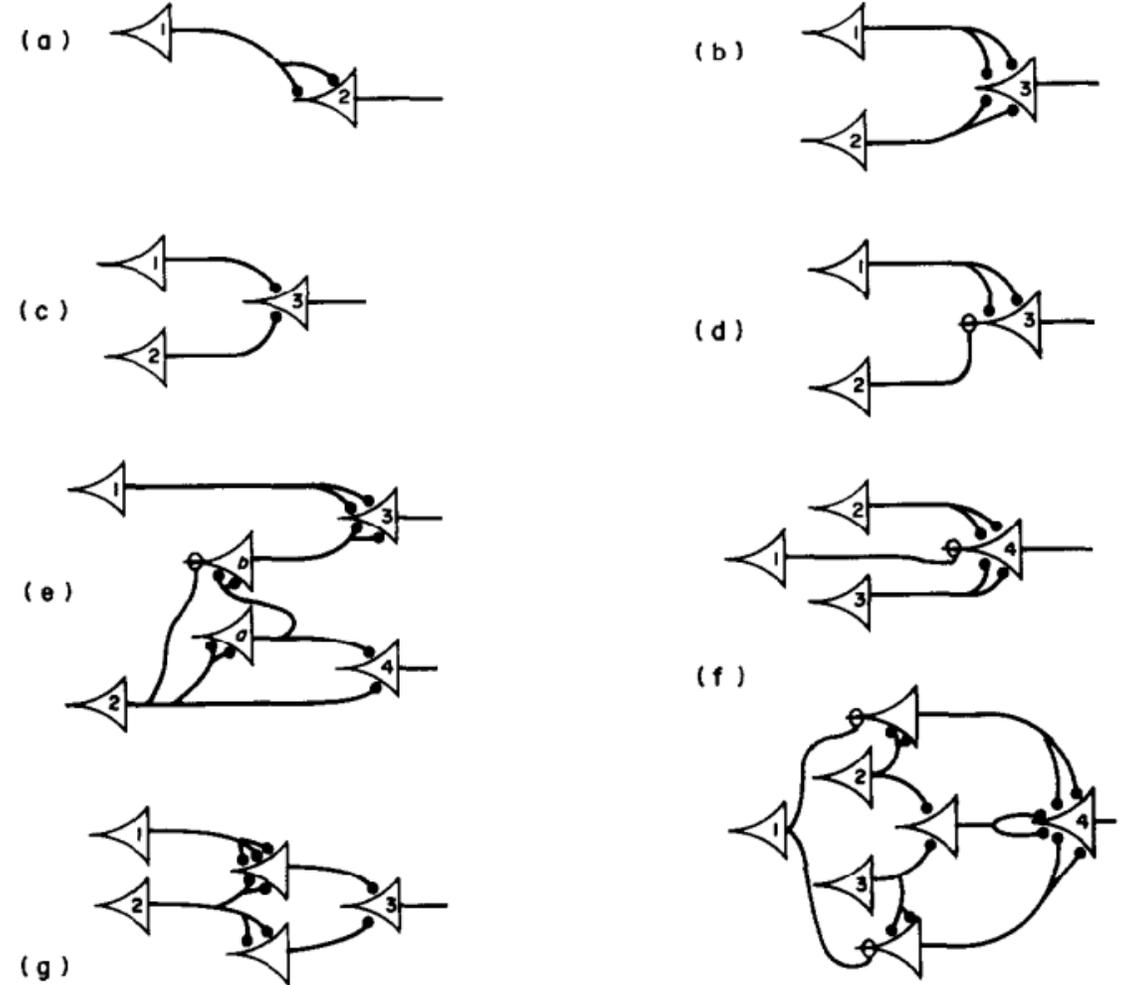
Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for nets containing circles; and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular devices among possible neurophysiological arrangements are equivalent, in the sense that for every set behaving under one assumption, there exists another set which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.

**1. Introduction.** Theoretical neurophysiology rests on certain cardinal assumptions. The nervous system is a set of neurons, each having a soma and an axon. Their adjunctions, or synapses, are always between the axon of one neuron and the soma of another. At any instant a neuron has some threshold, which excitation must exceed to initiate an impulse. This, except for the fact and the time of its occurrence, is determined by the neuron, not by the excitation. From the point of excitation the impulse is propagated to all parts of the neuron. The velocity along the axon varies directly with its diameter, from  $< 1 \text{ ms}^{-1}$  in thin axons, which are usually short, to  $> 150 \text{ ms}^{-1}$  in thick axons, which are usually long. The time for axonal conduction is consequently of little importance in determining the time of arrival of impulses at points unequally remote from the same source. Excitation across synapses occurs predominantly from axonal terminations to somata. It is still a moot point whether this depends upon irreversibility of individual synapses or merely upon prevalent anatomical configurations. To suppose the latter requires no hypothesis *ad hoc* and explains known exceptions, but any assumption as to cause is compatible with the calculus to come. No case is known in which excitation through a single synapse has elicited a nervous impulse in any neuron, whereas any neuron may be excited by impulses arriving at a sufficient number of neighboring synapses within the period of latent addition, which lasts  $< 0.25 \text{ ms}$ . Observed temporal summation of impulses at greater intervals

\* Reprinted from the *Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115-133 (1943).

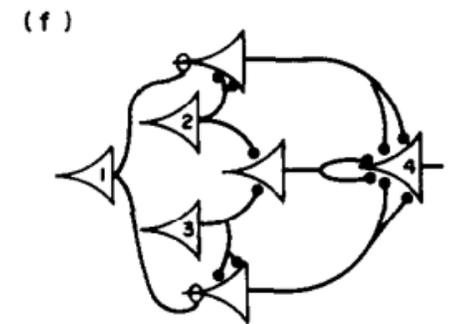
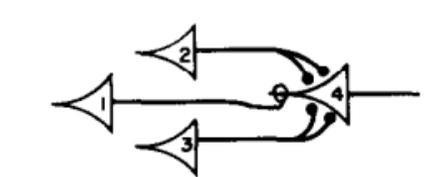
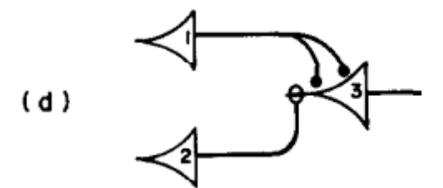
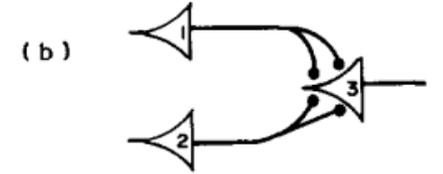
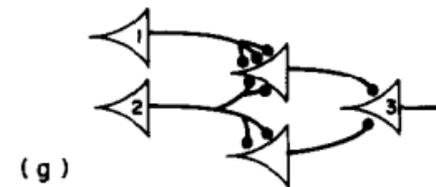
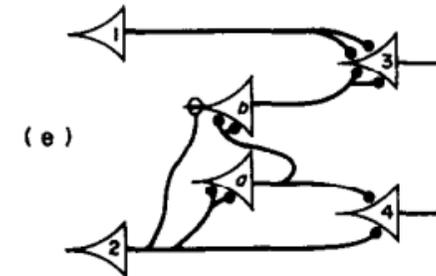
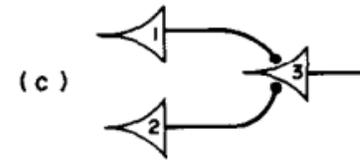
Los supuestos del modelo de MP<sup>[2]</sup> intentan reflejar las características de la fisiología neuronal:

- Cada neurona puede tener solamente **2 estados**: excitado y en reposo.
- Todas las neuronas tienen un valor de **voltaje umbral** fijo.

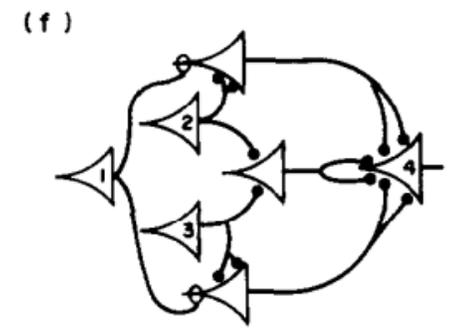
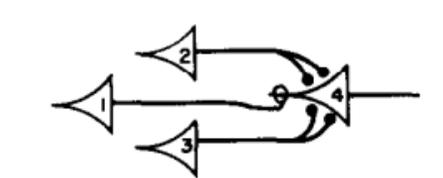
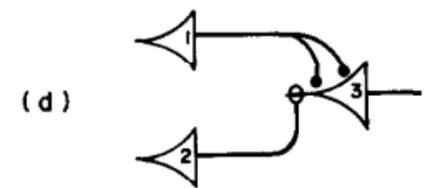
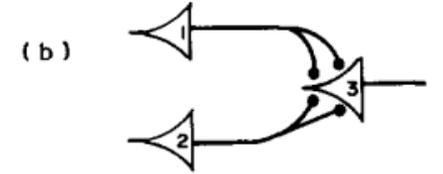
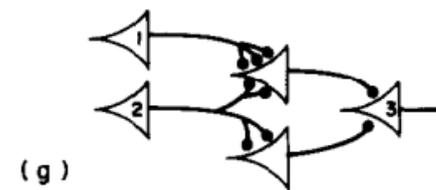
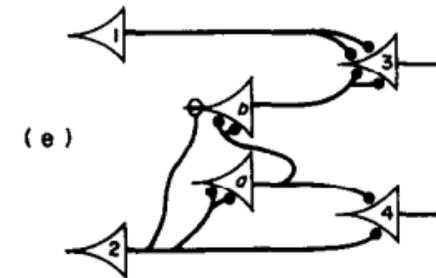
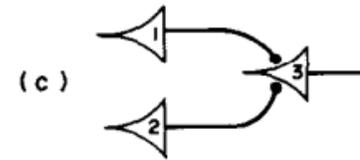


[2] McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115-133.

- Las neuronas reciben pulsos de entrada (excitatorios o inhibitorios) de otras neuronas, con diferentes **pesos sinápticos**.
- La respuesta sináptica está dada por la sumatoria de las señales recibidas consistente con el **Principio de Sumación**.



- Existe un *tiempo característico* consistente con el **Período Refractario** bajo el cual no es posible generar un nuevo impulso sin que haya terminado el impulso anterior.
- También existe un **Tiempo de Integración** de las señales recibidas que es consistente con las observaciones fisiológicas (*Retardo Sináptico*).



Definimos la **Actividad** de una neurona ( $e$ ) en el tiempo  $t + 1$  como una función ( $H(\Phi)$ ) que depende de la actividad de las neuronas a las que está conectada (*presinápticas*) y que recibe información de éstas en el tiempo  $t$ .

$$e(t + 1) = H(\Phi)$$

Esta función  $H(\Phi)$  toma valores de 0 o 1, y es conocida como **Función Escalón de Heaviside**.

$$H(\Phi) = \begin{cases} 0 & \text{si } \Phi < 0 \\ 1 & \text{si } \Phi \geq 0 \end{cases}$$

¿Es la función  $H(\Phi)$  una función booleana?

Notar que la función escalón intenta emular el Principio de “Todo o Nada” tomando solamente 2 valores posibles, correspondientes a los 2 estados posibles en los que puede estar la neurona<sup>[3]</sup>.

$$H(\Phi) = \begin{cases} 0 & \text{si } \Phi < 0 \\ 1 & \text{si } \Phi \geq 0 \end{cases}$$

Gráfico obtenido en Geogebra



[3] Notar la similitud de los estados binarios de la neurona con los estados binarios asociados a las celdas de los autómatas de Wolfram y Conway (Ver Práctico 08: *Autómatas Celulares*).

Los valores de  $\Phi$  dependen de la integración de la actividad de las neuronas presinápticas, sus respectivos pesos sinápticos y el umbral del potencial de acción.

$$\Phi(t) = \sum_{i=1}^n p_i e_i(t) - \theta$$

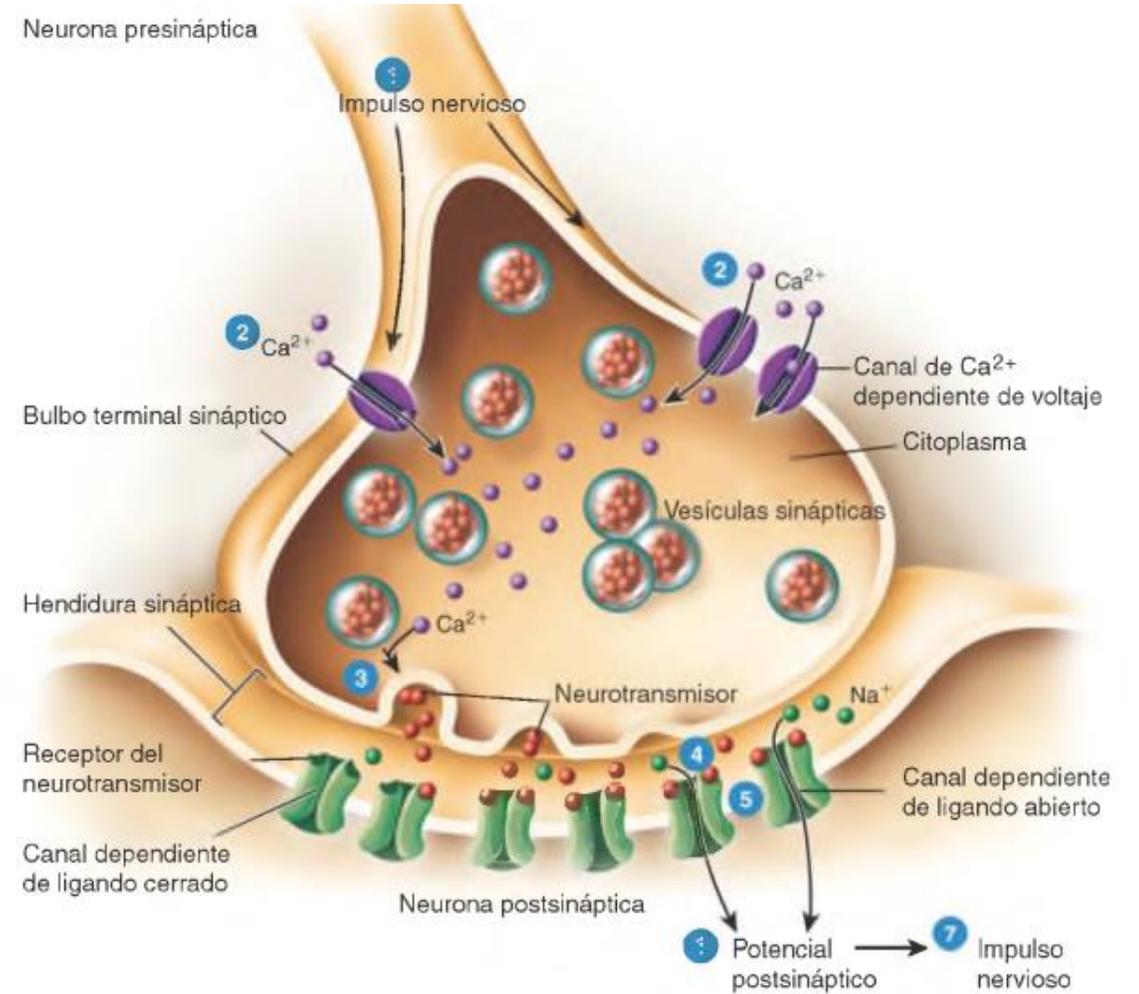
Donde,

$p_i$  es el peso sináptico de la  $i$ -ésima neurona.

$e_i$  es el valor de la actividad de la  $i$ -ésima neurona en tiempo  $t$ .

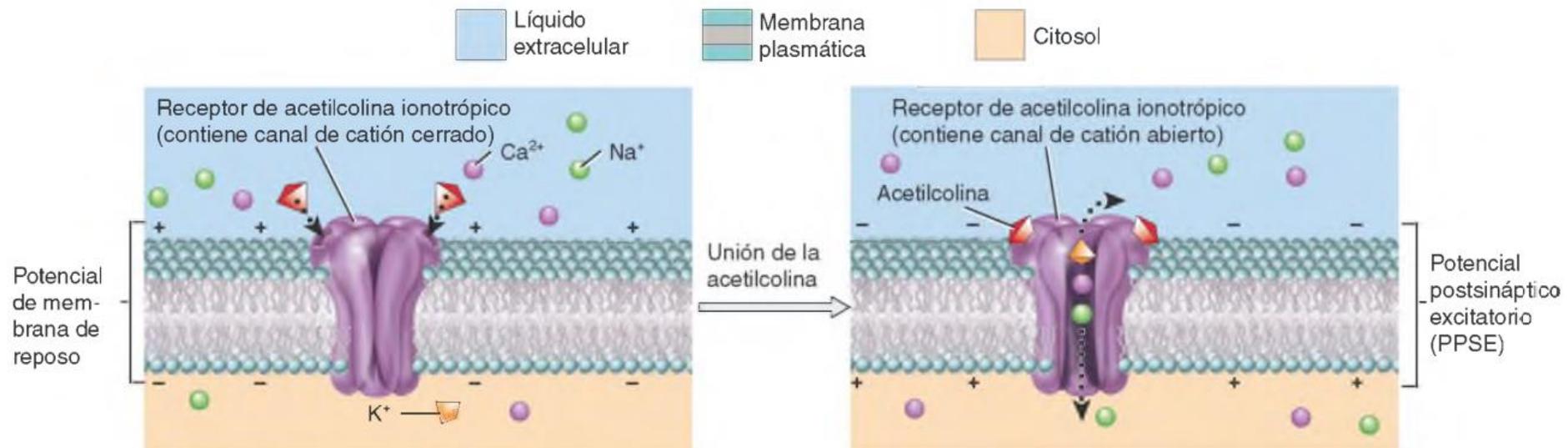
$\theta$  es el valor umbral de activación del potencial de acción.

El **peso sináptico** ( $p_i$ ) es un parámetro asociado al tipo de conexión (sinapsis) de una neurona con otra. Este tipo de conexión es variable y por lo tanto puede tomar diferentes valores (negativo, positivo o 0). Esta información puede codificarse como diferentes respuestas neurofisiológicas (**excitatoria, inhibitoria, moduladora**).

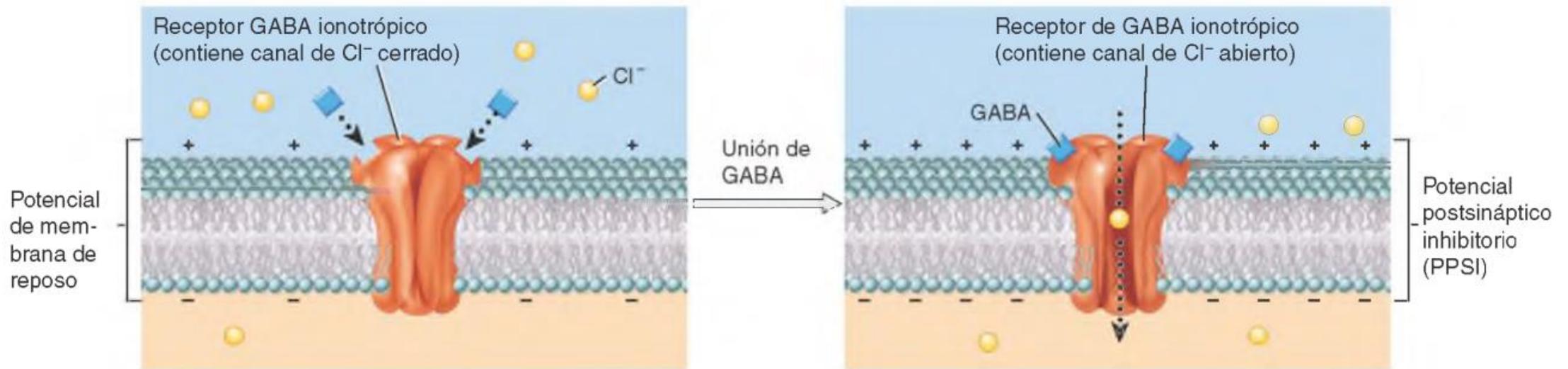


El peso sináptico es un parámetro que depende del tipo de **neurotransmisor**, su concentración y el **tipo de receptor** de la neurona postsináptica.

Ej.: Acetilcolina, Aminoácidos (Glutamato, Aspartato, GABA), Adrenalina, Noradrenalina, Dopamina, entre otros.

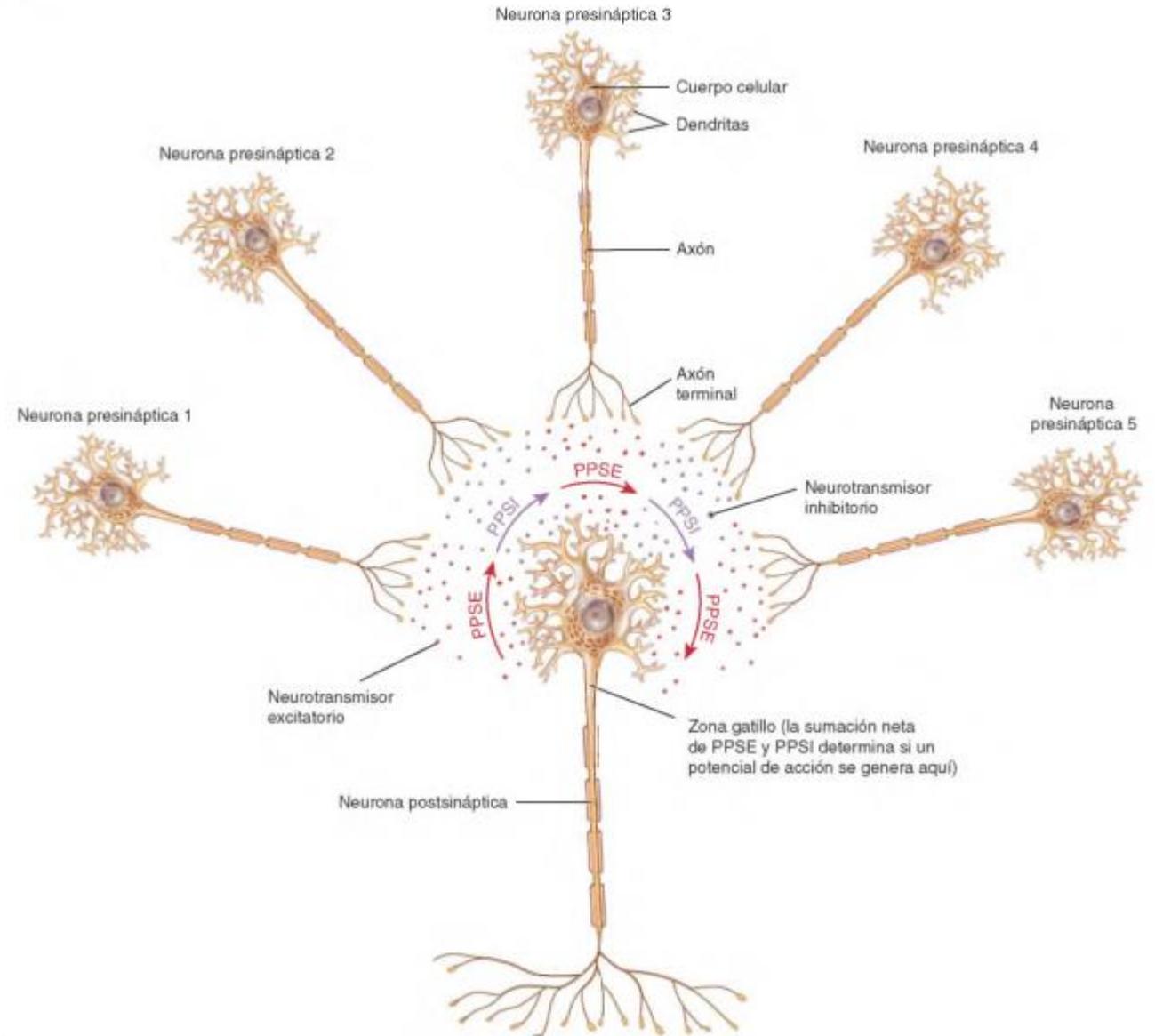


Los neurotransmisores que causan una sinapsis excitatoria generan una *despolarización de la membrana*, mientras que los neurotransmisores que causan una sinapsis inhibitoria generan una *hiperpolarización de la membrana*.



El **estímulo total** que recibe una neurona es, por tanto, la sumatoria de todos los estímulos que recibe de las neuronas presinápticas, comparadas con el propio **umbral** de la neurona ( $\theta$ ).

Esta sumación se produce en la **Zona Gatillo** de la neurona.



Interpretaremos que,

$$si \sum_{i=1}^n p_i e_i(t) - \theta \geq 0 \implies H(\Phi) = 1 \implies \text{hay } PA$$

$$si \sum_{i=1}^n p_i e_i(t) - \theta < 0 \implies H(\Phi) = 0 \implies \text{no hay } PA$$

La **lógica proposicional** es un sistema que permite establecer – mediante funciones lógicas – el estado de una afirmación (verdadero o falso). Debido a la naturaleza binaria de estos estados, es posible asociarlos a 0s y 1s, por lo que se pueden traducir como *funciones booleanas*.

Ej.: dos neuronas  $p, q$  pueden tener 2 estados cada una, por lo que existen  $2^2 = 4$  estados globales. Interpretamos “verdadero-falso” como 1s y 0s que a su vez se traducen en estados neurofisiológicos “excitado-reposo”.

$p$	$q$
V	V
F	V
V	F
F	F

Algunas de las funciones lógicas proposicionales más comunes son: AND, OR, XOR, NOT. Cada una genera una respuesta diferente, las cuales pueden representarse utilizando **Tablas de Verdad**.

$p$	$q$	AND
V	V	V
F	V	F
V	F	F
F	F	F

Ej.: la función lógica AND solamente es verdadera cuando  $p, q$  son verdaderas. La función AND dispara un potencial de acción cuando las respuestas recibidas de las neuronas presinápticas  $p, q$  envían señales excitatorias.

$p$	$q$	<b>OR</b>
V	V	V
F	V	V
V	F	V
F	F	F

Ej.: la **función lógica OR** codifica una respuesta verdadera cuando  $p$  y/o  $q$  son verdaderas. La función OR dispara un potencial de acción cuando las respuestas recibidas de las neuronas presinápticas  $p$  y/o  $q$  envían señales excitatorias. Basta con recibir una señal excitatoria de alguna de las dos neuronas para que se dispare el potencial de acción. Este tipo de función lógica se le conoce como **OR inclusivo**.

$p$	$q$	<b>XOR</b>
V	V	F
F	V	V
V	F	V
F	F	F

Ej.: la **función lógica XOR** codifica una respuesta verdadera cuando  $p$  o  $q$  son verdaderas. La función XOR dispara un potencial de acción cuando las respuestas recibidas de las neuronas presinápticas  $p$  o  $q$  envían señales excitatorias, pero no cuando ambas envían una señal excitatoria al mismo tiempo. Este tipo de función lógica se le conoce como **OR exclusivo**.

$p$	$q$	<b>EQ</b>
V	V	V
F	V	F
V	F	F
F	F	V

Ej.: la **función lógica EQ** codifica una respuesta verdadera cuando  $p$  y  $q$  tienen el mismo valor (son equivalentes). La función EQ dispara un potencial de acción cuando las respuestas recibidas de las neuronas presinápticas  $p$  y  $q$  envían el mismo tipo de señal (independientemente del valor). Notar que esta función lógica es contraria a la función XOR.

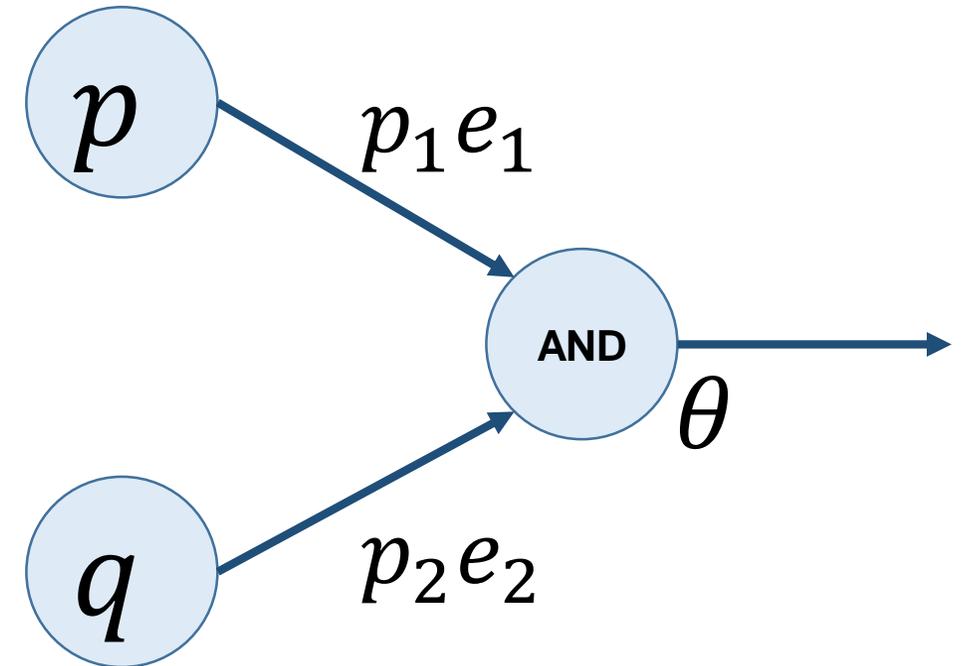
$p$	$q$	<b>NOT <math>q</math></b>
V	V	F
F	V	F
V	F	V
F	F	V

Una manera de simbolizar la función NOT es utilizando el símbolo  $\neg$  que implica negación. Por tanto NOT $q$  puede escribirse como  $\neg q$ .

Ej.: esta **función lógica NOT $q$**  codifica una respuesta verdadera cuando  $q$  es falsa, independientemente del valor de  $p$ . La función NOT $q$  dispara un potencial de acción cuando recibe la señal de que la neurona presináptica  $q$  ha enviado una respuesta inhibitoria. La función NOT puede ser utilizada como una proposición opuesta a otra función (Ej.: la función XOR puede escribirse como NOT(EQ)).

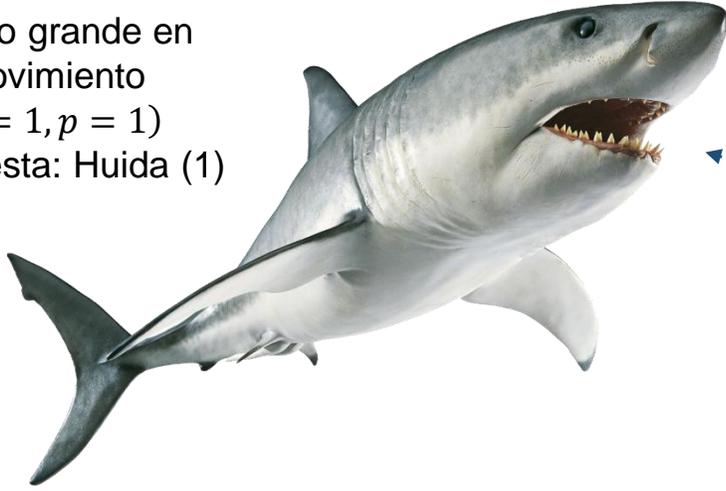
De esta forma podemos codificar este sistema lógico proposicional en una Neurona de tipo McCulloch-Pitts.

Para poder codificar una proposición lógica en una red neuronal MP, debemos conocer los valores de los pesos sinápticos y del umbral.



# Modelo de McCulloch-Pitts: Actividad

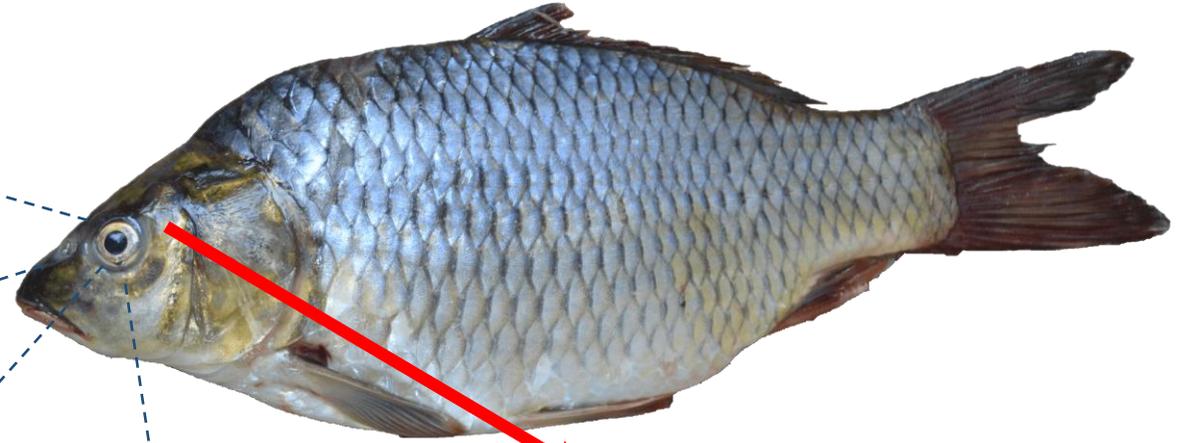
Objeto grande en movimiento  
( $q = 1, p = 1$ )  
Respuesta: Huida (1)



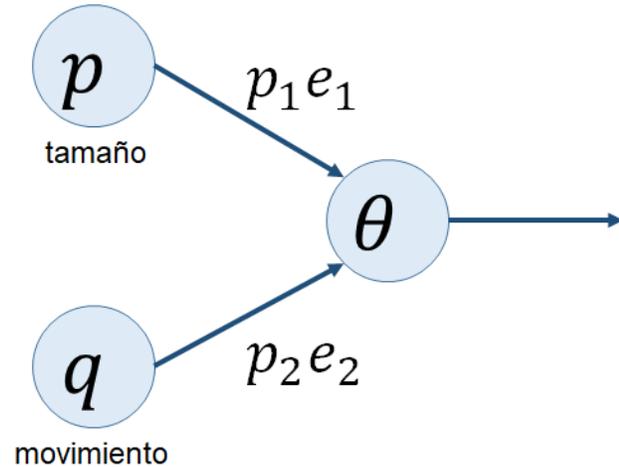
Objeto pequeño en movimiento  
( $q = 1, p = 0$ )  
Respuesta: Quieto (0)



Objeto grande y quieto  
( $q = 0, p = 1$ )  
Respuesta: Quieto (0)



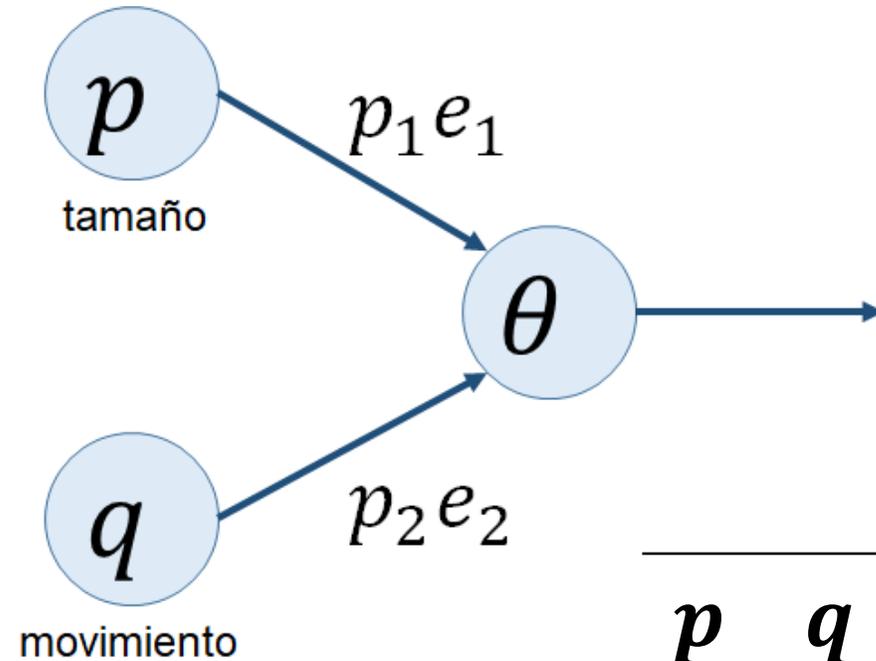
Objeto pequeño y quieto  
( $q = 0, p = 0$ )  
Respuesta: Quieto (0)



## Parte a)

La función lógica que corresponde al comportamiento observado es **AND**.

El pez solamente se mueve (huye) si el objeto que observa es grande ( $q = 1$ ) y se está moviendo ( $p = 1$ ).



$p$	$q$	<b>AND</b>
1	1	1
1	0	0
0	1	0
0	0	0

Aplicando las ecuaciones del modelo se obtiene un **sistema lineal de inecuaciones *indeterminado***:

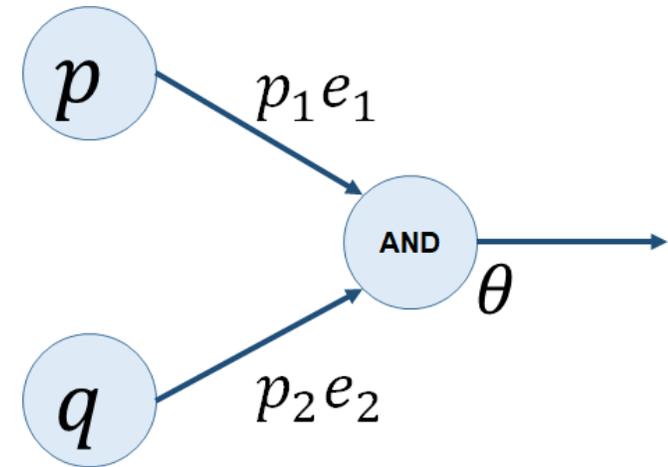
$$\Phi_1 = (1)p_1 + (1)p_2 - \theta \geq 0 \Rightarrow p_1 + p_2 \geq \theta$$

$$\Phi_2 = (1)p_1 + (0)p_2 - \theta < 0 \Rightarrow p_1 < \theta$$

$$\Phi_3 = (0)p_1 + (1)p_2 - \theta < 0 \Rightarrow p_2 < \theta$$

$$\Phi_4 = (0)p_1 + (0)p_2 - \theta < 0 \Rightarrow 0 < \theta$$

Notar que **existen infinitas soluciones** para este sistema. Un posible conjunto solución de este sistema es:  $\theta = 1, p_1 = p_2 = 0,5$ .



$p$	$q$	AND
1	1	1
1	0	0
0	1	0
0	0	0

## Parte b):

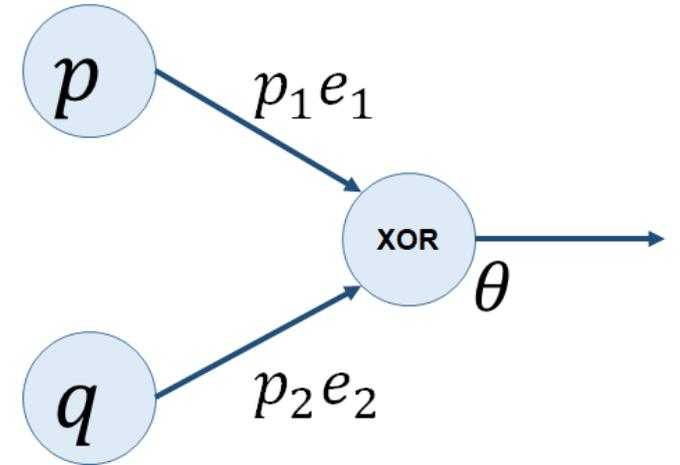
$$\Phi_1 = (1)p_1 + (1)p_2 - \theta < 0 \implies p_1 + p_2 < \theta$$

$$\Phi_2 = (1)p_1 + (0)p_2 - \theta \geq 0 \implies p_1 \geq \theta$$

$$\Phi_3 = (0)p_1 + (1)p_2 - \theta \geq 0 \implies p_2 \geq \theta$$

$$\Phi_4 = (0)p_1 + (0)p_2 - \theta < 0 \implies 0 < \theta$$

Notar que este sistema es **incompatible**: no es posible encontrar dos valores de  $p_1, p_2$  que sumados sean menores que  $\theta$  y simultáneamente que cada uno sea mayor que  $\theta$ .

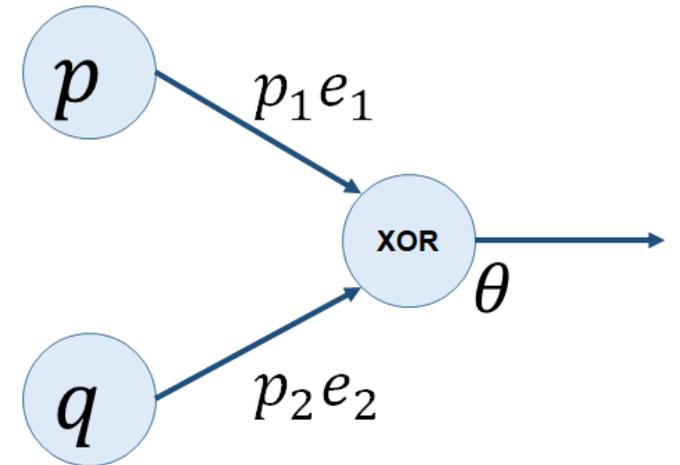


$p$	$q$	XOR
1	1	0
1	0	1
0	1	1
0	0	0

## Parte b)

El hecho de que el sistema de inecuaciones sea incompatible establece que **no es posible computar la función XOR a partir de una sola neurona.**

Por tanto, **se necesitan añadir más neuronas a la red.**



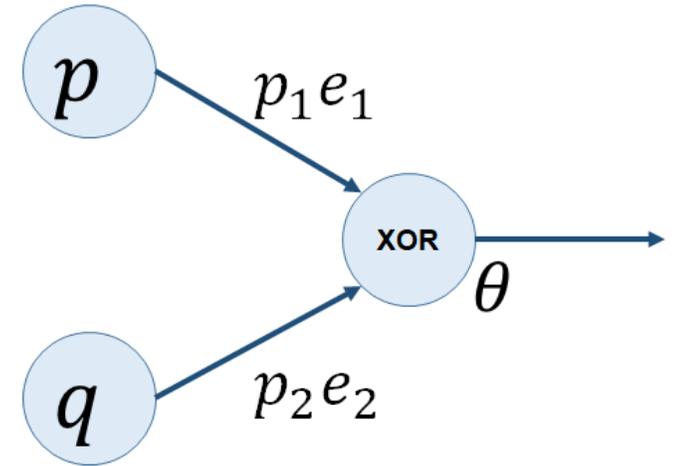
$p$	$q$	XOR
1	1	0
1	0	1
0	1	1
0	0	0

## Parte b)

La función lógica XOR se puede codificar a partir de otras funciones lógicas:

$$\neg[(p \text{ AND } q) \text{ OR } [\neg(p \text{ OR } q)]]$$

Ya sabemos que la función AND es computable por una sola neurona. Necesitamos saber si OR y  $\neg$ OR también pueden ser computables por una sola neurona.



<i>p</i>	<i>q</i>	<b>XOR</b>
1	1	0
1	0	1
0	1	1
0	0	0

**Parte b):**

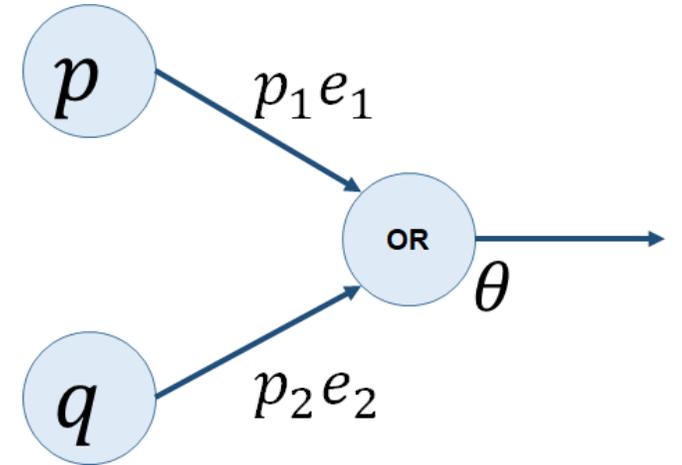
$$\Phi_1 = (1)p_1 + (1)p_2 - \theta \geq 0 \implies p_1 + p_2 \geq \theta$$

$$\Phi_2 = (1)p_1 + (0)p_2 - \theta \geq 0 \implies p_1 \geq \theta$$

$$\Phi_3 = (0)p_1 + (1)p_2 - \theta \geq 0 \implies p_2 \geq \theta$$

$$\Phi_4 = (0)p_1 + (0)p_2 - \theta < 0 \implies 0 < \theta$$

El sistema resultante es compatible, por lo que existen varios conjuntos de soluciones que lo satisfacen. Esto implica que **la función OR es computable con una única neurona.**



$p$	$q$	<b>OR</b>
1	1	1
1	0	1
0	1	1
0	0	0

**Parte b):**

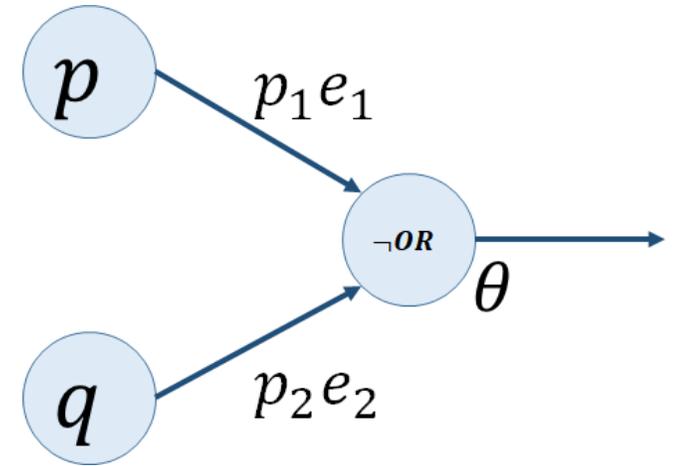
$$\Phi_1 = (1)p_1 + (1)p_2 - \theta < 0 \implies p_1 + p_2 < \theta$$

$$\Phi_2 = (1)p_1 + (0)p_2 - \theta < 0 \implies p_1 < \theta$$

$$\Phi_3 = (0)p_1 + (1)p_2 - \theta < 0 \implies p_2 < \theta$$

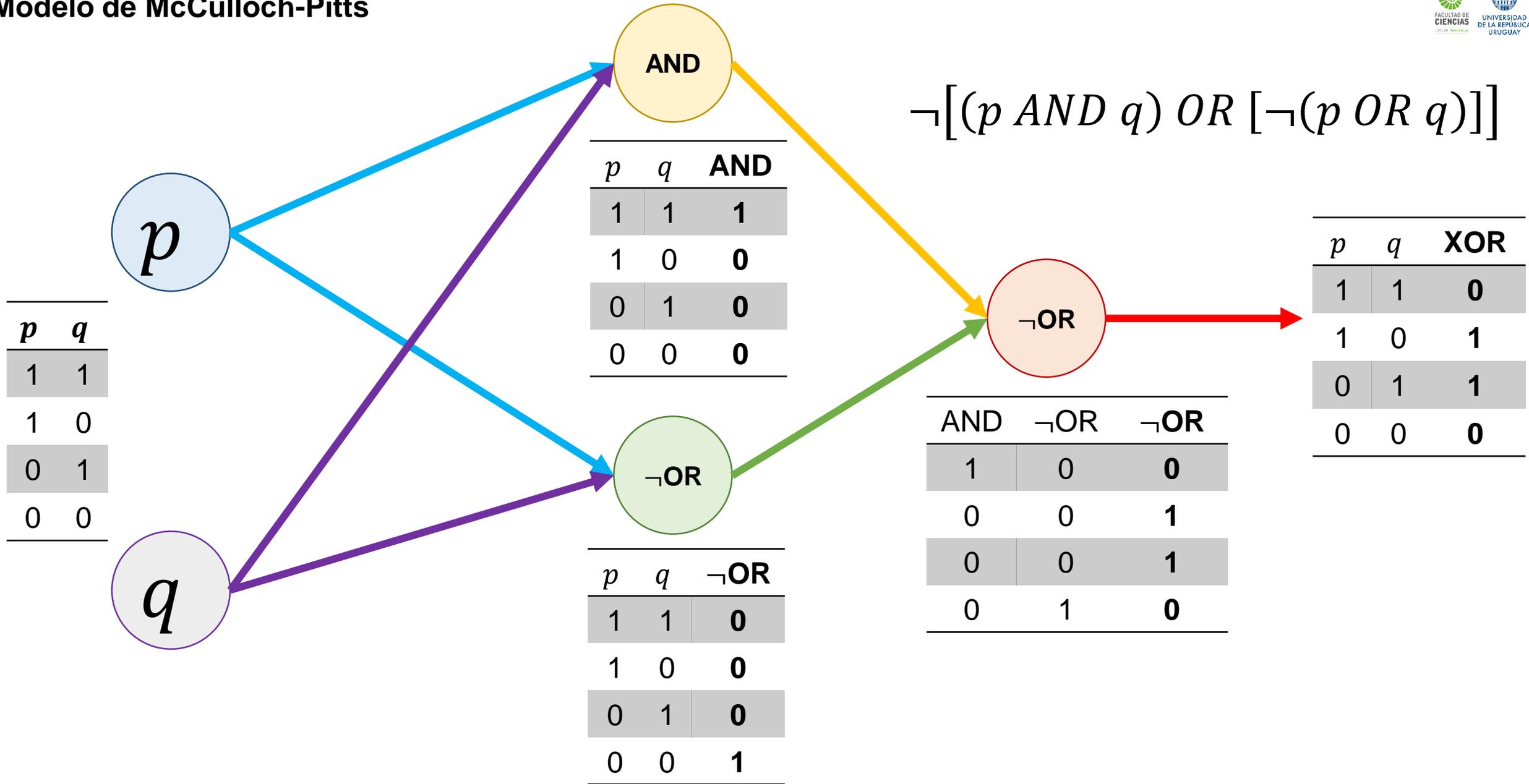
$$\Phi_4 = (0)p_1 + (0)p_2 - \theta \geq 0 \implies 0 \geq \theta$$

El sistema resultante es compatible, por lo que existen varios conjuntos de soluciones que lo satisfacen. Esto implica que **la función  $\neg\text{OR}$  es computable con una única neurona.**



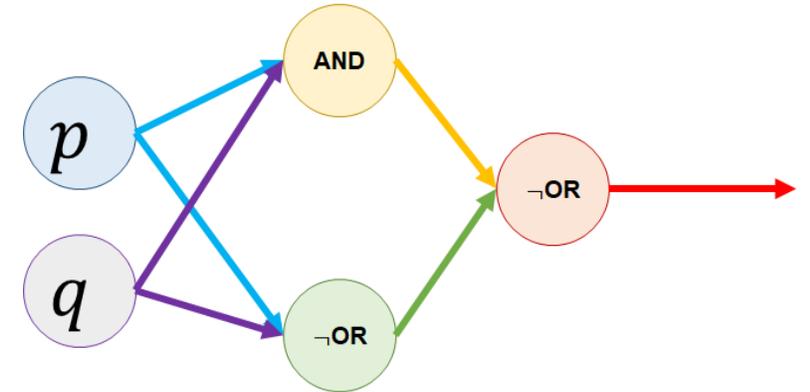
$p$	$q$	$\neg\text{OR}$
1	1	0
1	0	0
0	1	0
0	0	1

# Modelo de McCulloch-Pitts



## Parte b)

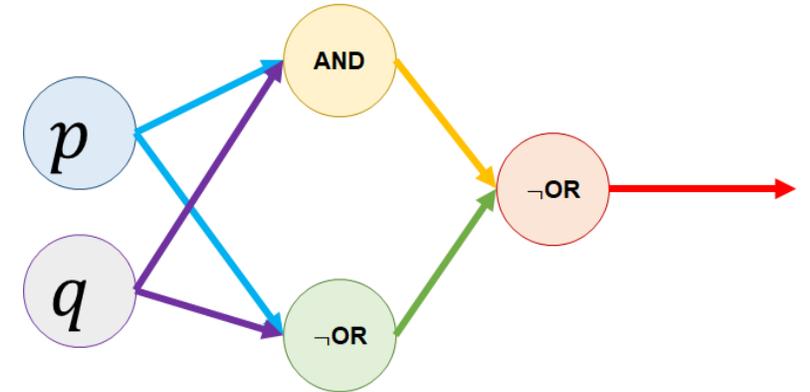
Para poder computar la función lógica XOR son necesarias al menos 3 neuronas: 2 neuronas capaces de computar la función NOT(OR) y 1 neurona capaz de computar la función AND.



$p$	$q$	<b>XOR</b>
<b>1</b>	<b>1</b>	0
<b>1</b>	<b>0</b>	1
<b>0</b>	<b>1</b>	1
<b>0</b>	<b>0</b>	0

**Parte c) ¿Qué sucedería si alguna de las conexiones sinápticas de la red se rompe? ¿Son las redes MP resistentes a daños locales?**

Si eliminamos alguna de las neuronas de la red, la función lógica ya no es computable. Con que eliminemos una de las neuronas el sistema ya no funciona. **Las redes MP no son resistentes a daños locales (i.e.: *neurodegeneración*).**



$p$	$q$	XOR
1	1	0
1	0	1
0	1	1
0	0	0

# Contenido de la clase

- Introducción: Teoría Computacional del Cerebro
- Modelo de McCulloch-Pitts
- Modelo de Memorias Asociativas

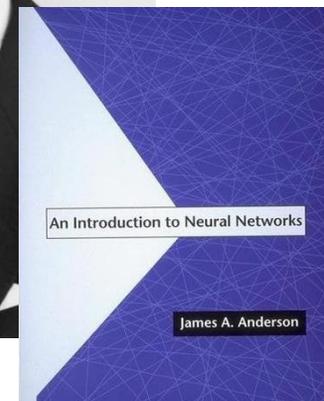
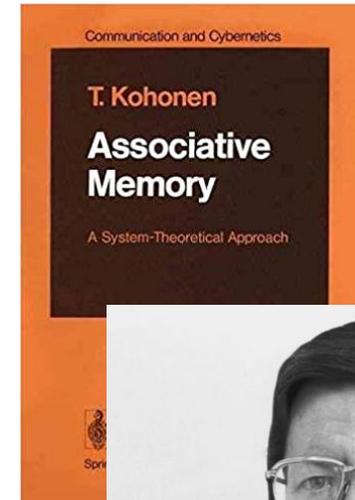
# Contenido de la clase

- Introducción: Teoría Computacional del Cerebro
- Modelo de McCulloch-Pitts
- **Modelo de Memorias Asociativas**

En la década de los 70s surgió una nueva escuela de pensamiento de las redes neuronales, la cual vendría a sustituir y mejorar los modelos de redes tipo MP. Los trabajos *independientes* de J. Anderson<sup>[4]</sup> y T. Kohonen<sup>[5]</sup> darían comienzo a las **Teorías de Memorias Asociativas**.

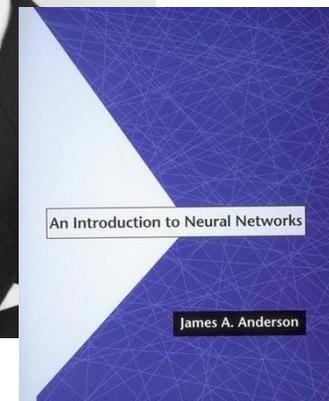
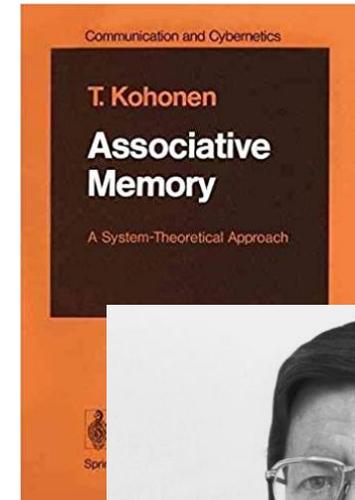
[4] Anderson, J. A. (1972). A simple neural network generating an interactive memory. *Mathematical Biosciences*, 14(3-4), 197-220.

[5] Kohonen, T. (1972). *Correlation matrix memories*. *IEEE Transactions on Computers*, 100(4), 353-359.



Algunos estudios antecedentes que marcaron la necesidad de nuevas teorías fueron:

- 1) Limitaciones del modelo McCulloch-Pitts.
- 2) Aprendizaje Hebbiano.
- 3) Perceptrón de Rosenblatt.
- 4) Distribución de la memoria.
- 5) Resistencia al daño neuronal.

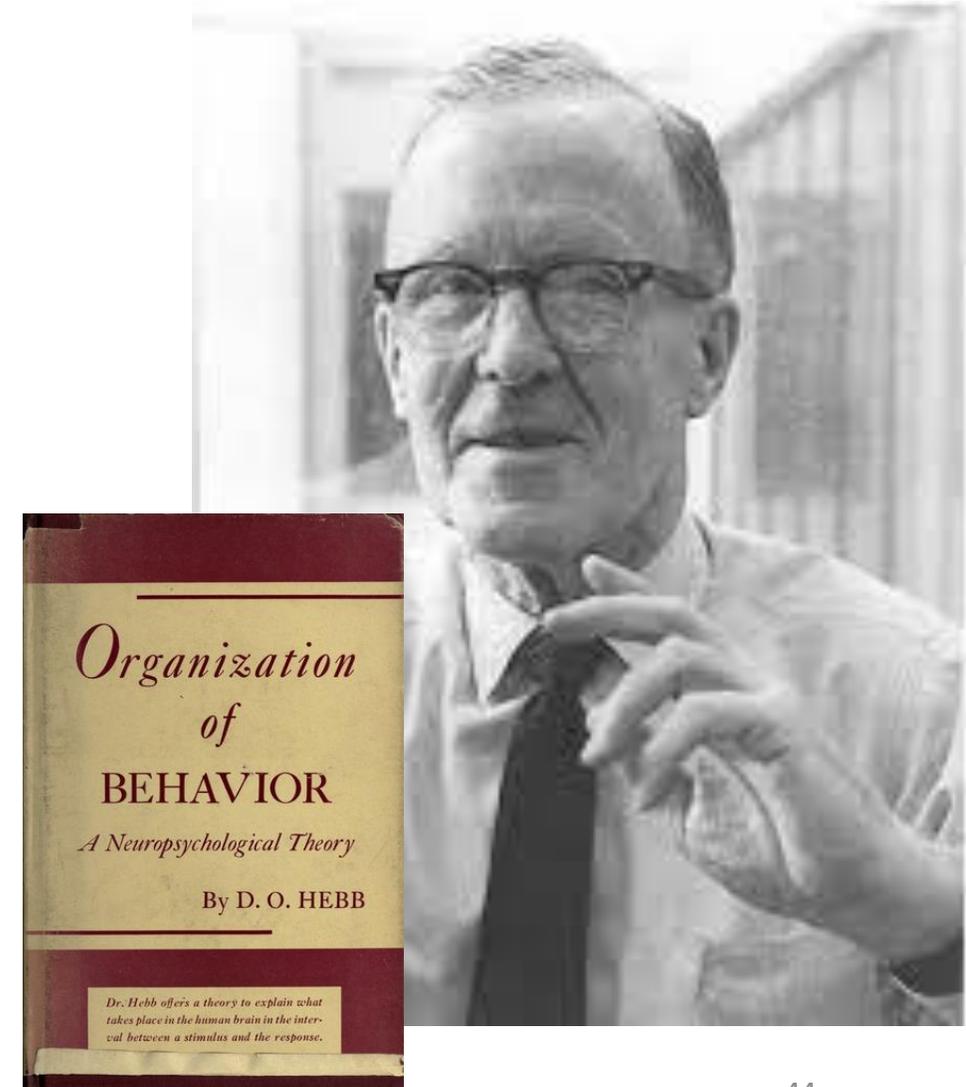


# Aprendizaje Hebbiano

En 1949, Hebb publica un trabajo pionero acerca del **aprendizaje, el *ensamble* neuronal y la plasticidad sináptica**<sup>[6]</sup>.

*“Cuando el axón de la célula A está lo suficientemente cerca para excitar la célula B de manera repetida y persistente, existirá un crecimiento progresivo y un cambio metabólico en ambas células de forma que la eficiencia de la célula A para excitar a la célula B se incrementa”.*

[6] Hebb, D. O. (1949). *The Organization of Behavior: A Neurophysiological Theory*.



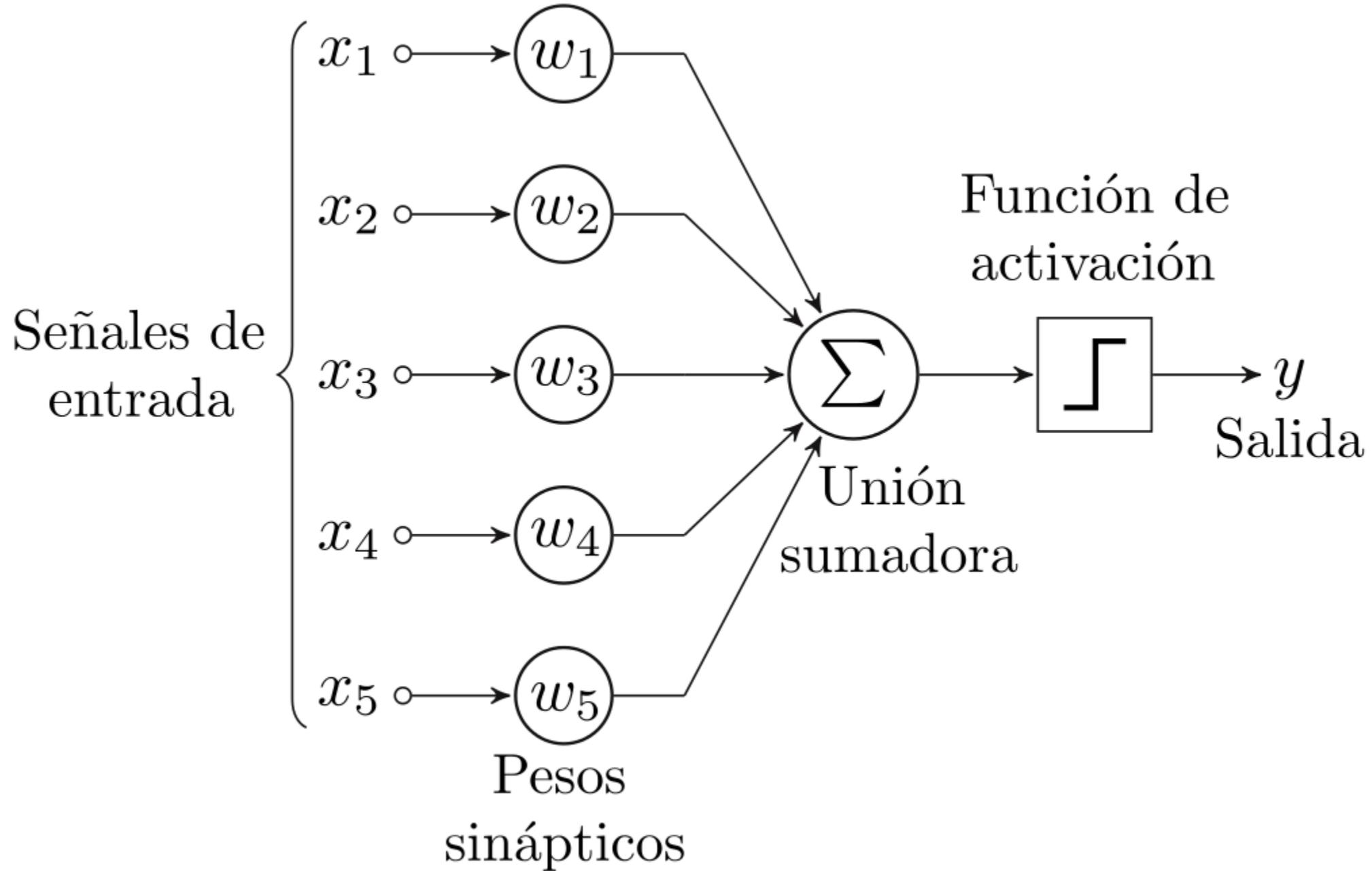
# Perceptrón de Rosenblatt

En 1958, F. Rosenblatt propone el **Modelo del Perceptrón**<sup>[7]</sup>: una neurona artificial capaz de generar una respuesta a partir de un conjunto de señales de entrada por medio de algoritmos.

**Las redes neuronales de MP son versiones biológicas de los Perceptrones de Rosenblatt.**

[7] Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.



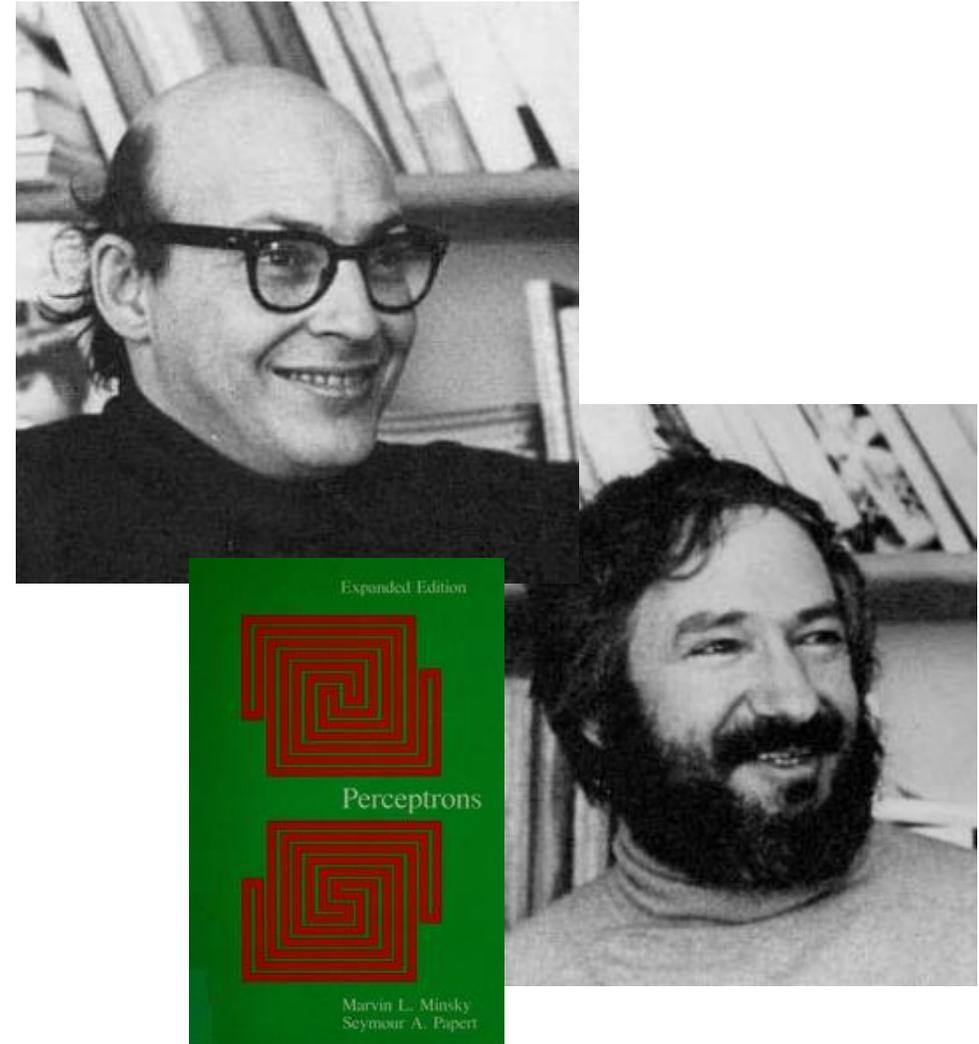


# Perceptrón de Rosenblatt

En 1969, Minsky y Papert publican un trabajo<sup>[8]</sup> sobre perceptrones en donde demuestran que existen ciertas relaciones lógicas que los perceptrones no pueden computar, entre ellas la **función XOR**.

La investigación en redes neuronales e inteligencia artificial se vio frenada frente a estas dificultades (*'AI Winter'*).

[8] Minsky, M., & Papert, S. (1969). *Perceptrons: an introduction to computational geometry*. MIT Press.



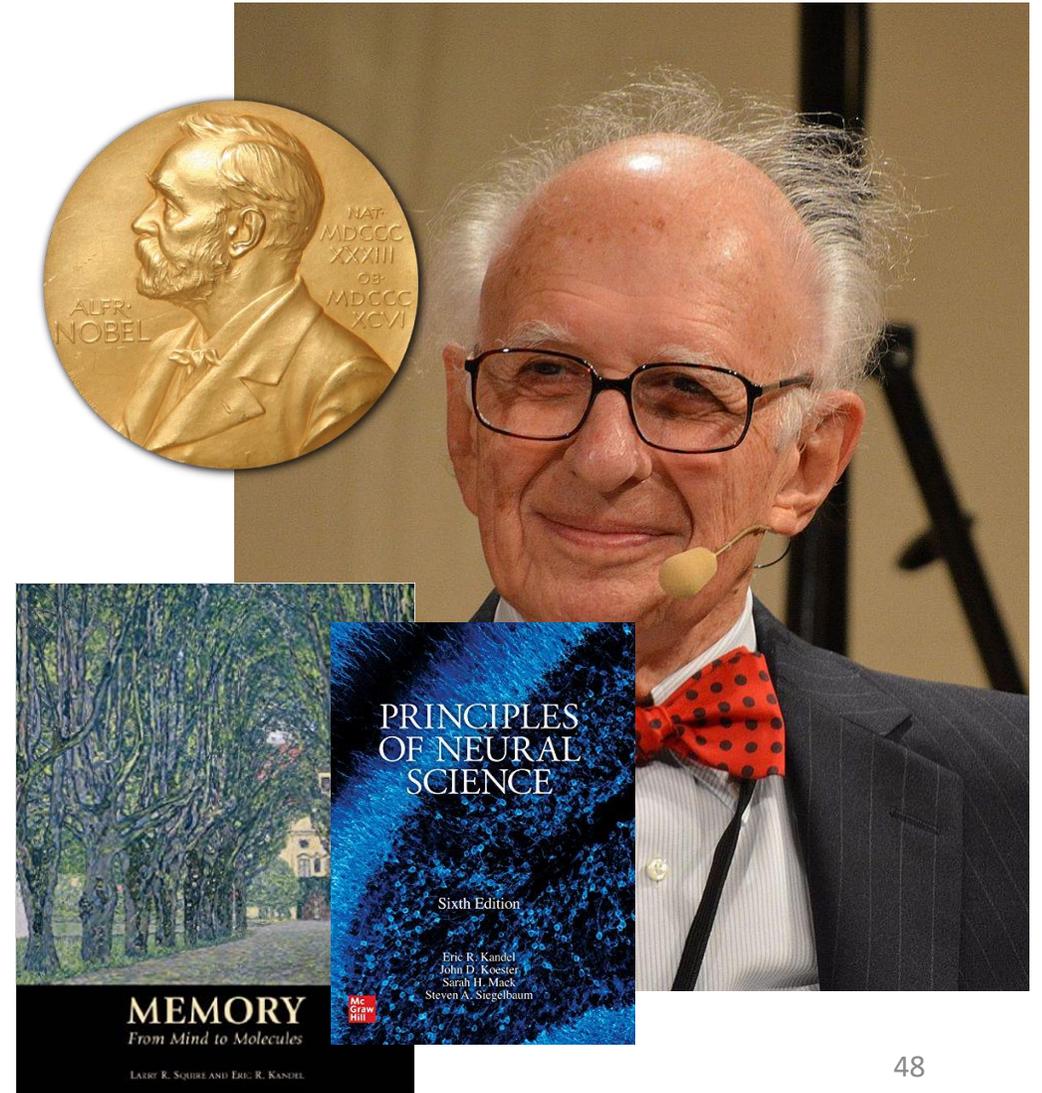
# Modelo de Memorias Asociativas

Por otro lado, los experimentos de Kandel y otros<sup>[9,10]</sup> confirmaron que:

- 1) La memoria no se encuentra en un lugar determinado del cerebro sino que está distribuida.
- 2) Afectar una porción del cerebro no necesariamente destruye por completo la capacidad de recordar.

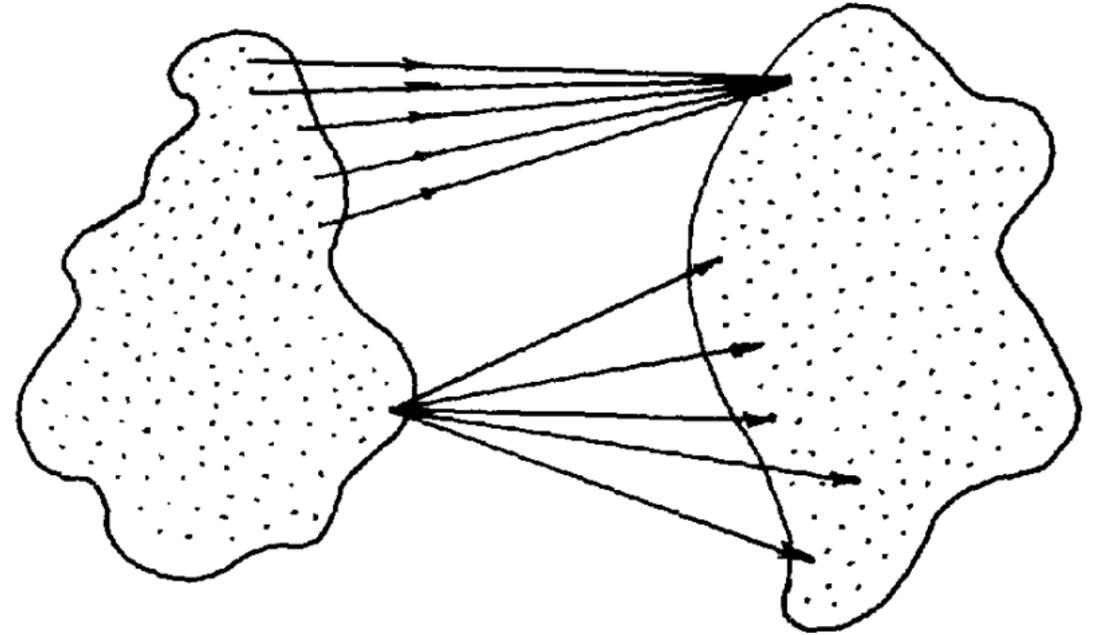
[9] Squire, L. R., & Kandel, E. R. (2003). *Memory: From mind to molecules*. Macmillan.

[10] Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2021). *Principles of neural science*. New York: McGraw-Hill, Health Professions Division.



La idea del modelo de Memorias Asociativas se establece a partir de dos conjuntos de neuronas: un conjunto que envía la información (Ej.: neuronas sensoriales) y otro que la recibe (Ej.: neuronas procesadoras).

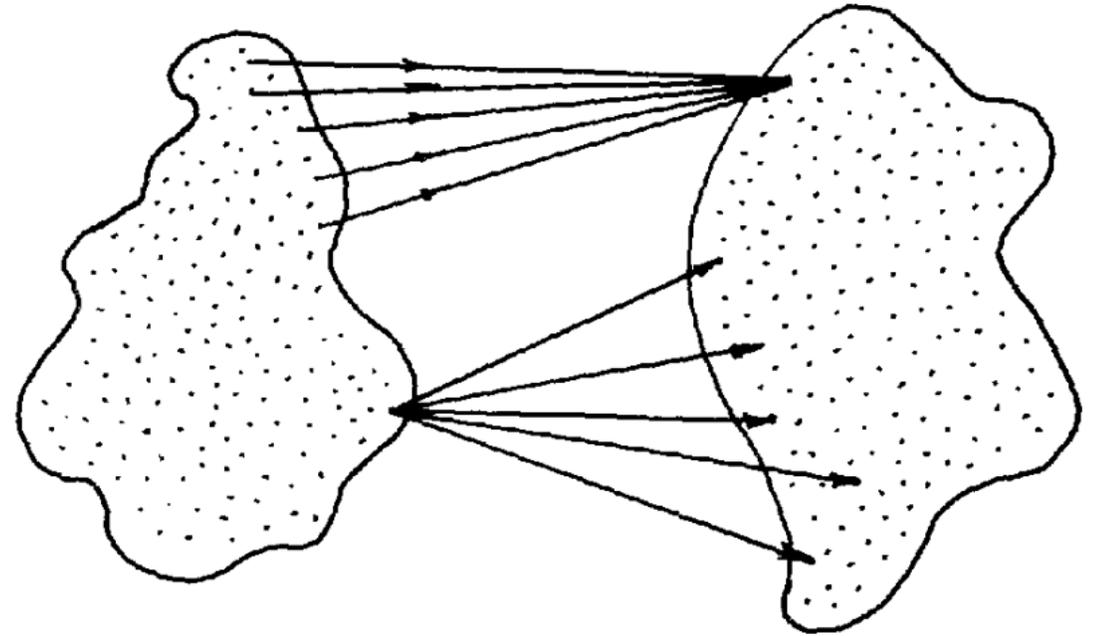
**Las actividades de cada neurona en cada uno de los conjuntos se pueden escribir a través de vectores columna.**



$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

Podemos considerar al vector entrada  $f_i$  como la detección/recepción de una  $i$ -ésima señal o patrón (Ej.: la señal visual de un rostro), y al vector salida  $g_i$  como la codificación/interpretación de dicha señal/patrón en el cerebro (Ej.: asignarle una identificación al rostro).



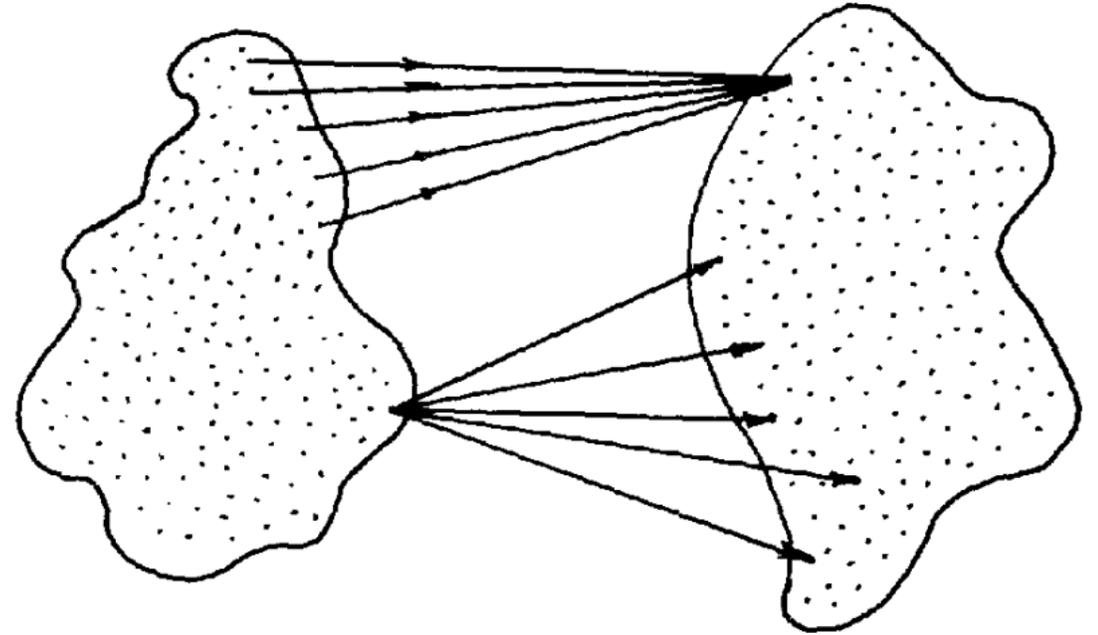
$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

La relación entre ambos vectores está dada por:

$$g_i = A_{ii} f_i$$

Donde  $A_{ii}$  es la **Matriz Asociativa** que vincula ambos conjuntos de neuronas y permite la codificación de la información. Esta matriz contiene todos los *pesos sinápticos*  $p_{ii}$  que relacionan a las neuronas de los dos conjuntos.



$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

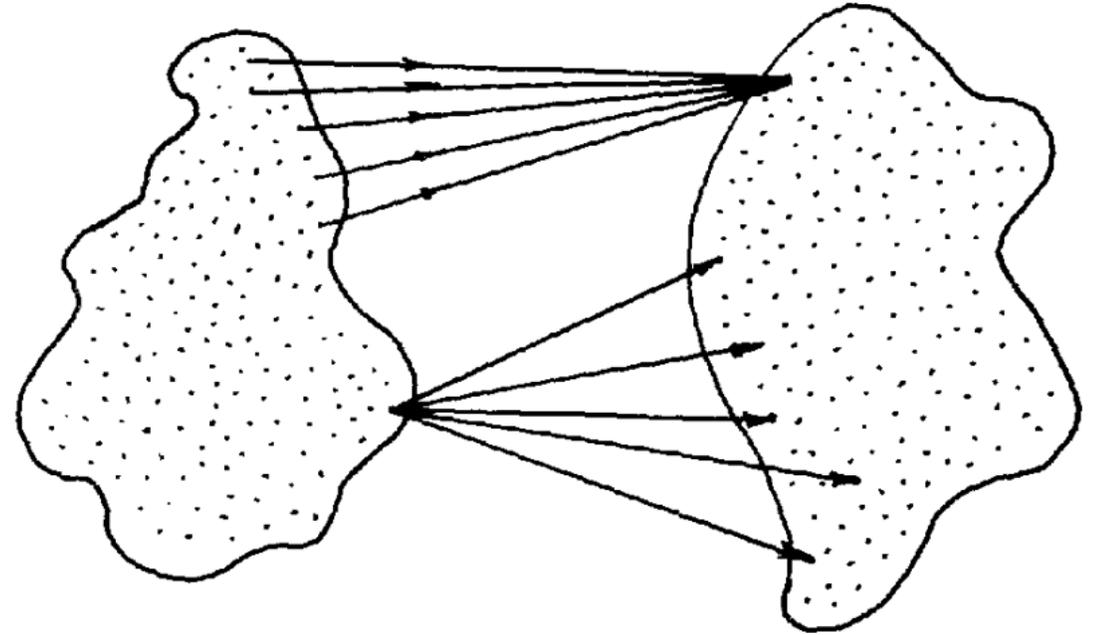
$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

La matriz  $A_{ii}$  puede obtenerse como el **producto vectorial**:

$$A_{ii} = g_i \times f_i^T$$

Donde  $f_i^T$  es el **vector traspuesto** de  $f_i$  por lo que es un vector fila. La matriz  $A_{ii}$  es:

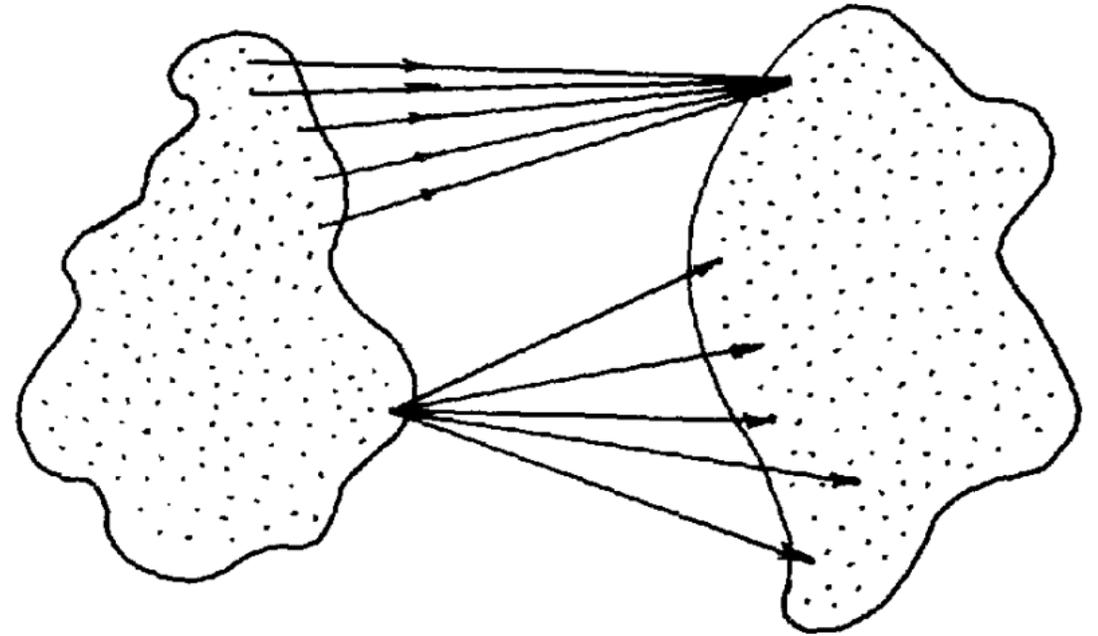
$$A_{ii} = \begin{pmatrix} g_1 f_1 & \cdots & g_1 f_n \\ \vdots & \ddots & \vdots \\ g_m f_1 & \cdots & g_m f_n \end{pmatrix}$$



$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

Notar que las dimensiones de la matriz es de  $m \times n$ . Para obtener una matriz cuadrada debemos tomar  $m = n$ , lo cual implica que los pares de vectores  $f_i, g_i$  tienen las mismas dimensiones (una aproximación que no siempre es realista).

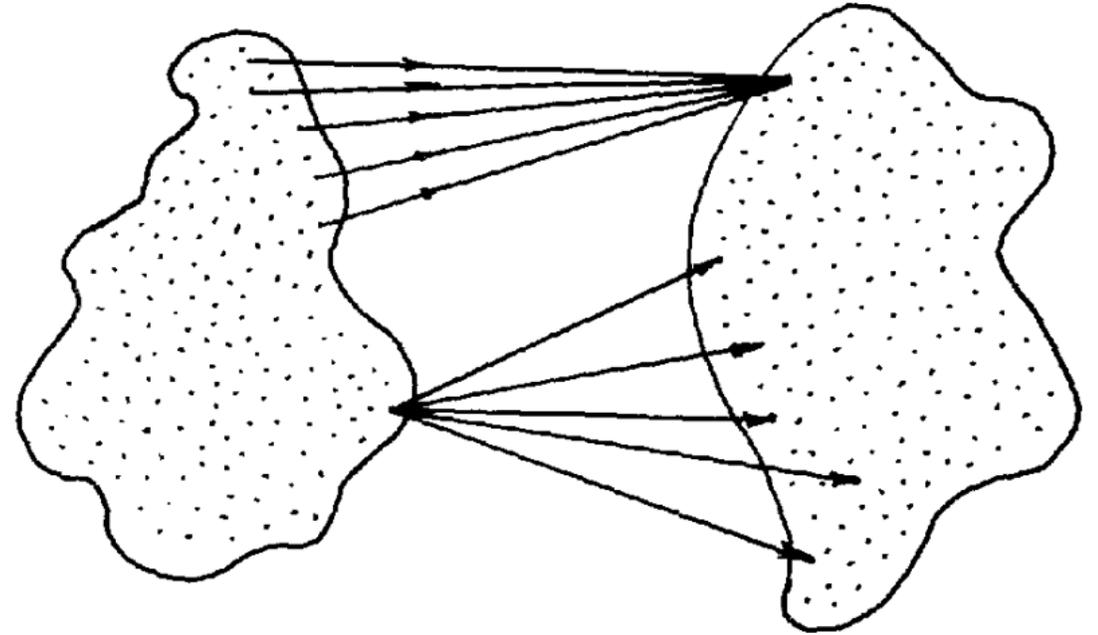


$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

La ecuación  $A_{ii} = g_i \times f_i^T$  se cumple tomando  $f_i$  como un **vector normalizado**, es decir, que su norma sea 1:

$$\|f_i\| = \sqrt{\sum_{i=1}^n f_i^2} = 1$$



$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

Podemos demostrar que dicha ecuación devuelve el vector  $g_i$ :

$$\text{si } A_{ii} = g_i \times f_i^T \implies A_{ii}f_i = (g_i \times f_i^T) \times f_i$$

**Propiedad asociativa**  
del producto vectorial:  
 $(a \times b) \times c = a \times (b \times c)$

$$(g_i \times f_i^T) \times f_i = g_i \times (f_i^T \times f_i)$$

$(f_i^T \times f_i) = (f_i \cdot f_i)$

$$(g_i \times f_i^T) \times f_i = g_i \times (f_i^T \times f_i) = g_i \times (f_i \cdot f_i) = g_i \|f_i\|^2$$

$(f_i \cdot f_i) = \|f_i\|^2$

Puesto que  $f_i$  está normalizado, entonces  $\|f_i\|^2 = 1$ ,

$$A_{ii}f_i = g_i$$

Si  $f_i$  no estuviera normalizado  $\|f_i\|^2 = \alpha$ , con  $\alpha > 0, \alpha \neq 1$  (magnitud escalar):

$$A_{ii}f_i = \alpha g_i$$

Por lo tanto, dada una matriz  $A_{ii}$  y un vector  $f_i$ :

$$A_{ii} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad f_i = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

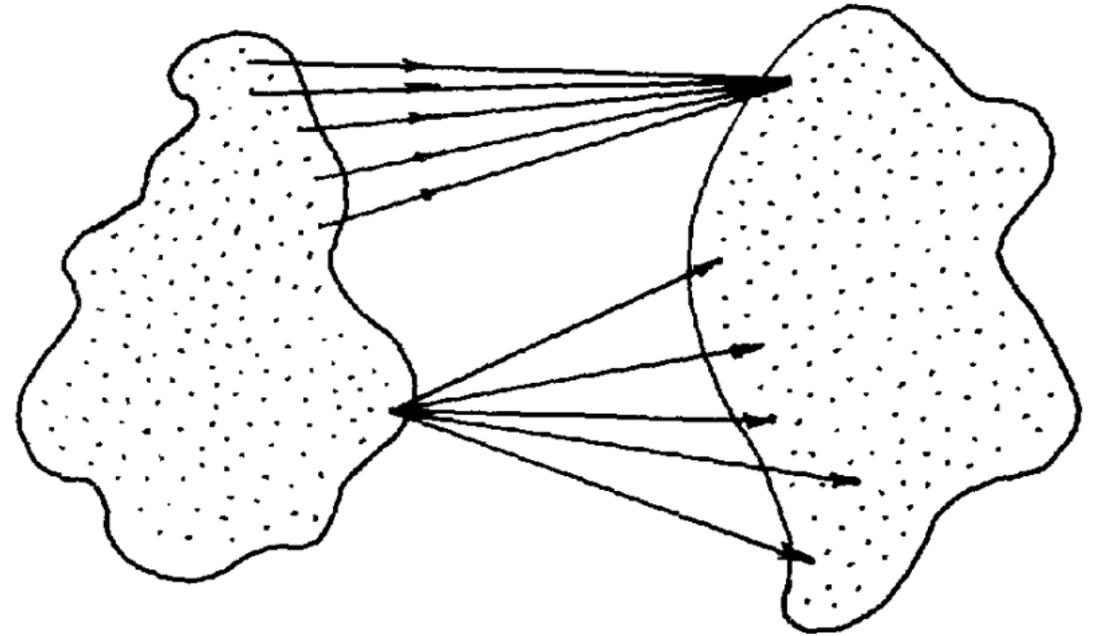
El vector  $g_i$  asociado es:

$$g_i = A_{ii}f_i = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = \begin{pmatrix} a_{11}f_1 + a_{12}f_1 + \cdots + a_{1n}f_1 \\ \vdots \\ a_{m1}f_n + a_{m2}f_n + \cdots + a_{mn}f_n \end{pmatrix}$$

$$\Rightarrow g_i = \begin{pmatrix} \sum_{i=1}^n a_{1i}f_i \\ \vdots \\ \sum_{i=1}^n a_{mi}f_i \end{pmatrix} = \begin{pmatrix} g_1 \\ \vdots \\ g_m \end{pmatrix}$$

Esto implica que el vector  $g_i$  es una **combinación lineal** del vector  $f_i$  asociado a los pesos sinápticos de la matriz asociativa.

La **linealidad** entre estos vectores es consistente con las observaciones experimentales en las frecuencias de disparo de potenciales de acción.

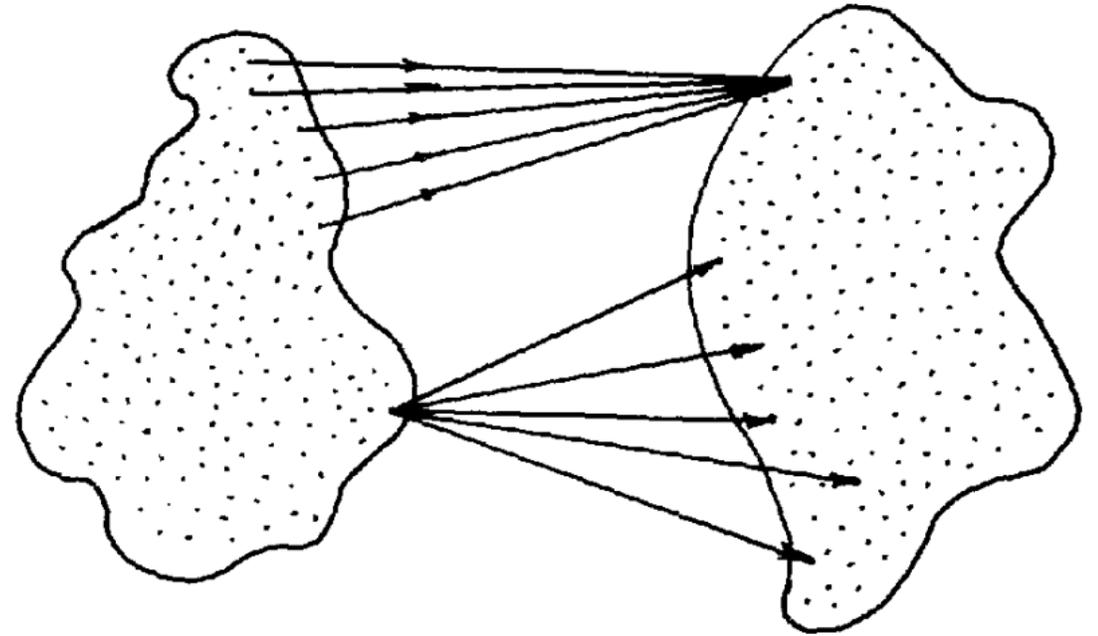


$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

El poder de este modelo implica no solo poder computar la memoria de un patrón asociando dos vectores  $f_i, g_i$ , sino también poder asociar varios patrones dentro de la misma matriz.

En nuestro ejemplo sobre rostros implica poder reconocer varios rostros entre sí.



$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

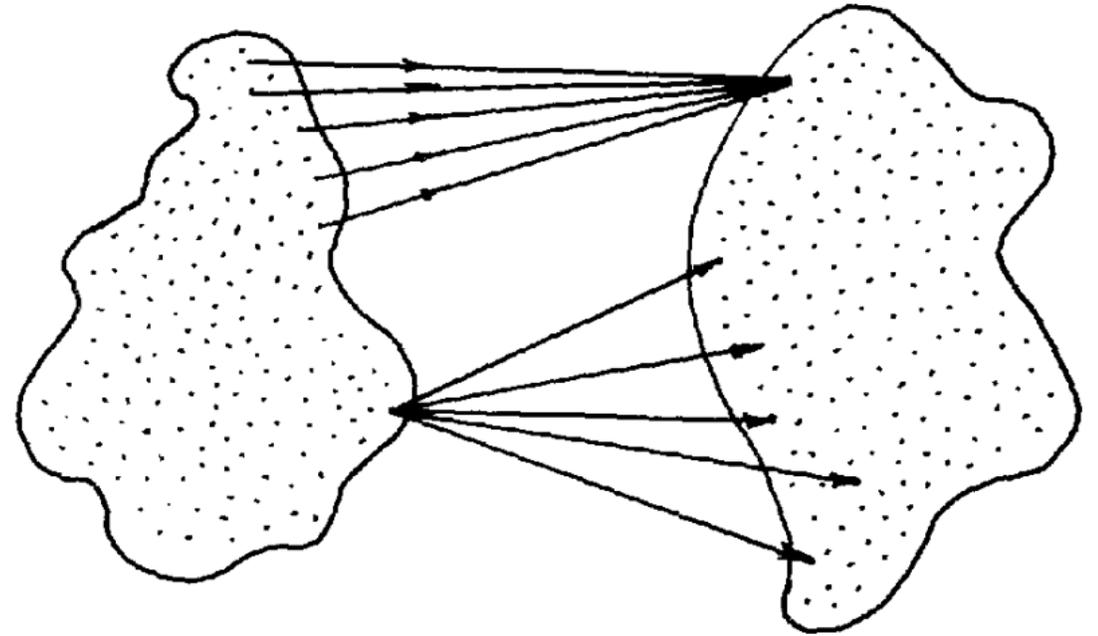
$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

Dados un conjunto de  $k$  patrones, cada uno con un par de vectores  $f_i, g_i$  asociados por una matriz asociativa  $A_{ii}$ :

$$k = \{(f_1, g_1), (f_2, g_2), \dots, (f_k, g_k)\}$$

La matriz asociativa que vincula todos los  $k$  pares  $f_i, g_i$  es:

$$A = \sum_{i=1}^k A_{ii}$$

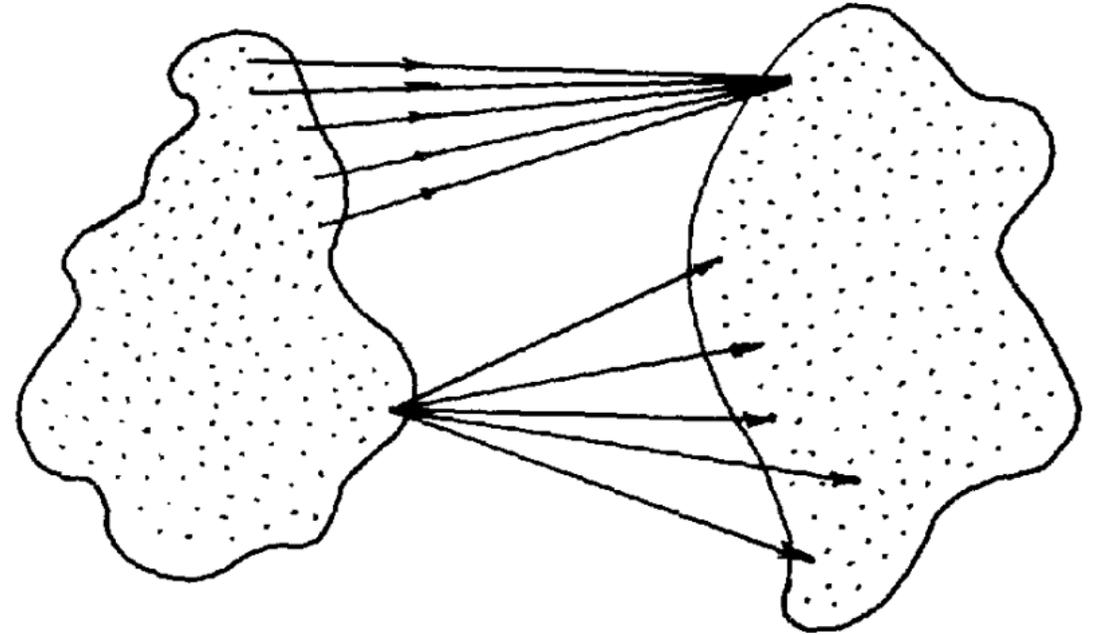


$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

Es decir, la nueva matriz  $A$  puede vincular todos los patrones  $f_i, g_i$  asociados a un rostro. Tiene la información codificada de todas las matrices  $A_{ii}$  generadas por todos los pares  $f_i, g_i$ .

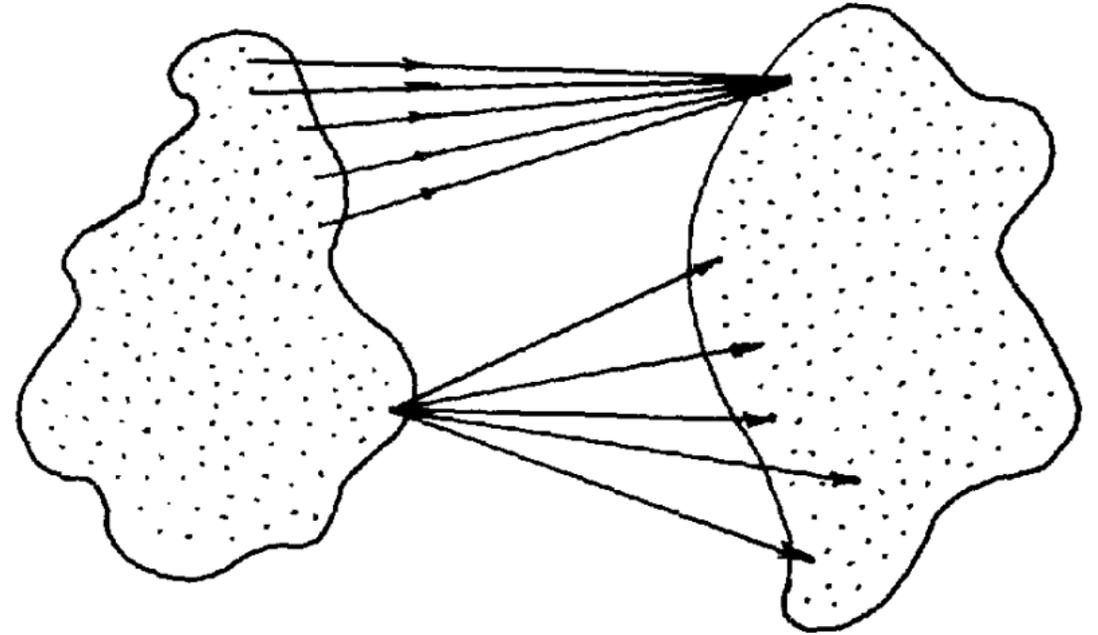
Una vez que la red recibe la información sensorial  $f_i$ , la red logra codificarla con el rostro adecuado ( $g_i$ ).



$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} \quad g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

Esta codificación es *exacta* cuando todos los vectores  $f_i$  son **ortonormales** entre sí. Podemos interpretarlo como si todos los rostros detectados fueran *inequívocamente* diferentes entre sí.

Si esto no se cumple, la matriz asociativa genera un vector resultado  $g_j$  que difiere del vector  $g_i$  original.



$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

$$g_i = Af_i$$

Definición de A

$$g_i = \left(\sum_{i=1}^k A_{ii}\right) f_i$$

Definición de  $A_{ii}$

$$g_i = \left(\sum_{i=1}^k g_i \times f_i^T\right) \times f_i$$

Expansión de la sumatoria

$$g_i = (g_1 \times f_1^T) \times f_i + \dots + (g_i \times f_i^T) \times f_i + \dots + (g_k \times f_k^T) \times f_i$$

$$g_i = g_1 \times (f_1^T \times f_i) + \dots + g_i \times (f_i^T \times f_i) + \dots + g_k \times (f_k^T \times f_i)$$

Propiedad asociativa del producto vectorial:  
 $(a \times b) \times c = a \times (b \times c)$

~~$$g_i = g_1 \times (f_1 \cdot f_i) + \dots + g_i \times (f_i \cdot f_i) + \dots + g_k \times (f_k \cdot f_i)$$~~

$(f_i^T \times f_i) = (f_i \cdot f_i)$

$$g_i = g_i \times (f_i \cdot f_i)$$

$$g_i = g_i \|f_i\|^2$$

$(f_i \cdot f_i) = \|f_i\|^2$

Condición de ortonormalidad:  
si  $f_j \perp f_i \Rightarrow (f_j \cdot f_i) = 0$

$$g_i = g_i$$

$\|f_i\|^2 = 1$  si está normalizado

*Quod erat demonstrandum*

$$f_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}, \quad g_1 = \begin{pmatrix} -4 \\ 7 \\ 2 \\ 0 \end{pmatrix}$$

La matriz asociativa  $A_1$  es:

$$A_1 = g_1 \times f_1^T = \begin{pmatrix} -4 \\ 7 \\ 2 \\ 0 \end{pmatrix} \times \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \end{pmatrix} = \begin{pmatrix} -2 & -2 & -2 & -2 \\ 7/2 & 7/2 & 7/2 & 7/2 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$f_2 = \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{pmatrix}, \quad g_2 = \begin{pmatrix} 5 \\ 1/2 \\ 0 \\ 3 \end{pmatrix}$$

La matriz asociativa  $A_2$  es:

$$A_2 = g_2 \times f_2^T = \begin{pmatrix} 5 \\ 1/2 \\ 0 \\ 3 \end{pmatrix} \times \begin{pmatrix} 1/2 & -1/2 & 1/2 & -1/2 \end{pmatrix} = \begin{pmatrix} 5/2 & -5/2 & 5/2 & -5/2 \\ 1/4 & -1/4 & 1/4 & -1/4 \\ 0 & 0 & 0 & 0 \\ 3/2 & -3/2 & 3/2 & -3/2 \end{pmatrix}$$

$$f_3 = \begin{pmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{pmatrix}, \quad g_3 = \begin{pmatrix} 1/3 \\ 2 \\ 2 \\ 2 \end{pmatrix}$$

La matriz asociativa  $A_3$  es:

$$A_3 = g_3 \times f_3^T = \begin{pmatrix} 1/3 \\ 2 \\ 2 \\ 2 \end{pmatrix} \times \begin{pmatrix} 1/2 & 1/2 & -1/2 & -1/2 \end{pmatrix} = \begin{pmatrix} 1/6 & 1/6 & -1/6 & -1/6 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix}$$

$$f_4 = \begin{pmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{pmatrix}, \quad g_4 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ -1 \end{pmatrix}$$

La matriz asociativa  $A_4$  es:

$$A_4 = g_4 \times f_4^T = \begin{pmatrix} 0 \\ 0 \\ -1 \\ -1 \end{pmatrix} \times \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{pmatrix}$$

La matriz asociativa  $A$  es:

$$A = \sum_{i=1}^{k=4} A_i = A_1 + A_2 + A_3 + A_4$$

$$A = \begin{pmatrix} -2 & -2 & -2 & -2 \\ 7/2 & 7/2 & 7/2 & 7/2 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 5/2 & -5/2 & 5/2 & -5/2 \\ 1/4 & -1/4 & 1/4 & -1/4 \\ 0 & 0 & 0 & 0 \\ 3/2 & -3/2 & 3/2 & -3/2 \end{pmatrix} + \begin{pmatrix} 1/6 & 1/6 & -1/6 & -1/6 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{pmatrix}$$

$$A = \begin{pmatrix} 2/3 & -13/3 & 1/3 & -14/3 \\ 19/4 & 17/4 & 11/4 & 9/4 \\ 5/2 & 3/2 & -1/2 & 1/2 \\ 3 & -1 & 0 & -2 \end{pmatrix}$$

Podemos demostrar que efectivamente esa matriz devuelve los vectores  $g_i$  cuando es multiplicada por  $f_i$ :

$$g_1 = Af_1 = \begin{pmatrix} 2/3 & -13/3 & 1/3 & -14/3 \\ 19/4 & 17/4 & 11/4 & 9/4 \\ 5/2 & 3/2 & -1/2 & 1/2 \\ 3 & -1 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \left( \frac{2}{3} - \frac{13}{3} + \frac{1}{3} - \frac{14}{3} \right) \\ \frac{1}{2} \left( \frac{19}{4} + \frac{17}{4} + \frac{11}{4} + \frac{9}{4} \right) \\ \frac{1}{2} \left( \frac{5}{2} + \frac{3}{2} - \frac{1}{2} + \frac{1}{2} \right) \\ \frac{1}{2} (3 - 1 + 0 - 2) \end{pmatrix}$$

$$g_1 = \begin{pmatrix} -4 \\ 7 \\ 2 \\ 0 \end{pmatrix}$$

Podemos demostrar que efectivamente esa matriz devuelve los vectores  $g_i$  cuando es multiplicada por  $f_i$ :

$$g_2 = Af_2 = \begin{pmatrix} 2/3 & -13/3 & 1/3 & -14/3 \\ 19/4 & 17/4 & 11/4 & 9/4 \\ 5/2 & 3/2 & -1/2 & 1/2 \\ 3 & -1 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{pmatrix} = \begin{pmatrix} \frac{2}{6} + \frac{13}{6} + \frac{1}{6} + \frac{14}{6} \\ \frac{19}{8} - \frac{17}{8} + \frac{11}{8} - \frac{9}{8} \\ \frac{5}{4} - \frac{3}{4} - \frac{1}{4} - \frac{1}{4} \\ \frac{3}{2} + \frac{1}{2} + 0 + 1 \end{pmatrix}$$

$$g_2 = \begin{pmatrix} 5 \\ 1/2 \\ 0 \\ 3 \end{pmatrix}$$

Podemos demostrar que efectivamente esa matriz devuelve los vectores  $g_i$  cuando es multiplicada por  $f_i$ :

$$g_3 = Af_3 = \begin{pmatrix} 2/3 & -13/3 & 1/3 & -14/3 \\ 19/4 & 17/4 & 11/4 & 9/4 \\ 5/2 & 3/2 & -1/2 & 1/2 \\ 3 & -1 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{pmatrix} = \begin{pmatrix} \frac{2}{6} - \frac{13}{6} - \frac{1}{6} + \frac{14}{6} \\ \frac{19}{8} + \frac{17}{8} - \frac{11}{8} - \frac{9}{8} \\ \frac{5}{4} + \frac{3}{4} + \frac{1}{4} - \frac{1}{4} \\ \frac{3}{2} - \frac{1}{2} + 0 + 1 \end{pmatrix}$$

$$g_3 = \begin{pmatrix} 1/3 \\ 2 \\ 2 \\ 2 \end{pmatrix}$$

Podemos demostrar que efectivamente esa matriz devuelve los vectores  $g_i$  cuando es multiplicada por  $f_i$ :

$$g_4 = Af_4 = \begin{pmatrix} 2/3 & -13/3 & 1/3 & -14/3 \\ 19/4 & 17/4 & 11/4 & 9/4 \\ 5/2 & 3/2 & -1/2 & 1/2 \\ 3 & -1 & 0 & -2 \end{pmatrix} \begin{pmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{pmatrix} = \begin{pmatrix} -\frac{2}{6} - \frac{13}{6} + \frac{1}{6} + \frac{14}{6} \\ -\frac{19}{8} + \frac{17}{8} + \frac{11}{8} - \frac{9}{8} \\ -\frac{5}{4} + \frac{3}{4} - \frac{1}{4} - \frac{1}{4} \\ -\frac{3}{2} - \frac{1}{2} + 0 + 1 \end{pmatrix}$$

$$g_4 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ -1 \end{pmatrix}$$

Notar que todos los vectores  $f_i$  son ortonormales:

$$(f_1 \cdot f_2) = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} \cdot \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{pmatrix} = \left( \frac{1}{4} - \frac{1}{4} + \frac{1}{4} - \frac{1}{4} \right) = 0$$

$$(f_1 \cdot f_3) = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} \cdot \begin{pmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{pmatrix} = \left( \frac{1}{4} + \frac{1}{4} - \frac{1}{4} - \frac{1}{4} \right) = 0$$

$$(f_1 \cdot f_4) = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} \cdot \begin{pmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{pmatrix} = \left( -\frac{1}{4} + \frac{1}{4} + \frac{1}{4} - \frac{1}{4} \right) = 0$$

Notar que todos los vectores  $f_i$  son ortonormales:

$$(f_2 \cdot f_3) = \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{pmatrix} \cdot \begin{pmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{pmatrix} = \left( \frac{1}{4} - \frac{1}{4} - \frac{1}{4} + \frac{1}{4} \right) = 0$$

$$(f_2 \cdot f_4) = \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{pmatrix} \cdot \begin{pmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{pmatrix} = \left( -\frac{1}{4} - \frac{1}{4} + \frac{1}{4} + \frac{1}{4} \right) = 0$$

$$(f_3 \cdot f_4) = \begin{pmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{pmatrix} \cdot \begin{pmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{pmatrix} = \left( -\frac{1}{4} + \frac{1}{4} - \frac{1}{4} + \frac{1}{4} \right) = 0$$

## Modelo de Memorias Asociativas: Actividad

Y además están todos normalizados:

$$\|f_1\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \sqrt{\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4}} = \sqrt{\frac{4}{4}} = \sqrt{1} = 1$$

$$\|f_2\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2} = \sqrt{\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4}} = \sqrt{\frac{4}{4}} = \sqrt{1} = 1$$

$$\|f_3\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2} = \sqrt{\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4}} = \sqrt{\frac{4}{4}} = \sqrt{1} = 1$$

$$\|f_4\| = \sqrt{\left(-\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2} = \sqrt{\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4}} = \sqrt{\frac{4}{4}} = \sqrt{1} = 1$$

¿Qué ocurre si una de las entradas está deteriorada? ¿Qué diferencias notaremos en  $g_i$  al modificar algún valor de los vectores de entrada  $f_i$ ?

$$f_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} \rightarrow f_1^d = \begin{pmatrix} \mathbf{1/3} \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}$$

$$g_1^d = Af_1^d = \begin{pmatrix} 2/3 & -13/3 & 1/3 & -14/3 \\ 19/4 & 17/4 & 11/4 & 9/4 \\ 5/2 & 3/2 & -1/2 & 1/2 \\ 3 & -1 & 0 & -2 \end{pmatrix} \begin{pmatrix} \mathbf{1/3} \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} \frac{2}{9} - \frac{13}{6} + \frac{1}{6} - \frac{14}{6} \\ \frac{19}{12} + \frac{17}{8} + \frac{11}{8} + \frac{9}{8} \\ 5 - \frac{3}{4} - \frac{1}{4} + \frac{1}{4} \\ 1 - \frac{1}{2} + 0 - 1 \end{pmatrix} = \begin{pmatrix} -4,111 \\ 6,208 \\ 1,375 \\ -0,5 \end{pmatrix}$$

¿Qué ocurre si una de las entradas está deteriorada? ¿Qué diferencias notaremos en  $g_i$  al modificar algún valor de los vectores de entrada  $f_i$ ?

$$f_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} \rightarrow g_1 = \begin{pmatrix} -4 \\ 7 \\ 2 \\ 0 \end{pmatrix}, \quad f_1^d = \begin{pmatrix} 1/3 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} \rightarrow g_1^d = \begin{pmatrix} -4,111 \\ 6,208 \\ 1,375 \\ -0,5 \end{pmatrix}$$

Los valores obtenidos no son exactos, pero son *relativamente* cercanos al vector original.

¿Qué ocurre si eliminamos una columna entera de la matriz asociativa? ¿Qué diferencias notaremos respecto al vector  $g_i$  obtenido?

$$A = \begin{pmatrix} 2/3 & -13/3 & 1/3 & -14/3 \\ 19/4 & 17/4 & 11/4 & 9/4 \\ 5/2 & 3/2 & -1/2 & 1/2 \\ 3 & -1 & 0 & -2 \end{pmatrix} \rightarrow A^d = \begin{pmatrix} \mathbf{0} & -13/3 & 1/3 & -14/3 \\ \mathbf{0} & 17/4 & 11/4 & 9/4 \\ \mathbf{0} & 3/2 & -1/2 & 1/2 \\ \mathbf{0} & -1 & 0 & -2 \end{pmatrix}$$

$$A^d f_1 = \begin{pmatrix} 0 & -13/3 & 1/3 & -14/3 \\ 0 & 17/4 & 11/4 & 9/4 \\ 0 & 3/2 & -1/2 & 1/2 \\ 0 & -1 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \left( 0 - \frac{13}{3} + \frac{1}{3} - \frac{14}{3} \right) \\ \frac{1}{2} \left( 0 + \frac{17}{4} + \frac{11}{4} + \frac{9}{4} \right) \\ \frac{1}{2} \left( 0 + \frac{3}{2} - \frac{1}{2} + \frac{1}{2} \right) \\ \frac{1}{2} (0 - 1 + 0 - 2) \end{pmatrix} = \begin{pmatrix} -4,333 \\ 4,625 \\ 0,75 \\ -1,5 \end{pmatrix}$$

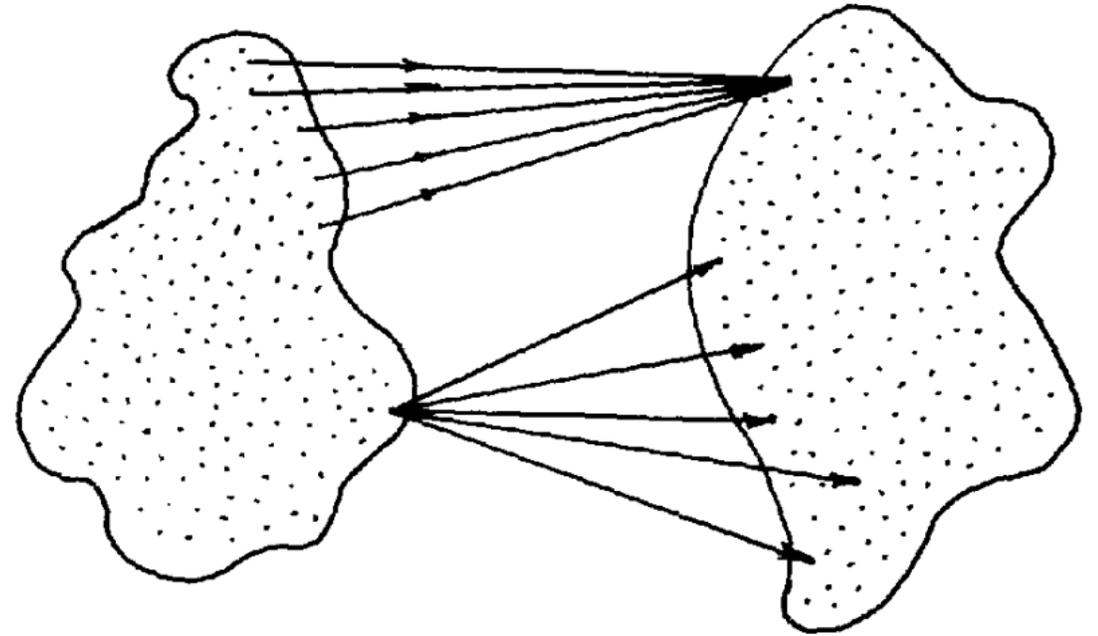
¿Qué ocurre si eliminamos una columna entera de la matriz asociativa? ¿Qué diferencias notaremos respecto al vector  $g_i$  obtenido?

$$A = \begin{pmatrix} 2/3 & -13/3 & 1/3 & -14/3 \\ 19/4 & 17/4 & 11/4 & 9/4 \\ 5/2 & 3/2 & -1/2 & 1/2 \\ 3 & -1 & 0 & -2 \end{pmatrix} \rightarrow A^d = \begin{pmatrix} 0 & -3 & -1 & -6 \\ 0 & 17/4 & 11/4 & 9/4 \\ 0 & 3/2 & -1/2 & 1/2 \\ 0 & -1 & 0 & -2 \end{pmatrix}$$

$$g_1 = \begin{pmatrix} -4 \\ 7 \\ 2 \\ 0 \end{pmatrix} \rightarrow g_1^d = \begin{pmatrix} -4,333 \\ 4,625 \\ 0,75 \\ -1,5 \end{pmatrix}$$

¡Estamos eliminando el 25% de las conexiones!

Puesto que al modificar una de las entradas del vector  $f_i$  obtuvimos un vector de salida  $g_i$  “parecido” al vector  $g_i$ , podemos decir que este modelo de redes neuronales es **menos susceptible a entradas deterioradas** (daños en la capacidad sensorial).

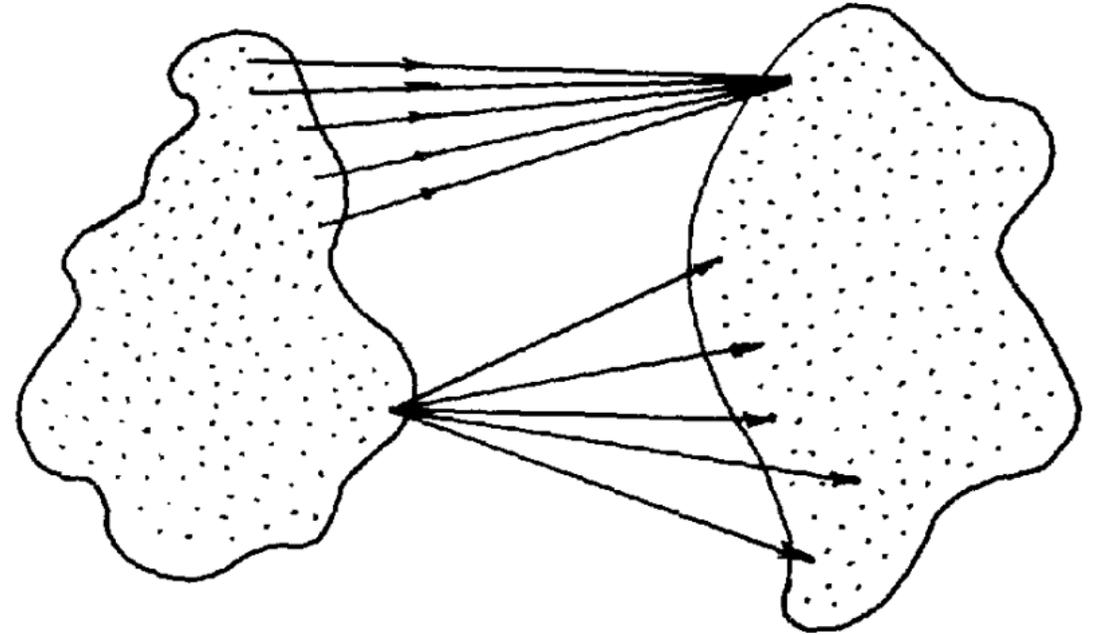


$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

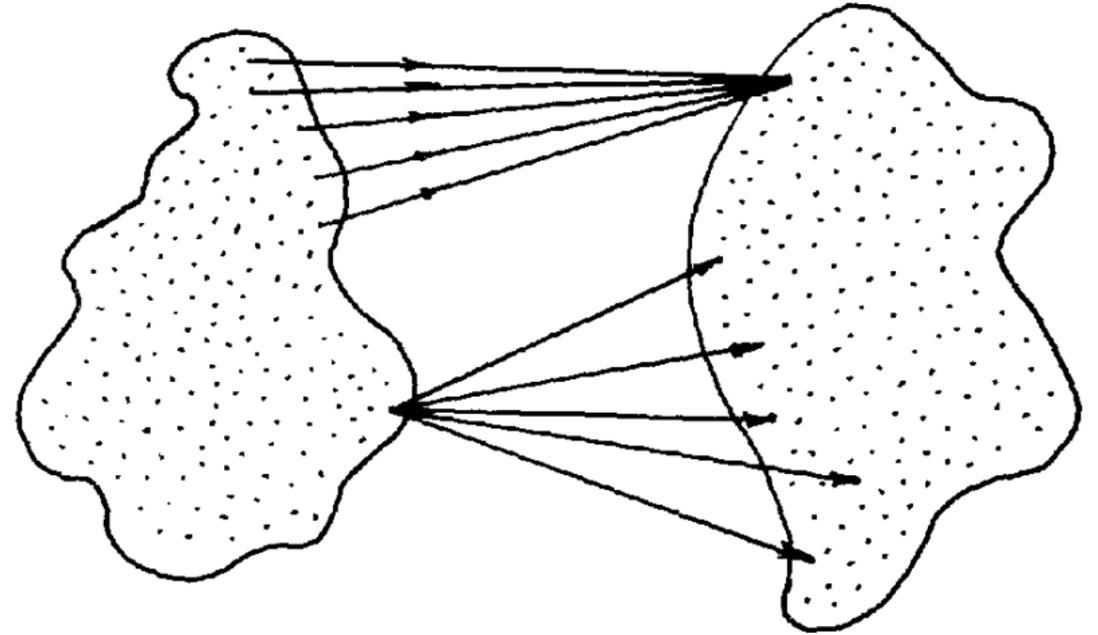
Incluso si la matriz asociativa  $A$  está dañada (Ej.: eliminando una columna), aún así es posible obtener vectores de salida  $g_i$  “similares” a los vectores originales.

Podemos decir que estos modelos son **resistentes a daños locales** del ensamble de la red neuronal.



$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} \quad g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

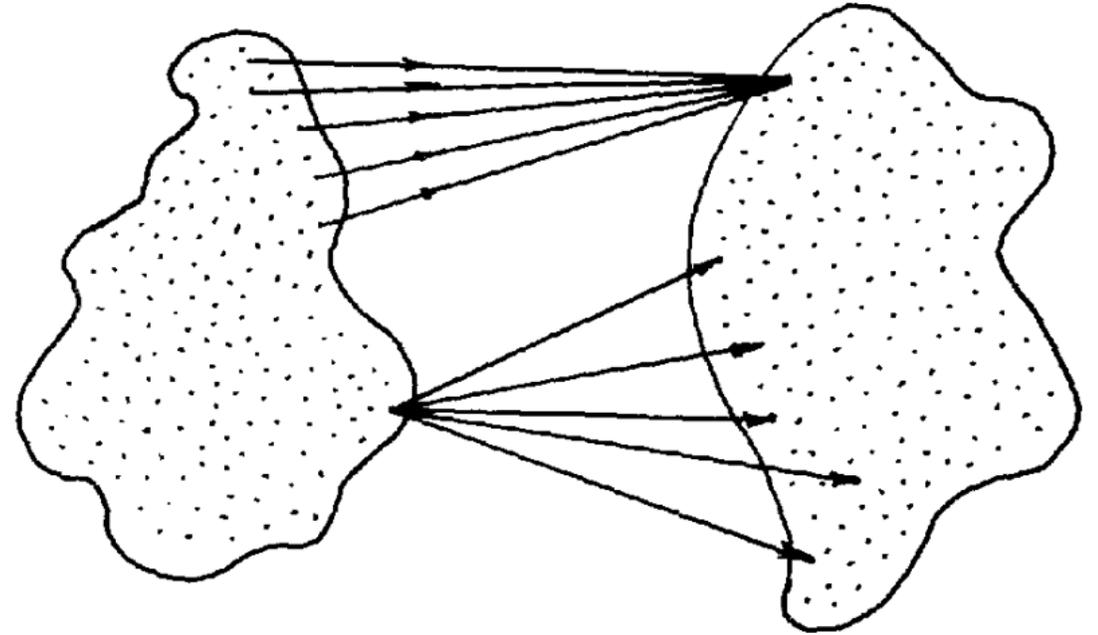
Si computamos un vector de entrada  $f_i$  que es una combinación lineal de otros dos vectores  $f_1, f_2$ , la matriz asociativa daría reconocimiento a un vector  $g_i$  que sería combinación lineal de los vectores de salida  $g_1, g_2$ . Decimos que la matriz asociativa es capaz de **Interpolar información.**



$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

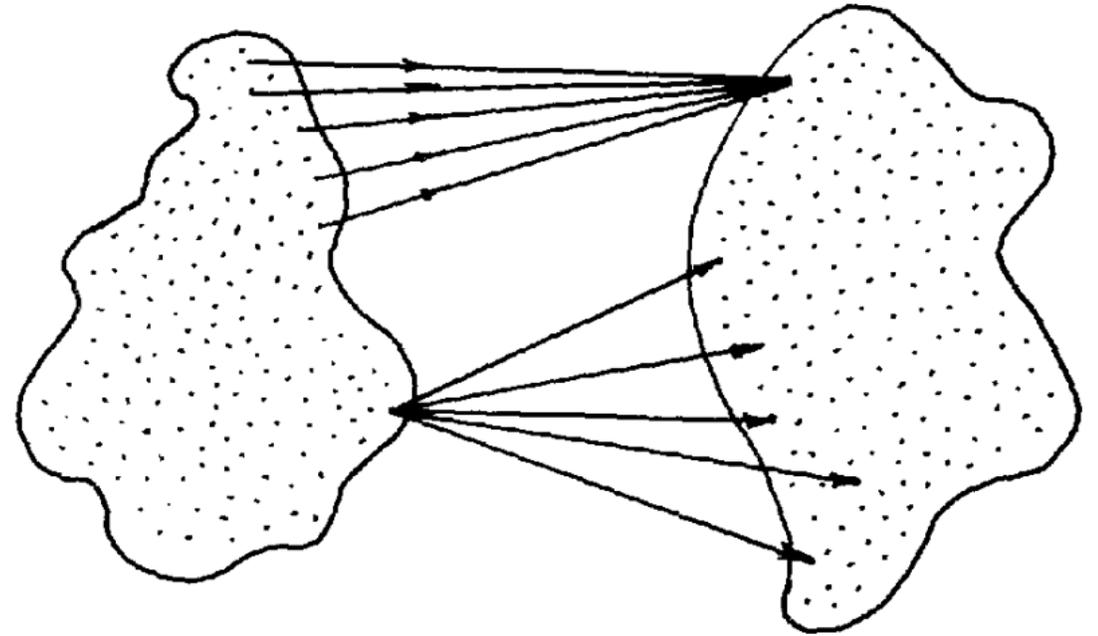
Una matriz asociativa construida a partir de vectores  $f_i$  que no son ortogonales entre sí, es capaz de generar vectores de salida  $g_i$  que poseen ciertas propiedades en común. Este tipo de **respuesta es ruidosa.**



$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

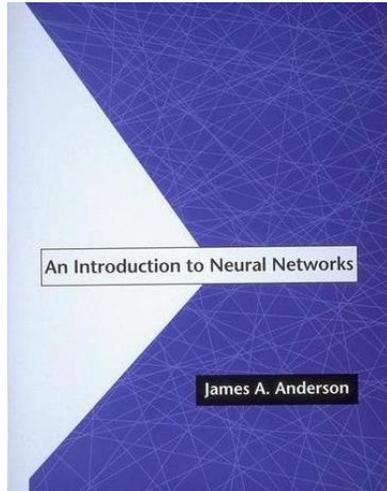
Puesto que los vectores  $f_i$  comparten cierta información entre ellos, a esta capacidad del modelo se la ha propuesto como un posible origen de la **Conceptualización**, particularmente para las redes neuronales en mamíferos.



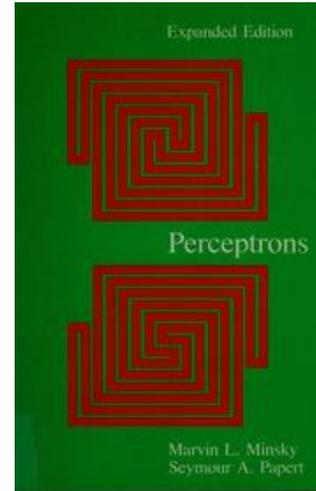
$$f_i = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$g_i = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

# Bibliografía



Anderson, J. A. (1995). *An introduction to neural networks*. MIT Press.  
 Cap. 2: *Synaptic integration and neuron models*. Págs.: 48-62.  
 Cap. 6: *The linear associator: background and foundations*. Págs.: 163-174.  
**[Disponible en Biblioteca]**



Minsky, M., & Papert, S. (1969). *Perceptrons: an introduction to computational geometry*. MIT Press.  
**[Disponible en Biblioteca]**



Mizraji, E. (2010). *En busca de las leyes del pensamiento. Una mirada desde la era de la información*. Ediciones Trilce.  
**[Disponible en Biblioteca]**

Este es un libro divulgativo y como tal no constituye un texto de estudio para esta temática. No obstante, puede resultarles una lectura interesante. Eduardo Mizraji fue docente de la Sección Biofísica y Biología de Sistemas, y ha dictado el curso de Biofísica en años anteriores.

# Anexo

**Demostrar:**

Sea  $f_{ij} = \alpha f_i + \beta f_j$  con  $\alpha, \beta \in \mathbb{R}$ , con vectores de salida  $g_i, g_j$  respectivamente. El vector de salida  $g_{ij}$  estará dado por:

$$g_{ij} = \alpha g_i + \beta g_j$$

Sean los pares de entrada-salida  $\{(f_i, g_i), (f_j, g_j)\}$  generadores de la matriz asociativa  $A_{ij}$ , tales que:

$$g_i = A_{ij}f_i$$

$$g_j = A_{ij}f_j$$

La combinación lineal  $f_{ij}$  es:

$$f_{ij} = \alpha f_i + \beta f_j$$

Con  $\alpha, \beta \in \mathbb{R}$ . Por tanto, el vector de salida correspondiente  $g_{ij}$  es:

$$g_{ij} = A_{ij}f_{ij}$$

$$g_{ij} = A_{ij}(\alpha f_i + \beta f_j)$$

$$g_{ij} = \alpha A_{ij}f_i + \beta A_{ij}f_j$$

$$g_{ij} = \alpha g_i + \beta g_j$$

*Quod erat demonstrandum.*