

Guía y ejemplo de la tercera parte de la Actividad 1 (modelado y prototipado de un sistema de IA bioinspirado)

Determinación del objetivo: ¿qué se os pide en esta parte de la Actividad 1?

Lectura: Lectura y práctica aplicada. Diferencias entre el sistema nervioso y los sistemas de computación artificiales. Prototipo de un sistema artificial

Objetivos

El objetivo de esta actividad es, por un lado, realizar una lectura sobre un texto técnico en el ámbito de la neurociencia cognitiva para sacar conclusiones sobre las principales diferencias entre los mecanismos de procesamiento de información en el cerebro y en los sistemas artificiales.

La segunda parte de la práctica tiene como objetivo mismo buscar, filtrar y encontrar un artículo desde la perspectiva simbólica para así poder desarrollar una visión crítica de las aplicaciones de la neurociencia cognitiva.

Deberás realizar una lectura comprensiva y detallada del siguiente artículo y encontrar un artículo desde la perspectiva simbólica que muestre una visión crítica desde el simbolismo.

Como segunda parte, se ha de buscar un artículo/texto, comprobar su fiabilidad y entender las diferencias y similitudes de la responsabilidad y el simbolismo.

Como tercera parte se propone un prototipo de sistema artificial, no una analogía completa del cerebro, sino se ha de escoger una analogía simbólica o corpórea y prototipar un modelo.

En esta tercera parte de la Actividad 1, más específicamente, lo que se nos pide es prototipar cómo implementaríamos, en un sistema de IA, alguno de los múltiples procesos que lleva a cabo el sistema cognitivo humano, dentro de cualquiera de sus funciones.

Para llevar esto a cabo, se sugiere los siguientes **pasos**:

- **Pensar en algún proceso cognitivo, propio del cerebro humano, que nos interese o nos llame la atención.** Ejemplos:
 - Perceptivos
 - Percepción (identificación y reconocimiento) de figuras u objetos.
 - Percepción (identificación y reconocimiento) de caras.
 - Mnésicos (relativos a la memoria)

- Memorización (estrategias mentales para ello) de cuantos estímulos sea posible dentro de un cierto elenco (por ejemplo, palabras de una lista o imágenes de una ilustración) y en un tiempo determinado.
- Categorización: capacidad de la memoria de distribuir la información almacenada de una manera eficiente, de cara a su posterior recuerdo (por ejemplo, almacenamiento de los conceptos de “elefante” y “perro” en la categoría “animales”, frente a “martillo” y “destornillador” en la categoría “herramientas”, etc).
- Memoria de trabajo: cuántos elementos de información distintos es capaz de retenerse en la memoria durante el tiempo en el que se realiza alguna actividad con ellos (por ejemplo, números al realizar operaciones matemáticas, o bien trazos distintos al tener que dibujar una cierta figura tras haberla observado previamente, etc.).
- Atencionales
 - Selección de información: capacidad de centrarse, de entre todos los estímulos que llegan perceptivamente al sistema, exclusivamente en aquellos que son relevantes para la tarea o actividad en curso.
 - Control ejecutivo de la atención: capacidad del sistema cognitivo humano de supervisar los resultados de sus propias acciones y aprender, con ello, a inhibir ciertas respuestas inmediatas, reflejas o *sobreaprendidas* ante ciertos estímulos cuando dichas respuestas no son adecuadas en la tarea en curso.
 - Atención dividida: estrategias para simultanear con el mayor éxito posible la realización de dos o más tareas a un mismo tiempo.
- Afectivos (emoción)
 - *Appraisal*: concepto que se refiere a cómo, ante un cierto estímulo o situación, el sistema cognitivo de las personas determina cómo de demandante es esa situación y de cuántos recursos se dispone ante ella para saber, entonces, qué emoción hacer sentir a la persona (estrés, miedo, motivación positiva, alegría, tristeza, etc.).
 - *Congruencia del estado de ánimo*: concepto que se refiere al hecho comprobado de que, estando en un cierto estado de ánimo, a las personas les cuesta menos recordar aquellas informaciones que aprendió o retuvo en el pasado encontrándose en ese mismo estado de ánimo, frente a información retenida cuando se encontraba en otros estados de ánimo distintos.
- Ejecutivos:
 - Toma de decisiones complejas y elaboradas en situaciones que las requieren.
 - Toma de decisiones rápidas en situaciones demandantes y con poco tiempo para elaborarlas (por ejemplo, situaciones peligrosas o biológicamente amenazantes, que implicarán decisiones rápidas, y en los que, además, la persona podrá encontrarse bajo el influjo de fuertes emociones).
 - Planificación: cómo las personas determinan cuál es la secuencia de pasos o acciones más eficiente para conseguir un determinado objetivo o meta (aunque sea simple, puntual y cotidiano).
 - Flexibilidad cognitiva: capacidad del sistema cognitivo humano para detectar que un cierto proceso o entorno ha dejado de funcionar conforme a las “reglas” habituales y es necesario adaptarse a él mediante otras nuevas.

- Motores (relativos al movimiento corporal):
 - Cómo decide el sistema cognitivo la forma adecuada de realizar un movimiento corporal para cualquier fin necesario.
 - Coordinación motora: estrategias del sistema cognitivo para simultanear, de manera lo más eficiente posible, dos o más movimientos simultáneos (tiene mucho que ver con la anteriormente mencionada atención dividida).
 - Sociales:
 - *Teoría de la Mente*: es la capacidad de la mente de una persona para formarse una idea de qué pensamientos experimentará otra persona en una determinada situación.
 - *Empatía*: similar a la anterior, pero desde el punto de vista emocional. Es la capacidad de la mente de una persona para formarse una idea de qué emociones y sentimientos experimentará otra persona en una determinada situación.
 - Creativos:
 - Cómo consigue el sistema cognitivo realizar razonamientos de abstracción y analogía.
 - Creatividad artística.
-
- **Buscar en fuentes bibliográficas el modelo cognitivo que la Psicología Cognitiva ha propuesto para describir ese proceso mental humano que hemos elegido. Estos modelos suelen tener, por su propia naturaleza, forma de esquema.**
 - **Pensar cómo lo implementaríamos en un sistema de IA. IMPORTANTE: se trata de un prototipo esquematizado, por lo que bastará con quedarnos en el nivel de algoritmo (en forma de diagrama de flujo, por ejemplo). No hay que llegar a proponer *scripts* concretos con código en ningún lenguaje de programación en particular. Para ello, traduciremos el esquema del modelo cognitivo a un esquema que resulte más implementable en una máquina (por ejemplo, pasando de una estructura formal a un diagrama de flujo o, incluso, una estructura conexionista).**
 - **Decidir qué tipo de modelo de alto nivel elegiremos (entidad-relación, jerárquico, de red, relacional...), qué elementos introduciremos y qué relaciones existirán entre ellos (en este punto es donde puede ayudarnos el haber planteado una estructura conexionista)**
 - **Qué recursos y metodologías propias de la IA pueden sernos útil plantear aquí. Por ejemplo, *machine learning*, *deep learning*, redes neuronales, etc.**

Ejemplo

Vamos a proponer aquí un ejemplo de resolución de esta tarea. Lo haremos siguiendo la propia estructura sugerida anteriormente.

1. **Pensar en algún proceso cognitivo, propio del cerebro humano, que nos interese o nos llame la atención.**

En este caso, hemos escogido como campo de estudio el de los procesos atencionales y, en particular, el **control ejecutivo de la atención**, que es la capacidad del sistema cognitivo humano de supervisar los resultados de sus propias acciones y aprender, con ello, a inhibir ciertas respuestas inmediatas, reflejas o *sobreaprendidas* ante ciertos estímulos cuando dichas respuestas no son adecuadas en la tarea en curso.

Existen multitud de métodos para poner a prueba las capacidades de una persona en lo relativo a este control ejecutivo de la atención. Una de las más conocidas es el denominado **Test de Stroop**. Se trata de una prueba o ejercicio en el que se le muestra a la persona una serie de palabras (en particular, nombres de colores, como “azul”, “verde” o “rojo”), de modo que la persona deba responder **diciendo en voz alta, lo más rápidamente que pueda, en qué color está escrita la palabra**. El “problema” es que **el color que se lee en la palabra no coincide con el de la tinta con la que aparece escrita**. La dificultad de la tarea reside en que la persona deberá, entonces, tratar de decir el color de la tinta lo más rápido posible sin caer en dejarse llevar por leer el contenido de las palabras como tal, ya que esta última es una conducta o acción que, sin embargo, está muy aprendida y es demasiado automática en las personas.

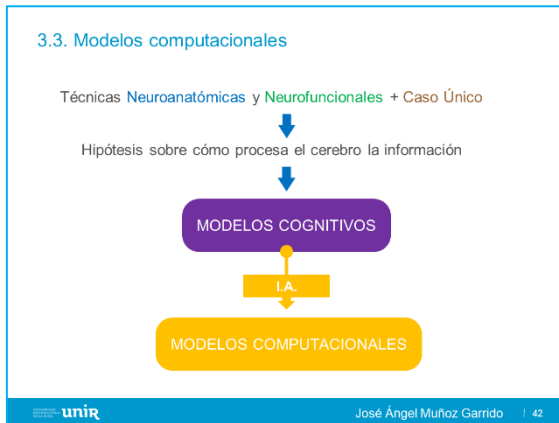
Por ejemplo:

ROJO	ROJO
VERDE	AZUL
AZUL	VERDE
VERDE	ROJO
AMARILLO	AMARILLO

En este ejemplo, lo ideal sería poder decir, lo más rápido posible (y yendo de arriba abajo en cada columna), “azul – amarillo – verde – rojo – azul”, etc.; en cambio, sería un error decir “rojo – verde – azul – verde – amarillo”, etc., ya que esto último sería leer, y es justo la respuesta cognitiva rápida que hay que tratar de evitar.

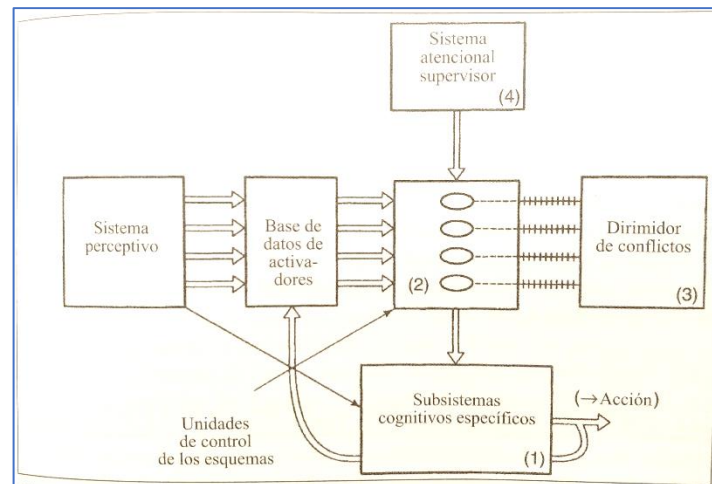
2. Buscar en fuentes bibliográficas el modelo cognitivo que la Psicología Cognitiva ha propuesto para describir ese proceso mental humano que hemos elegido. Estos modelos suelen tener, por su propia naturaleza, forma de esquema.

Un **modelo cognitivo** es la esquematización formal que los científicos (en este caso, los psicólogos cognitivos) han elaborado para describir cómo funciona un determinado proceso cognitivo, con base en las hipótesis y hallazgos científicos disponibles hasta la fecha, tal como se refleja en el esquema a continuación de estas líneas.



En el caso del ejemplo que hemos elegido, el proceso cognitivo involucrado es, como hemos dicho, el **control ejecutivo de la atención**. Revisando la bibliografía científica que existe al respecto, podemos ver que han sido varios los **modelos cognitivos** propuestos para describir cómo funciona este proceso. En nuestro caso, a modo de mero ejemplo, hemos escogido el de Norman y Shallice. Este modelo concibe que hay tareas que pueden desarrollarse de manera

relativamente automática (sin control consciente), mientras que otras sí necesitan una dedicación deliberada de los recursos atencionales, especialmente cuando: 1) suponen planificar o tomar decisiones; 2) implican resolver problemas; 3) todavía no han sido aprendidas y/o son novedosas; 4) resultan técnicamente difíciles o incluso peligrosas; 5) requieren superar o inhibir una respuesta habitual fuertemente establecida. **Esto último es justo lo que sucede en la tarea de Stroop.**



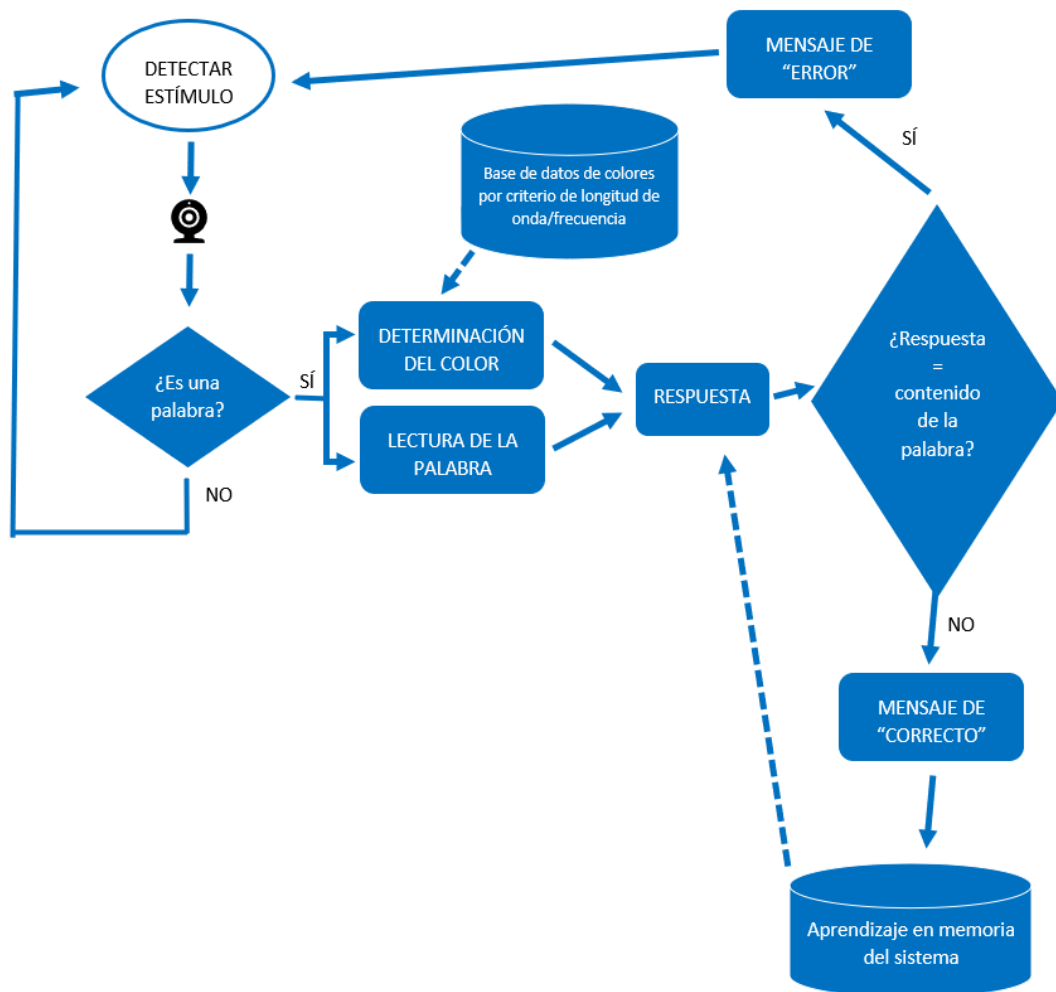
Tomado de Styles, E. (2010)

El modelo establece que, cuando llegan al sistema cognitivo los diversos estímulos a los que hay que responder, en seguida se activan los esquemas de las posibles respuestas que la persona posee ante ellos (en nuestro caso, en el test de Stroop, leer la palabra como tal, o bien decir el color en el que está escrita, etc.); sin embargo, no todas las respuestas están igual de aprendidas y automatizadas (leer la palabra es mucho más automático que decir su color, por ejemplo). En este sentido, el Sistema Atencional Supervisor (SAS) es el que permite a la persona aplicar un control consciente de las respuestas que escoge, mediante un “dirimidor de conflictos” que es el que resuelve, en cada caso, escoger la respuesta menos automática (decir el color de la palabra) antes que la más automatizada (leer la palabra como tal), ya que esto es lo que se demanda en la tarea de Stroop.

3. Pensar cómo lo implementaríamos en un sistema de IA. **IMPORTANTE:** se trata de un prototipo esquematizado, por lo que bastará con quedarnos en el nivel de algoritmo (en forma de diagrama de flujo, por ejemplo). No hay que llegar a proponer *scripts* concretos con código en ningún lenguaje de programación en

particular. Para ello, traduciremos el esquema del modelo cognitivo a un esquema que resulte más implementable en una máquina (por ejemplo, pasando de una estructura formal a un diagrama de flujo o, incluso, una estructura conexionista).

Ya tenemos el modelo cognitivo que queremos imaginar implementado en una máquina. Sin embargo, su esquematización es aún demasiado general o abstracta para un sistema de inteligencia artificial. Tal vez una mejor forma de aproximarnos al “lenguaje de las máquinas” sea representar lo mismo en términos de diagrama de flujo:



El funcionamiento de nuestra propuesta artificial se explica como sigue: nuestro sistema dispone de una cámara, con la finalidad de poder captar estímulos. Esto se corresponde con la fase más perceptiva del modelo de atención ejecutiva de Norman y Shallice; a continuación, el sistema habrá de determinar si el estímulo percibido es una palabra. En caso negativo, volverá a la fase inicial de detección de estímulos. En caso afirmativo, procederá a desplegar dos esquemas de acción simultáneos (en correspondencia con la base de activación de esquemas del modelo de Norman y Shallice): por un lado, la determinación del color de la palabra (a partir de una base de conocimiento previo), y por otro lado, la lectura de la misma; después, se dictamina y emite la respuesta definitiva (interpretable como la acción final en el esquema de Norman y Shallice). Entonces, el sistema determina si la respuesta dada coincide con el contenido de la palabra detectada. En caso afirmativo, ello constituye un error a efectos de la tarea de Stroop, por lo que el sistema emite un mensaje de “ERROR” y regresa a la fase inicial de detección de un nuevo estímulo; por el contrario, en caso negativo, el sistema

entenderá que ha sabido priorizar la respuesta de color por encima de la respuesta de lectura directa de la palabra, lo que constituye un acierto a efectos de la Tarea de Stroop. Ante este acierto, el sistema emite un mensaje de “CORRECTO” y, mediante técnicas de *machine learning*, lo almacena para contribuir a mejorar su propia capacidad de llevar a cabo correctamente la tarea de Stroop. En términos de Norman y Shallice, puede decirse que el sistema entrena su “sistema dirimidor de conflictos”.

4. **[OPCIONAL]** Decidir qué tipo de modelo de alto nivel elegiremos (entidad-relación, jerárquico, de red, relacional...), qué elementos introduciremos y qué relaciones existirán entre ellos (en este punto es donde puede ayudarnos el haber planteado una estructura conexionista).

Si escogemos, por ejemplo, un modelo de entidad-relación, podremos estructurar la tarea de Stroop mediante los conceptos de *objeto*, atributo y *relación*, como en la página siguiente:

- **Objetos:**
 - **Palabras mostradas**
 - **Atributos:**
 - **Contenido de la palabra:**
 - “Azul”
 - “Verde”
 - “Rojo”
 - “Amarillo”
 - **Color de la palabra:**
 - Azul
 - Verde
 - Rojo
 - Amarillo
 - **Relaciones entre ellas:**
 - Congruencia/incongruencia en contenido de las palabras / color en el que están escritas.
 - Aleatorización de los colores en su forma de aparición.
 - **Respuestas dadas**
 - **Atributos**
 - **Leer la palabra**
 - Acierto/Error
 - **Decir el color de la palabra**
 - Acierto/Error
 - **Relaciones entre ellas**
 - Acierto si el color dicho es el de la palabra
 - Fallo si el color dicho no es el de la palabra

Todo ello puede resumirse en la tabla de la página siguiente:

OBJETOS	ATRIBUTOS		RELACIONES
Palabras mostradas	Contenido de la palabra	“Azul”	Congruente con el color
			Incongruente con el color
		“Rojo”	Congruente con el color
			Incongruente con el color
		“Verde”	Congruente con el color
			Incongruente con el color
		“Amarillo”	Congruente con el color
			Incongruente con el color
	Color de la palabra	Azul	Congruente con el contenido
			Incongruente con el contenido
		Rojo	Congruente con el contenido
			Incongruente con el contenido
		Verde	Congruente con el contenido
			Incongruente con el contenido
		Amarillo	Congruente con el contenido
			Incongruente con el contenido
Respuestas dadas	Leer la palabra	Acierto/Error	Siempre erróneo
	Decir el color de la palabra	Acierto/Error	Acierto si el color dicho es el de la palabra
			Fallo si el color dicho no es el de la palabra

5. Qué recursos y metodologías propias de la IA pueden sernos útil plantear aquí. Por ejemplo, *machine learning*, *deep learning*, redes neuronales, etc.

Podemos emular un SAS (Sistema Atencional Supervisor) como el del modelo de Norman y Shallice si, por ejemplo, consideramos que registramos las respuestas del sistema de IA y las categorizamos en **Aciertos** y **Fallos**. Como en cualquier otro sistema de IA, esta distinción puede aprovecharse para ser implementada de cara a la auto-mejora de sí mismo que el sistema lleva a cabo.

- *Machine Learning*: si tomamos estadísticas de las respuestas, como el tiempo en el que son dadas y el hecho de que sean acertadas o fallidas, este tipo de metodología puede permitirnos, entonces, guiar al sistema en su aprendizaje.
- *Redes neuronales*: serían aplicables desde el mismo punto de vista, es decir, el de la supervisión de las propias respuestas para su posterior retroalimentación.