Instituto de Matemáticas

Estructuras lingüísticas, ejemplificaciones y morfismos

MARCO A. PÉREZ - IMATE

(junto con DAVID I. SPIVAK - MIT)

Noviembre 30, 2015

Introducción y motivación

- Problema general:
  Comunicación de un fenómeno o de un experimento.
  Suele hacerse por medio de artículos de investigación y/o libros (prosa).

- Pregunta:
  ¿Qué otras formas distintas a la prosa existen para representar el conocimiento?

- Idea:
  Considerar las oraciones del idioma inglés como morfismos.

Algo de historia

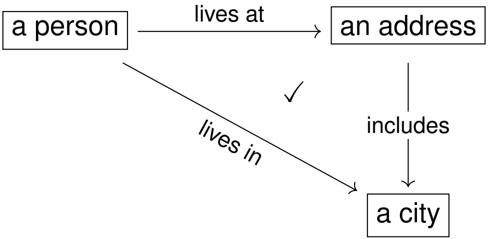
- [2012] Registros ontológicos: categorías y funtores.
  (D. I. Spivak, R. Kent). *Ologs: A categorical framework for knowledge representation*.

- [2015] Formalización de registros ontológicos: bicategorías, funtores laxos y transformaciones laxas.
  (-, D. I. Spivak). *Toward formalizing ologs: linguistic structures, instantiations, and mappings*.

Registros ontológicos: *à la Kent-Spivak*

Definición (informal)

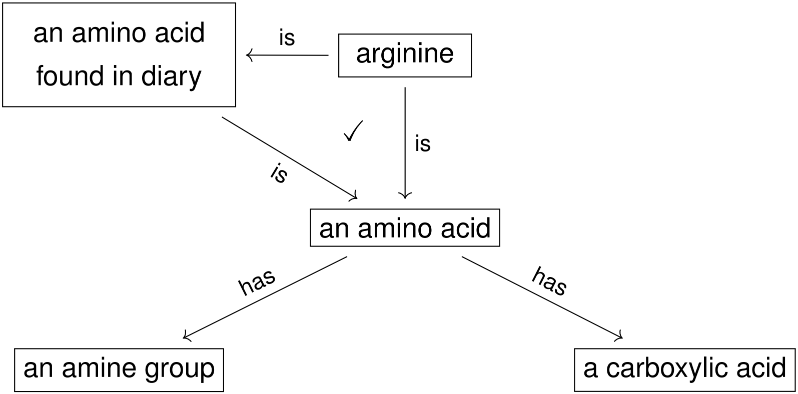
Un **registro ontológico** (también llamado **olog**, de "***ontology log***") es una categoría que modela una situación dada del mundo real.



Registros ontológicos: ventajas

- Dan una formulación precisa de una situación conceptual.

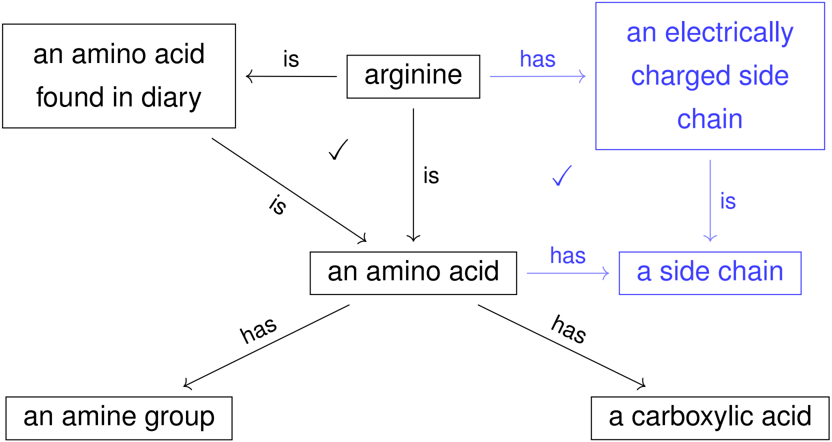
Registros ontológicos: ventajas



Registros ontológicos: ventajas

- Dan una formulación precisa de una situación conceptual.
- De puede extender a medida que se obtiene nueva información.

Registros ontológicos: ventajas



Registros ontológicos: ventajas

- Dan una formulación precisa de una situación conceptual.
- De puede extender a medida que se obtiene nueva información.
- Puede ser referido de forma fácil y precisa por otros autores.
- Se pueden comparar por medio de funtores.
- Se pueden convertir en un esquema de bases de datos, y almacenados en una computadora.

Registros ontológicos: ventajas

$$C : \boxed{\text{a moon}} \xrightarrow{\text{orbits}} \boxed{\text{a planet}}$$

a moon	**orbits a planet, namely**
The Moon	Earth
Phobos	Mars
Deimos	Mars
Ganymede	Jupiter
Titan	Saturn

Ejemplificación: un funtor $C \longrightarrow$ Set.

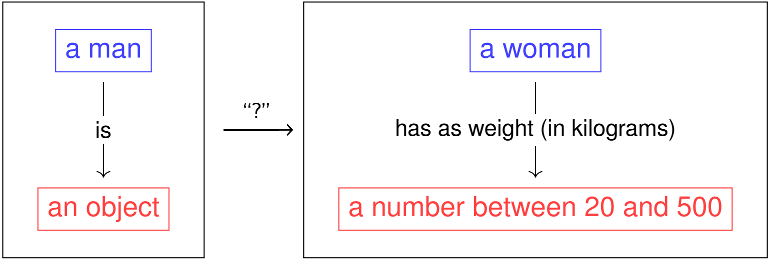
Registros ontológicos: desventajas

- Dada una ejemplificación $C \longrightarrow \mathrm{Set}$, no hay manera de asegurar que lo datos arrojados en Set correspondan con el significado de los conceptos descritos en C.

- Los autores que diseñan un registro ontológico no forman parte de su estructura.

Registros ontológicos: desventajas

- Puede darse el caso que los funtores entre registros no tengan un significado:



Esquema

1. Estructuras lingüísticas:
   Viendo al idioma inglés como una bicategoría.

2. Funtores lingüísticos:
   Morfismos entre registros ontológicos.

3. Ejemplificaciones sobre estructuras lingüísticas:
   Bases de datos.

4. Funtores ejemplificados:
   Preservación de bases de datos.

Esquema

1. **Estructuras lingüísticas:**
   Viendo al idioma inglés como una bicategoría.

2. Funtores lingüísticos:
   Morfismos entre registros ontológicos.

3. Ejemplificaciones sobre estructuras lingüísticas:
   Bases de datos.

4. Funtores ejemplificados:
   Preservación de bases de datos.

Expresiones lingüísticas

Definición

Una **frase nominal** está formada por un artículo indefinido “a/an” seguido por un sustantivo en Inglés, el cual refiere a un concepto para el cual un autor puede dar un conjunto de ejemplos.

Definición

Dadas dos frases nominales $\mathcal{N}_1$ y $\mathcal{N}_2$ y una **frase verbal** $\mathcal{V}$, diremos que $\mathcal{V}$ conecta **funcionalmente** $\mathcal{N}_1$ and $\mathcal{N}_2$ si la concatenación $\langle\!\langle \mathcal{N}_1 \mathcal{V} \mathcal{N}_2 \rangle\!\rangle$ refiere a una oración funcional.

- $$\boxed{\text{a person}} \xrightarrow{\text{has as mother}} \boxed{\text{a woman}}$$

  define una oración funcional.

- $$\boxed{\text{a woman}} \overset{\text{has as dog?}}{\dashrightarrow} \boxed{\text{a canine}}$$

  no es funcional.

Expresiones lingüísticas

Definición

Dos oraciones $\Sigma = (\mathcal{N}_1, \mathcal{V}, \mathcal{N}_2)$ y $\Sigma' = (\mathcal{N}_1, \mathcal{V}', \mathcal{N}_2)$ se dicen **equivalentes**, denotado $\Sigma \simeq \Sigma'$ (o también $\mathcal{V} \simeq \mathcal{V}'$), si refieren a la misma relación funcional.

$$
\begin{tikzcd}
\boxed{\text{an integer}} \arrow[r, bend left, "\text{has as successor}"] \arrow[r, bend right, "\text{yields, by adding 1,}"'] \arrow[r, phantom, "\checkmark"] & \boxed{\text{an integer}}
\end{tikzcd}
$$

Expresiones lingüísticas

Definición

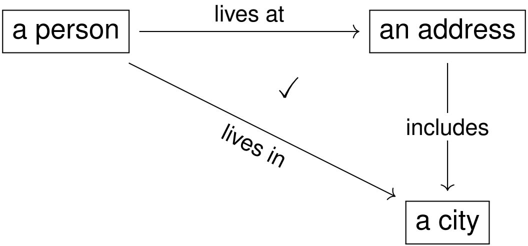
La **concatenación** de $\Sigma_1 = (\mathcal{N}_1, \mathcal{V}_1, \mathcal{N}_2)$ y $\Sigma_2 = (\mathcal{N}_2, \mathcal{V}_2, \mathcal{N}_3)$ se define como la oración

$$\Sigma_1; \Sigma_2 := (\mathcal{N}_1, \mathcal{V}_1; \mathcal{V}_2, \mathcal{N}_3)$$

donde

$$\langle\!\langle \mathcal{V}_1; \mathcal{V}_2 \rangle\!\rangle := \langle\!\langle \mathcal{V}_1 \mathcal{N}_2 \rangle\!\rangle \text{ “which” } \langle\!\langle \mathcal{V}_2 \rangle\!\rangle$$

Expresiones lingüísticas



Este diagrama (mejor conocido como un **hecho**) se lee como sigue:

For any person x, we know that x lives at an address which includes a city y_1, and we know that x lives in a city y_2; and the fact is, y_1 and y_2 are the same for any x.

Expresiones lingüísticas

El diagrama

$$\begin{array}{ccc} \boxed{\text{a person}} & \xrightarrow{\text{lives in}} & \boxed{\text{a city}} \\ \Big\downarrow{\scriptstyle\text{lives at}} & & \Big\downarrow{\scriptstyle\text{has its most affluent residence at}} \\ \boxed{\text{an address}} & \xrightarrow[\text{includes}]{} & \boxed{\text{a house number}} \end{array}$$

no es un hecho.

Autores y el concepto de aprobación

Sea s un autor.

Definición

Sea $\mathcal{N}$ una frase nominal. Diremos que s **aprueba** una frase nominal $\mathcal{N}$, denotado $s \models \mathcal{N}$, si el concepto al cual refiere $\langle\langle \mathcal{N} \rangle\rangle$ es reconocible por s, es decir, si s está de acuerdo que $\mathcal{N}$ refiere a un conjunto para el cual se pueden registrar ejemplos.

Definición

Sea $\Sigma = (\mathcal{N}_1, \mathcal{V}, \mathcal{N}_2)$ una oración tal que $s \models \mathcal{N}_1$ y $s \models \mathcal{N}_2$. Diremos que s **aprueba** $\mathcal{V}$ si s está de acuerdo que $\langle\langle \mathcal{N}_1 \mathcal{V} \mathcal{N}_2 \rangle\rangle$ refiere a una oración funcional, y entiende cómo los ejemplos de $\mathcal{N}_1$ corresponden a ejemplos de $\mathcal{N}_2$ a través de $\mathcal{V}$.

Autores y el concepto de aprobación

Sea s un autor.

Definición

Sean $\Sigma = (\mathcal{N}_1, \mathcal{V}, \mathcal{N}_2)$ y $\Sigma' = (\mathcal{N}_1, \mathcal{V}', \mathcal{N}_2)$ dos oraciones tales que $s \models \Sigma$ y $s \models \Sigma'$. Diremos que s **aprueba** que Σ y Σ' son equivalentes, denotado $s \models [\Sigma \simeq \Sigma']$, si s está de acuerdo que $\mathcal{V}$ y $\mathcal{V}'$ refieren a la misma relación funcional, y que $\mathcal{V}$ y $\mathcal{V}'$ actúan de la misma forma sobre los ejemplos de $\mathcal{N}_1$.

Postulados de aprobación

Postulado

Sean Σ_1 y Σ_2 dos oraciones concatenables. Si $s \models \Sigma_1$ y $s \models \Sigma_2$, entonces $s \models \Sigma_1; \Sigma_2$.

Postulado

Asumiremos que existe una única frase verbal e, que se lee

$$\langle\langle e \rangle\rangle = \text{“is of course”}$$

tal que si $\mathcal{N}$ es una frase nominal y s es un autor con $s \models \mathcal{N}$, entonces $s \models (\mathcal{N}, e, \mathcal{N})$.
A la frase e la llamaremos **frase verbal unidad**.

Postulados de aprobación

Postulado

Para toda oración $\Sigma = (\mathcal{N}_1, \mathcal{V}, \mathcal{N}_2)$ y para todo autor s, si $s \models \Sigma$ entonces

$$s \models [\Sigma; (\mathcal{N}_2, e, \mathcal{N}_2) \simeq \Sigma] \;\; \text{y} \;\; s \models [(\mathcal{N}_1, e, \mathcal{N}_1); \Sigma \simeq \Sigma].$$

Estructura de bicategoría del idioma Inglés

Definición (0-células)

Un **tipo**

$$N = (\mathcal{N}, T)$$

consiste en una frase nominal $\mathcal{N}$ y en un conjunto de autores T tal que $T \models \mathcal{N}$.

Denotamos

$$\mathrm{Auth}(N) := T.$$

Definición (1-células)

Dados dos tipos N_1 y N_2, un **aspecto** de N_1 a N_2, denotado $V\colon N_1 \to N_2$, consiste en un par

$$V = (\mathcal{V}, A)$$

donde $\mathcal{V}$ es una frase verbal, y A es un conjunto de autores tal que $A \subseteq \mathrm{Auth}(N_1) \cap \mathrm{Auth}(N_2)$ y $A \models \mathcal{V}$.
Denotamos

$$\mathrm{Auth}(V) := A.$$

Estructura de bicategoría del idioma Inglés

Definición (2-células)

Dados dos tipos N_1 y N_2, y dos aspectos $V_1, V_2 \colon N_1 \to N_2$, un **hecho** de V_1 a V_2, denotado

$$F \colon V_1 \Rightarrow V_2,$$

está dado por un conjunto de autores $F \subseteq \mathrm{Auth}(V_1) \cap \mathrm{Auth}(V_2)$ tal que $F \models [\mathcal{V}_1 \simeq \mathcal{V}_2]$.

$$\begin{tikzcd} N_1 \arrow[r, bend left, "V_1"] \arrow[r, bend right, "V_2"'] \arrow[r, phantom, "\Downarrow F"] & N_2 \end{tikzcd}$$

Estructura de bicategoría del idioma Inglés

Construcción de **Eng**:

- Denotaremos por Ob(**Eng**) la colección de todos los tipos.

- Para cada par de tipos $N_1, N_2 \in \mathrm{Ob}(\mathbf{Eng})$, existe una categoría $\mathbf{Eng}(N_1, N_2)$, cuyos objetos están dados por aspectos
$$V\colon N_1 \to N_2,$$
y los morfismos entre aspectos $V_1, V_2\colon N_1 \to N_2$ están dados por hechos
$$F\colon V_1 \Rightarrow V_2.$$

- Composiciones verticales:

$$\begin{tikzcd}[column sep=huge]
N_1 \arrow[r, bend left=70, "V_1"{name=A}] \arrow[r, "V_2" description, ""{name=B}] \arrow[r, bend right=70, "V_3"'{name=C}] & N_2
\arrow[from=A, to=B, phantom, "\Downarrow F_1"] \arrow[from=B, to=C, phantom, "\Downarrow F_2"]
\end{tikzcd}
\quad := \quad
\begin{tikzcd}[column sep=huge]
N_1 \arrow[r, bend left=70, "V_1"{name=A}] \arrow[r, bend right=70, "V_3"'{name=C}] \arrow[r, phantom, "\Downarrow F_1 \cap F_2"] & N_2
\end{tikzcd}$$

Postulado

Dadas tres frases verbales $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3 \colon \mathcal{N}_1 \to \mathcal{N}_2$. Si s es un autor tal que $s \models [\mathcal{V}_1 \simeq \mathcal{V}_2]$ y $s \models [\mathcal{V}_2 \simeq \mathcal{V}_3]$, entonces $s \models [\mathcal{V}_1 \simeq \mathcal{V}_3]$.

Estructura de bicategoría del idioma Inglés

- Hecho identidad:

  Para todo aspecto $V\colon N_1 \to N_2$, se tiene un hecho

$$\begin{tikzcd} N_1 \arrow[r, bend left, "V"] \arrow[r, bend right, "V"'] \arrow[r, phantom, "\Downarrow \mathrm{Auth}(V)"] & N_2 \end{tikzcd}$$

Estructura de bicategoría del idioma Inglés

- Composiciones horizontales:

$$\mathbf{Eng}(N_1, N_2) \times \mathbf{Eng}(N_2, N_3) \longrightarrow \mathbf{Eng}(N_1, N_3)$$
$$(V, V') \mapsto V; V' \text{ sobre aspectos,}$$
$$(F, F') \mapsto F \bullet F' \text{ sobre hechos.}$$

donde

$$V; V' := ((\mathcal{V}; \mathcal{V}'), \mathrm{Auth}(V) \cap \mathrm{Auth}(V')),$$

y para $F\colon V_1 \Rightarrow V_2$ y $F'\colon V_1' \Rightarrow V_2'$, $F \bullet F'\colon V_1; V_1' \Rightarrow V_2; V_2'$ es el hecho

$$F \cap F' \models [(\mathcal{V}_1; \mathcal{V}_2) \simeq (\mathcal{V}_1'; \mathcal{V}_2')].$$

Estructura de bicategoría del idioma Inglés

$$
\begin{tikzcd}
N_1 \arrow[r, bend left=50, "V_1"] \arrow[r, bend right=50, "V_1'"'] \arrow[r, phantom, "\Downarrow F"] & N_2 \arrow[r, bend left=50, "V_2"] \arrow[r, bend right=50, "V_2'"'] \arrow[r, phantom, "\Downarrow F'"] & N_3
\end{tikzcd}
\quad := \quad
\begin{tikzcd}
N_1 \arrow[r, bend left=50, "V_1\,;\,V_2"] \arrow[r, bend right=50, "V_1'\,;\,V_2'"'] \arrow[r, phantom, "\Downarrow F \cap F'"] & N_3
\end{tikzcd}
$$

Postulado

Sean $\mathcal{V}_1, \mathcal{V}_1' : \mathcal{N} \to \mathcal{N}'$ y $\mathcal{V}_2, \mathcal{V}_2' : \mathcal{N}' \to \mathcal{N}''$ frases verbales, y s un autor tal que $s \models [\mathcal{V}_1 \simeq \mathcal{V}_1']$ y $s \models [\mathcal{V}_2 \simeq \mathcal{V}_2']$.

Entonces

$$s \models [\mathcal{V}_1 ; \mathcal{V}_2 \simeq \mathcal{V}_1' ; \mathcal{V}_2'].$$

Estructura de bicategoría del idioma Inglés

Proposición (-, D. I. Spivak)

La colección **Eng** de todos los tipos, aspectos y hechos, equipada con las operaciones anteriores, define una bicategoría.

Registros ontológicos (redefinidos)

Definición

Dada una categoría pequeña $\mathcal{C}$. Una **estructura lingüística** sobre $\mathcal{C}$ es un funtor laxo

$$L : \mathcal{C} \longrightarrow \mathbf{Eng}.$$

Un **registro ontológico** es un par $(\mathcal{C}, L)$ donde $\mathcal{C}$ es una categoría pequeña y L es una estructura lingüística sobre $\mathcal{C}$.

Dada una estructura lingüística $L : \mathcal{C} \longrightarrow \mathbf{Eng}$:

- Cada objeto $c \in \mathrm{Ob}(\mathcal{C})$ es enviado a un tipo $L(c) = (\mathcal{L}(c), \mathrm{Auth}(L(c)))$, es decir $\mathrm{Auth}(L(c)) \models \mathcal{L}(c)$.

Registros ontológicos (redefinidos)

Dada una estructura lingüística $L : \mathcal{C} \longrightarrow \mathbf{Eng}$:

- Cada morfismo $f : c \to c'$ en $\mathcal{C}$ es enviado a un aspecto

$$L(f) = (\mathcal{L}(f), \mathrm{Auth}(f)) : L(c) \to L(c')$$

donde $\mathcal{L}(f) : \mathcal{L}(c) \to \mathcal{L}(c')$ es una frase verbal con

$$\mathrm{Auth}(L(f)) \models \mathcal{L}(f)$$

y

$$\mathrm{Auth}(L(f)) \subseteq \mathrm{Auth}(L(c)) \cap \mathrm{Auth}(L(c')).$$

Dada una estructura lingüística $L\colon \mathcal{C} \longrightarrow \mathbf{Eng}$:

- Cada par de morfismos $1 \xrightarrow{f} 2 \xrightarrow{g} 3$ en $\mathcal{C}$, con $h = f; g$, es enviado a una 2-célula (hecho) $F\colon L(f); L(g) \Rightarrow L(h)$, es decir

$$F \models [\mathcal{L}(f); \mathcal{L}(g) \simeq \mathcal{L}(h)].$$

$$\begin{tikzcd}
\mathcal{L}(1) \arrow[r, "\mathcal{L}(f)"] \arrow[dr, "\mathcal{L}(h)"'] & \mathcal{L}(2) \arrow[d, "\mathcal{L}(g)"] \arrow[dl, phantom, "\Swarrow"] \\
 & \mathcal{L}(3)
\end{tikzcd}$$

Registros ontológicos (redefinidos)

Sea C la categoría

$$C = \boxed{\begin{tikzcd} \bullet^1 \arrow[r, "f"] \arrow[dr, "h"'] & \bullet^2 \arrow[d, "g"] \\ & \bullet^3 \end{tikzcd}}$$

Registros ontológicos (redefinidos)

con la estructura lingüística

$$(C, L) = \begin{tikzcd}[column sep=huge, row sep=huge] \overset{1}{\boxed{\text{a person}}} \arrow[r, "\text{has as parents}"] \arrow[dr, "\text{has as mother}"'] & \overset{2}{\boxed{\begin{array}{c}\text{a pair } (w, m) \text{ where } w \text{ is} \\ \text{a woman and } m \text{ is a man}\end{array}}} \arrow[d, "\text{yields as } w"] \arrow[dl, phantom, "\Swarrow"] \\ & \overset{3}{\boxed{\text{a woman}}} \end{tikzcd}$$

Esquema

1. Estructuras lingüísticas:
   Viendo al idioma inglés como una bicategoría.
2. **Funtores lingüísticos:**
   Morfismos entre registros ontológicos.
3. Ejemplificaciones sobre estructuras lingüísticas:
   Bases de datos.
4. Funtores ejemplificados:
   Preservación de bases de datos.

Funtores lingüísticos

Definición

Sean $(\mathcal{C}, L)$ y $(\mathcal{D}, M)$ dos registros ontológicos. Un **funtor lingüístico** entre ellos, denotado

$$(F, F^\sharp)\colon (\mathcal{C}, L) \longrightarrow (\mathcal{D}, M),$$

consiste en un funtor $F\colon \mathcal{C} \longrightarrow \mathcal{D}$ junto con una transformación laxa $F^\sharp\colon L \rightsquigarrow M \circ F$.

$$\begin{tikzcd}
\mathcal{C} \arrow[rr, "F"] \arrow[dr, "L"'] & \arrow[d, phantom, "\overset{F^\sharp}{\rightsquigarrow}"] & \mathcal{D} \arrow[dl, "M"] \\
& \textbf{Eng} &
\end{tikzcd}$$

Funtores lingüísticos

El morfismo $F^\sharp$ incluye:

- Para cada $c \in \mathrm{Ob}(\mathcal{C})$, un aspecto

$$F^\sharp_c \colon L(c) \rightsquigarrow M(Fc),$$

llamado **c-componente** de $F^\sharp$, con un conjunto de autores $\mathrm{Auth}(F^\sharp_c) \models F^\sharp_c$.

- Para cada morfismo $f \colon c \to c'$, se tiene un hecho

$$\begin{tikzcd}
L(c) \arrow[r, rightsquigarrow, "F^\#_c"] \arrow[d, "L(f)"'] \arrow[dr, phantom, "F^\#_f \swarrow"] & M(Fc) \arrow[d, "M(Ff)"] \\
L(c') \arrow[r, rightsquigarrow, "F^\#_{c'}"'] & M(Fc')
\end{tikzcd}$$

Definición

Si $F^\sharp$ es la identidad, en cuyo caso tenemos

$$L = M \circ F,$$

diremos que L es el **pullback de M a lo largo de F**, y denotamos

$$L = F^*(M).$$

Funtores lingüísticos

$(\mathcal{C}, L) :=$

$$
\begin{tikzcd}
\underset{\text{1}}{\boxed{\text{a man}}} \arrow[d, "\text{is}"] \\
\underset{\text{2}}{\boxed{\text{an object}}}
\end{tikzcd}
$$

$(\mathcal{D}, M) :=$

$$
\begin{tikzcd}[column sep=huge, row sep=huge]
\overset{a}{\boxed{\text{a woman}}} \arrow[r, "\text{is}"] \arrow[d, "\substack{\text{has as weight}\\ \text{(in kilograms)}}"'] & \overset{c}{\boxed{\text{an animal}}} \arrow[d, "\substack{\text{has as weight}\\ \text{(in kilograms)}}"] \arrow[dl, phantom, "\swarrow"] \\
\overset{b}{\boxed{\substack{\text{a number}\\ \text{between}\\ \text{20 and 500}}}} \arrow[r, "\text{is}"'] & \overset{d}{\boxed{\text{a number}}}
\end{tikzcd}
$$

Funtores lingüísticos

Sean $F, G\colon \mathcal{C} \longrightarrow \mathcal{D}$ los siguientes funtores:

$$\mathcal{C} \xrightarrow{F} \mathcal{D} \qquad \mathcal{C} \xrightarrow{G} \mathcal{D}$$

$$\begin{array}{ccc} 1 & \longmapsto & c \\ \downarrow & & \downarrow \\ 2 & \longmapsto & d \end{array} \qquad \begin{array}{ccc} 1 & \longmapsto & a \\ \downarrow & & \downarrow \\ 2 & \longmapsto & b \end{array}$$

Pregunta

¿Se pueden extender F y G a funtores lingüísticos?

Funtores lingüísticos

$(\mathcal{C}, F^*(M)) :=$ $\quad$ $(\mathcal{C}, L) :=$ $\quad$ $(\mathcal{C}, G^*(M)) :=$

$F^\sharp$? $\longleftarrow$ $\quad$ $\xrightarrow{G^\sharp \ ?}$

$(\mathcal{C}, F^*(M))$	$(\mathcal{C}, L)$	$(\mathcal{C}, G^*(M))$
1 an animal	1 a man	1 a woman
has as weight (in kilograms)	is	has as weight (in kilograms)
2 a number	2 an object	2 a number between 20 and 120

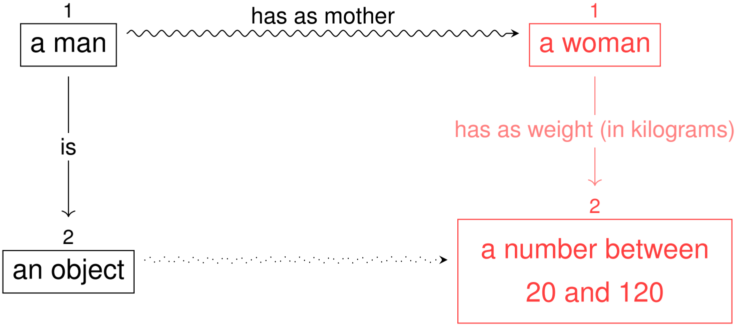
Funtores lingüísticos

<u>Construcción de $F^{\sharp} : L \rightsquigarrow M \circ F$</u>:

$$\begin{tikzcd}[column sep=huge, row sep=huge]
\overset{1}{\boxed{\text{a man}}} \arrow[r, rightsquigarrow, "\text{is}"] \arrow[d, "\text{is}"'] \arrow[dr, phantom, "\Swarrow"] & \overset{1}{\boxed{\text{an animal}}} \arrow[d, "\text{has as weight (in kilograms)}"] \\
\overset{2}{\boxed{\text{an object}}} \arrow[r, rightsquigarrow, "\text{has as weight (in kilograms)}"'] & \overset{2}{\boxed{\text{a number}}}
\end{tikzcd}$$

Funtores lingüísticos

¿Se puede construir $G^\sharp : L \rightsquigarrow M \circ G$?:



Esquema

1. Estructuras lingüísticas:
   Viendo al idioma inglés como una bicategoría.

2. Funtores lingüísticos:
   Morfismos entre registros ontológicos.

3. **Ejemplificaciones sobre estructuras lingüísticas:
   Bases de datos.**

4. Funtores ejemplificados:
   Preservación de bases de datos.

Definición

Sea $\mathcal{N}$ una frase nominal. Diremos que x es un **token** de $\mathcal{N}$ si x es un ejemplo para el cual la el concepto $\langle\!\langle \mathcal{N} \rangle\!\rangle$ aplica, es decir, si la siguiente oración es cierta:

$$\mathcal{N}(x) := x \text{ “is” } \langle\!\langle \mathcal{N} \rangle\!\rangle$$

Dado un autor s, si $s \models \mathcal{N}$ y $\mathcal{N}(x)$ es cierta según s, diremos que s **aprueba** a x como token de $\mathcal{N}$, denotado

$$s \models (x : \mathcal{N}).$$

Definición

Sea $(\mathcal{N}_1, \mathcal{V}, \mathcal{N}_2)$ una oración, x e y tokens de $\mathcal{N}_1$ y $\mathcal{N}_2$, respectivamente. Diremos que x **corresponde** a y a través de $\mathcal{V}$ si la siguiente oración es cierta:

$$\mathcal{V}(x,y) := x \text{ “is” } \langle\!\langle \mathcal{N}_1 \rangle\!\rangle \text{ “which” } \langle\!\langle \mathcal{V}\mathcal{N}_2 \rangle\!\rangle, \text{ “namely” } y$$

Si s es un autor tal que

$$s \models (\mathcal{N}_1, \mathcal{V}, \mathcal{N}_2),\ s \models (x : \mathcal{N}_1),\ s \models (y : \mathcal{N}_2),$$

y $\mathcal{V}(x,y)$ es cierta según s, entonces diremos que s **aprueba** la correspondencia $\mathcal{V}(x,y)$ entre x e y, denotado

$$s \models \mathcal{V}(x,y).$$

El idioma Inglés con ejemplos

$$\boxed{\text{a US president}} \xrightarrow[]{\substack{\text{can, using the usual chronological} \\ \text{ordering of presidents, be assigned}}} \boxed{\text{an integer}}$$

$$s \models (\text{Abraham Lincoln} : \text{a US president}),$$

$$s \models (16 : \text{an integer}).$$

Entonces $s \models \mathcal{V}(\text{Abraham Lincoln}, 16)$ si la siguiente oración es cierta:

> *Abraham Lincoln is a US president which can, using the usual chronological ordering of presidents, be assigned an integer, namely 16.*

Definición

Un **tipo ejemplificado** es un par $\overline{N} = (N, \mathbb{T})$, donde N es un tipo con frase nominal $\mathcal{N}$ y $\mathbb{T}$ es un conjunto tal que para todo $x \in \mathbb{T}$, tenemos

$$\mathrm{Auth}(N) \models (x : \mathcal{N}).$$

Denotamos el conjunto $\mathbb{T}$ por $\mathbb{T}(\overline{N}) := \mathbb{T}$.

El idioma Inglés con ejemplos

Definición

Sean $\overline{N}_1$ y $\overline{N}_2$ dos tipos ejemplificados. Un **aspecto ejemplificado** de $\overline{N}_1$ a $\overline{N}_2$, denotado

$$\overline{V} \colon \overline{N}_1 \to \overline{N}_2,$$

consiste en un par (V, f), donde $V \colon N_1 \to N_2$ es un aspeto, y f es una función $f \colon \mathbb{T}(\overline{N}_1) \to \mathbb{T}(\overline{N}_2)$, tal que para todo token $x \in \mathbb{T}(\overline{N}_1)$, se tiene

$$\mathrm{Auth}(V) \models \mathcal{V}(x, f(x)).$$

Denotaremos la función f por $\mathbb{T}(\overline{V}) := f$.

El idioma Inglés con ejemplos

Definición

Dados dos aspectos ejemplificados $\overline{V}_1, \overline{V}_2 \colon \overline{N}_1 \to \overline{N}_2$ tales que $\mathbb{T}(\overline{V}_1) = \mathbb{T}(\overline{V}_2)$. Un **hecho ejemplificado** de $\overline{V}_1$ a $\overline{V}_2$, denotado

$$\overline{F} \colon \overline{V}_1 \Rightarrow \overline{V}_2,$$

consiste en un hecho $\overline{F} = F$, donde $F \models [\mathcal{V}_1 \simeq \mathcal{V}_2]$.

El idioma Inglés con ejemplos

Proposición (-, D. I. Spivak)

La colección **InstEng** de todos los tipos, aspectos y hechos ejemplificados define una bicategoría.

Proposición (-, D. I. Spivak)

(a) Las aplicaciones

$$\overline{N} \mapsto N, \text{ para todo tipo ejemplificado } \overline{N},$$
$$\overline{V} \mapsto V, \text{ para todo aspecto ejemplificado } \overline{V},$$

definen un **funtor estricto** entre bicategorías

$$U\colon \textbf{InstEng} \longrightarrow \textbf{Eng}.$$

El idioma Inglés con ejemplos

Proposición (-, D. I. Spivak)

La colección **InstEng** de todos los tipos, aspectos y hechos ejemplificados define una bicategoría.

Proposición (-, D. I. Spivak)

(b) Las aplicaciones

$$\overline{N} \mapsto \mathbb{T}(\overline{N}), \text{ para todo tipo ejemplificado } \overline{N},$$
$$\overline{V} \mapsto \mathbb{T}(\overline{V}), \text{ para todo aspecto ejemplificado } \overline{V},$$

definen un **funtor estricto** entre bicategorías

$$\mathbb{T} \colon \mathbf{InstEng} \longrightarrow \mathbf{Set}.$$

El idioma Inglés con ejemplos

Proposición (-, D. I. Spivak)

La colección **InstEng** de todos los tipos, aspectos y hechos ejemplificados define una bicategoría.

Proposición (-, D. I. Spivak)

(c) Juntos, estos funtores definen una inclusión

$$(U, \mathbb{T}) \colon \mathbf{InstEng} \hookrightarrow \mathbf{Eng} \times \mathbf{Set}.$$

Estructuras lingüísticas ejemplificadas

Definición

Sea $\mathcal{C}$ una categoría pequeña. Una **estructura lingüística ejemplificada** sobre $\mathcal{C}$ es un funtor laxo

$$\overline{L} : \mathcal{C} \longrightarrow \textbf{InstEng}.$$

Un **registro ontológico ejemplificado** es un par $(\mathcal{C}, \overline{L})$ donde $\mathcal{C}$ es una categoría pequeña y $\overline{L} : \mathcal{C} \longrightarrow \textbf{InstEng}$ es una estructura lingüística ejemplificada sobre $\mathcal{C}$.

Estructuras lingüísticas ejemplificadas

Definición

Dada una estructura lingüística $L : \mathcal{C} \longrightarrow \mathbf{Eng}$ y un funtor $I \colon \mathcal{C} \longrightarrow \mathbf{Set}$, diremos que I **se ajusta** a L si el par $(L, I) \colon \mathcal{C} \longrightarrow \mathbf{Eng} \times \mathbf{Set}$ se factoriza a través de **InstEng**.

$$\begin{tikzcd}
\mathcal{C} \arrow[rr, "(L,I)"] \arrow[dr, "\overline{L}"'] & & \mathbf{Eng} \times \mathbf{Set} \\
& \mathbf{InstEng} \arrow[ur, hook, "(U,\mathbb{T})"'] &
\end{tikzcd}$$

Al funtor $\overline{L}$ se le denomina **ejemplificación** de L.

Observación

Toda estructura lingüística ejemplificada $\overline{L}: \mathcal{C} \longrightarrow \textbf{InstEng}$ está dada por un par (L, I), donde $L: \mathcal{C} \longrightarrow \textbf{Eng}$ es una estructura lingüística y $I: \mathcal{C} \longrightarrow \textbf{Set}$ es un funtor que se ajusta a L.

A saber:

$$L = \mathcal{C} \xrightarrow{\overline{L}} \textbf{InstEng} \xrightarrow{U} \textbf{Eng},$$

$$I = \mathcal{C} \xrightarrow{\overline{L}} \textbf{InstEng} \xrightarrow{\mathbb{T}} \textbf{Set}.$$

Estructuras lingüísticas ejemplificadas

Dado el registro ontológico $(\mathcal{C}, L)$:

$$\boxed{\text{a person}} \xrightarrow{\text{has}} \boxed{\text{a father}}$$

Consideremos un funtor $I\colon \mathcal{C} \longrightarrow \mathbf{Set}$ con los siguientes datos registrados:

a person	**has a father, namely**
George W. Bush	George H. W. Bush
Jeb Bush	George H. W. Bush
Emmy Noether	Max Noether

a father
George H. W. Bush
Max Noether
Bill Clinton

Estructuras lingüísticas ejemplificadas

Para que (L, I) sea válido como estructura lingüístuca ejemplificada, los autores deben aprobar los tokens anteriores y las siguientes oraciones:

- "George W. Bush is a person, which has a father, namely George H. W. Bush".
- "Jeb Bush is a person, which has a father, namely George H. W. Bush".
- "Emmy Noether is a person, which has a father, namely Max Noether".

Esquema

1. Estructuras lingüísticas:
   Viendo al idioma inglés como una bicategoría.

2. Funtores lingüísticos:
   Morfismos entre registros ontológicos.

3. Ejemplificaciones sobre estructuras lingüísticas:
   Bases de datos.

4. **Funtores ejemplificados:**
   Preservación de bases de datos.

Funtores lingüísticos ejemplificados

Definición

Dados dos registros ontológicos ejemplificados $(\mathcal{C}, \overline{L})$ y $(\mathcal{D}, \overline{M})$. Un **funtor lingüístico ejemplificado** entre ellos, denotado

$$(F, F^{\natural}) \colon (\mathcal{C}, \overline{L}) \longrightarrow (\mathcal{D}, \overline{M})$$

consiste en un funtor $F \colon \mathcal{C} \longrightarrow \mathcal{D}$ y en una transformación laxa $F^{\natural} \colon \overline{L} \rightsquigarrow \overline{M} \circ F$.

$$\begin{tikzcd}
\mathcal{C} \arrow[rr, "F"] \arrow[dr, "\overline{L}"'] & \arrow[d, phantom, "\overset{F^{\natural}}{\rightsquigarrow}"] & \mathcal{D} \arrow[dl, "\overline{M}"] \\
& \textbf{InstEng} &
\end{tikzcd}$$

Funtores lingüísticos ejemplificados

Observación

Sean $(\mathcal{C}, \overline{L})$ y $(\mathcal{D}, \overline{M})$ dos registros ontológicos ejemplificados, donde

$$\overline{L} = (L, I) \colon \mathcal{C} \longrightarrow \mathbf{Eng} \times \mathbf{Set},$$
$$\overline{M} = (M, J) \colon \mathcal{D} \longrightarrow \mathbf{Eng} \times \mathbf{Set}.$$

Un funtor lingüístico ejemplificado $(\mathcal{C}, \overline{L}) \longrightarrow (\mathcal{D}, \overline{M})$ está dado por:

- Un funtor $F \colon \mathcal{C} \longrightarrow \mathcal{D}$.
- Una transformación laxa $F^\sharp \colon L \rightsquigarrow M \circ F$.
- Una transformación natural $F^\flat \colon I \Rightarrow J \circ F$ que **se ajusta** a $F^\sharp$.

Observación

Esto último quiere decir que la transformación laxa

$$(F^\sharp, F^\flat)\colon \mathcal{C} \longrightarrow \mathbf{Eng} \times \mathbf{Set}$$

se factoriza a través de la inclusión **InstEng** $\hookrightarrow$ **Eng** $\times$ **Set**.

Explícitamente: Para cada $c \in \mathrm{Ob}(\mathcal{C})$ se tienen un aspecto y una función

$$F_c^\sharp\colon L(c) \rightsquigarrow M(Fc) \;\text{ y }\; F_c^\flat\colon I(c) \to J(Fc)$$

tales que

$$s \models F_c^\sharp(x, F_c^\flat(x)) \;\; \forall x \in I(c) \;\text{ y }\; \forall s \in \mathrm{Auth}(F_c^\sharp(x)).$$

Funtores lingüísticos ejemplificados

$$
\begin{tikzcd}[column sep=huge, row sep=large]
 & \mathbf{Eng} \\
\mathcal{C} \arrow[ur, bend left=40, "L"] \arrow[ur, bend right=20, "M \circ F"'] \arrow[ur, phantom, "F^{\sharp} \Rightarrow"] \arrow[dr, bend left=20, "I"] \arrow[dr, bend right=40, "J \circ F"'] \arrow[dr, phantom, "F^{\flat} \Leftarrow"] & \\
 & \mathbf{Set}
\end{tikzcd}
$$

Sean $I\colon \mathcal{C} \longrightarrow \mathbf{Set}$ y $J\colon \mathcal{D} \longrightarrow \mathbf{Set}$ dos ejemplificaciones (también llamadas **esquemas de bases de datos**).

Supongamos que estas dos bases de datos se quieren unir por medio de un funtor $F\colon \mathcal{C} \longrightarrow \mathcal{D}$, es decir, se quiere hallar

$$F^{\flat}\colon I \Rightarrow J \circ F.$$

Observación

Esta tarea es más fácil si se cuenta con una transformación laxa $F^{\sharp}$ para la cual se ajuste $F^{\flat}$.

Funtores lingüísticos ejemplificados

Supongamos

$(\mathcal{C}, L) = \boxed{\begin{array}{c} 1 \\ \boxed{\text{a human}} \end{array}} \xrightarrow{F} (\mathcal{D}, M) = \boxed{\begin{array}{c} a \\ \boxed{\text{a person}} \end{array}}$

$I(1) :=$

a human
Emmy Noether
George W. Bush

$J(a) :=$

a person
Emmy Noether
Max Noether
Bill Clinton
George H. W. Bush
George W. Bush

Funtores lingüísticos ejemplificados

Existen $5^2 = 25$ formas de definir $I \Rightarrow J \circ F$.

Supongamos que tenemos las siguientes opciones para $F^\sharp$:

$$F^\sharp = \boxed{\text{a human}} \xrightarrow{\text{is}} \boxed{\text{a person}}$$

$$F^* = \boxed{\text{a human}} \xrightarrow{\text{has as father}} \boxed{\text{a person}}$$

Así se tiene:

$(F^{\sharp}, F^{\flat}) :=$

a human	is a person, namely
Emmy Noether	Emmy Noether
George W. Bush	George W. Bush

$(F^{*}, F^{!}) :=$

a human	has as father a person, namely
Emmy Noether	Max Noether
George W. Bush	George H. W. Bush

¡Muchas gracias por su atención!

$\boxed{\text{a paper}} \xleftarrow{\text{is based on}} \boxed{\text{a contributed paper talk with slides}} \xrightarrow{\text{is stored on}} \boxed{\text{speaker website}}$

a contributed paper talk with slides	**is stored on speaker website, namely**
ENJIM2015marco.pdf	maperez.info

a contributed paper talk with slides	**is based on a paper, namely**
ENJIM2015marco.pdf	arxiv.org/pdf/1503.08326v2.pdf