

Representación del Conocimiento

Introducción

Dr. Alejandro Guerra-Hernández

Instituto de Investigaciones en Inteligencia Artificial
Universidad Veracruzana

*Campus Sur, Calle Paseo Lote II, Sección Segunda No 112,
Nuevo Xalapa, Xalapa, Ver., México 91097*

`mailto:aguerra@uv.mx`
`https://www.uv.mx/personal/aguerra/rc`

Maestría en Inteligencia Artificial 2024



Universidad Veracruzana

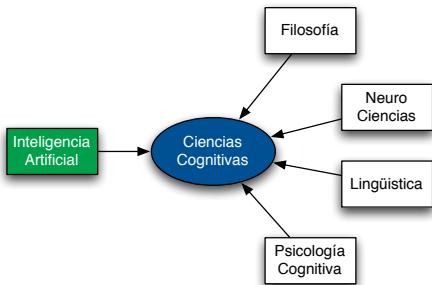
Índice

- 1 IA, agencia y representación
- 2 Conceptos básicos
- 3 Representación e IA
- 4 Sistemas basados en el conocimiento
- 5 Relevancia de una base de conocimientos
- 6 Relevancia del razonamiento
- 7 El papel de la lógica

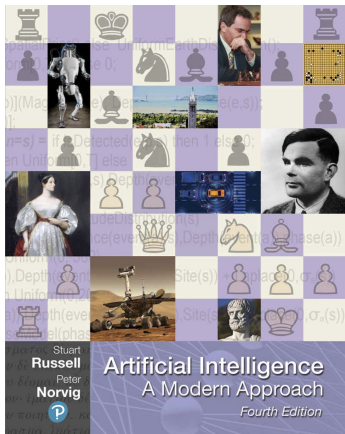


La Inteligencia Artificial en las Ciencias Cognitivas

- ▶ La IA tiene como objeto de estudio a las **entidades inteligentes** y su **comportamiento**; pero a diferencia de otras ciencias cognitivas (Varela [11]) que comparten este objeto de estudio, su meta no tiene que ver únicamente con la **comprensión** de tales entidades, sino con su **construcción**.



Un enfoque moderno a la IA

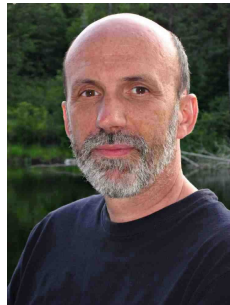
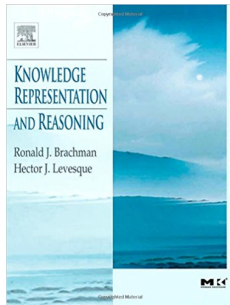


- ▶ La **construcción de agentes racionales** como idea central del enfoque adoptado en este curso, propuesto por Russell y Norvig [9], y curiosamente calificado de **moderno**.
- ▶ Usado en 1500 universidades de más de 100 países.
- ▶ Partes III y, parcialmente, la IV.
- ▶ <http://aima.cs.berkeley.edu/index.html>.



Un enfoque a la Representación del Conocimiento

► Libro de referencia:



Ceci n'est pas de l'informatique

- ▶ Aunque muchos practicantes de la IA se consideran a si mismos **científicos de la computación**, como bien señala Kayser [3] parafraseando al famoso cuadro de Magritte, este no es el caso.
- ▶ Reducir la IA a su componente computacional es **incorrecto** e **indeseable**.



Magritte y la representación

- ▶ A propósito del cuadro, Magritte nos ofrece un comentario sobre el mismo ideal para iniciar este curso:

« La famosa pipa. ¡Como la gente me reprochó por ello! Y sin embargo, ¿Podría usted **rellenar** mi pipa? No, sólo es una **representación** ¿No es así? ¡Así que si hubiera escrito en mi cuadro “Esto es una pipa”, habría estado **mintiendo!** »

Torkzyner [10]



Universidad Veracruzana

Cuestiones sobre la representación

- ▶ Algunas de las cuestiones planteadas por el cuadro de Magritte y su relación con el tema central de este curso, guiados por el texto sobre representación del conocimiento de Brachman y Levesque [1]:
 - ▶ ¿Qué queremos decir por **conocimiento**, **representación**, **razonamiento** y **acción**?
 - ▶ ¿Qué es la **verdad**?
 - ▶ ¿Por qué creemos que estos conceptos son **útiles** para la construcción de sistemas inteligentes.



Representación del conocimiento y razonamiento

- ▶ La **Representación del Conocimiento y el Razonamiento** es el área de la IA que estudia cómo el conocimiento puede ser representado y procesado **simbólicamente** mediante programas que razonan automáticamente.
- ▶ Simbólicamente, **HSSF** [7].
- ▶ ¿Cómo un agente usa lo que sabe para decidir qué hacer? –Estudiar el pensamiento como un proceso **computacional** [1].



Supuestos y preguntas

- ▶ Se asume que nuestro comportamiento inteligente se debe a que **sabemos** mucho, de muchas cosas, y que somos capaces de aplicar ese conocimiento apropiadamente para **adaptarnos** a nuestro medio ambiente y **satisfacer** nuestras metas.
- ▶ Nos preguntaremos:
 - ▶ ¿Qué necesita **saber** un agente –humano, animal, electrónico, ó mecánico, para **comportarse** inteligentemente?
 - ▶ ¿Qué clase de mecanismos computacionales le permiten **acceder** a ese conocimiento cuando así lo requiere.



Conocimiento

- ▶ Vergonzosamente **difícil de definir**, al igual que: inteligencia, artificial, autonomía, etc. Ver Nagel [5] y Pritchard [8].
- ▶ La cuestión sobre qué es el **conocimiento** es anterior a la IA y a la computación. Se ha abordado desde la antigua Grecia y sigue sin resolverse plenamente.
- ▶ Su definición como **creencias justificadas** ha sido confrontada desde la misma Grecia antigua, p. ej., los escépticos.
- ▶ No entraremos aquí en estos detalles filosóficos, pero si intentaremos ofrecer una versión **informal** de lo que se supone es el conocimiento.



Proposición

- ▶ Entre otras cosas, el conocimiento es una **relación** entre un agente y una **proposición** –Una idea expresada como un enunciado declarativo.
- ▶ Observen que es común decir “Ana sabe que...” donde a los puntos suspensivos sigue una proposición.
- ▶ **Ejemplo:** “Ana sabe que la clase de representación de conocimiento es martes y jueves”.



Valores de verdad

- ▶ Una parte del misterio acerca del conocimiento se debe a la naturaleza de las proposiciones.
- ▶ Por ahora, lo que nos interesa de las proposiciones es que son entidades abstractas sujetas a **valores de verdad**: Pueden ser verdaderas o falsas, correctas o incorrectas.
- ▶ Decir que “Ana sabe P ” equivale a decir que “Ana sabe que P es verdad”.
- ▶ Observen que esto refleja un **juicio** donde el agente se da cuenta de que su **medio ambiente** es de cierta forma (donde P es el caso) y no de otras (donde P no es el caso).

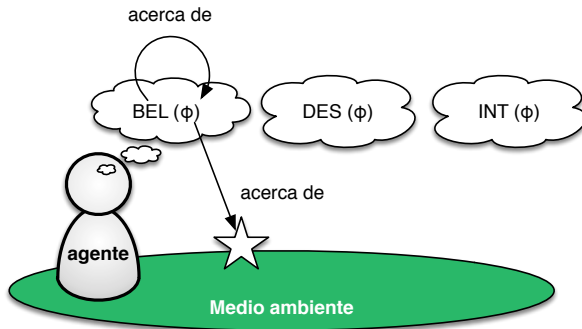


Actitudes proposicionales

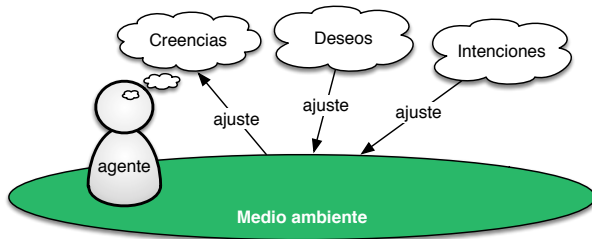
- ▶ Algo similar sucede con frases como “Ana **desea** que la clase de representación de conocimiento sea los martes y jueves”.
- ▶ Se trata de la misma proposición del ejemplo anterior, pero su **relación** con el agente es diferente.
- ▶ Los verbos como saber, creer, desear, intentar, etc., denotan lo que se conoce como **actitudes proposicionales**.
- ▶ Independientemente de la actitud, lo que importa con respecto a la proposición es su **valor de verdad**. Si “Ana desea que P ”, entonces Ana desea que el medio ambiente sea tal que P sea el caso.



Estructura de las actitudes proposicionales



Ajuste de las actitudes proposicionales



Creencias

- ▶ Las actitudes proposicionales **creer** y saber guardan cierta relación.
- ▶ Usamos la primera cuando deseamos expresar que el juicio del agente no es necesariamente **preciso**, o que no se sostiene por las razones adecuadas.
- ▶ A veces también se usa para expresar que el agente no está completamente **convencido** de que la proposición sea el caso.
- ▶ Lo importante aquí es que ambas actitudes comparten una idea básica sobre el conocimiento: Expresan que el agente **asume** que el mundo es de cierta forma y no de otra.



Representación

- ▶ El concepto de representación es tan escurridizo como el de conocimiento.
- ▶ En líneas generales, una representación es una **relación** entre dos dominios, donde se pretende que el primero **denote** o tome el lugar del segundo.
- ▶ Normalmente el primer dominio es más **concreto**, inmediato o accesible que el segundo.
- ▶ **Ejemplo:** Señales de tránsito.



Símbolo

- ▶ De especial interés para nosotros es la representación basada en **símbolos formales**, esto es, un carácter o secuencia de caracteres tomada de algún **alfabeto**.
- ▶ **Ejemplo:** El dígito 6 representa al número seis, lo mismo que la secuencia de letras VI y, como no, seis y six.
- ▶ Como en toda representación, se asume que es más fácil contender con lo símbolos, que con lo que representan.
- ▶ Observen que el caso de las **proposiciones** es una representación simbólica.



Representación del conocimiento

- ▶ La representación del conocimiento es entonces el área de estudio concerniente al **uso** de los símbolos formales para representar una colección de proposiciones creídas por un agente putativo.
- ▶ Observen que no estamos afirmando que los símbolos representarán **todas** las proposiciones creídas por el agente, podría haber un número infinito de proposiciones creídas y solo un número finito de ellas representadas.
- ▶ El **razonamiento** tiende un puente entre lo representado y lo creído.



Razonamiento

- ▶ Por razonamiento entendemos la **manipulación formal** de los símbolos que representan colecciones de proposiciones creídas, para producir representaciones de nuevas proposiciones.
- ▶ Explotar el hecho de que el símbolo sea más simple que lo que representa: Las representaciones simbólicas deben ser lo suficientemente concretas para **manipularlas** (cambiarlas de sitio, eliminarlas, copiarlas, concatenarlas, etc.) de tal manera, que sea posible construir **nuevas representaciones**.



Razonamiento como un cálculo

- ▶ Brachman y Levesque [1] nos recuerdan una analogía con la aritmética brillante: Podemos pensar en la suma como una **manipulación formal**, de los símbolos 103 y 8, podemos construir 111 para denotar la suma de ellos.
- ▶ La manipulación (suma por columnas, acarreo, etc.) define el **concepto** suma.
- ▶ El **razonamiento** es similar: De proposiciones como “Todos los días martes hay clase de representación de conocimiento” y “El día de hoy es martes” podemos construir el enunciado “Hay clase de representación de conocimiento”.
- ▶ ¿Qué manipulación define el razonamiento?



Inferencia lógica

- ▶ A esta forma de razonamiento le solemos llamar **inferencia lógica**, porque la proposición final es una **consecuencia lógica** de las primeras.
- ▶ De forma que, como fue por primera vez expuesto por **Leibniz**, el razonamiento es una forma de **cálculo** que en lugar de operar sobre números, como la aritmética, lo hace sobre proposiciones.



Relevancia de la representación en IA

- ▶ ¿Porqué es **relevante** el conocimiento para la IA?
- ▶ Una primer respuesta, como ya se sugirió, sería que pareciera que los humanos somos capaces de exhibir un **comportamiento inteligente**, gracias a lo que **sabemos**.
- ▶ Otra respuesta es que a veces resulta útil **describir** el comportamiento de sistemas que son suficientemente complejos, usando un vocabulario que incluye términos como creencias, deseos, metas, intenciones, etc., i.e., lo que Dennett [2] denomina la **postura intencional**.



El ajedrez desde la postura intencional

- ▶ **Ejemplo:** Al jugar ajedrez contra un programa podemos decir que nuestro contrincante movió su alfil porque creía que su reina era vulnerable, pero aún desea amenazar mi caballo.
- ▶ En términos de cómo el programa está construido, tal explicación se reduciría a que usando el procedimiento de evaluación P con la función de evaluación estática Q , se obtuvo un valor de +9 en una búsqueda alfa-beta minimax de profundidad 4.
- ▶ La segunda descripción, aunque precisa, está expresada en un **nivel de detalle** incorrecto y no nos ayuda a determinar que movimiento haré yo en consecuencia.
- ▶ La primera, que resulta más útil para este **fin**, se basa en la postura intencional.



Pertinencia de la postura intencional

- ▶ Lo anterior no significa que la postura intencional sea siempre la adecuada y la clave está en la **complejidad** del sistema que se quiere representar: Asumir la postura intencional para describir el apagador de luz resulta innecesario, inadecuado y caricaturesco.
- ▶ McCarthy [4] estableció que en **computación** la postura intencional es útil para:
 - ▶ Contender con nuestra limitación para adquirir conocimiento.
 - ▶ Predecir y establecer generalizaciones en términos de la estructura de un programa.

Siempre y cuando las actitudes proposicionales **reflejen el uso** que se les suele dar.



Limites de la postura intencional

- ▶ ¿Es esto todo lo que entendemos por representación de conocimiento? Es solo **hablar acerca** del conocimiento.
- ▶ La postura intencional no dice nada acerca de que está y que no está representado simbólicamente.
- ▶ **Ejemplo:** El programa de ajedrez representa simbólicamente la posición de las piezas en el tablero, pero su deseo de eliminar mi caballo lo más pronto posible puede no estarlo. Puede tratarse de una propiedad emergente de diferentes componentes del programa: sus funciones de evaluación, su librería de movimientos, etc. Y aún así, resulta útil para mi saber que el programa desea despachar a mi caballo.
- ▶ ¿Cual es el rol entonces de la representación de conocimiento?



Hipótesis de la Representación del Conocimiento

- ▶ Queremos construir sistemas que contengan representaciones simbólicas con dos **propiedades** importantes:
 - ▶ Que desde fuera, nosotros podamos entender que representan sus proposiciones;
 - ▶ y que desde adentro, el sistema haya sido diseñado para comportarse como lo hace, debido a estas representaciones.



Programa 1

- ▶ Usaremos dos programas para explicar en qué consiste un SBC. Aquí está el primero:

```
1 adivina_color(nieve) :- !, write("Es blanca.").
2 adivina_color(pasto) :- !, write("Es verde.").
3 adivina_color(cielo) :- !, write("Es amarillo.").
4 adivina_color(_) :- write("Me has ganado.").
```



Programa 2

► Aquí está el segundo:

```
1  adivina_color(X) :-
2      color(X,Y), !,
3      write("Es de color "), write(Y), write(".").
4  adivina_color(_) :- write("Me has ganado.").
5
6  color(nieve,blanco).
7  color(cielo,amarillo).
8  color(vegetacion,verde).
9  color(X,Y) :-
10     hecho_de(X,Z),
11     color(Z,Y).
12
13  hecho_de(pasto,vegetacion).
14  hecho_de(selva, vegetacion).
```



Conformidad con la hipótesis

- ▶ La cláusula `color(nieve,blanca)` es una estructura simbólica que **representa** la proposición “el color de la nieve es blanca”.
- ▶ Mejor aún, sabiendo Prolog, sabemos que el sistema imprimirá la frase **adecuada** debido a que el programa dará con la cláusula en el momento adecuado.
- ▶ Si eliminamos la cláusula el programa **dejará** de actuar de esa manera.
- ▶ Ahora bien, **no existe** una cláusula similar en el primer programa. Lo más cercano es la primer cláusula, pero no resulta natural interpretarla como una creencia.



Base de conocimientos

- ▶ Lo que hace que un sistema esté basado en conocimiento, **no es**:
 - ▶ El uso de un formalismo lógico (cláusulas definitivas y resolución-SLD en este caso);
 - ▶ El hecho de que el sistema sea lo suficientemente complejo como para justificar el uso de la postura intencional y sus descripciones sobre el conocimiento;
 - ▶ El hecho de lo que creemos que es verdadero.
- ▶ Lo importante es la presencia de una **base de conocimientos**: una colección de estructuras simbólicas que representan lo que el agente cree y razona, durante su operación.



Ejemplos de sistemas en la IA

- ▶ Los **sistemas expertos** son un claro ejemplo de ello.
- ▶ Pero se pueden encontrar bases de conocimiento en sistemas para el procesamiento de **lenguaje natural**, **planeación**, **diagnóstico** y **aprendizaje**.
- ▶ Otros sistemas incluyen en menor grado conocimiento, como algunos juegos y sistemas de visión de alto nivel (reconocimiento de objetos).
- ▶ Y por supuesto, hay una parte de la IA que no está basada en el conocimiento: Reconocimiento de habla de bajo nivel, visión, control motor, entre otros.



¿Porqué una base de conocimientos?

- ▶ ¿Qué ventajas tiene, si acaso tiene alguna, usar una base de conocimiento?
- ▶ Los defensores del llamado **conocimiento procedimental** defenderían al primer programa: Una especie de **compilación** de la base de conocimientos del segundo programa que distribuye lo que debe saberse para actuar en cada módulo que así lo requiere.
- ▶ De hecho, el **desempeño** del sistema será mejor en este caso, ya que buscar los hechos en la base de conocimientos para posteriormente, decidir qué hacer, solo puede hacer que el sistema sea más lento.



Propiedades deseables en un SBC I

- ▶ Agregar **nuevas tareas** dependientes del conocimiento previo. **Ej.** Agregar la tarea de enumerar todos los objetos de un color dado e incluso pintar un dibujo, usando la BC para determinar los colores a usar.
- ▶ **Extender** el comportamiento del programa agregando nuevas creencias. **Ej.** Agregar una cláusula que especifique que el color de los canarios es amarillo, para propagar esa información a cualquier rutina que la necesite.
- ▶ **Depurar** el comportamiento erróneo del sistema localizando creencias erróneas. **Ej.** Corregir la cláusula que especifica que el color del cielo es amarillo.



Propiedades deseables en un SBC II

- ▶ **Explicar** y **justificar** el comportamiento del sistema. ¿Porqué dice mi programa que el pasto es verde?.



Penetrabilidad cognitiva

- ▶ La ventaja de los SBC es que, por diseño, pueden ser informados de nuevos hechos sobre el mundo y ajustar su comportamiento en consecuencia.
- ▶ La propiedad de algunas de nuestras acciones de ser dependientes de lo que creemos, es lo que Zenon Pylyshyn llama **Penetrabilidad cognitiva**.
- ▶ **Ejemplo:** Nuestra respuesta normal al oír una alarma contra incendios sería levantarnos y salir del lugar donde nos encontramos para ir a un área segura. Sin embargo, la respuesta es diferente si sabemos que la alarma está siendo probada.
- ▶ Nuestro reflejo de retirar la mano del calor, por el contrario, no parece ser cognitivamente penetrable.



¿Porqué el razonamiento?

- ▶ La motivación detrás del razonamiento tiene que ver con el hecho de que nos gustaría que la actuación del sistema dependiera no únicamente de lo que está **explícitamente** representado, sino de lo que el sistema cree como **consecuencia de ello**.
- ▶ **Ejemplo:** En el segundo programa Prolog, no hay una cláusula que diga que el pasto es verde, y aún así queremos que el sistema sepa esto.



Base de datos vs conocimiento

- ▶ Mucho del conocimiento tiene que ver con hechos muy generales que necesitarán posteriormente aplicarse en situaciones particulares.
- ▶ Ejemplo:
 1. Paciente X alérgico al medicamento M .
 2. Todo alérgico al medicamento M también es alérgico al medicamento M' .
- ▶ ¿Es apropiado darle el medicamento M' para el paciente X ? Estos y otros hechos acerca de las alergias, son suficientes para desaconsejar tal medicación.
- ▶ Que el sistema vaya más allá de los hechos que podamos coleccionar a manera de una base de datos.



Más allá, ¿Donde?

- ▶ Existe una respuesta simple a esta cuestión, aunque no siempre sea **práctica**.
- ▶ El sistema debe creer P si, conforme a las creencias que tiene representadas, el mundo que se imagina es uno donde P es verdadera. Si las cláusulas (1) y (2) están representadas en el sistema y el mundo es tal que son verdaderas, entonces en ese mundo
 3. El paciente X es alérgico al medicamento M' .
también es una cláusula verdadera, aún cuando este hecho está representado solo **implícitamente**.



Consecuencia lógica

- ▶ Las proposiciones representadas por un conjunto de enunciados Δ tienen como consecuencia la proposición representada por el enunciado P , si la verdad de P está **implícita** en la verdad de los enunciados de Δ .
- ▶ En otras palabras, si el mundo es tal que si cada enunciado en Δ es verdadero, entonces este es el caso también para P .
- ▶ Todo lo que necesitamos para obtener una noción de consecuencia es la noción de **valor de verdad**.
- ▶ Todo lenguaje de representación de conocimiento, sin importar su sintaxis y sus procedimientos de razonamiento, debe tener una noción de consecuencia **bien definida**.



¿Qué creencias debe tener un sistema?

- ▶ Todas aquellas y solo aquellas que son **consecuencia lógica** de sus representaciones explícitas.
- ▶ La tarea de un **mecanismo de razonamiento** es entonces computar las consecuencias lógicas de una base de conocimiento.
- ▶ La razón por la cual tal respuesta no es siempre práctica tiene que ver con el **costo computacional** de calcular qué enunciados son consecuencia de la base de conocimiento que nos interesa usar.



Completitud y solidez: Ya saben quién...

- ▶ Pero como lo estableció ya saben quién: Todo procedimiento que siempre nos de una respuesta en un tiempo razonable, ocasionalmente perderá algunas consecuencias o nos dará una respuesta equivocada.
- ▶ En el primer caso, decimos que el proceso de razonamiento es lógicamente **incompleto**; en el segundo decimos que el procedimiento es **incorrecto**.
- ▶ Lo hemos visto en el uso de Prolog, ¿Ejemplos?



¿Vale la pena considerar tal razonamiento?

- ▶ **Ejemplo:** Supongamos que todo lo que sé acerca de Piolín es que es un ave. Podría tener un cierto número de hechos acerca de las aves en mi base de conocimiento, pero difícilmente estas tendrían como consecuencia que Piolín vuela (podría ser el caso que Piolín fuese una avestruz).
- ▶ Y sin embargo, es razonable creer que Piolín vuela. Estamos ante un razonamiento **no sólido** (*sound*), donde las cláusulas de la base de datos pueden ser verdaderas en el mundo, mientras que Piolín no vuela.



Inconsistencia

- ▶ **Ejemplo:** Un sistema que integra en su base de conocimiento con hechos de varias fuentes que resultan **inconsistentes**, es decir, no pueden ser verdaderos juntos.
- ▶ En tal caso, es inapropiado computar la consecuencia lógica **completa** del sistema, porque entonces todo enunciado sería creído: Puesto que no hay mundos donde la base de conocimientos sea verdadera, el consecuente se vuelve trivialmente verdadero.
- ▶ En este caso es evidente que una forma de razonamiento **incompleto** sería más conveniente, al menos hasta que se pueda resolver la contradicción.



Ser naïve

- ▶ En algunos casos la respuesta simple es la más adecuada y aunque sea un error equiparar razonamiento en una base de conocimientos con el concepto de inferencia lógica completa y sólida, tal aproximación puede ser un buen **comienzo**.



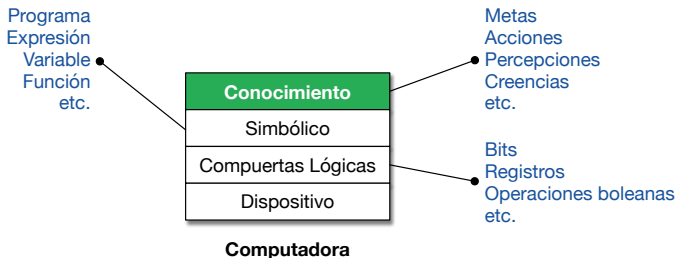
Lógica

- ▶ La razón por la cual la lógica es **relevante** para la representación de conocimiento y el razonamiento es que, al menos conforme a un punto de vista, ésta es el estudio de las relaciones de consecuencia –lenguajes, condiciones de verdad y reglas de inferencia.
- ▶ Es por tanto normal, que usemos una gran cantidad de herramientas y técnicas provenientes de la lógica simbólica formal. Particularmente, haremos uso de la **lógica de primer orden** (FOL, por sus siglas en inglés), aunque claro, esto solo es un punto de partida y consideraremos también lenguajes bien distintos en forma y significado.



El nivel del conocimiento

- Donde la lógica reditúa realmente es en lo que Newell [6] llama **the knowledge level**: Todo SBC puede entenderse en al menos dos niveles diferentes:



Punto de partida

- ▶ Las herramientas de la lógica parecen **adecuadas** para el análisis a nivel del conocimiento.
- ▶ Comenzaremos por abordar la lógica proposicional y de primer orden, dejando de lado por ahora sus aspectos computacionales.
- ▶ Pero antes, una breve introducción al concepto de **agente**.



Referencias I

- [1] RJ Brachman y HJ Levesque. *Knowledge representation and reasoning*. San Francisco, CA, USA: Morgan Kaufmann Publishers, 2004.
- [2] D Dennett. *The Intentional Stance*. Cambridge, MA, USA: The MIT Press, 1987.
- [3] D Kayser. *La représentation des connaissances*. Collection Informatique. Paris, France: Editions Hermes, 1997.
- [4] J McCarthy. *Ascribing Mental Qualities to Machines*. Inf. téc. Stanford, CA, USA: Computer Science Department, Stanford University, 1979.
- [5] J Nagel. *Knowledge: A Very Short Introduction*. Very short introductions. New York, NY, USA: Oxford University Press, 2014.
- [6] A Newell. "The Knowledge Level". En: *AI Magazine* 2 (1981), págs. 1-20.
- [7] A Newell y HA Simon. "Computer Science As Empirical Inquiry: Symbols and Search". En: *Commun. ACM* 19.3 (1976), págs. 113-126. ISSN: 0001-0782. URL: <http://doi.acm.org/10.1145/360018.360022>.
- [8] D Pritchard. *What is this thing called knowledge?* Fourth. What is this thing called? Oxon, UK: Routledge, 2018.
- [9] SJ Russell y P Norvig. *Artificial Intelligence: A Modern Approach*. Third. Prentice Hall Series in Artificial Intelligence. USA: Prentice Hall, 2009.



Referencias II

- [10] H Torkzyner. *Magritte: Ideas and Images*. New York, NY, USA: Harry N. Abrams, Inc., 1979.
- [11] F Varela. *Invitation aux Science Cognitives*. Paris, France.: Editions du Seuil, 1989.

