

Lógica e Inteligencia Artificial

Inteligencia artificial simbólica. Agentes Lógicos

Prof. Dra. Claudia Pons

LIFIA - Universidad Nacional de La Plata
CONICET - Consejo Nacional de investigaciones Científicas y Técnicas

Clase 10

Inteligencia artificial simbólica. Agentes basados en conocimiento

Temas:

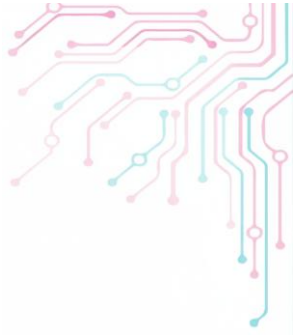
- Definición de un agente inteligente usando lógica simbólica.
- Programación lógica
- Definición de lenguajes de programación lógica

Bibliografía:

Inteligencia Artificial. Un enfoque moderno. Russell y Norvig. Capítulo 7.



¿Qué son los Agentes?



Un **agente** es un sistema que:

1. **Percibe** su entorno (datos de entrada).
2. **Decide** acciones en base a objetivos y reglas.
3. **Actúa** en el entorno.
4. **Evalúa** el resultado para ajustar su estrategia (**Aprende**).



¿Qué son los Agentes? *Un ejemplo*

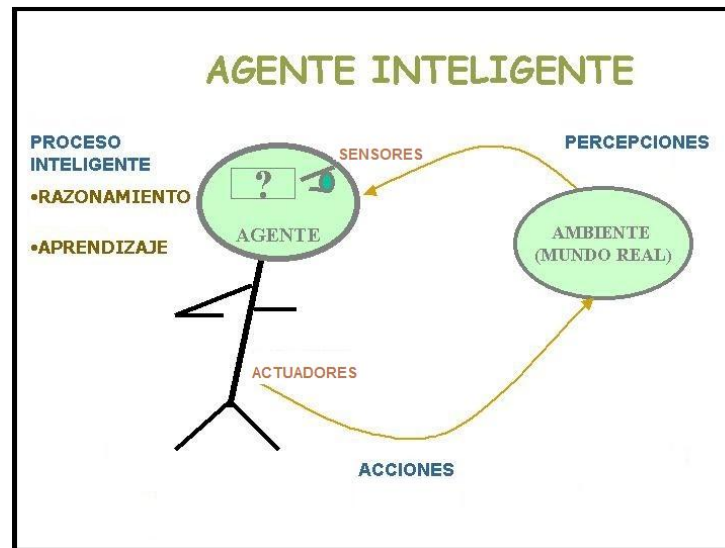


Ejemplo Cuando combinamos un **LLM** con herramientas externas y memoria, obtenemos **agentes cognitivos** capaces de:

- Llamar a APIs (por ejemplo, consultar una base de datos clínica).
- Usar razonamiento iterativo (planificar pasos antes de responder).
- Recordar interacciones previas (memoria a corto/largo plazo).
- Tomar decisiones autónomas para cumplir una meta.

¿Qué son los Agentes?

Un **agente** es una entidad que puede **percibir** su entorno mediante **sensores** y **actuar** sobre dicho entorno a través de **actuadores**



Agente humano: ojos, oídos, y otros órganos como sensores; manos, pies, boca y otras partes del cuerpo como actuadores.

Agente robótico: cámaras, infrarrojo, como sensores ; motores como actuadores.

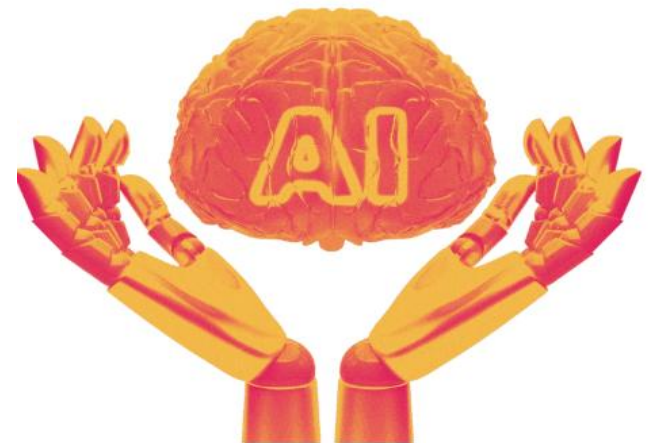
¿Dónde está el “cerebro” del Agente?

La función del agente mapea historias de percepciones a acciones:

$$[f: \mathcal{P}^* \rightarrow \mathcal{A}]$$

El programa del agente corre sobre una arquitectura física para ejecutar f

Agent = architecture + program



¿Qué es Inteligencia Artificial?

¿Son IA todos los sistemas computacionales que aprenden y realizan tareas con optimalidad y consistencia?

Es IA cuando la máquina no sigue un algoritmo (una receta) para resolver un problema, sino que encuentra su propia forma de resolverlo.

Dos enfoques diferentes

IA Simbólica.

EL PROGRAMADOR CONOCE LAS REGLAS

Ej. Sistemas expertos basados en inferencia lógica (razonan por deducción), Agentes basados en conocimiento.

IA no-Simbólica.

EL PROGRAMADOR DESCONOCE LAS REGLAS

Pero tiene DATOS

Machine learning (razonan por inducción)

Hoy vamos a conversar sobre la IA Simbólica

El programador conoce algunas REGLAS y el sistema logrará aprender otras:

Agentes basados en conocimiento

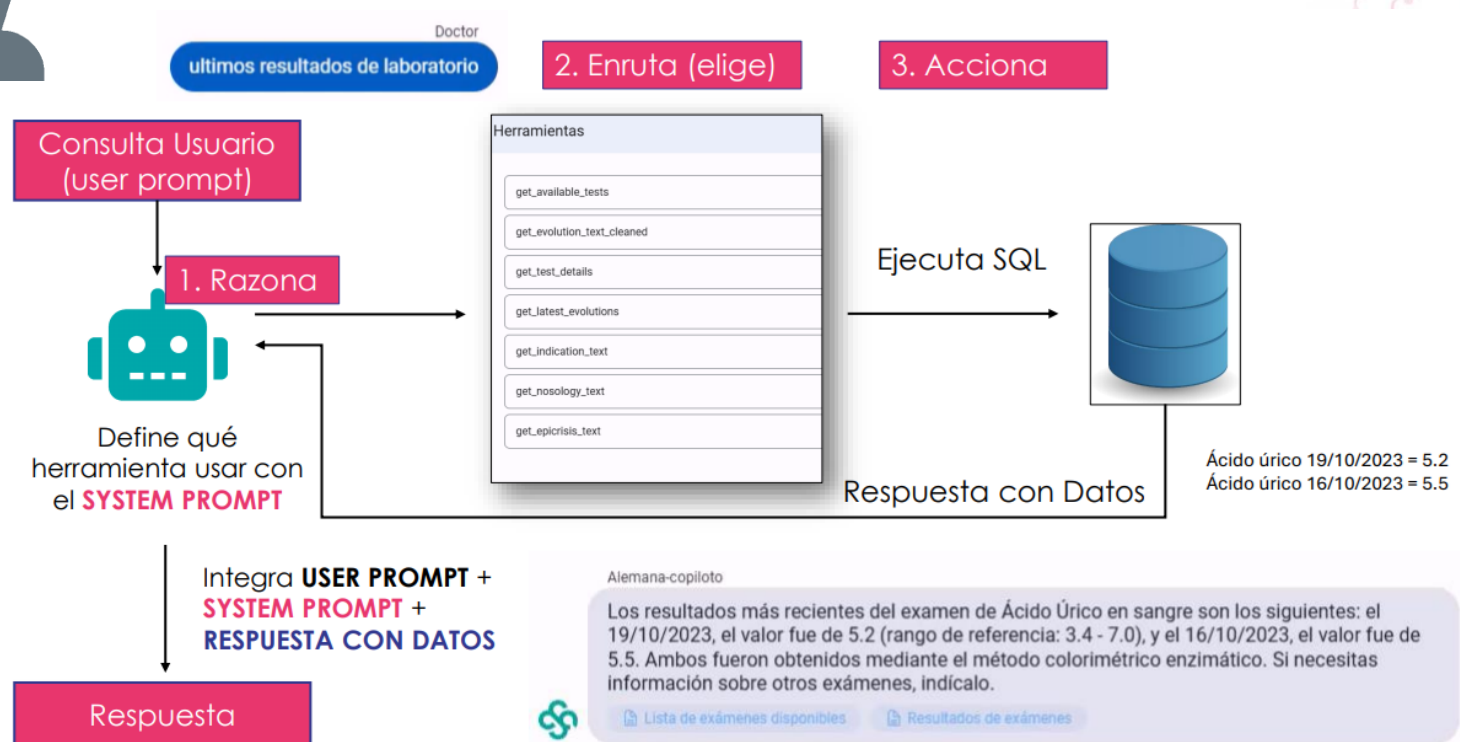
Agentes basados en conocimiento

Diseñaremos agentes que puedan:

- ✓ Construir representaciones del mundo,
- ✓ Utilizar un proceso de inferencia para derivar nuevas representaciones del mundo,
- ✓ Y emplear estas para decidir que hacer.

El conocimiento y el razonamiento juegan un papel importante cuando se trata con entornos parcialmente observables. Un agente basado en conocimiento puede combinar el conocimiento general con las percepciones reales para inferir aspectos ocultos del estado del mundo, antes de seleccionar cualquier acción. Por ejemplo, un médico diagnostica a un paciente (es decir, infiere una enfermedad que no es directamente observable) antes de seleccionar un tratamiento. Parte del conocimiento que utiliza el médico está en forma de reglas que ha aprendido de los libros de texto y sus profesores, y parte en forma de patrones de asociación que el médico no es capaz de describir explícitamente. Si este conocimiento está en la cabeza del médico, es su conocimiento.

Ejemplo de agente



Agentes basados en conocimiento

*Diseñamos agentes que pueden formar **representaciones de un mundo** complejo, usamos un proceso de inferencia para **derivar nuevas representaciones** sobre el mundo y usamos estas nuevas representaciones para deducir **qué hacer**.*

La lógica es el instrumento principal para la representación del conocimiento

Los humanos saben cosas; y lo que saben les ayuda a hacer cosas. En la IA, los agentes basados en el conocimiento utilizan un proceso de **razonamiento** sobre una representación interna del conocimiento para decidir qué acciones tomar.

Base de conocimiento

El componente principal de un agente basado en conocimiento es su **Base de Conocimiento o Knowledge Base (BC o KB)**.

Una **KB** es un conjunto de “sentencias” escritas en algún lenguaje de representación del conocimiento y representa aserciones acerca del mundo del agente.

La **KB** inicialmente contiene conocimiento básico y el agente la enriquece a medida que interactúa con el mundo y logra aprender nuevas aserciones.

función AGENTE-BC(*percepción*) **devuelve** una *acción*

variables estáticas: *BC*, una base de conocimiento

t, un contador, inicializado a 0, que indica el tiempo

DECIR(*BC*, CONSTRUIR-SENTENCIA-DE-PERCEPCIÓN(*percepción*, *t*))

acción ← PREGUNTAR(*BC*, PEDIR-Acción(*t*))

DECIR(*BC*, CONSTRUIR-SENTENCIA-DE-ACCIÓN(*acción*, *t*))

t ← *t* + 1

devolver *acción*

Figura 7.1 Un agente basado en conocimiento genérico.

Ejemplo: el mundo de Wumpus

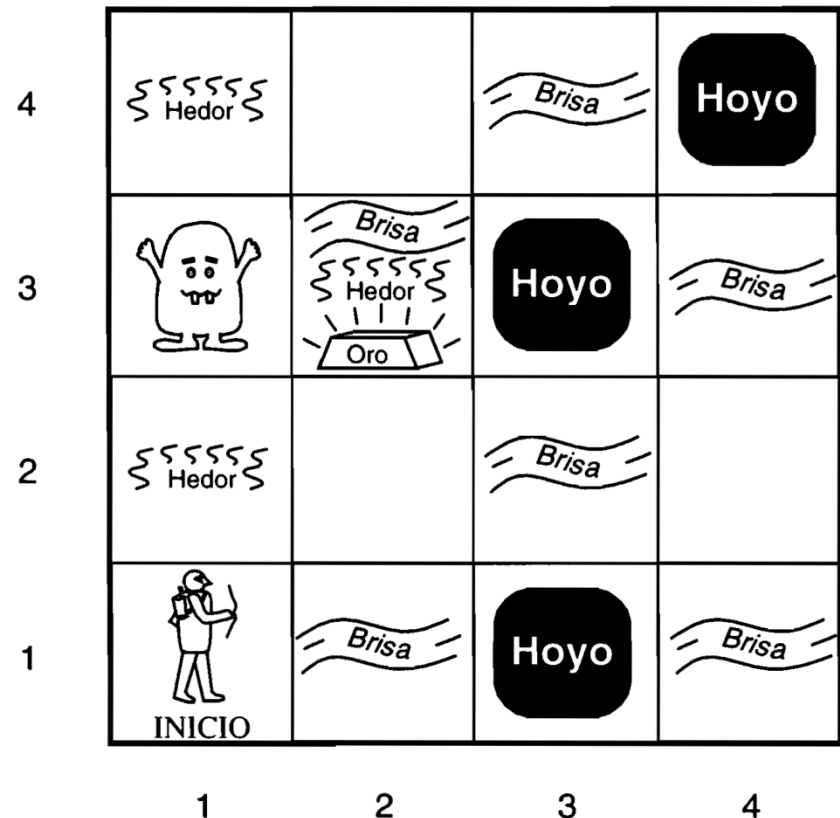
Agente = **R**endimiento, **E**ntorno, **A**ctuadores, **S**ensores

R: -1 por cada acción. -1000 por caer en un hoyo o ser comido por el wumpus, +1000 por encontrar el oro.

E: una matriz de 4x4 habitaciones donde puede haber hoyos y estar el wumpus y el oro.






A: actuadores para moverse a otra casilla, agarrar el oro, disparar una flecha.

S: tiene 5 sensores:
(hedor, brisa, resplandor, muro, grito)



Agente basado en conocimiento

El mundo real

 Hedor		Brisa	Hoyo
 Hedor	Brisa  Hedor	Hoyo	Brisa
 Hedor		Brisa	
 INICIO	Brisa	Hoyo	Brisa

El mundo como lo ve el agente

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2	2,2	3,2	4,2
1,1	2,1	3,1	4,1

OK

A

OK

OK

(a)

A = Agente
B = Brisa
G = Resplandor, Oro
OK = Casilla segura
P = Hoyo
S = Mal hedor
V = Visitada
W = Wumpus

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2	2,2 ¿P?	3,2	4,2
1,1	2,1	3,1 ¿P?	4,1

OK

V

OK

A

B

OK

(b)

Agente basado en conocimiento

A simple knowledge base:

$P_{x,y}$ is true if there is a pit in $[x,y]$.

$W_{x,y}$ is true if there is a wumpus in $[x,y]$, dead or alive.

$B_{x,y}$ is true if there is a breeze in $[x,y]$.

$S_{x,y}$ is true if there is a stench in $[x,y]$.

$L_{x,y}$ is true if the agent is in location $[x,y]$.

Base de conocimiento

La **base de conocimiento inicial** contiene las siguiente sentencias:

- No hay ningún hoyo en la casilla [1, 1].

$$R_1: \neg H_{1,1}$$

- En una casilla se siente una brisa si y sólo si hay un hoyo en una casilla vecina. Esta regla se ha de especificar para cada casilla; por ahora, tan sólo incluimos las casillas que son relevantes:

$$R_2: B_{1,1} \Leftrightarrow (H_{1,2} \vee H_{2,1})$$

$$R_3: B_{2,1} \Leftrightarrow (H_{1,1} \vee H_{2,2} \vee H_{3,1})$$

- Las sentencias anteriores son verdaderas en todos los mundos de *wumpus*. Ahora incluimos las percepciones de brisa para las dos primeras casillas visitadas en el mundo concreto en donde se encuentra el agente, llegando a la situación que se muestra en la Figura 7.3(b).

$$R_4: \neg B_{1,1}$$

$$R_5: B_{2,1}$$

Inferencia lógica

Recordemos que el objetivo de la inferencia lógica es decir si:

$BC \vdash A$, para alguna sentencia A .

Por ejemplo, queremos saber si hay un hoyo en la casilla 1,2.

Es decir:

$\{BC\} \vdash H_{1,2}$????

Inferencia lógica

$$BC \vdash \neg H_{1,2}$$

Eliminar bicondicional de R2:

$$R_6: (B_{1,1} \Rightarrow (H_{1,2} \vee H_{2,1})) \wedge ((H_{1,2} \vee H_{2,1}) \Rightarrow B_{1,1})$$

Eliminar el and de R6:

$$R_7: ((H_{1,2} \vee H_{2,1}) \Rightarrow B_{1,1})$$

Axioma L3 y MP, o contraposición:

$$R_8: (\neg B_{1,1} \Rightarrow \neg(H_{1,2} \vee H_{2,1}))$$

Modus ponens R8 con la percepción agregada como R4

$$R_9: \neg(H_{1,2} \vee H_{2,1})$$

Leyes de de morgan:

$$R_{10}: \neg H_{1,2} \wedge \neg H_{2,1}$$

Eliminar el and de R10: $\neg H_{1,2}$