

# TP4

September 3, 2025

## 1 Temas Tratados en el Trabajo Práctico 4

- Representación del Conocimiento y Razonamiento Lógico.
- Estrategias de resolución de hipótesis: Encadenamiento hacia Adelante, Encadenamiento hacia Atrás y Resolución por Contradicción.
- Representación basada en circuitos.

## 2 Ejercicios Teóricos

### 2.1 ¿Qué es una inferencia?

La inferencia es el proceso de usar un modelo de IA ya entrenado para hacer predicciones o tomar decisiones sobre datos nuevos que no ha visto antes.

### 2.2 ¿Cómo se verifica que un modelo se infiere de la base de conocimientos?

Verificar que la inferencia de un modelo de IA se base en su conocimiento es, en esencia, confirmar que sus predicciones son correctas y fiables. Este proceso no se hace mirando dentro del modelo como una “caja transparente”, sino evaluando su desempeño con datos que nunca ha visto. Después de entrenar un modelo, se utilizan conjuntos de datos de validación y prueba que contienen preguntas con respuestas conocidas. Si el modelo hace inferencias precisas sobre estos nuevos datos, se considera que ha aprendido a generalizar de su base de conocimiento en lugar de simplemente memorizar. Para cuantificar este rendimiento, se usan métricas como la precisión, la exactitud y la matriz de confusión, que muestran con claridad dónde el modelo acierta y dónde falla.

### 2.3 Observe la siguiente base de conocimiento:

$R1 : b \wedge c \rightarrow a$

$R2 : d \wedge e \rightarrow b$

$R3 : g \wedge e \rightarrow b$

$R4 : e \rightarrow c$

$R5 : d$

$R6 : e$

$R7 : a \wedge g \rightarrow f$

De  $R5$  y  $R6$  se sabe que tanto  $d$  como  $e$  son tautologías, se podría decir que es un axioma de la Base de Conocimientos. A partir de esto, se pueden desarrollar las demás reglas.

**2.3.1 ¿Cómo se puede probar que  $a = True$  a través del encadenamiento hacia adelante? Este método solamente usa reglas ya incorporadas a la base de conocimiento para inferir la hipótesis, ¿qué propiedad debe tener el algoritmo para asegurar que esta inferencia sea posible?**

Con el encadenamiento hacia adelante podemos probar  $a = True$  con  $a$  a partir de las cosas que sabemos, en este caso  $R5$  y  $R6$  las cuales son tautologías. Así, se razona en el orden  $R2$  y  $R4$  primero, luego  $R1$ :

$R2: (d \wedge e \rightarrow b) \quad \text{entonces} \quad b = True$

- Ya que  $d$  y  $e$  siempre son verdaderas.

$R4: (e \rightarrow c) \quad \text{entonces} \quad c = True$

- Ya que  $c$  siempre es verdadera

Entonces, puedo plantear:

$R1: (b \wedge c \rightarrow a) \quad \text{entonces} \quad a = True$

Por lo tanto, la hipótesis  $a$  puede inferirse correctamente a partir de las reglas y hechos de la base de conocimiento.

La propiedad que debe tener el algoritmo para asegurar que esta inferencia sea posible es la completitud: que todo enunciado que se sigue lógicamente de la base de conocimiento pueda, efectivamente, ser derivado por el procedimiento de inferencia.

**2.3.2 ¿Cómo se puede probar que  $a = True$  a través del encadenamiento hacia atrás? Este método asigna un valor de verdad a la hipótesis y deriva las sentencias de la base de conocimiento, ¿qué propiedad debe tener el algoritmo para asegurar que esta derivación sea posible?**

Primeramente, se hace la hipótesis  $a = True$  y se sigue con:

$R1:$

$(b \wedge c) \rightarrow a$

Entonces debo probar que  $b$  y  $c$  son verdaderos. ¿Cómo pruebo que son verdaderos? Partiendo de  $R2$  para  $b$  y  $R4$  para  $c$ :

$R2:$

$(d \wedge e) \rightarrow b$

Como  $d$  y  $e$  son verdades de BC (axiomas), entonces  $b = True$  en BC.

$R4:$

$e \rightarrow c$

Como  $e$  es una verdad de BC (axioma), entonces  $c = True$  en BC.

Acabamos de probar que  $b$  y  $c$  son verdaderos en BC, por lo que  $a$  deberá ser también verdadero en BC.

$a = \text{True}$

### 2.3.3 Exprese la base de conocimiento en su Forma Normal Conjuntiva. A continuación, demuestre por contradicción que $a = \text{True}$ .

Ahora, utilizamos la regla de resolución para encontrar una contradicción, que se manifiesta como la derivación de la cláusula vacía  $\{\}$ .

- Paso 1: Resolvemos C5 ( $b$ ) con C2 ( $\neg b \neg a$ ).

Al resolver  $b$  con  $\neg b$ , eliminamos ambos.

Resultado:  $\neg a$

Esta es nuestra nueva cláusula, C7. Esto no nos ayuda a avanzar, ya que es lo mismo que C6.

- Paso 2: Intentemos resolver C5 ( $b$ ) con C1 ( $\neg b c$ ).

Al resolver  $b$  con  $\neg b$ , el resultado es  $c$ .

Resultado:  $c$

Esta es nuestra nueva cláusula, C8.

- Paso 3: Resolvemos C8 ( $c$ ) con C3 ( $\neg c a$ ).

Al resolver  $c$  con  $\neg c$ , el resultado es  $a$ .

Resultado:  $a$

Esta es nuestra nueva cláusula, C9.

- Paso 4: Ahora, resolvemos C9 ( $a$ ) con nuestra suposición inicial, C6 ( $\neg a$ ).

El resultado es la cláusula vacía,  $\{\}$ .

- Conclusión

Hemos llegado a una contradicción al derivar la cláusula vacía. Esto demuestra que nuestra suposición de que  $\neg a$  era verdadera es falsa. Por lo tanto, se ha probado que  $a$  es verdadera.

## 2.4 Diseñe con lógica proposicional basada en circuitos las proposiciones *OrientadoDerecha* y *Agente ubicado en la casilla [1,2]* para el mundo de wumpus de 4x4. Dibuje el circuito correspondiente.

Orientación:



2.5 El nonograma es un juego en el cual se posee un tablero en blanco y cada fila y columna presenta información sobre la longitud de un bloque en dicha fila/columna. Además, la leyenda puede indicar más de un número, indicando esto que existen varios bloques de las longitudes mostradas por la leyenda y en el mismo orden, separados por al menos un espacio vacío.

Resuelva el nonograma de la imagen de abajo escribiendo en primer lugar cada regla que puede incorporarse a la base de conocimientos inicial e incorporando cada inferencia que realice.

```
[1]: import requests
from PIL import Image
from io import BytesIO
import matplotlib.pyplot as plt

# URL directa de Google Drive
url = "https://drive.google.com/uc?
      ↪export=view&id=1SKiXvrI_TX-U4sbw60TYSRmaNYyFixmI"

# Descargar la imagen
response = requests.get(url)
img = Image.open(BytesIO(response.content))

# Mostrar la imagen
plt.imshow(img)
plt.axis('off') # Ocultar ejes
plt.show()
```

		3	3	2 1
1 1				
4				
2 1				
3				

Vamos a identificar el tablero como una matriz de 4X4, donde cada casilla se identificara con  $i,j$ , donde  $i$  es el numero de fila (de arriba hacia abajo) y  $j$  el número de columna (de izquierda a derecha).

$P_{i,j}$  = casilla  $i,j$  pintada.

$\neg P_{i,j}$  = casilla  $i,j$  en blanco.

Partimos de las filas y columnas que tienen una única posibilidad de resolución. Fila 2: indica que los 4 de la fila deben pintarse.

R1:  $(P_{2,1} \wedge P_{2,2} \wedge P_{2,3} \wedge P_{2,4})$ .

Fila 3: Indica que primero se pintan dos casillas de la fila, y luego 1. La única posibilidad es que la casilla (3,3) quede en blanco.

R2:  $(P_{3,1} \wedge P_{3,2} \wedge \neg P_{3,3} \wedge P_{3,4})$

Columna 3: Indica que se deben pintar dos casillas y luego una, por lo que la unica posibilidad es que la casilla (3,3) quede sin pintar.

R3:  $(P_{1,3} \wedge P_{2,3} \wedge \neg P_{3,3} \wedge P_{4,3})$

		3	3	2 1	3
1 1					
4					
2 1					
3					

Seguimos nuestra deducción desde la regla 3. Si en la fila 1, la casilla (1,3) ya está pintada, sus casillas adyacentes en la fila no pueden estar pintadas ya que la regla de la fila indica que deben pintarse de a una.

R4:  $P_{1,3} \rightarrow (\neg P_{1,2} \wedge \neg P_{1,4})$

Luego, si la casilla (1,2) está en blanco según la regla 4, significa que las casillas de al lado, (1,1) y (1,3) deben pintarse, por las condiciones dadas para la fila 1.

R5:  $\neg P_{1,2} \rightarrow (P_{1,1} \wedge P_{1,3})$

		3	3	2 1	3
1 1					
4					
2 1					
3					

En base a la regla 5, si la casilla (1,2) no está pintada, significa que el resto de la columna 2 debe pintarse, ya que la condición de la columna 2 es que deben pintarse 3 casillas seguidas.

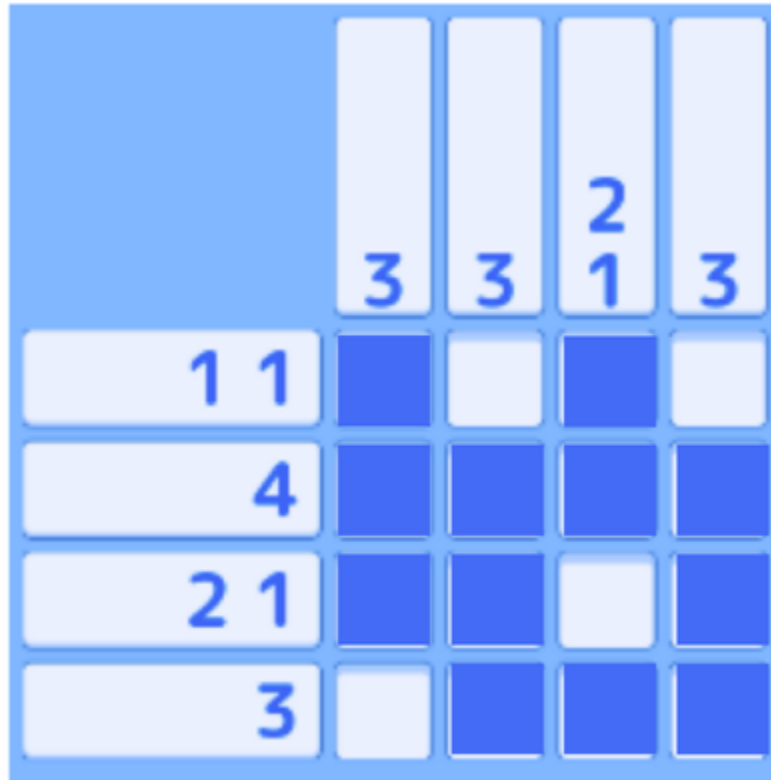
R6:  $\neg P_{1,2} \rightarrow (P_{2,2} \wedge P_{3,2} \wedge P_{4,2})$

		3	3	2 1	3
1 1					
4					
2 1					
3					

En base a la regla 4, si la casilla (1,4) está en blanco, el resto de la columna 4 debe estar pintado, ya que la condición de esa columna es que se pinten tres casillas seguidas.

R7:  $\neg P_{1,4} \rightarrow (P_{2,4} \wedge P_{3,4} \wedge P_{4,4})$





En la fila 4, si las casillas (4,2), (4,3) y (4,4) ya están pintadas, la casilla (4,1) debe quedarse en blanco ya que la condición de esa fila indica que solo se pinten 3 seguidas.

R8:  $(P_{4,2} \wedge P_{4,3} \wedge P_{4,4}) \rightarrow \neg P_{1,4}$

De esta forma, ya se ha analizado el estado de todas las casillas, por lo que queda resuelto en nanograma:

			2	
	3	3	1	3
1 1				
4				
2 1				
3				

### 3 Ejercicios de Implementación

#### 3.1 Implementar un motor de inferencia con encadenamiento hacia adelante. Pruébalo con las proposiciones del ejercicio 3.

```
[ ]: def forward_chaining(reglas, hechos_iniciales, objetivo):
    """
    Implementa un motor de inferencia con encadenamiento hacia adelante.

    Args:
        reglas (list): Una lista de reglas en formato (premisa, conclusion).
                       Las premisas son un conjunto, y la conclusion es un
    ↪ string.
        hechos_iniciales (set): Un conjunto de hechos conocidos.
        objetivo (str): La proposición que se quiere demostrar.

    Returns:
        bool: True si el objetivo se puede derivar, False en caso contrario.
    """
    hechos_conocidos = set(hechos_iniciales)

    # Bucle que se ejecuta mientras se puedan derivar nuevos hechos.
    while True:
```

```

nuevos_hechos_derivados = False

# Iterar a través de cada regla en la base de conocimiento.
for premisa, conclusion in reglas:
    # Comprobar si todas las proposiciones en la premisa son hechos
    ↪ conocidos.
    if premisa.issubset(hechos_conocidos):
        # Si la conclusión no es un hecho conocido, la añadimos.
        if conclusion not in hechos_conocidos:
            hechos_conocidos.add(conclusion)
            nuevos_hechos_derivados = True
            print(f"Se ha inferido '{conclusion}' a partir de
    ↪ '{premisa}'")

    # Si no se ha derivado ningún nuevo hecho, el bucle se detiene.
    if not nuevos_hechos_derivados:
        break

# Comprobar si el objetivo está en el conjunto de hechos conocidos.
return objetivo in hechos_conocidos

# --- Prueba con las proposiciones del ejercicio 3 ---

# 1. Definir la base de conocimiento como reglas
# R1: b AND c -> a
# R2: d AND e -> b
# R3: g AND e -> b
# R4: e -> c
# R7: a AND g -> f
# Las premisas son conjuntos, para manejar AND.

reglas_ejercicio3 = [
    (frozenset(['b', 'c']), 'a'),
    (frozenset(['d', 'e']), 'b'),
    (frozenset(['g', 'e']), 'b'),
    (frozenset(['e']), 'c'),
    (frozenset(['a', 'g']), 'f')
]

# 2. Definir los hechos iniciales (del ejercicio 3)
# R5: d
# R6: e
# g no se menciona, por lo que no es un hecho.
hechos_ejercicio3 = {'d', 'e'}

# 3. Definir el objetivo
objetivo_a = 'a'

```

```

# 4. Ejecutar el motor de inferencia
print("Iniciando motor de inferencia con encadenamiento hacia adelante...")
resultado = forward_chaining(reglas_ejercicio3, hechos_ejercicio3, objetivo_a)

# 5. Imprimir el resultado
print("\n--- Resultado ---")
if resultado:
    print(f"Éxito: Se ha demostrado que '{objetivo_a}' es verdadero.")
else:
    print(f"Fallo: No se pudo demostrar que '{objetivo_a}' es verdadero.")

```

### 3.2 Implementar un motor de inferencia con encadenamiento hacia atrás. Pruébalo con las proposiciones del ejercicio 3.

```

[ ]: def backward_chaining(reglas, hechos_iniciales, objetivo):
    """
    Implementa un motor de inferencia con encadenamiento hacia atrás.

    Args:
        reglas (list): Una lista de reglas en formato (premisa, conclusion).
                       Las premisas son un conjunto, y la conclusion es un
        ↪ string.
        hechos_iniciales (set): Un conjunto de hechos conocidos.
        objetivo (str): La proposición que se quiere demostrar.

    Returns:
        bool: True si el objetivo se puede derivar, False en caso contrario.
    """
    # 1. Caso base: El objetivo ya es un hecho conocido.
    if objetivo in hechos_iniciales:
        print(f"Éxito: El objetivo '{objetivo}' ya es un hecho conocido.")
        return True

    # 2. Búsqueda de reglas que pueden probar el objetivo.
    reglas_posibles = [regla for regla in reglas if regla[1] == objetivo]

    # 3. Si no hay reglas para probar el objetivo, la prueba falla.
    if not reglas_posibles:
        print(f"Fallo: No hay reglas para probar el objetivo '{objetivo}'.")
        return False

    # 4. Intentar probar el objetivo a través de cada regla posible.
    for premisa, _ in reglas_posibles:
        print(f"Intentando probar '{objetivo}' usando la regla con premisa
        ↪ '{premisa}'")

```

```

    todos_subobjetivos_probados = True

    # 5. Intentar probar cada subobjetivo en la premisa.
    for sub_objetivo in premisa:
        # Llamada recursiva para probar el subobjetivo.
        if not backward_chaining(reglas, hechos_iniciales, sub_objetivo):
            todos_subobjetivos_probados = False
            break

    # 6. Si todos los subobjetivos se probaron, el objetivo es verdadero.
    if todos_subobjetivos_probados:
        print(f"Éxito: Todos los subobjetivos para '{objetivo}' fueron_
↪probados. Conclusión alcanzada.")
        return True

    # 7. Si ninguna regla pudo probar el objetivo, la prueba falla.
    return False

# --- Prueba con las proposiciones del ejercicio 3 ---

# 1. Definir la base de conocimiento como reglas
# R1: b AND c -> a
# R2: d AND e -> b
# R3: g AND e -> b
# R4: e -> c
# R7: a AND g -> f
# Las premisas son conjuntos, para manejar AND.
reglas_ejercicio3 = [
    (frozenset(['b', 'c']), 'a'),
    (frozenset(['d', 'e']), 'b'),
    (frozenset(['g', 'e']), 'b'),
    (frozenset(['e']), 'c'),
    (frozenset(['a', 'g']), 'f')
]

# 2. Definir los hechos iniciales (del ejercicio 3)
# R5: d
# R6: e
hechos_ejercicio3 = {'d', 'e'}

# 3. Definir el objetivo
objetivo_a = 'a'

# 4. Ejecutar el motor de inferencia
print("Iniciando motor de inferencia con encadenamiento hacia atrás...")
resultado = backward_chaining(reglas_ejercicio3, hechos_ejercicio3, objetivo_a)

```

```

# 5. Imprimir el resultado
print("\n--- Resultado ---")
if resultado:
    print(f"Éxito: Se ha demostrado que '{objetivo_a}' es verdadero.")
else:
    print(f"Fallo: No se pudo demostrar que '{objetivo_a}' es verdadero.")

```

### 3.3 Implementar un motor de inferencia por contradicción que detecte si el conjunto de proposiciones del ejercicio 3 es inconsistente.

```

[ ]: import itertools

def to_cnf(regla):
    """
    Convierte una regla simple (implicación) a FNC.
    Asume que la premisa es una conjunción de literales.
    """
    premisa, conclusion = regla
    literales_premisa = [f"¬{p}" for p in premisa]

    # La implicación (A -> B) se convierte en ¬A ∨ B
    # donde A es la conjunción de la premisa.
    return frozenset(literales_premisa + [conclusion])

def resolver_clausulas(c1, c2):
    """
    Aplica la regla de resolución a dos cláusulas.
    Retorna la cláusula resuelta o None si no se pueden resolver.
    """
    for literal in c1:
        negado = f"¬{literal}" if not literal.startswith("¬") else literal[1:]
        if negado in c2:
            nueva_clausula = (c1 - {literal}) | (c2 - {negado})
            return nueva_clausula
    return None

def motor_inconsistencia_resolucion(clausulas):
    """
    Motor de inferencia que detecta inconsistencias usando resolución.

    Args:
        clausulas (list): Una lista de conjuntos inmutables (frozenset) de
        literales.

    Returns:
    """

```

```

    bool: True si el conjunto es inconsistente, False en caso contrario.
    """
    clausulas_conocidas = set(clausulas)

    while True:
        nuevas_clausulas = set()
        pares_clausulas = itertools.combinations(clausulas_conocidas, 2)

        for c1, c2 in pares_clausulas:
            resultado_resolucion = resolver_clausulas(c1, c2)

            if resultado_resolucion is not None:
                print(f"Resolviendo {c1} y {c2} -> {resultado_resolucion}")

                if not resultado_resolucion: # La cláusula vacía {}
                    print("\n;Inconsistencia detectada! Se ha derivado la_
↪cláusula vacía.")
                    return True

                # Si la nueva cláusula no es trivial y no se ha visto antes.
                if resultado_resolucion not in clausulas_conocidas:
                    nuevas_clausulas.add(resultado_resolucion)

                # Si no se generan nuevas cláusulas, no hay contradicción.
                if not nuevas_clausulas:
                    print("\nNo se encontraron nuevas cláusulas. El conjunto de_
↪proposiciones es consistente.")
                    return False

        clausulas_conocidas.update(nuevas_clausulas)

# --- Prueba con un conjunto inconsistente ---
# Para probar la inconsistencia, añadiremos una regla que contradice
# las premisas. Por ejemplo, sabemos que 'c' es True, así que si agregamos
# la negación de 'c', causaremos una contradicción.

# Base de conocimiento del ejercicio 3 convertida a FNC:
# R1: b AND c -> a --> ¬b v ¬c v a
# R2: d AND e -> b --> ¬d v ¬e v b
# R3: g AND e -> b --> ¬g v ¬e v b
# R4: e -> c --> ¬e v c
# R5: d --> d
# R6: e --> e

# Convertimos las proposiciones a un formato de conjunto.
    clausulas_base = [
        frozenset(['¬b', '¬c', 'a']),

```

```

    frozenset(['¬d', '¬e', 'b']),
    frozenset(['¬g', '¬e', 'b']),
    frozenset(['¬e', 'c']),
    frozenset(['d']),
    frozenset(['e'])
]

# Añadimos una proposición que haga el conjunto inconsistente.
# Sabemos que de 'e' (R6) y '¬e v c' (R4) se deriva 'c'.
# Por lo tanto, si añadimos '¬c', el conjunto será inconsistente.
clausulas_inconsistentes = clausulas_base + [frozenset(['¬c'])]

# Ejecutar el motor
print("Analizando la base de conocimiento para detectar inconsistencia...")
es_inconsistente = motor_inconsistencia_resolucion(clausulas_inconsistentes)

print("\n--- Resultado ---")
if es_inconsistente:
    print("El conjunto de proposiciones es inconsistente. ¡Prueba exitosa! ")
else:
    print("El conjunto de proposiciones es consistente.")

```

## 4 Bibliografía

Russell, S. & Norvig, P. (2004) *Inteligencia Artificial: Un Enfoque Moderno*. Pearson Educación S.A. (2a Ed.) Madrid, España

Poole, D. & Mackworth, A. (2023) *Artificial Intelligence: Foundations of Computational Agents*. Cambridge University Press (3a Ed.) Vancouver, Canada