Facsímil 12

Lógica e incertidumbre

Capítulo 07

Lo que deberías saber: lógica, conocimiento e incertidumbre

686f6c61 · 2026

Capítulo 07: Lo que deberías saber: lógica, conocimiento e incertidumbre

Entrando en el tema

Este facsímil ha recorrido la otra mitad de la inteligencia artificial: no la que aprende de millones de ejemplos, sino la que **razona con reglas, probabilidades y grados**. Empezamos con la lógica, que decide con verdades tajantes; vimos que el mundo real rara vez es tan limpio y pasamos a la probabilidad y a Bayes, que miden la creencia; admitimos que muchos conceptos no son ni ciertos ni falsos, sino graduales, y entró la lógica difusa; juntamos las piezas en los sistemas expertos, que codifican el saber de un dominio en reglas que un motor de inferencia recorre y explica; y terminamos con el razonamiento causal, el salto de «qué va con qué» a «qué pasa si intervengo».

Este capítulo de cierre hace dos cosas: recapitula las ideas para que las tengas juntas, mostrando cómo las cinco herramientas trabajan sobre un mismo problema, y te entrega los cuadernos para ejecutar cada pieza con tus propias manos.

Recapitulación activa del facsímil

1. Lógica para máquinas: proposicional, primer orden y resolución

La lógica es la forma más antigua de hacer que una máquina deduzca. Vimos la **lógica proposicional** (enunciados verdaderos o falsos combinados con conectivas), las **tablas de verdad** para decidir validez y satisfacibilidad, y la **regla de resolución** con su prueba por refutación: para demostrar algo, niégalo y deriva una contradicción. Subimos a la **lógica de primer orden**, con predicados, cuantificadores y **unificación**, que es lo que permite a un mini-Prolog encadenar reglas. Y enlazamos con el SAT del facsímil 2: la misma satisfacibilidad booleana, vista desde la lógica.

2. Razonamiento bajo incertidumbre: Bayes y redes bayesianas

Cuando los hechos son inciertos, la verdad tajante no sirve. La **probabilidad** mide el grado de creencia y el **teorema de Bayes** dice cómo actualizarla al ver evidencia. Aprendimos por qué un test positivo en una enfermedad rara deja la sospecha sorprendentemente baja (la tasa base manda), montamos un **clasificador naive Bayes** y subimos a las **redes bayesianas**: un grafo que factoriza una distribución conjunta enorme en tablas pequeñas y permite preguntar cosas como «¿cuál es la probabilidad de robo si suena la alarma?».

3. Cuando la verdad es difusa: factores de certeza y lógica difusa

No toda incertidumbre es probabilística: hay conceptos **vagos** («calentito», «alto», «caro») que se capturan con grados de pertenencia, no con probabilidades. La **lógica difusa** de Zadeh y los **factores de certeza** de MYCIN modelan ese terreno. Recorrimos la inferencia de **Mamdani** de principio a fin —fuzzificar, evaluar reglas, agregar y **defuzzificar** por centroide — y vimos que un controlador difuso produce respuestas suaves donde un sistema de umbrales daría saltos bruscos.

4. Sistemas expertos y soporte a la decisión

Un **sistema experto** codifica el saber de un dominio en una base de reglas que un **motor de inferencia** recorre, sea **hacia delante** (de los datos a las conclusiones) o **hacia atrás** (de la hipótesis a los hechos que la sostienen). Vimos su arquitectura, la resolución de conflictos, la idea del algoritmo Rete, los sistemas clásicos (MYCIN, DENDRAL, XCON) y su fragilidad, y la diferencia con un **sistema de soporte a la decisión**, que asiste a una persona en vez de decidir por ella.

5. Razonamiento causal: de la correlación a la intervención

Dimos el salto que la probabilidad sola no da: de la correlación a la causa. Con la **paradoja de Simpson** vimos cómo una tendencia se invierte al estratificar por un confusor; con la **escalera de la causalidad** separamos ver, hacer e imaginar; y con el **operador do** y la **fórmula de ajuste por puerta trasera** aprendimos a estimar el efecto real de una intervención, no solo la asociación observada. Es la herramienta que distingue un modelo que predice de uno que entiende qué pasa al actuar.

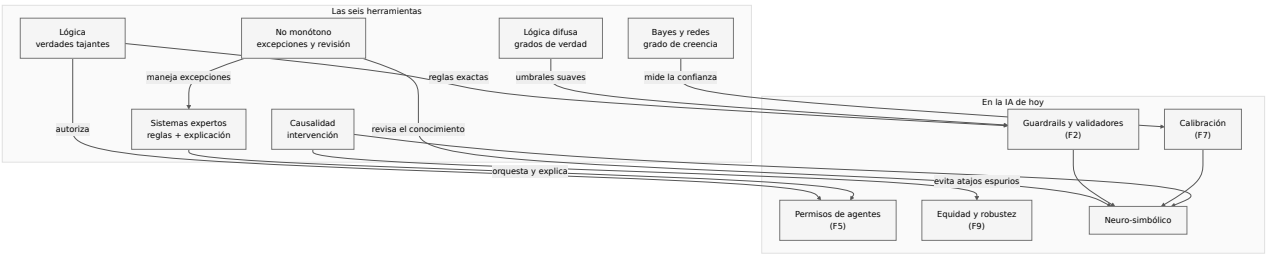
6. Razonamiento no monótono y revisión de creencias

Y cerramos con la lógica que **se retracta**. La lógica clásica del capítulo 1 es monótona: una vez que concluye algo, ninguna información nueva se lo quita. El sentido común no funciona así: concluyes que «Tweety vuela» porque es un ave, y te retractas al saber que es un pingüino. Vimos la **negación por fallo**, la **lógica por defecto** de Reiter (reglas que valen salvo excepción), la **revisión de creencias** AGM (cómo una base de conocimiento incorpora un hecho que contradice lo anterior sin volverse inconsistente) y la **argumentación**. Es lo que permite a un sistema razonar con información incompleta y corregirse, justo lo que un mundo cambiante exige.

Cómo encaja todo

Estas seis piezas no son una reliquia de museo: son la otra mitad que, sin que la veas, sostiene la inteligencia artificial que usas a diario. Cuanto más potente y más opaco es un modelo, más falta hacen las herramientas que saben decir, con claridad, «esto es seguro», «esto es probable», «esto es cuestión de grado» y «esto causa aquello». Un modelo de

lenguaje es, por dentro, una máquina **probabilística**; su **calibración** (facsímil 7) es Bayes aplicado a sus salidas. Los **guardrails** y validadores (facsímil 2) son **lógica**: reglas exactas que aceptan o rechazan. Los **permisos** de un agente (facsímil 5) son un pequeño **sistema experto**. Y la promesa a medio plazo —sistemas que aprenden *y* explican— es el matrimonio **neuro-simbólico**: la red percibe y propone, la lógica y la causalidad comprueban y justifican.



Cuadernos para practicar

Has modelado un problema con las cinco herramientas del facsímil; estos cuadernos te dejan ejecutar cada pieza. Son *notebooks* que se abren en Google Colab —gratis, en el navegador— o te puedes descargar. Cada uno está explicado paso a paso, con salidas reales que se generan al pulsar el botón de ejecutar, y dice de qué capítulo sale.

Lógica y resolución: deducir como una máquina

Qué practicas: construir tablas de verdad, demostrar argumentos por resolución y unificar términos de primer orden. **Dónde encaja:** capítulo 1 (lógica proposicional, primer orden y resolución).

Montas las herramientas de la lógica desde cero. Evalúas una fórmula sobre sus mundos posibles, demuestras un argumento clásico **por refutación** derivando la cláusula vacía, y compruebas el coste real de cada enfoque: con 18 variables, la tabla de verdad tiene **262.144** filas que mirar, mientras que un mini-DPLL decide lo mismo con **8** decisiones. Para rematar, implementas la **unificación** y un mini-Prolog que responde una consulta encadenando reglas.

► Abrir en Google Colab **↓ Descargar el cuaderno (.ipynb)**

Bayes y redes bayesianas: actualizar creencias con datos

Qué practicas: aplicar el teorema de Bayes, entrenar un naive Bayes y consultar una red bayesiana. **Dónde encaja:** capítulo 2 (probabilidad, Bayes y redes bayesianas).

Empiezas con la trampa de la tasa base: un test positivo para una enfermedad rara deja la probabilidad de estar enfermo en apenas un **16,7 %**, y ves cómo cambia al variar la prevalencia. Luego entrenas un **naive Bayes** y montas la **red bayesiana de la alarma**: la probabilidad de robo cuando llaman los dos vecinos sube del **0,1 %** de partida al **28,4 %**; y compruebas el *explaining away*, cuando un terremoto «explica» la alarma y la sospecha de robo se desploma del **37,4 %** al **0,33 %**.

► **Abrir en Google Colab** ↓ **Descargar el cuaderno (.ipynb)**

Lógica difusa: cuando la verdad es cuestión de grado

Qué practicas: definir conjuntos difusos y ejecutar un controlador de Mamdani de principio a fin. **Dónde encaja:** capítulo 3 (factores de certeza y lógica difusa).

Defines funciones de pertenencia y ves que **26 grados** pertenecen a la vez a «templado» con grado **0,607** y a «caliente» con grado **0,333**. Combinas evidencias con factores de certeza y montas un **controlador difuso** completo (la propina de un restaurante según servicio y comida): fuzzificas, evalúas las reglas, agregas y **defuzzificas** por centroide hasta un número concreto —una propina del **14,81 %**—. Y lo comparas con un controlador de umbrales para ver la diferencia entre una respuesta suave y una a saltos.

► **Abrir en Google Colab** ↓ **Descargar el cuaderno (.ipynb)**

Encadenamiento: un sistema experto que decide y explica

Qué practicas: construir un motor de inferencia con encadenamiento hacia delante y hacia atrás, y que explique sus conclusiones. **Dónde encaja:** capítulo 4 (sistemas expertos y soporte a la decisión).

Construyes un sistema experto que decide si conceder un préstamo a partir de una base de reglas. Lo resuelves de las dos maneras y cuentas el esfuerzo: **hacia delante** (de los datos) deriva los hechos intermedios revisando las reglas hasta el punto fijo; **hacia atrás** (desde la meta «aprobar») sigue solo la cadena que la sostiene, con

muchas menos comprobaciones. Añades **factores de certeza** y propagas la confianza hasta la conclusión: se aprueba, pero con una certeza de alrededor del **53 %**, no del 100 %. Y el sistema **explica** la cadena completa de reglas y hechos.

► **Abrir en Google Colab** ↓ **Descargar el cuaderno (.ipynb)**

Razonamiento causal: de la correlación a la intervención

Qué practicas: distinguir correlación de causa, ver la paradoja de Simpson y ajustar por un confusor. **Dónde encaja:** capítulo 5 (razonamiento causal).

Generas datos con un confusor y ves nacer una **correlación espuria**; reproduces la **paradoja de Simpson**, donde la tendencia agregada se invierte al estratificar; y calculas, sobre un modelo causal, la diferencia entre **observar** $P(Y \mid X)$ y **intervenir** $P(Y \mid do(X))$. Con la **fórmula de ajuste por puerta trasera** recuperas el efecto causal correcto y compruebas, con números, cuánto engañaba la correlación de partida.

► **Abrir en Google Colab** ↓ **Descargar el cuaderno (.ipynb)**

Razonamiento no monótono: la lógica que se retracta

Qué practicas: negación por fallo, lógica por defecto con excepciones y revisión de creencias. **Dónde encaja:** capítulo 6 (razonamiento no monótono y revisión de creencias).

Construyes un mini-razonador que concluye «Tweety vuela» porque es un ave y **se retracta** al saber que es un pingüino: la no monotonía en acción. Implementas la **negación por fallo** (concluir «no vuela» cuando no se puede demostrar «vuela»), calculas la **extensión** de una base de reglas por defecto y montas una **revisión de creencias** al estilo AGM, donde la base incorpora un hecho que contradice lo anterior y restaura la consistencia quitando lo menos atrincherado. Acabas con un pequeño grafo de **argumentación**: argumentos que se atacan y cuáles quedan aceptados.

► **Abrir en Google Colab** ↓ **Descargar el cuaderno (.ipynb)**

En resumen

- La IA simbólica y el razonamiento bajo incertidumbre son **la otra mitad** del campo: lógica para lo tajante, probabilidad para lo incierto, difusa para lo gradual, sistemas expertos para orquestar con explicación y causalidad para distinguir correlación de causa.[1]
- Estas herramientas no compiten con el aprendizaje profundo: lo **complementan**. Donde hace falta auditar, justificar, garantizar una regla o saber qué pasa al intervenir, lo simbólico y lo causal siguen ganando.
- Juntar las cinco sobre un mismo problema muestra la idea central del facsímil: **cada parte de una decisión merece la herramienta adecuada**, y un sistema que lo respeta es más robusto y más defendible que cualquier caja negra.

Para saber más

- Garcez, A. d'Avila, Lamb, L. C. y Gabbay, D. M. (2009). *Neural-Symbolic Cognitive Reasoning*. Springer.
- Jackson, P. (1998). *Introduction to Expert Systems* (3.ª ed.). Addison-Wesley.
- Pearl, J. y Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. Basic Books.
- Russell, S. y Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4.ª ed.). Pearson.
- Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, 8(3), 338-353.

Notas

1. Pearl, J. y Mackenzie, D. (2018). *The Book of Why*. Basic Books. Una defensa accesible de por qué la estadística sola no responde preguntas causales y hace falta el lenguaje de las intervenciones. ↩