

Unidad II: Lógica de Primer Orden

Lógica de primer orden: Sintaxis y semántica

Clase 10 - Lógica para Ciencia de la Computación/Teoría de la Computación

Prof. Miguel Romero

¿Qué hace la lógica?

Lógica formal

¿Es este argumento válido?

Todas las personas son mortales.

Sócrates es persona.

Por lo tanto, Sócrates es mortal.

La lógica nos debería decir que sí.

Lógica formal

¿Es este argumento válido?

Algunas personas son mortales.

Sócrates es persona.

Por lo tanto, Sócrates es mortal.

La lógica nos debería decir que no.

¿Es este argumento válido?

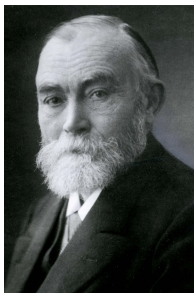
Todas las calilas son temporales.

Mcfly es calila.

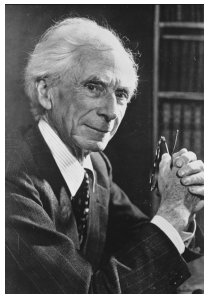
Por lo tanto, Mcfly es temporal.

Aristóteles (384 a.C - 322 a.C):

La correctitud de los argumentos depende de la **forma**.



Gottlob Frege (1848 - 1925)



Bertrand Russell (1872 - 1970)

Frege (1879): Un sistema formal para las matemáticas.

Russell (1902): El sistema de Frege es inconsistente (paradoja de Russell).

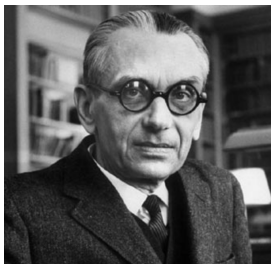
Programa de Hilbert

Programa de Hilbert (1922-1930):

Encontrar un sistema formal para las matemáticas y demostrar que:

- El sistema es **consistente**: no se puede demostrar una proposición matemática y su negación, al mismo tiempo.
- El sistema es **completo**: toda proposición matemática verdadera se puede demostrar en el sistema.
- El sistema es **decidible**: Existe una **forma mecánica** de determinar si una proposición matemática en el sistema es verdadera o falsa.

Teoremas de incompletitud



Kurt Gödel (1906 - 1978)

Teoremas de incompletitud (1931):

- Ningún sistema formal es completo.
- Ningún sistema formal puede demostrar su propia consistencia.

Los dos primeros puntos del programa de Hilbert no se pueden alcanzar.

Entscheidungsproblem

Entscheidungsproblem (El problema de decisión) [Hilbert-Ackermann 1928]:
Dada una fórmula en la lógica de primer orden, decidir si es **válida**, es decir,
si es siempre verdadera.



Alonzo Church (1903 - 1995)



Alan Turing (1912 - 1954)

Church y Turing (1936):

El problema de decisión es **indecidable**: no tiene solución algorítmica.

Fuerte conexión entre lógica y computación:

- La Ciencia de la Computación nace a partir del estudio de la lógica.

La lógica ha encontrado diversas aplicaciones en Ciencia de la Computación:

- Inteligencia Artificial
- Bases de Datos
- Verificación formal
- Teoría de la Computación
- Lenguajes de Programación
- ...

“El cálculo de la Ciencia de la Computación”

Tipos de lógicas

Existen diversos tipos de lógicas:

- Lógica proposicional
- Lógica de primer orden
- Lógica de segundo orden
- ...

Noa enfocaremos en la lógica de primer orden (LPO).

Definición:

Un **vocabulario** es un conjunto $\mathcal{L} = \{\mathcal{C}, \mathcal{F}, \mathcal{R}\}$ donde:

- $\mathcal{C} = \{c_1, c_2, \dots\}$ son los **símbolos de constantes**.
- $\mathcal{F} = \{f_1, f_2, \dots\}$ son los **símbolos de funciones**.
Cada símbolo f_i tiene asociada una aridad $n_i > 0$.
- $\mathcal{R} = \{R_1, R_2, \dots\}$ son los **símbolos de relaciones**.
Cada símbolo R_i tiene asociada una aridad $m_i \geq 0$.

El vocabulario determina los símbolos que podemos usar en las fórmulas.

Vocabulario: ejemplos

- Vocabulario para los números naturales:

$$\mathcal{L} = \{0, 1, s, +, \cdot, <\}$$

donde:

- 0 y 1 son constantes.
- s es función unaria, + y · son funciones binarias.
- < es una relación binaria.

- Vocabulario para grafos:

$$\mathcal{L} = \{E\}$$

donde E es una relación binaria.

Sintaxis de LPO: idea

Las fórmulas de la lógica de primer orden se construyen usando:

- Símbolos del vocabulario \mathcal{L} .
- Un símbolo especial de **relación binaria** $=$.
- Conectivos lógicos: \neg , \vee , \wedge , \rightarrow , \leftrightarrow .
- Paréntesis (y).
- **Variables**.
- **Cuantificadores** \exists y \forall .

Asumiremos que tenemos un conjunto infinito de variables $\mathcal{V} = \{x, y, \dots\}$.

Sea \mathcal{L} un vocabulario.

Definición:

El conjunto de \mathcal{L} -términos se define como el menor conjunto tal que:

- Cada constante $c \in \mathcal{L}$ es un \mathcal{L} -término.
- Cada variable $x \in \mathcal{V}$ es un \mathcal{L} -término.
- Si t_1, \dots, t_n son \mathcal{L} -términos y $f \in \mathcal{L}$ es un símbolo de función n -aria, entonces $f(t_1, \dots, t_n)$ es un \mathcal{L} -término.

Ejemplo:

Para el vocabulario de los naturales $\mathcal{L} = \{0, 1, s, +, \cdot, <\}$, son \mathcal{L} -términos:

0 $s(s(1))$ $+(\cdot(x, y), 1)$ $\cdot(+ (s(0), y), s(x))$

Definición:

El conjunto de \mathcal{L} -fórmulas se define como el menor conjunto tal que:

- Si t_1, t_2 son \mathcal{L} -términos, entonces $t_1 = t_2$ es \mathcal{L} -fórmula.
- Si t_1, \dots, t_m son \mathcal{L} -términos y $R \in \mathcal{L}$ es un símbolo de relación m -aria, entonces $R(t_1, \dots, t_m)$ es \mathcal{L} -fórmula.
- Si φ, ψ son \mathcal{L} -fórmulas, entonces $(\neg\varphi)$, $(\varphi \vee \psi)$, $(\varphi \wedge \psi)$, $(\varphi \rightarrow \psi)$ y $(\varphi \leftrightarrow \psi)$ son \mathcal{L} -fórmulas.
- Si φ es \mathcal{L} -fórmula y $x \in \mathcal{V}$ es una variable, entonces $(\exists x \varphi)$ y $(\forall x \varphi)$ son \mathcal{L} -fórmulas.

Decimos que las \mathcal{L} -fórmulas de la forma
 $t_1 = t_2$ y $R(t_1, \dots, t_m)$ son **fórmulas atómicas**.

Sintaxis de LPO: convenciones

Algunas convenciones:

- Utilizamos notación **infija** para funciones y relaciones conocidas.
 - $0 < x$ en vez de $<(0, x)$ $x + 1$ en vez de $+(x, 1)$.
- Omitimos paréntesis innecesarios, si no hay ambigüedad.

■ Algunas fórmulas para $\mathcal{L} = \{0, 1, s, +, \cdot, <\}$:

- $1 = s(0)$.
- $\forall x (x < s(x))$.
- $\forall x \exists y (x = y + y)$.
- $\forall x \forall y (s(x) = s(y) \rightarrow x = y)$.

■ Algunas fórmulas para $\mathcal{L} = \{E\}$:

- $\forall x \neg E(x, x)$.
- $\forall x \forall y (E(x, y) \rightarrow E(y, x))$.
- $\forall x \exists y E(x, y)$.

¿Cuál es el valor de verdad de las fórmulas anteriores?

- Depende de la **interpretación** que se les da a las constantes, funciones y relaciones del vocabulario.
- Esto viene dado por la noción de **estructura**.

Definición:

Una \mathcal{L} -estructura \mathfrak{A} es una tupla que contiene:

- Un dominio A no vacío.
- Para cada símbolo de constante $c \in \mathcal{L}$, un elemento $c^{\mathfrak{A}} \in A$.
- Para cada símbolo de función n -aria $f \in \mathcal{L}$, una función $f^{\mathfrak{A}} : A^n \rightarrow A$.
- Para cada símbolo de relación m -aria $R \in \mathcal{L}$, una relación $R^{\mathfrak{A}} \subseteq A^m$.

Notación:

- $c^{\mathfrak{A}}$, $f^{\mathfrak{A}}$ y $R^{\mathfrak{A}}$ son las interpretaciones de sus símbolos respectivos.
- Escribimos una estructura como $\mathfrak{A} = (A, c^{\mathfrak{A}}, \dots, f^{\mathfrak{A}}, \dots, R^{\mathfrak{A}}, \dots)$.

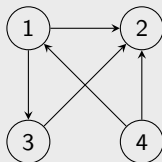
Ejemplo:

Sea el vocabulario $\mathcal{L} = \{E\}$, donde E es una relación binaria.

Una posible \mathcal{L} -estructura puede ser $\mathfrak{A} = (A, E^{\mathfrak{A}})$ con:

$$\begin{aligned} A &= \{1, 2, 3, 4\} \\ E^{\mathfrak{A}} &= \{(1, 2), (1, 3), (3, 2), (4, 1), (4, 2)\} \end{aligned}$$

Notar que esta estructura corresponde al grafo dirigido



Ejemplo:

Sea el vocabulario $\mathcal{L} = \{0, 1, s, +, \cdot, <\}$ anterior.

- Los números naturales se pueden representar con la \mathcal{L} -estructura

$$\mathfrak{N} = (\mathbb{N}, 0^{\mathfrak{N}}, 1^{\mathfrak{N}}, s^{\mathfrak{N}}, +^{\mathfrak{N}}, \cdot^{\mathfrak{N}}, <^{\mathfrak{N}}).$$

donde:

- $0^{\mathfrak{N}}$ y $1^{\mathfrak{N}}$ son el cero y el uno de \mathbb{N} .
- $s^{\mathfrak{N}}$ es la función *sucesor* en \mathbb{N} (es decir, $s(n) = n + 1$).
- $+^{\mathfrak{N}}$ y $\cdot^{\mathfrak{N}}$ es la suma y multiplicación usual en \mathbb{N} .
- $<^{\mathfrak{N}}$ es la relación “menor que” usual en \mathbb{N} .

- Similarmente, los reales se pueden representar con la \mathcal{L} -estructura

$$\mathfrak{R} = (\mathbb{R}, 0^{\mathfrak{R}}, 1^{\mathfrak{R}}, s^{\mathfrak{R}}, +^{\mathfrak{R}}, \cdot^{\mathfrak{R}}, <^{\mathfrak{R}}).$$

¿Es verdadera la fórmula $\forall x \exists y (x = y + y)$ sobre \mathfrak{N} ? ¿Sobre \mathfrak{R} ?

¿Es verdadera la fórmula $1 < x$ sobre \mathfrak{N} ?

Definición:

El conjunto de **variables** de un \mathcal{L} -término t , denotado por $V(t)$, se define inductivamente como:

- Si t es un símbolo de constante, entonces $V(t) = \emptyset$.
- Si $t = x$ es una variable, entonces $V(t) = \{x\}$.
- Si $t = f(t_1, \dots, t_n)$ para t_1, \dots, t_n \mathcal{L} -términos, entonces $V(t) = V(t_1) \cup \dots \cup V(t_n)$.

Ejemplo:

$\mathcal{L} = \{s, f, g\}$, donde s es función unaria, f y g son funciones binarias.

$$\begin{aligned} V(f(g(x, y), s(0))) &= V(g(x, y)) \cup V(s(0)) \\ &= V(x) \cup V(y) \cup V(0) \\ &= \{x\} \cup \{y\} \cup \emptyset \\ &= \{x, y\} \end{aligned}$$

Definición:

El conjunto de **variables** para una \mathcal{L} -fórmula φ , denotado por $V(\varphi)$, se define inductivamente como:

- Si $\varphi = t_1 = t_2$, entonces $V(\varphi) = V(t_1) \cup V(t_2)$.
- Si $\varphi = R(t_1, \dots, t_n)$, entonces $V(\varphi) = V(t_1) \cup \dots \cup V(t_n)$.
- Si $\varphi = (\neg\psi)$, entonces $V(\varphi) = V(\psi)$.
- Si $\varphi = (\psi_1 \star \psi_2)$, entonces $V(\varphi) = V(\psi_1) \cup V(\psi_2)$.
- Si $\varphi = (\exists x \psi)$ o $\varphi = (\forall x \psi)$, entonces $V(\varphi) = \{x\} \cup V(\psi)$.

Ejemplo:

$\mathcal{L} = \{s, P, Q\}$, donde s es función unaria, P y Q relaciones unarias.

$$\begin{aligned} V((\exists x P(x)) \vee (\forall y Q(s(y)))) &= V(\exists x P(x)) \cup V(\forall y Q(s(y))) \\ &= \{x\} \cup V(P(x)) \cup \{y\} \cup V(Q(s(y))) \\ &= \{x\} \cup V(x) \cup \{y\} \cup V(s(y)) \\ &= \{x\} \cup \{x\} \cup \{y\} \cup V(y) \\ &= \{x\} \cup \{x\} \cup \{y\} \cup \{y\} \\ &= \{x, y\} \end{aligned}$$

Definición:

El conjunto de **variables libres** de una \mathcal{L} -fórmula φ , denotado por $VL(\varphi)$, se define inductivamente como:

- Si φ es atómica, entonces $VL(\varphi) = V(\varphi)$.
- Si $\varphi = (\neg\psi)$, entonces $VL(\varphi) = VL(\psi)$.
- Si $\varphi = (\psi_1 \star \psi_2)$, entonces $VL(\varphi) = VL(\psi_1) \cup VL(\psi_2)$.
- Si $\varphi = (\exists x \psi)$ o $\varphi = (\forall x \psi)$, entonces $VL(\varphi) = VL(\psi) \setminus \{x\}$.

Las variables libres son las variables que **no están cuantificadas**.

Ejemplos:

Sea $\mathcal{L} = \{P, Q, R\}$ con P y Q relaciones unarias, y R relación binaria.

¿Cuáles son las variables libres de las siguientes fórmulas?

- $\exists x \forall y (R(x, y) \rightarrow Q(y)).$
- $P(x) \wedge \exists y R(x, y).$
- $P(z) \wedge \exists z Q(z).$

Notación:

- Si φ es una \mathcal{L} -fórmula, entonces escribimos $\varphi(x_1, \dots, x_k)$ para indicar que $VL(\varphi) = \{x_1, \dots, x_k\}$.
- Decimos que φ es una **oración** si $VL(\varphi) = \emptyset$.

- ¿Es verdadera la fórmula $1 < x$ sobre \mathfrak{N} ?
- Depende del valor que le asignamos a la variables libre x .
 - Si $x = 10$ es verdadera sobre \mathfrak{N} .
 - Si $x = 1$ es falsa sobre \mathfrak{N} .
- Necesitamos la noción de **asignación**.

Definición:

Dada una \mathcal{L} -estructura \mathfrak{A} con dominio A , una **asignación** es una función $\sigma : \mathcal{V} \rightarrow A$ que asigna a cada variable un valor en A .

Dada asignación σ , definimos la función $\hat{\sigma}$ que da valores a los \mathcal{L} -términos:

- Si $t = c$ es una constante, entonces $\hat{\sigma}(t) = c^{\mathfrak{A}}$.
- Si $t = x$ es una variable, entonces $\hat{\sigma}(t) = \sigma(x)$.
- Si $t = f(t_1, \dots, t_n)$, entonces $\hat{\sigma}(t) = f^{\mathfrak{A}}(\hat{\sigma}(t_1), \dots, \hat{\sigma}(t_n))$.

Ejemplo:

Para $\mathcal{L} = \{0, 1, s, +, \cdot, <\}$ y la estructura \mathfrak{N} , considere una asignación tal que $\sigma(x) = 7$ y evalúe el siguiente término:

$$\hat{\sigma}(s(1) \cdot s(x))$$

Por simplicidad usaremos σ en lugar de $\hat{\sigma}$.

Sean \mathcal{L} un vocabulario, \mathfrak{A} una \mathcal{L} -estructura con dominio A y σ una asignación para \mathfrak{A} .

Definición:

Decimos que (\mathfrak{A}, σ) **satisface** una \mathcal{L} -fórmula φ , denotado como $(\mathfrak{A}, \sigma) \models \varphi$ si:

- $\varphi = t_1 = t_2$ y $\sigma(t_1) = \sigma(t_2)$.
- $\varphi = R(t_1, \dots, t_n)$ y $(\sigma(t_1), \dots, \sigma(t_n)) \in R^{\mathfrak{A}}$.
- $\varphi = (\neg\psi)$ y $(\mathfrak{A}, \sigma) \not\models \psi$.
- $\varphi = (\psi_1 \vee \psi_2)$ y $(\mathfrak{A}, \sigma) \models \psi_1$ o $(\mathfrak{A}, \sigma) \models \psi_2$.

Definición:

- $\varphi = (\psi_1 \wedge \psi_2)$, $(\mathfrak{A}, \sigma) \models \psi_1$ y $(\mathfrak{A}, \sigma) \models \psi_2$.
- $\varphi = (\psi_1 \rightarrow \psi_2)$ y $(\mathfrak{A}, \sigma) \not\models \psi_1$ o $(\mathfrak{A}, \sigma) \models \psi_2$.
- $\varphi = (\psi_1 \leftrightarrow \psi_2)$ y ambos $(\mathfrak{A}, \sigma) \models \psi_1$, $(\mathfrak{A}, \sigma) \models \psi_2$ o ambos $(\mathfrak{A}, \sigma) \not\models \psi_1$, $(\mathfrak{A}, \sigma) \not\models \psi_2$.
- $\varphi = (\exists x \psi)$ y existe $a \in A$ tal que $(\mathfrak{A}, \sigma[x/a]) \models \psi$, donde

$$\sigma[x/a](y) = \begin{cases} a & y = x \\ \sigma(y) & y \neq x \end{cases}$$

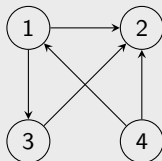
- $\varphi = (\forall x \psi)$ y para todo $a \in A$ se tiene que $(\mathfrak{A}, \sigma[x/a]) \models \psi$.

Si φ es oración, decimos simplemente $\mathfrak{A} \models \varphi$.

Ejemplo:

Sea $\mathcal{L} = \{E\}$ y la \mathcal{L} -estructura $\mathfrak{A} = (A, E^{\mathfrak{A}})$ con

$$\begin{aligned} A &= \{1, 2, 3, 4\} \\ E^{\mathfrak{A}} &= \{(1, 2), (1, 3), (3, 2), (4, 1), (4, 2)\} \end{aligned}$$



¿Cuáles de las siguientes oraciones son ciertas en \mathfrak{A} ?

- $\forall x \neg E(x, x).$
- $\forall x \exists y E(x, y).$
- $\exists x \forall y \neg E(y, x)$
- $\exists x \forall y (y \neq x \rightarrow E(y, x)).$

Ejemplo:

Sea $\mathcal{L} = \{0, 1, s, +, \cdot, <\}$ y considere las estructuras \mathfrak{N} y \mathfrak{R} .

De ejemplos de asignaciones tal que las siguientes fórmulas son verdaderas:

- $\forall y (x = y \wedge x < y).$
- $\exists z (x < z \wedge z < y).$
- $\forall x \forall y (x < y \rightarrow \exists z (x < z \wedge z < y)).$

Definición:

Dos \mathcal{L} -oraciones φ_1, φ_2 son **equivalentes**, denotado por $\varphi_1 \equiv \varphi_2$, si para toda \mathcal{L} -estructura \mathfrak{A} se tiene que

$$\mathfrak{A} \models \varphi_1 \quad \text{si y solo si} \quad \mathfrak{A} \models \varphi_2.$$

Ejemplos:

Equivalencias derivadas de la lógica proposicional:

- $\varphi \wedge \psi \equiv \psi \wedge \varphi$
- $\varphi \wedge (\psi \wedge \theta) \equiv (\varphi \wedge \psi) \wedge \theta$
- $\varphi \wedge \varphi \equiv \varphi$
- $\neg(\neg\varphi) \equiv \varphi$
- $\varphi \wedge (\psi \vee \theta) \equiv (\varphi \wedge \psi) \vee (\varphi \wedge \theta)$
- $\neg(\varphi \wedge \psi) \equiv \neg\varphi \vee \neg\psi$
- $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi$
- ...

Ejemplos:

Más equivalencias:

- $\neg(\exists x \varphi) \equiv \forall x(\neg\varphi)$
- $\neg(\forall x \varphi) \equiv \exists x(\neg\varphi)$
- $\exists x(\varphi \vee \psi) \equiv (\exists x \varphi) \vee (\exists x \psi).$
- $\forall x(\varphi \wedge \psi) \equiv (\forall x \varphi) \wedge (\forall x \psi).$

¿Qué sucede con las siguientes equivalencias?

- $\exists x(\varphi \wedge \psi) \equiv (\exists x \varphi) \wedge (\exists x \psi).$
- $\forall x(\varphi \vee \psi) \equiv (\forall x \varphi) \vee (\forall x \psi).$

Propuesto: demuestre las equivalencias.

Consecuencia lógica

Notación:

Decimos que una \mathcal{L} -estructura \mathfrak{A} satisface un conjunto Σ de oraciones, denotado por $\mathfrak{A} \models \Sigma$, si para cada oración $\varphi \in \Sigma$ se cumple $\mathfrak{A} \models \varphi$.

Definición:

Sean una \mathcal{L} -oración φ y un conjunto de oraciones Σ . Decimos que φ es **consecuencia lógica** de Σ , denotado por $\Sigma \models \varphi$, si para toda \mathcal{L} -estructura \mathfrak{A} se tiene

$$\text{si } \mathfrak{A} \models \Sigma \quad \text{entonces} \quad \mathfrak{A} \models \varphi.$$

Ejercicio:

¿Cuáles de las siguientes consecuencias lógicas se cumplen?

- $\{\exists x(\varphi \wedge \psi)\} \models (\exists x \varphi) \wedge (\exists x \psi)$
- $\{(\exists x \varphi) \wedge (\exists x \psi)\} \models \exists x(\varphi \wedge \psi)$
- $\{\forall x \varphi\} \models \exists x \varphi$
- $\{\forall x(P(x) \rightarrow R(x)), P(c)\} \models R(c)$

Definición:

Una \mathcal{L} -fórmula φ es **satisfacible** si existe una \mathcal{L} -estructura \mathfrak{A} y una asignación σ para \mathfrak{A} tal que $(\mathfrak{A}, \sigma) \models \varphi$.

Si φ es una \mathcal{L} -oración, entonces φ es satisfacible si existe \mathfrak{A} tal que $\mathfrak{A} \models \varphi$.

Definición:

Una \mathcal{L} -fórmula φ es **válida** si para toda \mathcal{L} -estructura \mathfrak{A} y toda asignación σ para \mathfrak{A} se cumple que $(\mathfrak{A}, \sigma) \models \varphi$.

Si φ es una \mathcal{L} -oración, entonces φ es válida si para toda \mathfrak{A} se cumple que $\mathfrak{A} \models \varphi$.

Problemas de decisión para LPO

Problemas de decisión fundamentales para LPO:

SAT = $\{\varphi \mid \varphi \text{ es una oración satisfacible}\}.$

VAL = $\{\varphi \mid \varphi \text{ es una oración válida}\}.$

¿Son decidibles estos problemas?