FUNDAMENTOS DE BASES DE DATOS

oswaldo.solarte@correounivalle.edu.co

Clases

Sesiones: Viernes 3:00 - 6:00 p.m.

oswaldo.solarte@correounivalle.edu.co

Teléfono:3212100 Ext 2784

Objetivos

Objetivo General

Estudiar los fundamentos de los modelos de datos

Objetivos Específicos

Presentar y analizar los modelos de datos vigentes y las bases conceptuales y teóricas subyacentes

Objetivos

Objetivos Específicos

Propiciar una revisión de la literatura especializada con el propósito de analizar los diferentes modelos de datos

Identificar algunas de las nuevas tendencias en bases de datos

Contenido

Recursos

Materiales: campusvirtual.univalle.edu.co

Texto Guía

•Abiteboul S., Hull R., Vianu V. *Foundations of Databases*. Addison-Wesley Publishing Company, 1995

•Millán Martha. *Notas de Referencia para la Asignatura Fundamentos de Bases de Datos.* EISC 2008

Otros libros

- Hellestein J y StoneBraker (2005) *Readings in Data Bases systems* Four Edition, MIT Press
- Ramakrishnan R., Gehrke J., (2003*) DATABASE MANAGEMENT SYSTEMS THIRD EDITION*, McGraw Hill
- Date C.J (2006) Date on Data Base, Apress.
- Garcia-Molina H, Ullman J. , Widom J (2002). *Database Systems The Complete Book* Prentice Hall
- Mordechai Ben-A. (2012) *Mathematical Logical for Computer Science*, Springer Third Edition

Programación Asignatura

Evaluación:

- Reseñas de lecturas (2 artículos) (60%)
- Examen final: 30%
- Exposiciones (10%)

Criterios evaluación artículos

- Revisión de la literatura especializada
- Originalidad
- Coherencia en la argumentación
- Claridad y legibilidad

PRELIMINARES

Nociones Básicas

¿Qué es una Base de Datos?

- Gran cantidad de datos a gestionar
- Colección integrada de datos

¿Qué es in SGBD?

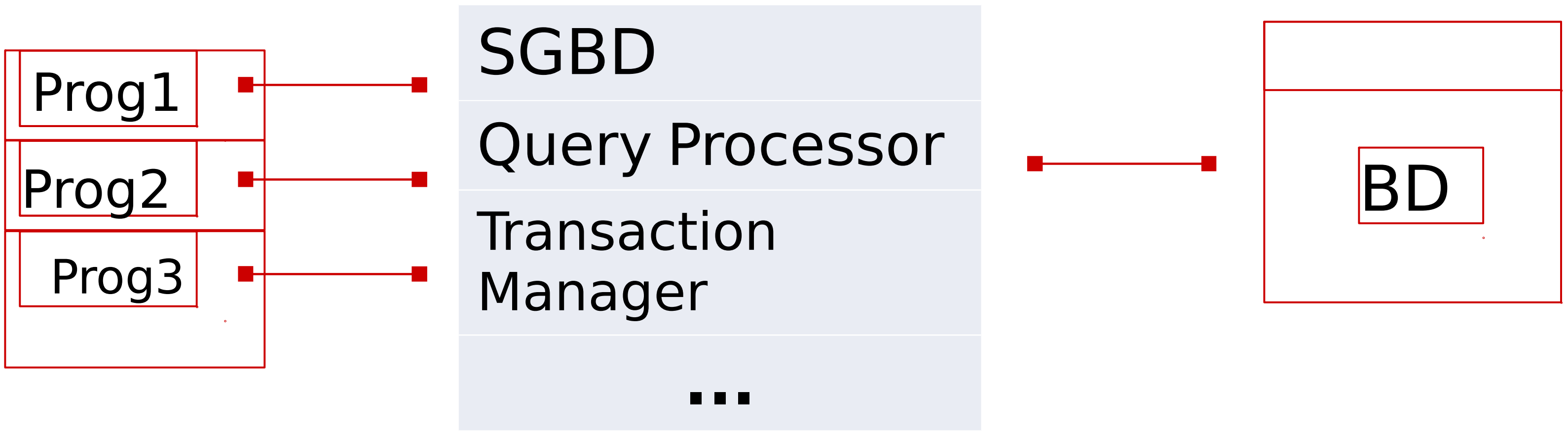
- Software básico que soporta la gestión de los datos

- Sistema de software que ofrece un entorno eficiente para acceder y modificar información almacenada

Principios Básicos

- SGBD es el mediador usuario-dispositivo

- Usuario concentrado en la representación lógica de datos. (Separación lógico-física. Principio de Independencia)



Modelo de datos (DDL, DML)

Independencia de Datos

Capacidad de cambiar esquema de bd sin tener que hacerlo en el siguiente nivel más alto

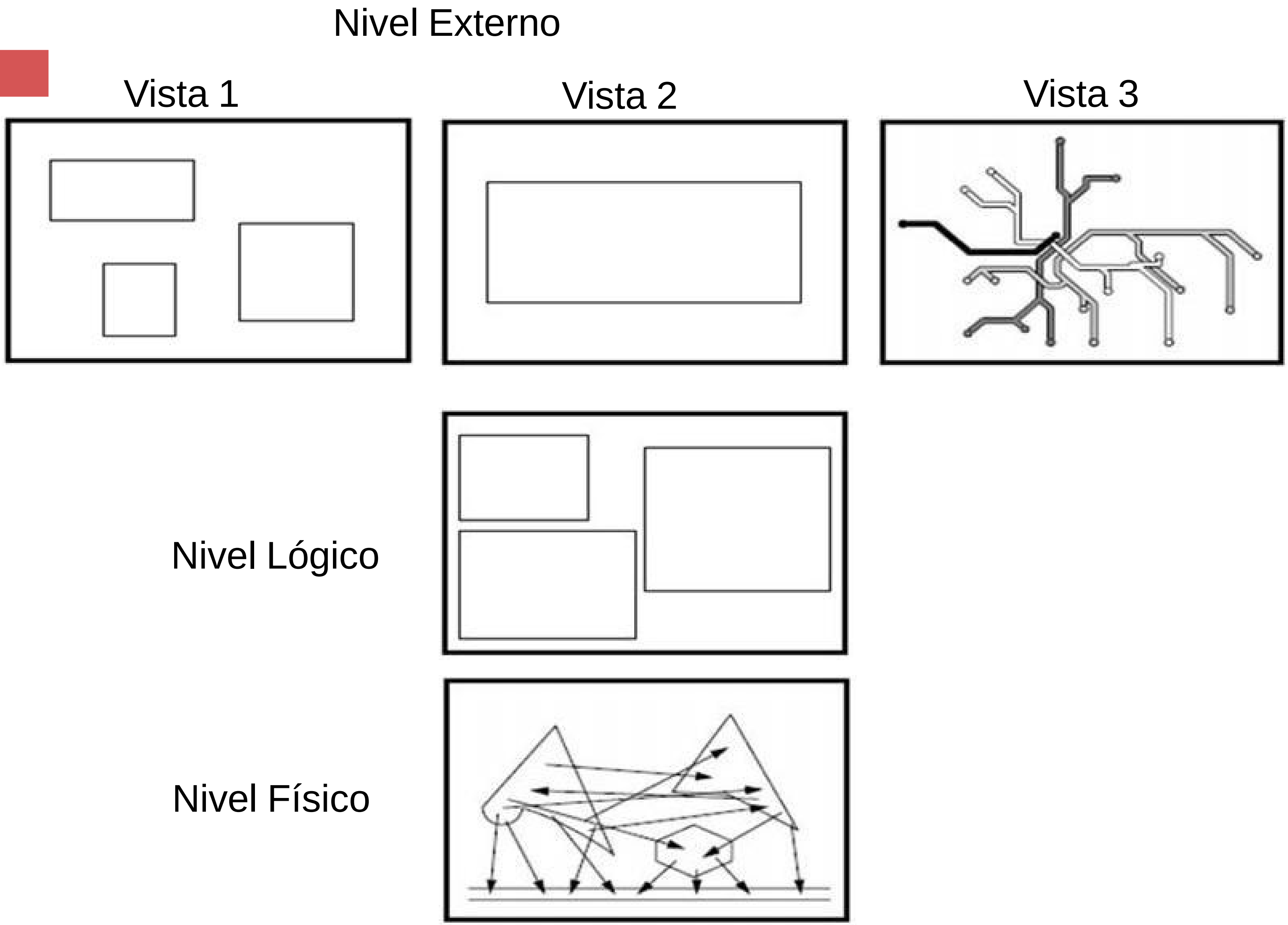
Nivel Externo

Independencia lógica de datos

Nivel Lógico

Independencia física de datos

Nivel Físico



Arquitectura de tres niveles de sistema de base de datos

Tomada de Abiteboul et al. Foundations of Databases

Funcionalidades

- Gestión de almacenamiento secundario
- Persistencia
- Control de concurrencia
- Protección de datos
- Interfaces humano-máquina
- Distribución
- Compilación y optimización

Componentes de un SMBD

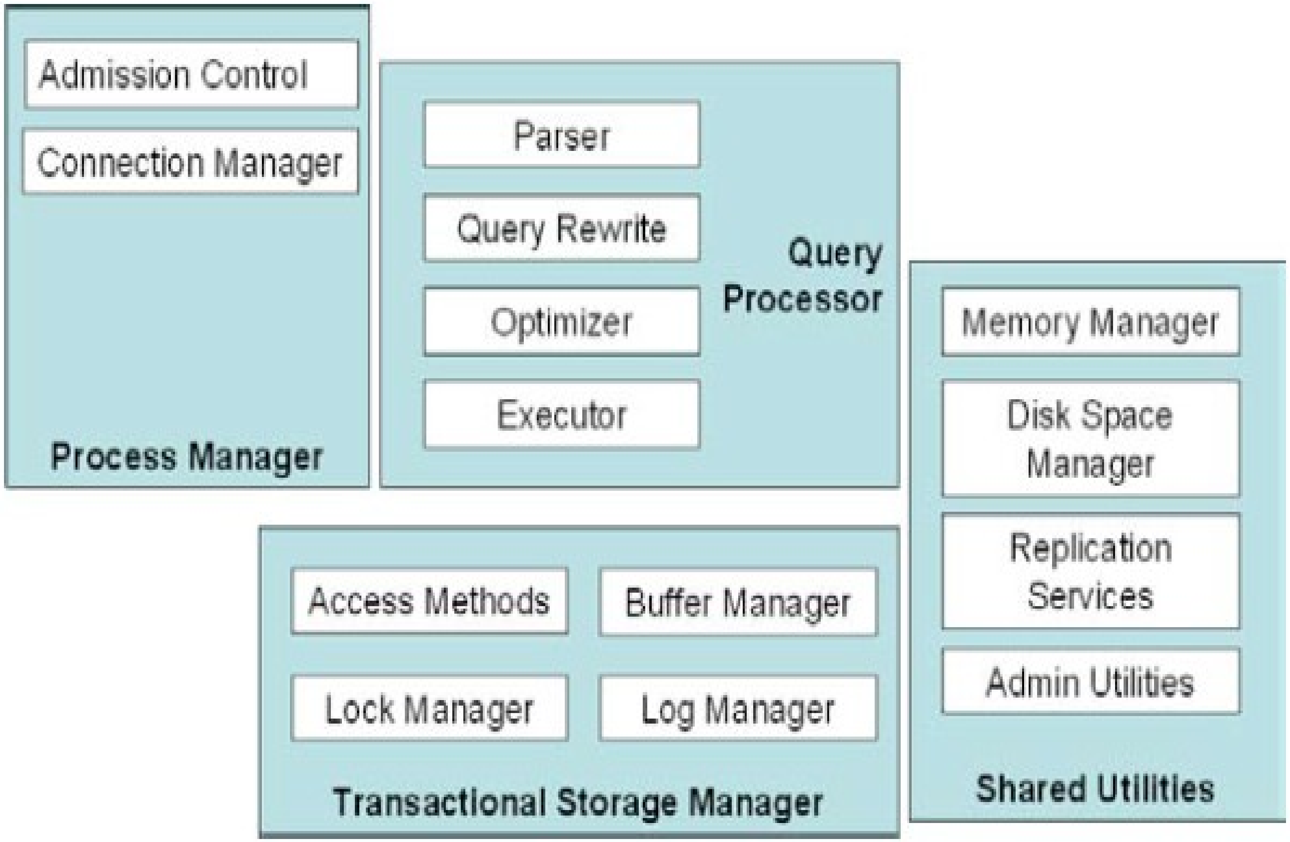


Figure 1: Main Components of a DBMS

Modelo de Datos

Herramienta conceptual para describir:

- Datos
- Relaciones entre datos
- Restricciones de datos

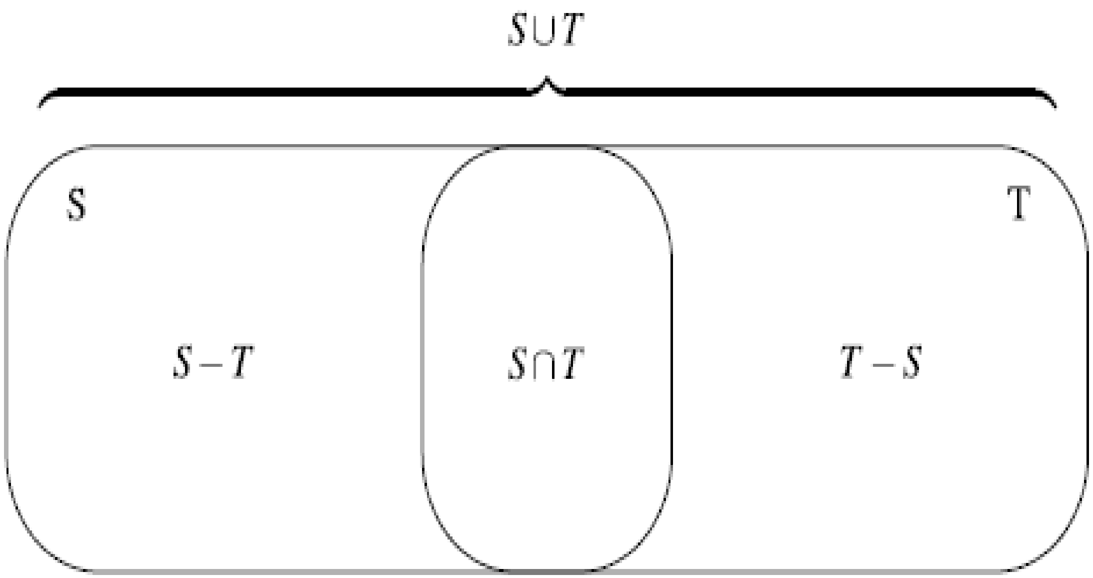
Modelos

- ER
- RED (Conjunto de enlaces)
- Jerárquico (Conjunto de árboles)
- Relacional (Conjunto de tablas)

CONCEPTOS PRELIMINARES

Teoría de conjuntos

Sean T y S dos conjuntos, y algunas operaciones entre estos:



El producto cartesiano $S \times T$ es el conjunto de todos los pares (s,t) tal que $s \in S$ y $t \in T$

Teoría de conjuntos

Relaciones entre conjuntos

Una relación n-aria *R* es un subconjunto de S_1 x ... x S_n.
Una relación unaria es simplemente un subconjunto.

- El conjunto de números primos
- La relaciones <, >

$\mathscr{Q} = \{(n_1, n_2) \mid n_2 = n_1^2\}$ Ej. (4, 16), (7, 49)

Teoría de conjuntos

Relación Binaria

Sobre un conjunto S es un subconjunto R de SxS
R(x,y) *ó* x ***R*** y *denota* $(x,y) \in$ ***R***

$\mathbf{Z}$, $\subseteq$ relación binaria sobre P($\mathbf{Z}$), conjunto de partes de $\mathbf{Z}$

Relación n-aria sobre S; subconjunto de S^n

Relación de Equivalencia

Relación binaria **R** sobre **S** es:

- Reflexiva: Si $(x,x) \in$ **R** para cada $x \in$ **S**
- Simétrica: Si $(x,y) \in$ **R** implica que $(y,x) \in$ **R** para cada $x, y \in$ **S**
- Transitiva: Si $(x,y) \in$ **R** y $(y,z) \in$ **R** implica que $(x,z) \in$ **R** para cada $x, y, z \in$ **S**

R *es una Relación de equivalencia si cumple las propiedades anteriores.*

Partición

- Partición de un conjunto S
- Familia de conjunto $\{S_i \mid i \in \mathbf{I}\}$ tal que:
  - $\bigcup_{i \in \mathbf{I}} S_i = S$
  - $S_i \cap S_j = \varnothing$ para $i \neq j$
  - $S_i \neq \varnothing$ para $i \in \mathbf{I}$

Si R es una relación de equivalencia sobre S entonces la familia de clases de equivalencia sobre R es una partición de S

Preliminares

Relación binaria **R** sobre **S** es:

- Irreflexiva:
  Si $(x,x) \notin$ **R** para cada $x \in$ **S**

- Antisimétrica:
  Si $(y,x) \notin$ **R** siempre que $x \neq y$, $(x,y) \in$ **R**

Lógica Básica

Lógica Proposicional

- Conjunto de variables proposicionales: *p*,*q*,*r*,...
- Constantes proposicionales: true, false

- Estudio de las combinaciones entre proposiciones mediante operadores booleanos.

Lógica Básica

Lógica Proposicional

Fórmula proposicional bien formada construidas con variables y constantes usando conectivos y unarios (¬) y binarios.

$\wedge, \uparrow$

$\vee, \downarrow$

$\rightarrow$

$\leftrightarrow, \oplus$

Operadores con orden de precedencia

Lógica Básica

Asignaciones de verdad
V: Conjunto de variables proposicionales
ξ: V $\rightarrow$ {*true*, *false*}

El valor de verdad $\varphi[\xi]$ de una fórmula proposicional φ bajo la asignación ξ para las variables apareciendo en φ se define inductivamente sobre la estructura de φ

Lógica Básica

Ejemplo

- true $[\xi]$ = *true*
- Si $\varphi = p$ para alguna variable, entonces

$$\varphi[\xi] = \xi(p);$$

- Si $\varphi = (\neg\ \psi)$ entonces $\psi\ [\xi]$=*true* sii

$$\varphi\ [\xi] = \textit{false};$$

- $(\psi 1\ \vee\ \psi\ 2)[\xi]$ = *true* si al menos

$$\psi 1\ [\xi] = \textit{true} \text{ o } \psi\ 2[\xi] = \textit{true};$$

Lógica Básica

Tablas de Verdad

x	y	$\wedge$	$\vee$	$\oplus$	$\rightarrow$	$\leftrightarrow$
T	T	T	T	F	T	T
T	F	F	T	T	F	F
F	T	F	T	T	T	F
F	F	F	F	F	T	T

Fórmulas bien formadas

$$fml \ ::= \ p$$
$$fml \ ::= \ \neg fml$$
$$fml \ ::= \ fml \ op \ fml$$
$$op \ ::= \ \vee \mid \wedge \mid \rightarrow \mid \leftrightarrow \mid \oplus \mid \uparrow \mid \downarrow$$

Lógica Básica

Satisfactibilidad

La fórmula φ es satisfactible si existe al menos una asignación de verdad que la hace verdadera.

Es insatisfactible en otro caso.

La fórmula φ es válida si cada asignación de verdad a las variables de φ la hacen verdadera

Equivalencia lógica

Proposiciones compuestas que tienen los mismos valores de verdad en todos los casos posibles se denominan *lógicamente equivalentes*

$\mathscr{I}(p)$	$\mathscr{I}(q)$	$v_{\mathscr{I}}(p \vee q)$	$v_{\mathscr{I}}(q \vee p)$
T	T	T	T
T	F	T	T
F	T	T	T
F	F	F	F

Equivalencia lógica

$\mathscr{I}(p)$	$\mathscr{I}(q)$	$v_{\mathscr{I}}(p \vee q)$	$v_{\mathscr{I}}(q \vee p)$
T	T	T	T
T	F	T	T
F	T	T	T
F	F	F	F

$A_1 \equiv A_2$ si y solo si $A_1 \leftrightarrow A_2$ es verdadero en cada interpretación

Ejercicios: $p \rightarrow q \equiv \neg q \rightarrow \neg p$

$p \rightarrow q \equiv \neg p \vee q$

$\neg (p \vee q) \equiv \neg p \wedge \neg q$

Verificar si estas fórmulas son lógicamente equivalentes

Lógica Básica

Consecuencia lógica

Una fórmula φ implica lógicamente a la fórmula ψ (ψ es una consecuencia lógica de φ) $\varphi \models \psi$ si para cada asignación de valores de verdad ξ, si $\varphi[\xi] = \mathit{true}$, entonces $\psi[\xi] = \mathit{true}$.

φ y ψ son lógicamente equivalentes ($\varphi \equiv \psi$) si $\varphi \models \psi$ y $\psi \models \varphi$

Consecuencia lógica

Sean $A_1, A_2, A_3, \ldots A_n$ y B Fórmulas.

B es una consecuencia lógica de $A_1, A_2, A_3, \ldots A_n$
Si y solo si, todo modelo de $A_1 \wedge A_2, \wedge A_3 \wedge \ldots A_n$,
es modelo de B.

$A_1 \wedge A_2 \wedge A_3 \wedge A_n \models B$, si y solo si la fórmula

$A_1 \wedge A_2 \wedge A_3 \wedge A_n \rightarrow B$ es Válida

Consecuencia lógica

Sea:

$$F_1 : \quad \big(p \to (q \vee r)\big) \wedge \neg r$$

$$F_2 : \quad p \to q$$

$$F_1 \models F_2 \,?$$

Consecuencia lógica

Sea:

$F_1: \quad (p \to (q \vee r)) \wedge \neg r$

$F_2: \quad p \to q$

$F_1 \models F_2$?

p	q	r	$q \vee r$	$p \to (q \vee r)$	$\neg r$	$p \to (q \vee r)) \wedge \neg r$	$p \to q$
T	T	T	T	T	F	F	T
T	T	F	T	T	T	T	T
T	F	T	T	T	F	F	F
T	F	F	F	F	T	F	F
F	T	T	T	T	F	F	T
F	T	F	T	T	T	T	T
F	F	T	T	T	F	F	T
F	F	F	F	T	T	T	T

Lógica Básica

Una literal es una fórmula de la forma p o ¬p (o *true* o *false*) para alguna variable proporcional p.

Una fórmula proposicional está en (FNC) Forma Normal <u>Conjuntiva</u> si tiene la forma

$$\psi 1 \wedge \psi 2 \wedge \ldots \wedge \psi n$$

donde cada fórmula ψi es una disjunción de literales

Lógica de Primer Orden (LPO)

(First order logic)

Generalización lógica proposicional

- Variables -*símbolos de predicado* rango-relaciones n-arias sobre un subconjunto subyacente
- Variables en LPO varían sobre un conjunto (universo del discurso) y cuantificadores ($\forall$, $\exists$)
- Símbolos de función se incorporan

Lógica de Primer Orden (LPO)

Generalmente, los ***predicados*** se utilizan para describir ciertas propiedades o relaciones entre las personas u objetos.

- Pedro y Juan <u>son hermanos.</u>
- Todos los perros <u>ladran</u>

Las entidades conectadas de esta manera, “Pedro y Juan” , se llaman términos.

Lógica de Primer Orden (LPO)

- Marta es la madre de Luis
- $x + y > 100$

Qué valor de verdad se puede dar a estas sentencias?

Lógica de Primer Orden (LPO)

- Marta es la madre de Luis
- $x + y > 100$

- Qué valor de verdad se puede dar a estas sentencias?

Depende del dominio o Universo de discurso….

El **dominio** es el conjunto de todas personas, ideas, símbolos, estructuras de datos, etc., que afectan el argumento lógico que se está considerando. A los elementos del dominio se les puede llamar individuos.

Lógica de Primer Orden (LPO)

Ejemplos:

James es un futbolista
futbolista(James)

Marta es la madre de Luis
esMadre (Marta, Luis)
M(x, y)

3 > 2
mayorQue(3,2)

Lógica de Primer Orden (LPO)

Ejercicios:

- Todas las personas tienen una madre

- Algunas personas son vegetarianas

- Nadie es perfecto

Lógica de Primer Orden (LPO)

Ejercicios:

Todas las personas tienen una madre

$\forall x \, TieneMadre(x)$

Algunas personas son vegetarianas

$\exists y \;\; esVegetariano(y)$

Nadie es perfecto

$\neg \exists x P(x)$

$\forall x \neg P(x)$

Lógica de Primer Orden (LPO)

Lenguaje de Primer Orden L incluye:

- Un conjunto de símbolos de constantes
- Para cada $n \geq 0$, un conjunto de símbolos de predicados n-arios
- Para cada $n \geq 1$ un conjunto posible de símbolos de función n-aria
- Símbolo de igualdad $\approx$ (predicado binario)
- Constantes true, false

Lógica de Primer Orden (LPO)

Términos de L
Construídos a partir de símbolos constantes, variables y funciones.
Átomo: true,false

Expresion de la forma $R(t_1, \ldots, t_n)$

R: Símbolo de predicado n-ario t_1: términos

Lógica de Primer Orden (LPO)

Gramática para las formulas LPO

$$
\begin{array}{lcl}
argument & ::= & x \\
argument & ::= & a \\
argument\_list & ::= & argument \\
argument\_list & ::= & argument, argument\_list \\
atomic\_formula & ::= & p\,(argument\_list) \\
\\
formula & ::= & atomic\_formula \\
formula & ::= & \neg formula \\
formula & ::= & formula \vee formula \\
formula & ::= & \forall\, x\; formula \\
formula & ::= & \exists\, x\; formula
\end{array}
$$

Lógica de Primer Orden (LPO)

Fórmulas bien formadas sobre L:

- Átomos y conectivos
- Si φ es una fórmula y x una variable
- $\exists x \varphi$
- $\forall x \varphi$ son fórmulas
- $\forall x (0 \leq x), \neg(x \approx S(x))$
- $\neg \exists x (\forall y (y \leq x))$

Un término o fórmula es básico (ground) si no contiene variables

Alcance cuantificador (Variables libre y acotadas)

- Cada ocurrencia de una variable es un átomo libre
- Si φ es $(\psi 1 \vee \psi 2)$ entonces una ocurrencia de la variable x en φ es libre si ésta es libre como ocurrencia de $\psi 1$ o $\psi 2$ (Se extiende a otros conectivos)

Alcance cuantificador (Variables libre y acotadas)

Las variables que aparecen en un cuantificador $(\exists, \forall)$, se consideran ligadas.

Cualquier variable que no está ligada es libre.

Ejemplo: $\forall x\ (P(x) \rightarrow Q(x))$. La variable X está ligada

Ejercicio: Identificar variables libres y ligadas en :

$$\forall z(P(z) \wedge Q(x)) \vee \exists y Q(y).$$

Interpretación en LPO

Forma de dar significado

I= (U,C,P,F)

U: Conjunto no vacío de elementos abstractos, del universo del discurso

C, P y F dan significado a los conjuntos de símbolos de constantes, símbolos de predicado y símbolos de función:

C: Función de símbolos de constantes en U
P: Transforma cada símbolo de predicado n-ario en P en una relación n-aria sobre U (Subconjunto de U)

Interpretación

Sea **I** una interpretación para el lenguaje **L**. Si **C** es un símbolo de constante en **L**.

C$_{\mathbf{I}}$ denota el elemento en universo asociado con **C** por medio de **I**.

Ejemplo: Una interpretación **I**$_{\mathbf{N}}$ de **L**$_{\mathbf{N}}$ (enteros no negativos) donde:

Universo N

El símbolo de constante **O** es 0, El símbolo de predicado binario ≤

Los símbolos de funciones binaria +, × El símbolo de función unaria S (sucesor) $\mathbf{[\,S\,(\,S\,(\,O\,)+O\,)\,]}_{I_N} \approx \mathbf{2}$

Modelo

Una interpretación $\mathbf{I}$ es un modelo de un conjunto ϕ de sentencias si $\mathbf{I}$ satisface cada fórmula en ϕ.

ϕ es satisfactible si tiene un modelo