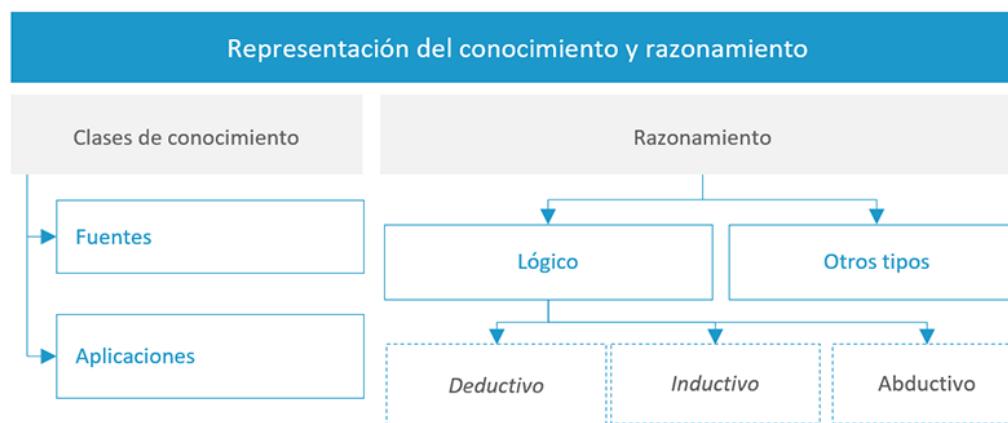


Tema 2: Representación del conocimiento y Razonamiento

Índice de la clase

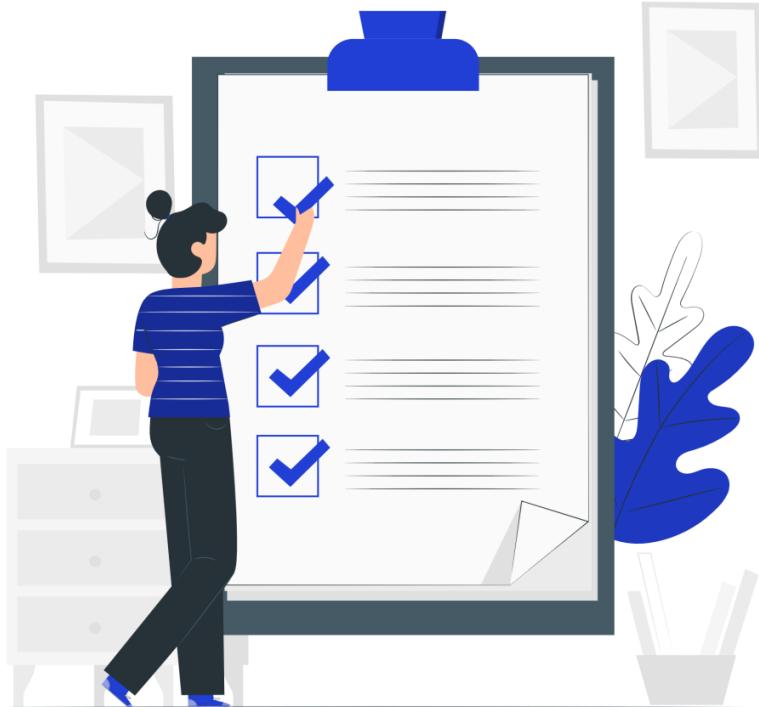
1. Razonamiento y tipos de razonamiento
2. Representación y tipos de representación
3. Resumen ideas principales
4. Tarea para casa

Tema 2



Encuesta

Input para el profesor, ¿han podido preparar la clase?



Fuente: <https://teaminsights.io/es/blog/noticias/6-errores-frecuentes-en-las-preguntas-de-encuesta-y-como-superarlos/>

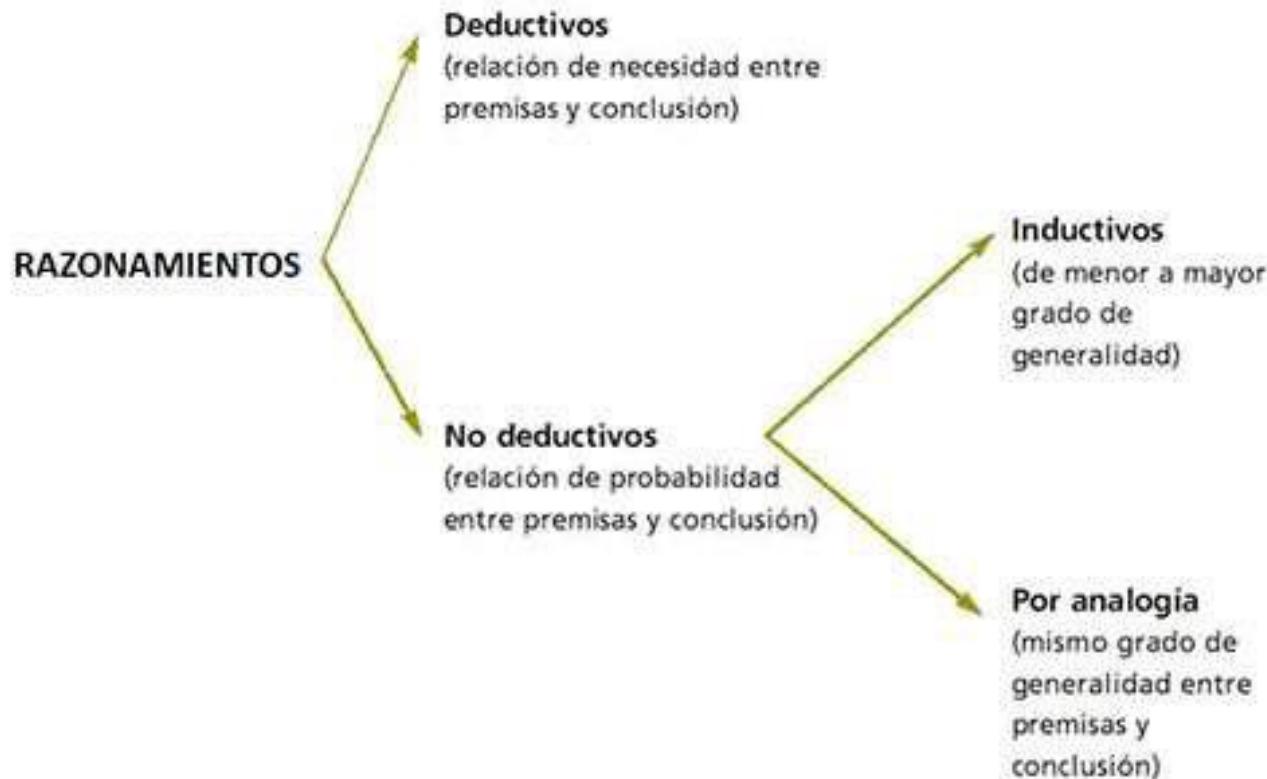


Razonamiento y Tipos de Razonamiento

Tipos de razonamiento

Múltiples clasificaciones en la literatura

lógicos (formalismos) vs no-lógicos (sin formalismos)



Tipos de razonamiento

Elementos que conforman un razonamiento



Contenido: La lógica representa un hecho del mundo (o una creencia). A este significado la podremos asignar el **valor de verdad** que se cumpla en la situación representada (valor de Verdadero o Falso).

Forma: determina si un **razonamiento es válido o inválido**:

Ejemplo: "Si A implica B, y no se da B, entonces seguro que no se da A"
(Modus Tollens, o Tollendo Tollens)

Los razonamientos son válidos o inválidos **independientemente del significado de sus elementos**.

Problema 1 para la sesión

¿Pueden los modelos de IA basada en grandes modelos de lenguaje como ChatGPT (y otros) realizar razonamientos? ¿De qué tipo?

Un ejemplo: *Todos los cuervos vuelan, y todos los pájaros vuelan; en consecuencia, ¿podemos decir que el cuervo es un pájaro?*

<https://zoom.us/wb/doc/fqfVNmlGQayVhQye7D2L0g/p/265197336920064>

Problema 1 para la sesión (cont)



Todos los cuervos vuelan, y todos los pájaros vuelan; en consecuencia, ¿podemos decir que el cuervo es un pájaro?

- Se nos está preguntando si este razonamiento es válido. Esto depende de qué tipo razonamiento sea: cada tipo de razonamiento tiene su propio criterio de validación.

RESPUESTA 1

ChatGPT (2/2024): Por lo tanto, podemos concluir que los cuervos son una subclase de los pájaros, pero no todos los pájaros son cuervos. En resumen, mientras que todos los cuervos son pájaros, no todos los pájaros son cuervos.

Dice que no es válido, pero no por las razones correctas. En este caso premisas y conclusión son verdaderas, pero la forma de razonamiento NO es válida. ¿Podemos poner un caso formalmente idéntico en el que no se cumpla esa condición de validez?

Problema 1 para la sesión (cont)



RESPUESTA 2

BING (11/2024): Interesante pregunta. Apliquemos un poco de lógica.

- **Premisa 1:** Todos los cuervos vuelan.
- **Premisa 2:** Todos los pájaros vuelan.

La conclusión que se propone es que, si ambos vuelan, entonces los cuervos son pájaros. Pero, ¿realmente la conclusión se desprende de estas premisas?

En este caso, no necesariamente. Ambas premisas comparten una característica común (volar), pero esto no es suficiente para afirmar que los cuervos son pájaros, aunque en la realidad sí lo sean. Por lo tanto, faltaría una premisa adicional:

- **Premisa 3:** Todos los cuervos son pájaros.

Esta tercera premisa nos permite concluir que, efectivamente, los cuervos son pájaros. Sin esta información, solo sabemos que ambos vuelan, pero no tenemos suficiente para clasificar taxonómicamente a los cuervos.

Si necesitas más aclaraciones, estoy aquí para ayudarte

Problema 1 para la sesión (cont)

¿Puedes decirme qué regla de la lógica de predicados puede extraer la conclusión 3 a partir de las premisas 1 y 2 anteriores?

RESPUESTA 3



BING (11/2024): Para extraer la conclusión 3 a partir de las premisas 1 y 2, podemos utilizar la regla del silogismo hipotético en lógica de predicados. Esta regla establece que si tenemos dos implicaciones del tipo $(A \rightarrow B)$ y $(B \rightarrow C)$, podemos concluir $(A \rightarrow C)$

...



En resumen, con las premisas dadas, no podemos directamente concluir que “todo cuervo es un pájaro” sin una premisa adicional que conecte el hecho de volar con ser un pájaro. La regla del silogismo hipotético sería la adecuada si tuviéramos esa premisa adicional.

Razonamiento lógico deductivo

- ▶ **Deductivo:** Sobre la representación lógica se define el razonamiento deductivo
- ▶ A la representación con lógica matemática se le asocian mecanismos de razonamiento deductivo



Premisa(regla): Si X es un pájaro, X vuela

Premisa (hecho): Tito es un gorrión

Premisa (regla): Si X es un gorrión, X es un pájaro



Conclusión (hecho): Tito vuela

Deducción: si el valor de todas las **premisas es verdadero** y el **razonamiento es válido**, sabemos que el **valor** de las **conclusiones** es también **verdadero**

Problema 2 para la sesión

Extrae una conclusión válida razonando sobre el siguiente caso: el alumno 1 aprueba, el alumno 2 aprueba, el alumno 3 aprueba.

- Se nos está preguntando cuál es la respuesta a un razonamiento de tipo inductivo.

RESPUESTA 2

ChatGPT (2/24): Dado que todos los alumnos (1, 2 y 3) aprueban, podemos extraer la conclusión de que "todos los alumnos aprueban".

Esto se puede expresar formalmente de la siguiente manera:

$\forall x (\text{Alumno}(x) \rightarrow \text{Aprueba}(x))$

Esto significa que para cualquier valor de x (que representaría a cualquier alumno), si x es un alumno, entonces x aprueba.

En inducción, no tenemos una regla de validez, sino una regla que invalida: el razonamiento no es válido si existe algún contraejemplo.

¿Podríamos pensar en un contrajemplo? **El razonamiento no es válido porque puede existir un contraejemplo.**

Razonamiento lógico inductivo

- ▶ **Inductivo:** Las observaciones particulares se pueden generalizar a veces en leyes. Suele entenderse que se usa un criterio probabilístico.
- ▶ **Contraejemplo:** la existencia de un solo contraejemplo invalida la hipótesis



Falsabilidad (Popper, 1934): las hipótesis falsificables (sobre las cuales podemos diseñar un experimento que pueda derivar en un contraejemplo) son preferibles

Problema para la sesión (cont)

Extrae una conclusión válida razonando sobre el siguiente caso: Pedro aprueba RyP, José suspende RyP, María suspende RyP, Paloma aprueba RyP

RESPUESTA 3

ChatGPT: Dado el caso presentado, podemos extraer la conclusión de que no todos los estudiantes aprueban la asignatura "RyP" (Razonamiento y Programación). Esto se puede expresar formalmente de la siguiente manera:

Aquí la respuesta es correcta desde el punto de vista deductivo; pero se podría extraer una hipótesis que explique los aprobados. Hay múltiples hipótesis que pueden explicar esas observaciones, podemos pensar en una: si el nombre de X empieza por P, X aprueba RyP.

No podemos garantizar que nuestra hipótesis (absurda) sea válida en general, pero podemos admitirla provisionalmente (con poco fundamento).

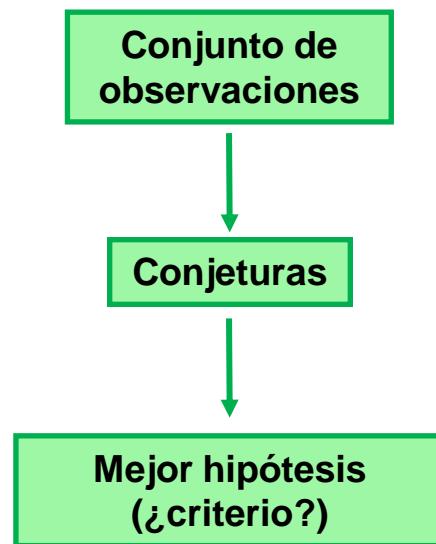
Razonamiento lógico abductivo

- ▶ Hay cierto debate sobre si es un tipo de inductivo
- ▶ **Hipótesis razonada:** ante unos hechos, se busca la hipótesis que mejor los explica

Pedro aprueba RyP
José suspende RyP
María suspende RyP
Paloma aprueba RyP



Conjetura (regla generalizada): si el nombre de X empieza por P, X aprueba RyP



Observación: aparece un sandwich medio comido en la cocina.

Hipótesis: tu hija lo preparó esta mañana, iba tarde al trabajo y se lo dejó al salir corriendo.

Criterio de evaluación: ¿navaja de Occam?

Ver: <https://www.youtube.com/watch?v=Jcw5KBZ5Y28>

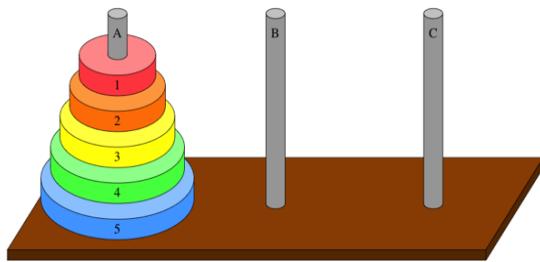


Representación y tipos de representación

Problema del conocimiento

- ▶ Ni humanos ni máquinas pueden acceder a la realidad directamente.
- ▶ En todos los casos se genera una representación (una imagen, quizá no exacta) interna de la realidad, con la que trabajar.
- ▶ Esta representación puede ser:
completa/incompleta, precisa/ambigua, correcta/incorrecta, relevante/irrelevante.
- ▶ Aunque la realidad no cambie, la representación sí puede modificarse
- ▶ Podemos representar las situaciones (**estados**) pero también los **procesos** (por ejemplo usando reglas)

Completa/Precisa/Correcta/Relevante



Objetos: discos, ejes

Discos: tamaño, posición

Posición: eje, altura

Ejes: capacidad

El juego consiste en pasar todos los discos al eje C, con restricciones

Restricción 1: sólo se puede mover el disco superior

Restricción 2: un disco
siempre sobre mesa u otro
mayor

¿Son relevantes el color, la mesa y la posición de los ejes?

¿Es precisa la restricción?

Requisitos de la representación en IA

- ▶ **Formal:** sin ambigüedades
- ▶ **Expresiva:** suficientemente rica para capturar los aspectos relevantes
- ▶ **Natural:** análoga a las formas naturales de expresión del conocimiento (esto no ocurre en Machine Learning)
- ▶ **Tratable:** se debe poder manipular para reflejar los procesos de razonamiento y generar respuestas



Resumen ideas principales

Resumen de las Ideas Principales del Tema 2

Resumen de las Características Clave

- Razonamiento lógicos vs no-lógicos
- Razonamientos lógicos: deductivo e inductivo (abductivo)
 - En RyP los algoritmos emplearán razonamientos deductivos para llegar a la solución
- En IA las representaciones deben cumplir una serie de características
 - En RyP utilizaremos la lógica de predicados

Nota: ¡No olviden leer las notas para afianzar los conceptos!

Construyendo el castillo de RyP

Resumen de las Características Clave



Nota: ¡Mostrar código clase anterior!

En RyP los algoritmos emplearán razonamientos deductivos para llegar a la solución

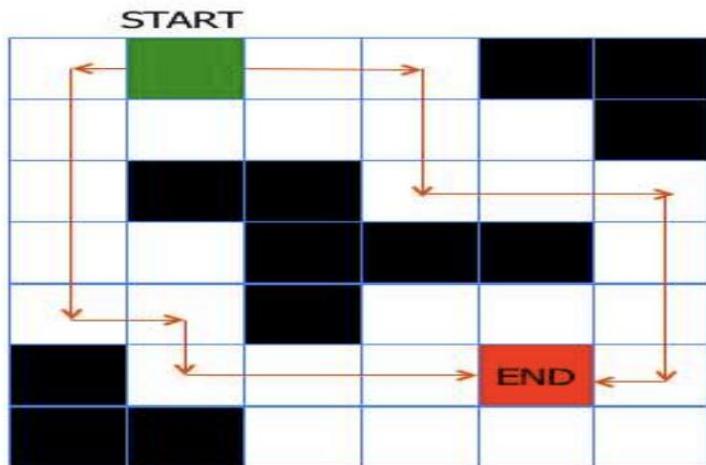
En RyP utilizaremos la lógica de predicados que cumple las características de **Formal, Expresiva, Natural y Tratable**



Tarea para casa

Ejemplos de representación

Path finding



Fuente: An example grid that shows multiple solutions to a path planning problem. Source: C.J. Taylor, University of Pennsylvania

¿Cómo representar cada posible **situación inicial** de este problema?

¿Cómo representar cada posible **solución** a este problema?

<https://zoom.us/wb/doc/5mmuHtwGTC625GfYVVJpow>

Encuesta

Input para el profesor, ¿cómo ha ido la clase?



Fuente:

https://www.emtusahuelva.com/index.php?option=com_content&view=article&id=1848:encuesta-de-satisfaccion-de-clientes-2023&catid=17&Itemid=160

Enlace encuesta:

https://grupoproeduca.zoom.us/survey/oRletlKhEDwYkJL0RCpFYc2E7PViKNUKoMhCX69NNV-HkkkZMW8.IW3UgbpKcmnFvI1/view?id=kGv_2buPTSCmXtxBtNBIIA#/sharePreview

QR para acceder a la encuesta



unir
LA UNIVERSIDAD
EN INTERNET

www.unir.net