

# Resumen

## Inteligencia Artificial

Emiliano Rios Gavagnin  
emiliano.rios@est.fi.uncoma.edu.ar  
FAI-1113

2018

## 1. Inteligencia Artificial

### 1.1. ¿Qué es la Inteligencia Artificial?

Existen variadas definiciones según el autor y su perspectiva, pero todas ellas se basan en los mismos conceptos: según la forma de procesar el pensamiento y razonamiento o según el comportamiento. La otra forma de analizar las definiciones es si se basan en cuanto a la similitud que existe con el humano o con la racionalidad. Se asume que algo es racional cuando hace lo “correcto” según la información que tiene.

Piensan como humanos	Piensan racionalmente
<<El nuevo y excitante esfuerzo de hacer que los computadores piensen... Maquinas con mentes, en el mas amplio sentido literal>> <<La automatización de actividades que vinculamos con procesos de pensamiento humano, actividades como la toma de decisiones, resolución de problemas, aprendizaje>>	<<El estudio de las facultades mentales mediante el uso de modelos computacionales>> <<El estudio de los cálculos que hacen posibles percibir, razonar y actuar>>
Actúan como humanos	Actúan racionalmente
<<El arte de desarrollar maquinas con capacidad de desarrollar funciones que cuando son realizadas por personar requieren inteligencia>> <<El estudio de cómo lograr que los computadores realicen tareas que, por el momento, los humanos hacen mejor>>	<<La inteligencia computaciones es el estudio del diseño de agentes inteligentes>> <<IA... esta relacionada con conductas inteligentes en artefactos>>

Figura 1: Algunas definiciones de inteligencia artificial, agrupadas en cuatro categorías.

#### 1.1.1. Actuar como humanos

El **Test de Turing** fue diseñado para proveer una definición satisfactoria y operaciones de inteligencia. Una computadora pasa el test si un interrogador humano, luego de un conjunto de preguntas, no puede distinguir si las respuestas fueron de una persona o una computadora. Por lo tanto, para que una computadora pase la prueba, debe poseer las siguientes capacidades:

- Procesamiento de lenguaje natural
- Representación de conocimiento

- Razonamiento automático
- Aprendizaje (Machine learning)

También existe el denominado **Test de Turing Total** el cual involucra el aspecto físico de la entrevista. Incluye una señal de vídeo para que el interrogador pueda evaluar las capacidades perceptivas. Para pasar este test la computadora debe tener las siguientes capacidades:

- Visión artificial
- Robótica

### 1.1.2. Pensar como humanos

Para poder decir que un programa piensa como humano, primero deberíamos definir como piensa un humano. Nos deberíamos meter dentro de la mente humana para ver como trabaja. Existen 3 formas de hacer esto:

- A través de introspección, tratando de ver nuestros pensamientos.
- A través de experimentos psicológicos, viendo a una persona en acción.
- A través de la visualización cerebral.

Si las “entradas/salidas” del pensamiento humano coinciden con la del programa, esto sería evidencia de que parte del mecanismo del programa podría también operar en los humanos.

### 1.1.3. Pensar racionalmente

Su razonamiento debe ser correcto. Debe obtener conclusiones correctas partiendo de premisas correctas.

### 1.1.4. Actuar racionalmente

Un agente es algo que actúa. Pero los agentes se suponen que tienen atributos diferentes que los distinguen de un simple programa, por ejemplo operar autonomamente, percibir su ambiente, adaptarse al cambio, etc. Un agente racional es aquel que actúa para lograr el mejor resultado o, cuando hay incertidumbre, el mejor esperado.

Parte de ser un agente racional es hacer inferencias correctas, razonar lógicamente. Por otra parte existen situaciones donde no hay algo correcto que hacer pero algo debe hacerse. Hay formas de actuar racionalmente que no involucran inferencia.

## 1.2. Objeciones a Turing

### 1.2.1. Teológica

- **Objeción:** *“Pensar es una función del alma del hombre. Dios ha dado un alma a cada hombre y mujer, pero no así a animales y maquinas. Por lo tanto, estos no pueden pensar.”*
- **Respuesta:** Turing responde de dos formas. En primer lugar, si Dios fuera omnipotente, entonces debería poder darle alma a las maquinas. Además, los humanos somos capaces de procrearnos y dar a luz a humanos con alma, por lo que no estamos robando su poder, solo somos instrumentos para la creación de “envases de almas”, por lo que crear máquinas con almas no sería muy distinto. En segundo lugar, Turing marca como se dio un argumento similar a las teorías de Copernico y sin embargo estaba en lo cierto. Rechazamos ese argumento ante evidencia empírica, y lo mismo debería suceder aquí.

### 1.2.2. “Cabeza en la arena”

- **Objeción:** *“Las consecuencias de crear maquinas que piensan serian horribles, por lo que deberíamos esperar que no suceda y suponer que algo así nunca sucederá.”*
- **Respuesta:** Turing simplemente marca que, al igual que la objeción teológica, la objeción esta basada en la idea de que los humanos somos necesariamente superiores a todos los demás, y eso es una tesis empírica que todavía no ha sido probada.

### 1.2.3. Matemática

- **Objeción:** *“El teorema de Gödel demuestra que ningún sistema formal o máquina puede generar todas las verdades matemáticas. Además, la inteligencia humana no tiene límite, por lo tanto las máquinas no pueden ser inteligentes.”*
- **Respuesta:** Turing responde los dos planteos de la objeción. Primero, explica que en realidad no sabemos si la inteligencia es ilimitada o no, y la forma de determinar si una máquina está limitada de la misma forma que la inteligencia humana sería a través del Test de Turing. Por otro lado, si una máquina fuera capaz de inventar sus propios métodos de prueba, o sus propias reglas, entonces sí sería capaz de generar todas las verdades matemáticas. Obviamente, puede cometer errores, pero también los cometemos los humanos.

### 1.2.4. Conciencia

- **Objeción:** *“Las máquinas no pueden tener emociones, sentir placer, depresión, etc. Los humanos inteligentes sí las tienen, por lo tanto las máquinas no pueden ser inteligentes.”*
- **Respuesta:** Turing responde que decir esto es invalidar el test de Turing. Si una máquina puede hablar inteligentemente sobre, por ejemplo, una sonata que compuso, entonces debe ser inteligente. Al fin y al cabo, el hablar de nuestros sentimientos es la única evidencia que tenemos que otros humanos los tienen, por lo que la objeción entonces también objetaría la conciencia de otros humanos.

### 1.2.5. Incapacidades

- **Objeción:** *“No podemos decir que una máquina es inteligente hasta que logre hacer X.”*
- **Respuesta:** Turing dice que muchas habilidades no podían ser realizadas en el pasado y ahora sí, que fue una cuestión de desarrollo humano hasta que las máquinas logren hacerlo, por lo que es cuestión de tiempo para algunas cosas. Además argumenta que muchas cosas no deberían de porque ser necesarias de realizar por una máquina, por ejemplo, ¿Por qué deberíamos esperar que una máquina tenga la capacidad de cometer errores? ¿Por qué castigar a una máquina porque es mejor que un humano en matemáticas?

### 1.2.6. Lady Lovelace

- **Objeción:** *“Los humanos somos creativos y originales, nunca una máquina podría serlo.”*
- **Respuesta:** Turing argumenta, pobremente, que muchas veces las máquinas sorprenden a sus creadores con su comportamiento, y se podría tomar como pensamiento original. En la actualidad se ha dicho que las máquinas suelen utilizar algún tipo de cálculo probabilístico, permitiendo cierta originalidad, a través de resultados inesperados.

### 1.2.7. Continuidad del Sistema Nervioso

- **Objeción:** *“La disanalogía entre las máquinas de Turing y el sistema nervioso es tan grande, que ninguna máquina jamás podrá imitarlo y por lo tanto ser inteligente.”*
- **Respuesta:** Turing confirma que esto es cierto, pero que si alguna máquina pasa el test de Turing, esto no sería relevante.

### 1.2.8. Informalidad del comportamiento

- **Objeción:** *“No conocemos las reglas y leyes que gobiernan el comportamiento humano, por lo tanto no podríamos crear una máquina inteligente sin dichas reglas.”*
- **Respuesta:** Turing responde haciendo notar la diferencia entre conocer dichas reglas que realmente existen.

### 1.2.9. Percepción extrasensorial

- **Objeción:** “Los humanos inteligentes poseemos ESP (Percepción extrasensorial), cosa que las máquinas claramente no.”
- **Respuesta:** La única manera de probar la ESP sería a través de una prueba empírica, y tal prueba también podría ser realizada por una IA.

## 2. Agentes

### 2.1. ¿Qué es un agente?

- Un *agente* es cualquier cosa que percibe un ambiente mediante sensores y actúa en ese ambiente a través de efectores.
- Decimos que el comportamiento de un agente está descrito por la *función del agente* que mapea cada secuencia de percepciones a una acción.
- La *medida de performance* establece el criterio del éxito del comportamiento de un agente.
- Un *agente racional* es aquel que hace lo correcto según lo que conoce. La acción correcta sería aquella que hiciera que el agente tenga mayor éxito. Una definición más precisa dependerá de que es racional en cada momento dado, lo que lleva a definirlo de la siguiente manera: *Por cada posible secuencia de percepciones, un agente racional debería elegir una acción de la que se espera que maximice su medida de performance, dada la evidencia provista por la secuencia de percepciones y el conocimiento que tiene el agente en el momento.*
- Un *agente omnisciente* es un agente que sabe el resultado de sus acciones y puede actuar acorde a ello, en la realidad es imposible.
- Un *agente inteligente* es aquel que es capaz de realizar acciones autónomas flexibles, es decir:
  - Reactividad: son capaces de percibir su ambiente y responder a cambios que ocurren en él.
  - Pro-actividad: son capaces de exhibir funcionamiento orientado a un objeto, tomando la iniciativa.
  - Habilidad social: son capaces de interactuar con otros agentes con el fin de satisfacer sus objetivos.

#### 2.1.1. Tipos de agentes

- Agente reflexo simple
- Agente reflexo basado en modelos
- Agente basado en metas
- Agente basado en utilidades
- Agentes que aprenden

### 2.2. Ambiente

*Ambiente de tarea*, se podría ver como el problema para el cual el agente es la solución.

<b>P</b> erformance	Medición del rendimiento, a que debería aspirar nuestro agente.
<b>E</b> nviroment	Cuál será el ambiente en el que se realizara la tarea, que condiciones presenta. Puede definirnros que tan complejo será nuestro problema.
<b>A</b> ctuators	Con que interactúa el agente con el ambiente.
<b>S</b> ensors	Con que analiza el agente el ambiente.

Figura 2: Especificación del ambiente de tareas (PEAS)

#### Propiedades del ambiente

- Totalmente observable vs. Parcialmente observable
- Determinístico vs. Estocástico
- Episódico vs. Secuencial
- Estático vs. Dinámico
- Discreto vs. Continuo
- Un agente vs. Múltiples agentes

### 3. RRS

#### 3.1. ¿Qué es la representación de conocimiento?

Es el estudio concerniente con el uso formal de símbolos para representar una colección de proposiciones creídas por un agente.

#### 3.2. ¿Qué es una ontología?

Una ontología representa formalmente el conocimiento como un conjunto de conceptos en un dominio y las relaciones entre estos conceptos. Así determinamos, qué es lo que el agente puede percibir y por lo tanto, qué parte del mundo va a representar y razonar a partir de ellos para lograr sus objetivos.

#### 3.3. ¿Qué es el razonamiento?

Manipulación de símbolos que codifican proposiciones para producir representaciones de nuevas proposiciones.

#### 3.4. Sistema de Representación y Razonamiento (RRS)

- **Lenguaje Formal:** especifica las sentencias legales.
- **Semántica:** especifica el significado de los símbolos.
- **Procedimiento de Prueba o Teoría de Razonamiento:** una especificación no determinista de cómo una respuesta puede ser producida.

#### 3.5. Implementación de RRS

- **Analizador del Lenguaje:** distingue las sentencias legales del lenguaje formal y produce estructuras de almacenamiento.
- **Procedimiento de razonamiento:** implementación de una teoría de razonamiento y una estrategia de búsqueda.

*Nota: ¡La semántica no está reflejada en la implementación!*

### 3.6. Uso de RRS

1. Comenzar con un dominio de trabajo.
2. Distinguir aquellos elementos relevantes (Ontología).
3. Elegir símbolos para nombrar objetos y relaciones.
4. Darle al sistema conocimiento del dominio.
5. Realizar preguntas al sistema.

### 3.7. Agentes basados en conocimiento

La inteligencia de los humanos puede obtenerse por procesos de razonamiento que operan sobre representación del conocimiento. El principal componente es su **Base de conocimiento (KB)**.

#### 3.7.1. Características

- Puede aceptar nuevas tareas en forma de metas descritas explícitamente.
- Pueden adquirir competencias rápidamente o aprender nuevo conocimiento del ambiente.
- Pueden adaptarse a cambios en el ambiente, actualizando el conocimiento relevante.

## 4. Búsqueda

### 4.1. Búsqueda ciega

### 4.2. Búsqueda informada (heurística)

### 4.3. Búsqueda adversaria

### 4.4. CSP - Constraint Satisfaction Problem

## 5. Reiter - Lógica Default

### 5.1. Introducción

La Lógica Default consta de dos tipos de conocimientos, las formulas lógicas normales (llamadas de *axiomas* o *hechos*) y reglas llamadas *defaults*. Una *teoría default* consiste de un conjunto de hechos que representan conocimiento certero pero usualmente incompleto sobre el mundo; y un conjunto de defaults que establecen posibles conclusiones pero no necesariamente verdaderas. Las lógicas default son *no monotónicas*, es decir, agregar información a nuestro conocimiento a veces provoca que dejemos de inferir lo que ya podíamos inferir.

Algunas teorías default pueden no tener extensiones, las cuales son inútiles porque no proveen ninguna información útil. Por lo tanto, existen teorías default con ciertas restricciones que nos aseguran una utilidad. Por ejemplo, las *teorías default normales* tienen la característica que siempre tienen extensiones.

#### 5.1.1. Ejemplo - Trabajo en colectivo

Suponga que me preguntan como pretendo ir al trabajo. Mi respuesta es “en colectivos” porque suelo ir al trabajo en colectivos. La regla default para representar esto seria:

$$\frac{irTrabajo : usarColectivo}{usarColectivo}$$

Entonces salgo de casa, voy hasta la parada y leo una noticia que dice “No hay colectivos hoy por el paro”. Ahora no puedo ir al trabajo en colectivo, por lo que mi regla default ya no es aplicable (ya no puedo asumir que usare el colectivo). Debo revisar mi conclusión anterior, por lo que mi comportamiento es no monotónico (agrega información a mi conocimiento y cambiaron mis conclusiones o lo que podía inferir).

Tratar de imitar este comportamiento con la logica clasica no es muy apropiado. Claro, podria usar una regla de la forma:

$$irTrabajo \wedge \neg paro \rightarrow usarColectivo$$

Pero esta regla seguiria siendo insuficiente, porque si por otra razon no puedo usar el colectivo entonces deberia agregar a la regla que no sea verdad dicha condición, y por lo tanto deberiamos listar *todas* las posibles razones por las que no podría o no usaría el colectivo.

### 5.1.2. Formas de uso

Una forma de usar reglas defaults es en la forma de *razonamiento prototipico*, donde la mayoría de las intancias de un concepto cumplen con cierta propiedad. Por ejemplo, “Típicamente, un niño tiene sus padre vivos”, o “En general, si es el cumpleaños de alguien, sus amigos le dan regalos”.

Otra forma es en el *razonamiento sin riesgos*. Se presenta en situaciones donde llegamos a una conclusión incluso si no es la mas probable, por el simple hecho que otra decisión puede llegar a un desastre. El ejemplo clasico de esto es el conocido principio de justicia: “En la ausencia de evidencia que demuestren su culpabilidad, el acusado es inocente”.

El *razonamiento de la mejor conjetura* es la forma en la que concluimos algo por el simple hecho que resulta mas conveniente, aunque no tengamos evidencia que realmente sea asi.

### 5.1.3. Ejemplo - Ley Alemana

Según la ley Alemana, un extranjero suele ser exiliado si ha cometido algún crimen. Una excepción a esta regla es con refugiados politicos. Esta información es representada por el siguiente default:

$$\frac{criminal(X) \wedge extranjero(X) : exiliar(X)}{exiliar(X)}$$

en combinación con el siguiente hecho:

$$refugiadoPolitico(X) \rightarrow \neg exiliar(X)$$

## 5.2. Sintaxis

Una *teoría default*  $T$  es un par  $(W, D)$  que consiste de un conjunto  $W$  de formulas de predicado logicas (llamadas *hechos* o *axiomas* de  $T$ ) y un conjunto contable  $D$  de reglas default. Un *default*  $\delta$  tiene la forma:

$$\frac{\varphi : \psi_1, \dots, \psi_n}{\chi}$$

donde  $\varphi, \psi_1, \dots, \psi_n, \chi$  son formulas de predicado logicas **cerradas**, y  $n > 0$ . Los elementos son llamados de la siguiente forma:

- $\varphi$  es el *prerequisito* (A veces denotado  $pre(\delta)$ )
- $\psi_1, \dots, \psi_n$  son las *justificaciones* (A veces denotado  $just(\delta)$ )
- $\chi$  es el consecuente de  $\delta$  (A veces denotado  $cons(\delta)$ )

Notese que:

$$\frac{pajaro(X) : vuela(X)}{vuela(X)}$$

en realidad **no es una regla default**. Por definición, las formulas tienen que ser cerradas (no contener variables libres). Es en realidad un *default abierto* o *esquema default*, y representa un *conjunto de defaults* que incluso podría ser infinito.

### 5.3. Semántica

Una regla default puede ser interpretada como:

*Si  $\varphi$  es conocida de momento, y si todas las  $\psi_i$  son consistentes con la base de conocimiento actual, entonces concluimos  $\chi$ . La base de conocimiento actual  $E$  es obtenida de los hechos y los consecuentes de los default que hayan sido aplicados previamente.*

## 6. Planning