

# CI-0129 Inteligencia artificial

## Inteligencia más allá de la mente: ¿hasta dónde?

Charla preparada para el curso del Prof. Dr. Markus Eger por el

Prof. Dr. Alvaro de la Ossa Osegueda

[alvaro.delaossa@ucr.ac.cr](mailto:alvaro.delaossa@ucr.ac.cr)

Escuela de Ciencias de la Computación e Informática  
y Programa de Posgrado en Ciencias Cognoscitivas

18 de junio de 2020



**ECCI**

Escuela de  
Ciencias de la  
Computación e  
Informática

Esta charla es acerca de los *límites* de la inteligencia artificial

¿Cuáles son? ¿Cómo son?

¿Qué tan cerca estamos?

.

Para poder describir esos límites,  
es necesario identificar los objetivos de la IA

# Objetivos de la inteligencia artificial: puntos de vista diversos

## *Ciencias Cognoscitivas*

¿Mente = Máquina?

La prueba de Turing, el argumento del cuarto chino, ¿otros experimentos posibles?

→ Explicación de las capacidades cognoscitivas

## *Ciencias de la Computación*

¿Qué dispositivos inteligentes podemos construir?

¿Otros paradigmas sin las limitaciones de la Máquina de Turing?

→ Construcción de capacidades cognoscitivas

## *Inteligencia Artificial*

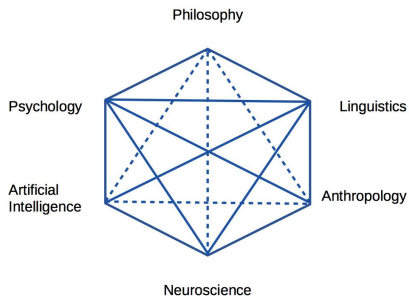
¿Qué capacidades computacionales y de otros tipos son necesarias para construir sistemas inteligentes?

## El marco de trabajo de la IA en las Ciencias Cognoscitivas

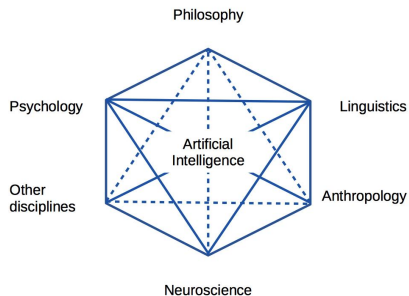
Construcción de *modelos plausibles*  
de las *capacidades mentales*.

# Cambio de enfoque de las Ciencias Cognoscitivas

*Ciencias cognoscitivas*: "el estudio interdisciplinario de la mente y de la inteligencia, incluyendo la filosofía, la psicología, la inteligencia artificial, la neurociencia, la lingüística y la antropología" (Stanford Encyclopedia of Philosophy).



1985



2020

# Objetivos de la inteligencia artificial: la perspectiva cognoscitiva

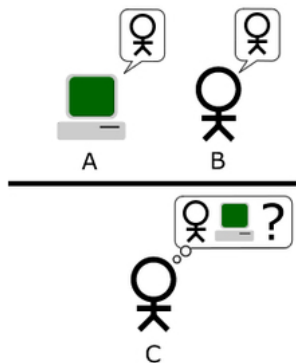
## *Plataforma de experimentación*

Servir de plataforma para verificar teorías de las demás disciplinas cognoscitivas

## *Dispositivos inteligentes*

Construir dispositivos cuya inteligencia (capacidad para resolver problemas) se confunda con la inteligencia humana

# Inteligencia artificial



La prueba de Turing

- Nombre acuñado por John McCarthy (1956, el taller de Dartmouth College).
- Basado en la predicción (¿reto?) de Alan Turing (1950):

*“Creo que en cerca de cincuenta años será posible programar computadoras con una capacidad de almacenamiento de cerca de  $10^9$ , para que jueguen el juego de la imitación tan bien, que un interrogador promedio no tendrá más de 70 por ciento de probabilidades de hacer la identificación correcta después de cinco minutos de interrogatorio.”*

# La noción de inteligencia

- Término comúnmente asociado con otros como *comprensión, entendimiento, capacidad mental, agilidad mental, aprendizaje, memoria, conciencia, categorización, planificación, creatividad, resolución de problemas, ...*
- $\nexists$  definición de consenso.





# Visiones de la inteligencia artificial en las Ciencias Cognoscitivas

## Dos preguntas separadas

### Inteligencia

¿Cómo es su naturaleza *biológica*?

¿Cómo se desarrolla?

¿Cómo la usamos?

¿Interesa *solo* la humana?

### Metáfora cerebro - máquina

¿Podemos replicar el cerebro?

¿Existen otros paradigmas de procesamiento que lo hagan posible?

# Visiones de la inteligencia artificial en las Ciencias Cognoscitivas

## *Funcionalista*

- Modelo del comportamiento inteligente de un agente, con base en observaciones de sus acciones.
- Problema: el modelo no dice nada de la estructura interna del agente
- Dificultad para replicar el comportamiento inteligente.

## *Cognitivista*

- Modelo de procesamiento de información que hace a un agente computacional exhibir el comportamiento inteligente de uno o más agentes.
- Ventaja: Mecanismo comparable computacionalmente (tiempo, espacio, precisión) con sujetos humanos.

# Enfoques de la inteligencia artificial en las Ciencias Cognoscitivas

Dos enfoques separados  
pero potencialmente complementarios

## **Explicar la cognición**

Modelos computacionales  
de las capacidades cognoscitivas

Dos enfoques: procesamiento simbólico,  
conexionismo

## **Construir dispositivos inteligentes**

Complementar (mejorar, extender) las  
capacidades cognitivas humanas,  
mejorar la calidad de vida.

# Simbolismo vrs. conexionismo

## Modelos simbólicos

Representaciones simbólicas explícitas,  
procesadas como hileras

Pueden ser aprendidas

Usadas para hacer inferencias y tomar  
decisiones de acción

## Modelos conexionistas

Redes de activación,  
procesadas en paralelo

Pueden ser aprendidas (proceso iterativo  
de ajuste de la activación de la red)

Sirven como predictores o clasificadores

# Un marco de trabajo para el estudio de las capacidades cognoscitivas<sup>1</sup>

<b>Dominio</b>	Las capacidades cognoscitivas
<b>Preguntas</b>	Asociadas a esas capacidades
<b>Supuestos sustantivos</b>	Supuestos teóricos (para la creación de modelos computacionales)
<b>Supuestos metodológicos</b>	Para delimitar las acciones del investigador

---

<sup>1</sup>von Eckardt, B. (1993). *What is Cognitive Science?*. MIT Press, Cambridge, Massachusetts, EE.UU.

# Supuestos acerca de las capacidades cognoscitivas

<b>Identificación</b>	El dominio de investigación es el de las capacidades cognoscitivas
<b>Propiedad</b>	Las capacidades cognoscitivas son intencionales, pragmáticamente evaluables, coherentes, confiables y productivas.
<b>Agrupamiento</b>	Las capacidades cognoscitivas de un ATN <sup>2</sup> conforman un <i>sistema</i> .

---

<sup>2</sup>Adulto típico normal

## Pregunta central

¿Es posible, con la investigación suficiente, dar respuesta a las preguntas básicas del marco de investigación de la IA en las Ciencias Cognoscitivas, y construir con ella una teoría unificada, empírica y conceptualmente aceptable?

# Preguntas acerca de las capacidades cognoscitivas

¿En qué consiste?

¿En virtud de qué se posee?

¿Cómo se ejercita?

¿Cómo interactúa con otras capacidades?



# Supuestos computacionales

**De enlace**

Cerebro = Computador

**Sistémico**

Dispositivo computacional, modelable

**Metodológicos**

Enfoque individual, autonomía y modularidad de las capacidades, diferencias individuales.

# El marco de trabajo de la IA en las Ciencias de la Computación

Construcción de *dispositivos inteligentes*

# La noción de inteligencia

- En las ciencias de la computación, la visión tradicional de la inteligencia se asocia casi exclusivamente a la noción de *resolución de problemas*.
  - Tampoco hay una definición de consenso, prevalece la idea general de que *un dispositivo inteligente debe resolver una clase de problemas mejor o igual a un sujeto humano experto*.
- Métricas: *efectividad, desempeño, precisión*, o alguna combinación de estas.

# Problemas fundamentales

## Representación

Lenguajes formales; estructuras complejas; algoritmos eficientes

## Organización, acceso

De las redes y sistemas neuronales biológicas, a representaciones gráficas *plausibles*

## Control, atención

Mecanismos de coordinación

## Aprendizaje

Mecanismos de adaptación

## Resolución de problemas

Comportamiento inteligente

# Visiones de la inteligencia artificial en las Ciencias de la Computación

## Dos preguntas separadas

### Usabilidad

¿Qué capacidad se pretende mejorar,  
cómo?

¿El dispositivo es autónomo?

¿Requiere de entrenamiento?

¿A qué población está dirigido?

### Desempeño

¿Tiempos de respuesta razonables?

¿Requiere de grandes cantidades de  
recursos computacionales?

¿Es posible construir un dispositivo  
similar más eficiente?

# Dos visiones de la inteligencia artificial

## *Funcionalista*

Resolución de problemas  
basada en el uso intensivo de  
heurísticas y métodos numéricos.

## *Computacional*

Modelos de procesamiento que requieren  
de grandes capacidades computacionales  
para modelar sistemas complejos.

# Inteligencia artificial: criterio de validez

<b>Desempeño</b>	Igual o más <i>efectivo</i> , más <i>preciso</i> , más <i>eficiente</i> (o una combinación de esos criterios) que un ser humano
<b>Usabilidad</b>	Dispositivos que <i>extienden</i> nuestras capacidades para resolver problemas y que nos resulta sencillo utilizar
<b>Apropiación</b>	Aceptación y adopción por las personas en sus tareas cotidianas por algún valor agregado
<b>Adaptación</b>	Mecanismos de aprendizaje para la personalización, la adaptación al entorno, etc.

# Enfoques de la inteligencia artificial en las Ciencias de la Computación

Dos enfoques separados  
pero potencialmente complementarios

## Construir dispositivos inteligentes

Replicar o extender las  
capacidades humanas

Dos enfoques: prototipado rápido,  
modelado cognitivo

## Optimizar el desempeño

Mejorar los dispositivos a través de la  
adaptación continua

Enfoque: experimentación y ajuste de  
parámetros



# Necesidad de grandes capacidades computacionales

## Desempeño (el problema)

Representaciones simbólicas explícitas,  
procesadas como hileras

Pueden ser aprendidas

Usadas para hacer inferencias y tomar  
decisiones de acción

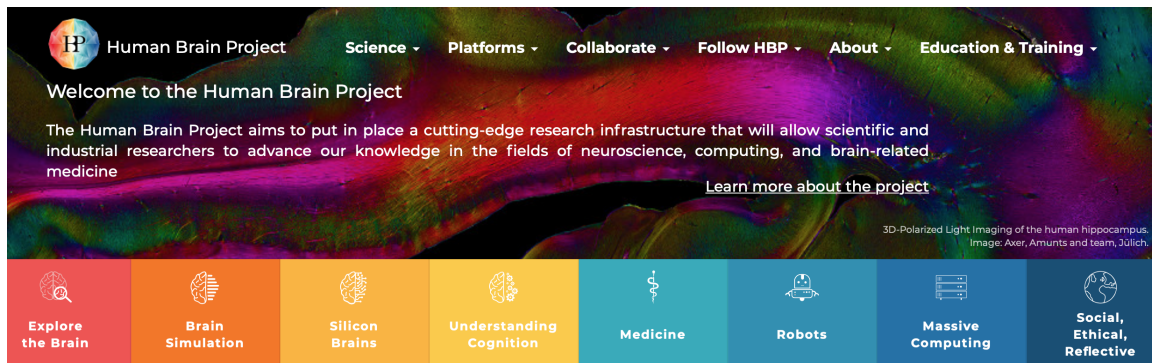
## Comp. de alto rendimiento (la solución)

Redes de activación, procesadas en  
paralelo

Aprendizaje: ajuste

Sirven como predictores o  
clasificadores

# Ejemplo de investigación de punta



Human Brain Project: <https://www.humanbrainproject.eu/>

# Los límites de la inteligencia artificial

Preguntas abiertas, retos, amenazas y oportunidades

# Límites de la inteligencia artificial, 1 de 2

**Mente = Computador**

Necesidad de nuevos y mejores experimentos para juzgar la inteligencia de un dispositivo computacional.

**Plausibilidad**

Falta de una definición operativa general de *plausibilidad*: criterios estructurales y funcionales de expertos.

**Dispositivos inteligentes**

Escalabilidad, desempeño temporal y espacial.

**Computabilidad  
y tratabilidad**

¿Nuevos paradigmas? ¿La computación cuántica puede ofrecer algo?

# Límites de la inteligencia artificial, 2 de 2

## Limitaciones de los modelos funcionalistas

Ausencia de conocimiento sobre la estructura del dispositivo

## Modelos simbólicos

Limitaciones de la Máquina de Turing (desempeño, (in)decidibilidad)

## Modelos conexionistas

Falta de integración de modelos de diversas capacidades (mucho trabajo en coordinación visión-movimiento, muy escaso en interacción memoria-emociones)

## Desempeño

Cambio de enfoque de *computólogos* de la IA hacia la CAR, alejamiento de las Ciencias Cognoscitivas.

# Preguntas abiertas: ¿dónde se integran la IA y la CAR?

## Máquinas

¿Es posible construir arquitecturas específicas enfocadas en el procesamiento simbólico, en lugar del uso intensivo de métodos numéricos?

## Problemas

¿Es posible construir una representación jerárquica que relacione representaciones subsimbólicas con representaciones simbólicas, para facilitar el proceso de abstracción de conceptos a partir de representaciones conexionistas?

## Explicaciones

Las representaciones subsimbólicas son difíciles de entender, y no proveen explicaciones. ¿Qué hay en una representación subsimbólica que pueda ser asociado a una explicación?

# Preguntas abiertas: ¿qué hay con la ética de la IA?

## Seguridad e integridad

¿Quién y cómo debe velar por la seguridad e integridad de las personas que utilizan dispositivos inteligentes?

## Responsabilidad social

¿Existe algún beneficio social del desarrollo de cierto dispositivo inteligente?

¿Cuál es el impacto de la incorporación de dispositivos inteligentes en la vida de las personas?

¡¡ Muchas gracias !!

Estoy a su disposición para responder sus consultas:

`alvaro.delaossa@ucr.ac.cr`