



Universidad Veracruzana

Seminario de semántica

Doctorado en estudios del
lenguaje y lingüística aplicada

César Aguilar

Martes 9 de abril de 2024

Síntesis de la clase anterior (1)

En nuestra última clase vimos una panorámica general sobre lo que es la semántica léxica, desde sus conceptos básicos hasta el tipo de problemas que estudia.

Sus resultados tienen un impacto muy importante en muchas áreas de investigación. P. e., hacer que un sistema inteligente infiera una relación de antonimia entre lexemas como:

King + (Woman - Man) = ??

King + (Woman - Man) = Queen

Síntesis de la clase anterior (2)

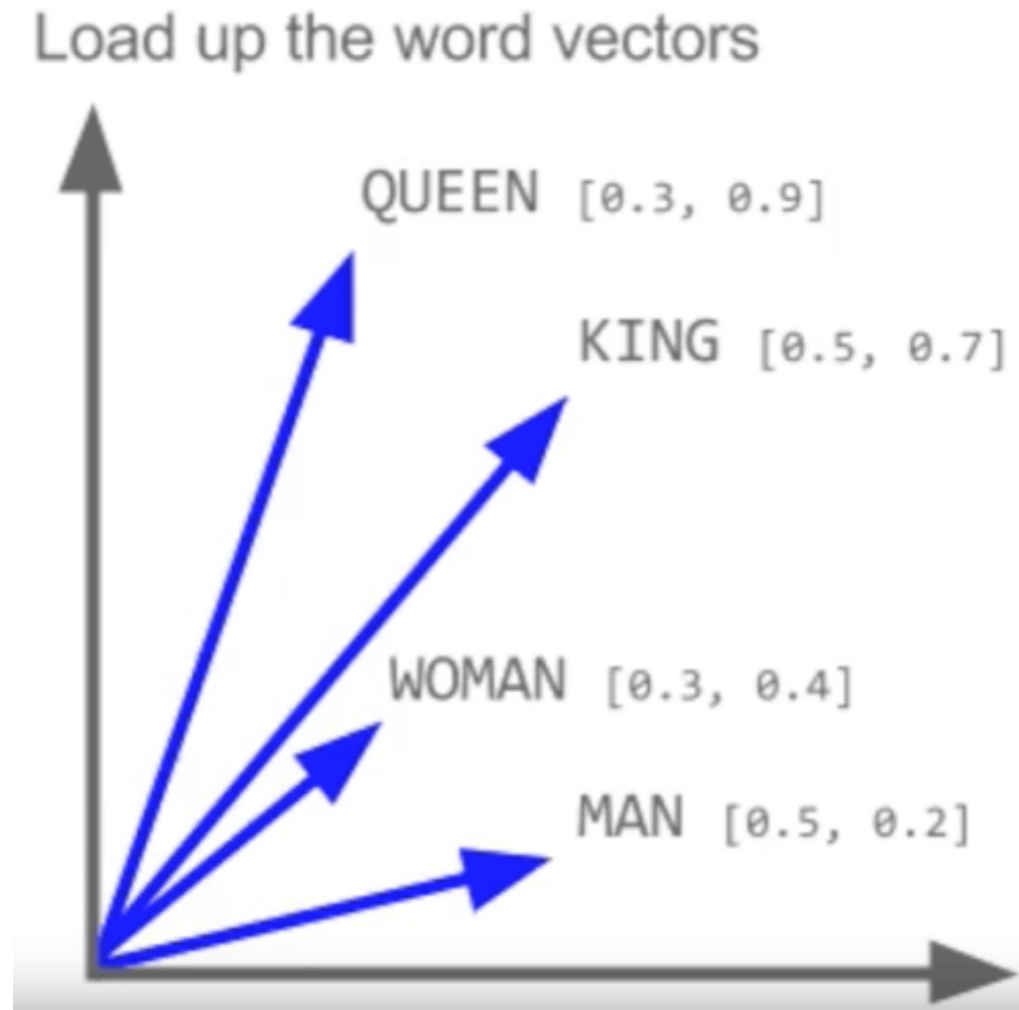
Si piensan un poco en la relación que hay entre *King*, *Man*, *Woman* y *Queen*, podemos deducir que estos lexemas comparten rasgos de significado entre sí (y eso les permite ser antónimos). Esto se ve mejor si lo graficamos de la siguiente manera:



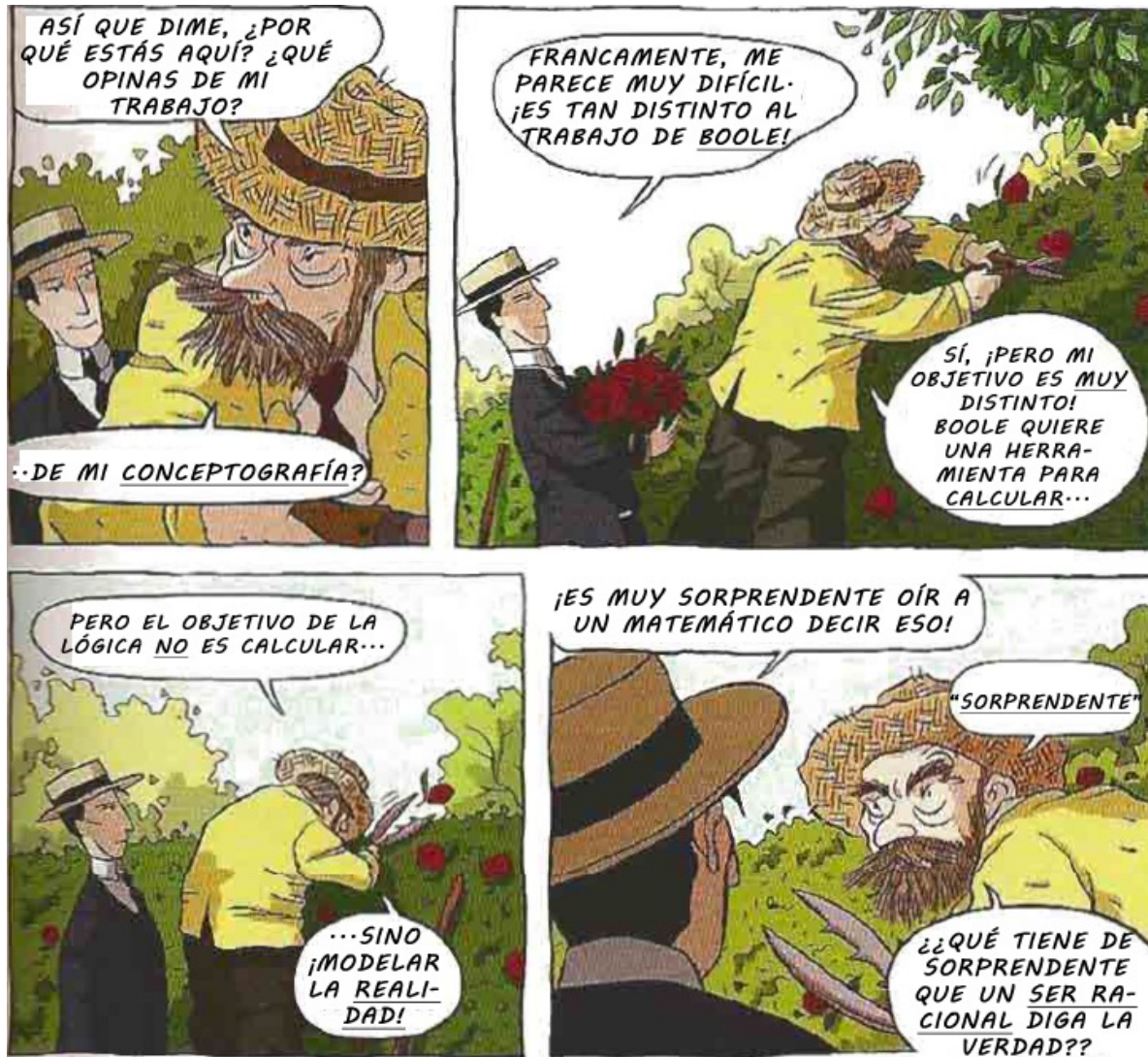
Síntesis de la clase anterior (3)

Ahora bien, si comparan los vectores que aparecen en el cuadro, intuitivamente, ¿qué lexemas están más cercanos, y cuáles son los más alejados?

¿Qué otros lexemas pueden añadir a esta relación, pensando en una red léxica?



The Turn of the Screw (1)



Pasemos al siguiente tema del curso, que consiste en analizar el significado de frases y oraciones, lo que se conoce como **semántica composicional**. Como saben, esto es otro resultado de la discusión entre **Gotlob Frege** y **Bertrand Russell**.

The Turn of the Screw (2)

Y para entrar de lleno en este tema, recordemos nuestra definición sobre **composicionalidad semántica**:

El significado de una construcción compleja es el resultado de combinar el significado de cada una de sus partes, dependiendo siempre de su orden sintáctico.

Esta definición se la debemos a una de nuestras semantistas favoritas, la **Dra. Barbara Partee** (1940). Si les interesa, su sitio WEB es:

<http://people.umass.edu/partee/>



The Turn of the Screw (3)

Ahora bien, ¿cómo identificamos estos significados básicos útiles para hacer estas combinaciones?

Partamos de lo siguiente: todas las palabras tienen asociado un significado específico reconocible. Así, como lo hemos comentado antes, este significado es registrado en los diccionarios:

LONGMAN Dictionary of Contemporary English

Longman **DICTIONARY** **ACTIVATOR** **EXERCISES**

computer noun

computer *noun* **W1 S1**

Menu | Usage note | Word origin | Verb form | Word set

com-put-er /kəm'pjutə \$ -ər/ [countable]

an electronic machine that stores information and uses programs to help you find, organize, or change the information

- A message flashed up on my **computer screen**.
- recent problems with hospital **computer systems**

on computer


- The information is stored on computer.

by computer

- Shoppers can send in their orders by computer and pick up their goods later.

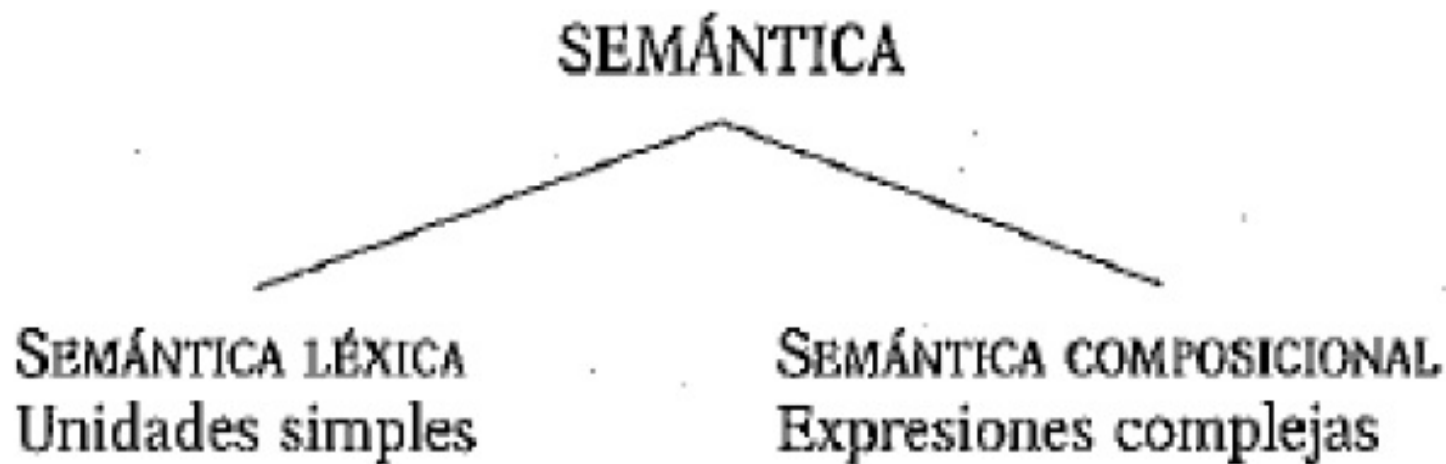
computer software/program/application

- the latest computer software
- **computer games** (=games that you play on a computer) for children
- the use of **computer graphics** (=pictures and images created by computers) in film
- the development of **computer technology**
- the **computer industry**
- a huge global **computer network**



The Turn of the Screw (4)

Para facilitarnos la vida, vamos a considerar algo que plantea nuestra otra semantista favorita, Escandell Vidal, al describir dos líneas de investigación para la semántica: una orientada a analizar el comportamiento de las palabras, y otra para analizar el comportamiento de las oraciones. Esto es:



Semántica y lógica (1)

Ahora abordaremos una línea de investigación con mayor solidez en los estudios de semántica: la representación formal del significado, y su posterior análisis en términos lógicos.

Desde sus inicios, la lingüística ha tenido una estrecha relación con la lógica, ya que ambas se interesan por problemas que son comunes a ambas áreas, entre ellas la codificación y representación de significado en las lenguas humanas.

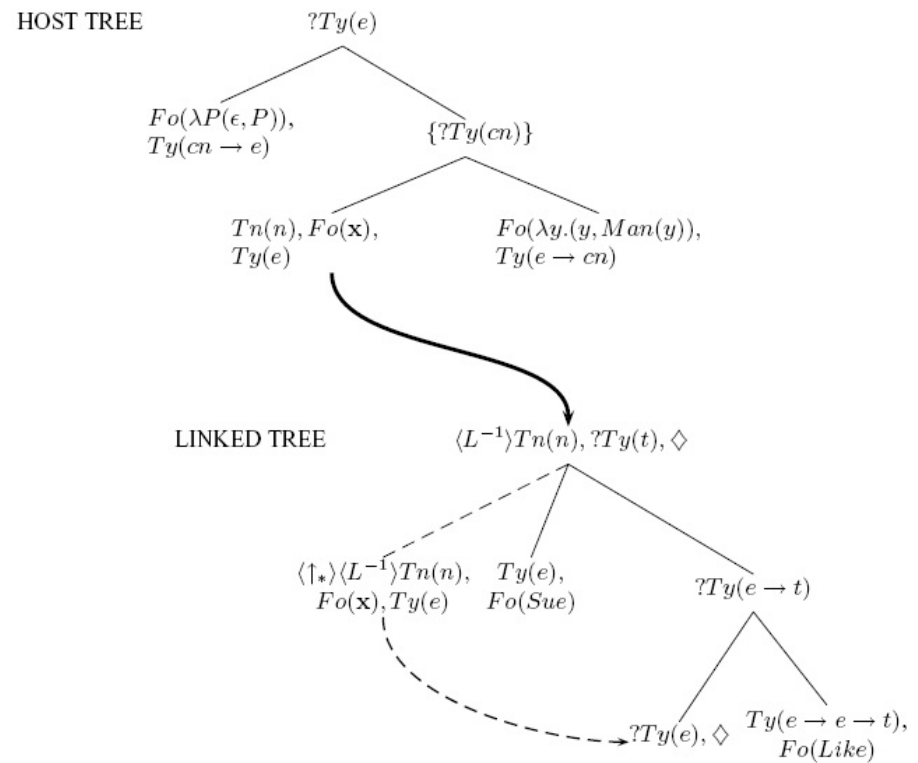


Figure 2.8. Parsing A man who

Semántica y lógica (2)

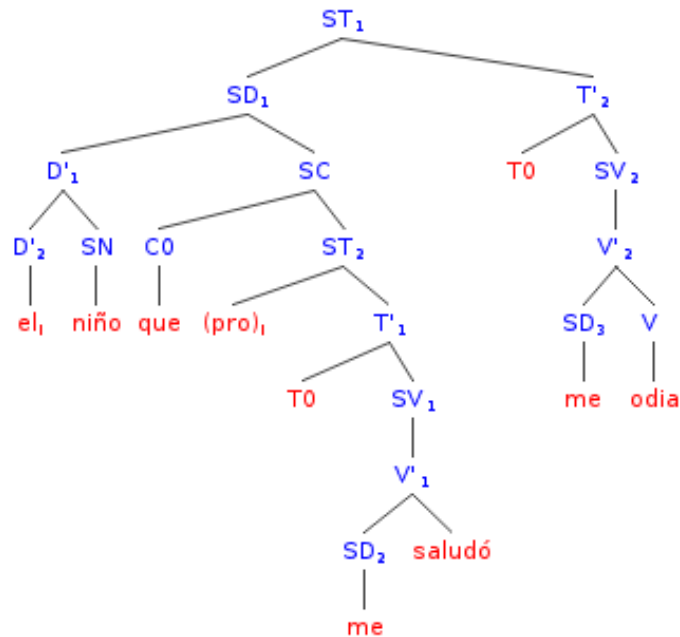
En el caso de esta clase, nos vamos a concentrar en un fenómeno específico, la estructuración de proposiciones como representaciones del significado de una oración.

Vamos a ver con detalle qué quiere decir esto, tomando como guía en nuestro estudio las ideas que desarrolla Escandell Vidal en su libro sobre semántica composicional.



Semántica y lógica (3)

Para empezar con nuestro tema, tenemos que hacer una distinción importante entre dos objetos de análisis: **oraciones** y **proposiciones**.



Una **oración** la podemos entender como una estructura sintáctica, la cual se genera cuando aplicamos un conjunto de reglas gramaticales a una secuencia de palabras (o lexemas).

La idea aquí es que estas reglas nos ayuden lo que queremos expresar en una situación dada.

Semántica y lógica (4)

En contraste, una proposición, de acuerdo con Escandell, es la representación de un estado de cosas (esto es, un evento cualquiera), las cuales podemos contrastar y validar con un referente en el mundo, determinando si lo que expresan es verdadero o falso respecto al referente mencionado, p. e.:

"No tengo hambre" = $\neg c$

"Mañana iré de excursión y llevaré una rica tortilla de patata" = $d \wedge e$

"O vienes conmigo o te quedas en casa" = $f \vee g$

"Vemos una película o, quizá, salimos más tarde" = $h \vee i$

"Si llegas pronto, haces tú la comida" = $j \rightarrow k$

"Si y solo si apruebo el Bachillerato visitaré Nueva York" = $p \leftrightarrow m$

Proposiciones declarativas

Usualmente, cuando pensamos en esta diferencia entre oraciones y proposiciones, tendemos a pensar que las únicas estructuras sintácticas que le interesan a la lógica son las de tipo declarativo, esto es, aquellas estructuras que hacen una descripción concreta de evento. El ejemplo que nos da Escandell es:

Oración:



María es traductora de inglés

Proposición:



María es traductora de inglés

Construyendo un lenguaje formal (1)

Una de las razones por las cuales tenemos esta creencia es porque la lógica trabaja a partir de lenguajes que tienen que ser precisos en lo que expresan. La idea entonces es crear un lenguaje formal que se caracterice por:

1. No generar, hasta donde sea posible, construcciones ambiguas, p. e.: Mi hija mira a los niños del parque con su pelota; el hombre bajo toca el bajo bajo las escaleras
2. Revisar qué tan consistentes son las proposiciones que se construyen con sus reglas. A veces se les denomina como gramáticas de juguete (*Toy Grammars*), ya que la idea es que sean fáciles de manejar.
3. El hecho de que sean fáciles de manejar no significa que no ocurran en la realidad. Más bien quiere decir que, con miras a explicar un fenómeno, primero empezamos por lo más sencillo, y luego llegamos a lo más complejo. Usamos entonces un método deductivo.

Construyendo un lenguaje formal (2)

Si vamos a trabajar con una gramática de juguete, tenemos que razonar lo siguiente:

Podemos tener enunciados básicos que den origen a enunciados complejos:

Gerardo es mi tutor doctoral; estudió lingüística computacional =

(1) Gerardo es mi tutor + (2) Gerardo estudió lingüística computacional

Para generar enunciados complejos, necesitamos algunos símbolos que ligen enunciados simples:

Mi madre trabajó mucho para comprar su casa =

(1) Mi madre trabajó mucho Y ["Con el dinero que ganó trabajando"]
ENTONCES (2) Mi madre compró su casa.

Construyendo un lenguaje formal (3)

Ahora, para saber si nuestras proposiciones son buenas o malas, necesitamos establecer un mecanismo de evaluación. Por ejemplo, podríamos analizar su referencia:

a) Mi mamá estudió Letras, no terminó la carrera y se puso a trabajar como maestra. [Es cierto]

b) Mi tía vivió en México. Trabajó muchos años como enfermera en un hospital público. [Es cierto]

¿Tu mamá vive en Veracruz Y se puso a trabajar en el Seguro?

[Es falso]

Construyendo un lenguaje formal (4)

Si observan, cada una de estas proposiciones es divisible en componentes, p. e., recordemos:

- (1) Gerardo es mi tutor +
- (2) Gerardo estudió lingüística computacional

A los componentes (1) y (2) vamos a denominarlos proposiciones atómicas (o básicas), mientras que la proposición resultante la vamos a denominar proposición molecular (o compuesta).

Construyendo un lenguaje formal (5)

Si vamos a evaluar nuestras proposiciones verificando su relación de verdad o falsedad respecto a un referente específico, necesitamos entonces un instrumento que nos permita realizar dicha verificación. A este instrumento se le conoce como **tabla de verdad**:

Una tabla de verdad no es más que un mecanismo formal que determina si una proposición molecular es correcta o no, de acuerdo con el tipo de proposiciones atómicas que la conforman, así como del operador que conecte a tales enunciados básicos.

p	q	r	$(p \wedge q) \vee (r \rightarrow q)$		
V	V	V	V	V	V
V	V	F	V	V	V
V	F	V	F	F	F
V	F	F	F	V	V
F	V	V	F	V	V
F	V	F	F	V	V
F	F	V	F	F	F
F	F	F	F	V	V

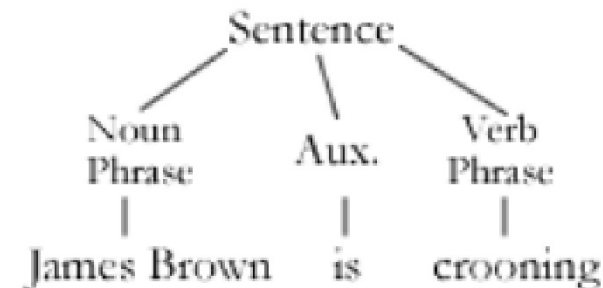
Representación formal del significado (1)

A los lingüistas que analizan las relaciones entre oraciones y proposiciones les gusta que la representación del significado sea lo más precisa posible. El ejemplo de la lámina anterior es una muestra de ello.

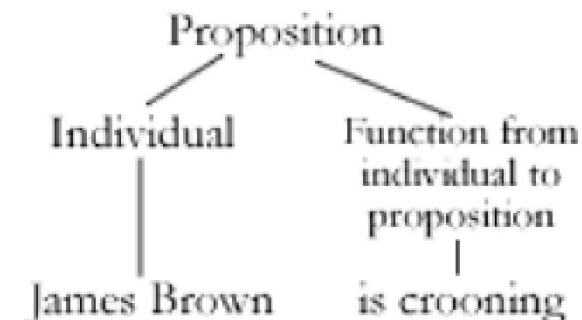
Para lograr esta precisión, aprovechan algunos trucos venidos de la lógica, como el uso de operadores que indican relaciones entre **proposiciones simples**, generando así **proposiciones complejas**.

Recordemos las relaciones que estamos considerando.

Example of a syntactic tree



Example of (one kind of) semantic tree



Representación formal del significado (2)

A	B	$A \wedge B$
V	V	V
V	F	F
F	V	F
F	F	F

Conjunción

A	B	$A \vee B$
V	V	V
V	F	V
F	V	V
F	F	F

Disyunción

A	B	$A \rightarrow B$
V	V	V
V	F	F
F	V	V
F	F	V

Condicional

A	B	$A \leftrightarrow B$
V	V	V
V	F	F
F	V	F
F	F	V

Bicondicional

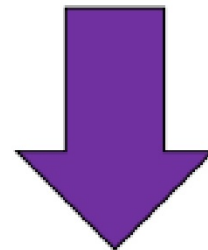
Negación

A	$\neg A$
V	F
F	V

Representación formal del significado (3)

Tomando en cuenta esta clase de operadores y relaciones, podemos generar proposiciones que representen el significado de oraciones como:

[Juan habló rápidamente con María Proposición simple 1], y
[quedaron de verse después Proposición simple 2]



$p \wedge q$

Ejercicio 1

Basándonos en el ejemplo anterior, tratemos de ver si podemos construir algunas proposiciones basadas en el significado de las siguientes oraciones:

1. Julia estaba cantando mientras Juan tocaba el teclado.
2. ¿Terminamos la tarea hoy, o la dejamos para mañana?
3. ¡Si regresas aquí después de las doce, mejor ni entres!
4. Sólo si tengo dinero haré el viaje a Estados Unidos
5. No hubo un tsunami en Valparaíso
6. Hubo un sismo de 8 grados, pero no se presentó un tsunami en Valparaíso
7. Si paga de contado, no le cargamos el costo de traslado a su domicilio

Ejercicio 2

Para construir nuestras proposiciones, vamos a seguir las siguientes reglas:

1. Primero, trabajar en equipos
2. Segmentar las oraciones en proposiciones simples, de modo que podamos saber cómo se configura la proposición compleja
3. Finalmente, representar con los símbolos p y q estas proposiciones simples, y luego utilizar alguno de estos operadores para establecer la relación correspondiente:

\wedge (conjunción) | \vee (disyunción) | \rightarrow (condicional)
 \leftrightarrow (bicondicional) | \neg (negación)

Proposiciones y estructura argumental (1)

Una vez que hemos hecho esta clase de representación del significado a partir de una proposición, podemos considerar que la lógica nos ha ofrecido una herramienta bastante útil para concretar esta representación de una forma clara y precisa.

¿Qué otras cuestiones podemos abordar tomando esta clase de herramientas lógicas?

Para responder a esta pregunta, retomemos un tema que hemos visto en sesiones anteriores: ¿recuerdan la relación que establecimos entre predicados y argumentos?

Proposiciones y estructura argumental (2)

La idea de que los eventos que ocurren en el mundo pueden ser descritos tomando en cuenta el total de actantes que participa en ellos viene también de Frege, en concreto de un método que propuso para resolver el problema de diferencia referencia de sentido: necesitamos dividir estas expresiones en sus partes mínimas, esto es, en predicados y argumentos. Mejor aún: predicados y argumentos equivale a **proposiciones atómicas que construyen proposiciones moleculares:**

X es un planeta

X = Venus

X = "El lucero del alba"

X = "El lucero vespertino"

A = (Planeta)

A = A [Relación de referencia]

A = B [Relación de sentido]



Proposiciones y estructura argumental (3)

En pocas palabras, Frege considera que en el cálculo de proposiciones es un mecanismo que permite representar la relación que puede establecer un predicado con sus respectivos argumentos.

Tratemos de explicar esto del siguiente modo: supongamos que tenemos un predicado del tipo *X es un satélite de Y*. Más o menos entendemos, por extensión, una relación como: *la Luna es un satélite de la Tierra*.



...Ser un satélite de...



Proposiciones y estructura argumental (4)

Como se puede observar, existen dos rasgos particulares que sobresalen en estos ejemplos:

1. Que el predicado ... *Ser un satélite de...* requiere de dos argumentos, y que éstos tienen un orden concreto.
2. Que tal predicado se puede aplicar a otros objetos que puedan operar como argumentos similares. Si esto es así, entonces podemos formar un conjunto de lunas y planetas ligados por el predicado ...*Ser un satélite de...*

Una forma muy simple de representar esta relación es la siguiente:

<Luna, Tierra>

Proposiciones y estructura argumental (5)

Frege consideró que las relaciones entre predicados y argumentos podía representarse como una **función**.

¿Y qué es una función? Digamos que se trata de *una relación o un vínculo que se establece entre dos elementos (aunque pueden ser más) pertenecientes a dos conjuntos distintos, de modo que se corresponden uno con otro*. Veamos el siguiente ejemplo:

$f(z)$ = "SER el planeta llamado Venus"

x = "Venus"

y = "Lucero del alba"

Tenemos entonces una función que es expresable con el predicado *ser el planeta llamado Venus*, y dos argumentos posibles que pueden establecer una relación con tal predicado, *Venus* y *Lucero del alba*. Esto nos da como resultado:

$$f(x, y) = z$$

Que equivale a decir: *los argumentos 'Venus' y 'Lucero del alba' tienen una relación con la función del predicado 'SER el planeta llamado Venus'*.

Algunas deducciones (1)

En resumen, podemos considerar que la relación de predicados y argumentos es un tipo de función semántica que permite vincular elementos que se corresponden entre sí. Esta función puede ser representada como una proposición.

En otras palabras, si tenemos una oración como:

Miguel está contento hoy

Podemos representarla como una función:

$f(ECH(m))$

En donde:

EC [Predicado] = *Estar contento*

m [Argumento] = *Miguel*

Algunas deducciones (2)

Finalmente, esta función se puede traducir en la siguiente proposición:

ECH(m)

O lo que es lo mismo: *Miguel* es un argumento que está dentro del alcance del predicado *estar contento hoy*.

¿Este tipo de proposición nos es útil para representar predicados y argumentos diferentes? Analicen estos casos:

Juan come un sandwich

CS(j, s)

Juan come y bebe mucho

CM(j) \wedge BM(j)

Pregunta: ¿cuál es la diferencia entre ambas oraciones? ¿resulta pertinente o no hacer esta diferencia?

¡Hagamos un grupo! (1)

Pasemos ahora a otro tema que nos va a servir de mucho para profundizar en nuestro análisis semántico, y para ello, vamos a aplicar la disyuntiva que empleó *The Doors* en 1965 para constituirse como una banda: *o creamos una nueva religión, o creamos un grupo de rock.*

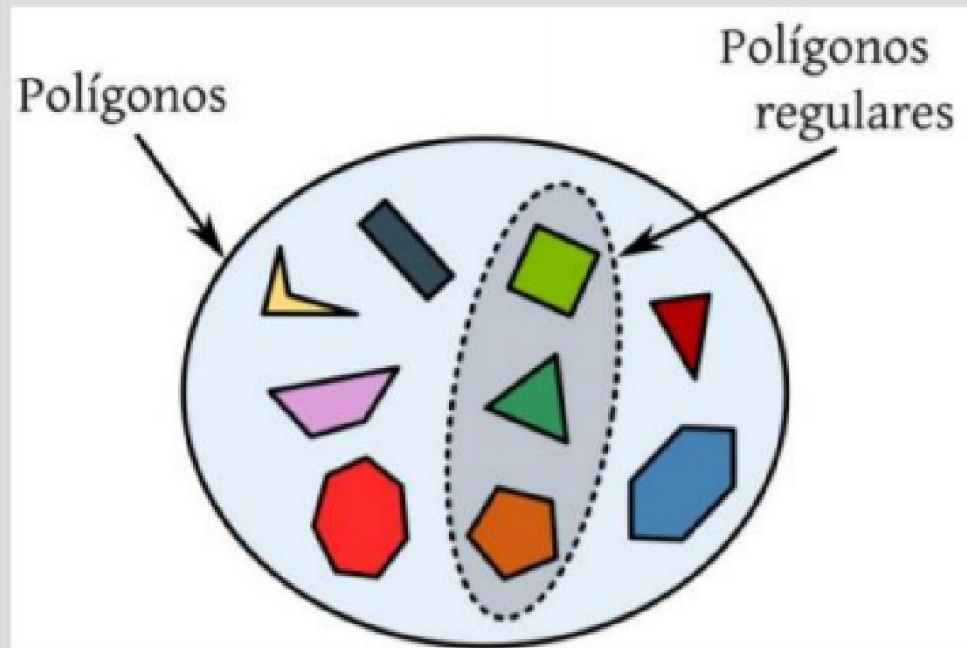


Hablemos entonces de lo que es la **teoría de conjuntos**.

¡Hagamos un grupo! (2)

¿Y por qué tenemos que hablar de esto? Veamos:

Empecemos entonces por definir qué es un conjunto, y lo más sencillo que podemos decir es que se trata de **una colección de objetos**, la cual a su vez puede ser considerada como un objeto también.



¡Hagamos un grupo! (3)

¿Cómo podemos establecer cuáles son los rasgos que caracterizan a un conjunto? Básicamente, tenemos 3 relaciones para esto:

Pertenencia: se dice que un objeto x pertenece a un conjunto C si x es realmente parte de los objetos que lo conforman:
 $x \in C$.

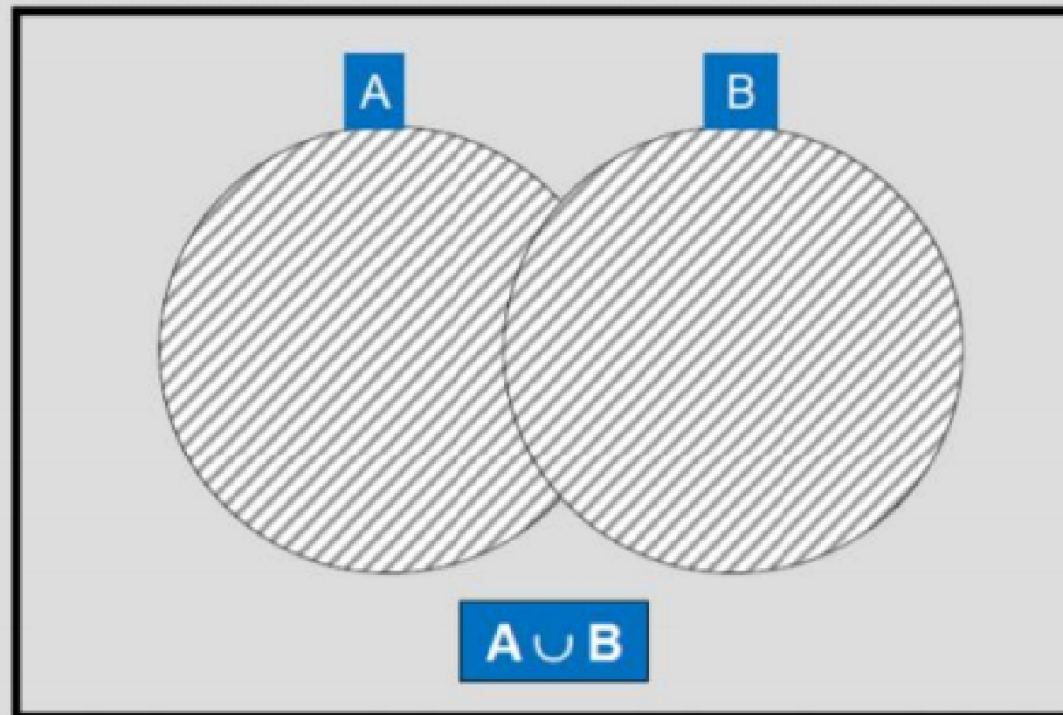
Igualdad: Dos conjuntos A y B son iguales si tienen exactamente los mismos elementos: $A = B$.

Inclusión: Un conjunto como A puede ser parte de otro conjunto B , en el caso de A comparta los mismos elementos que B : $A \subseteq B$.

¡Hagamos un grupo! (4)

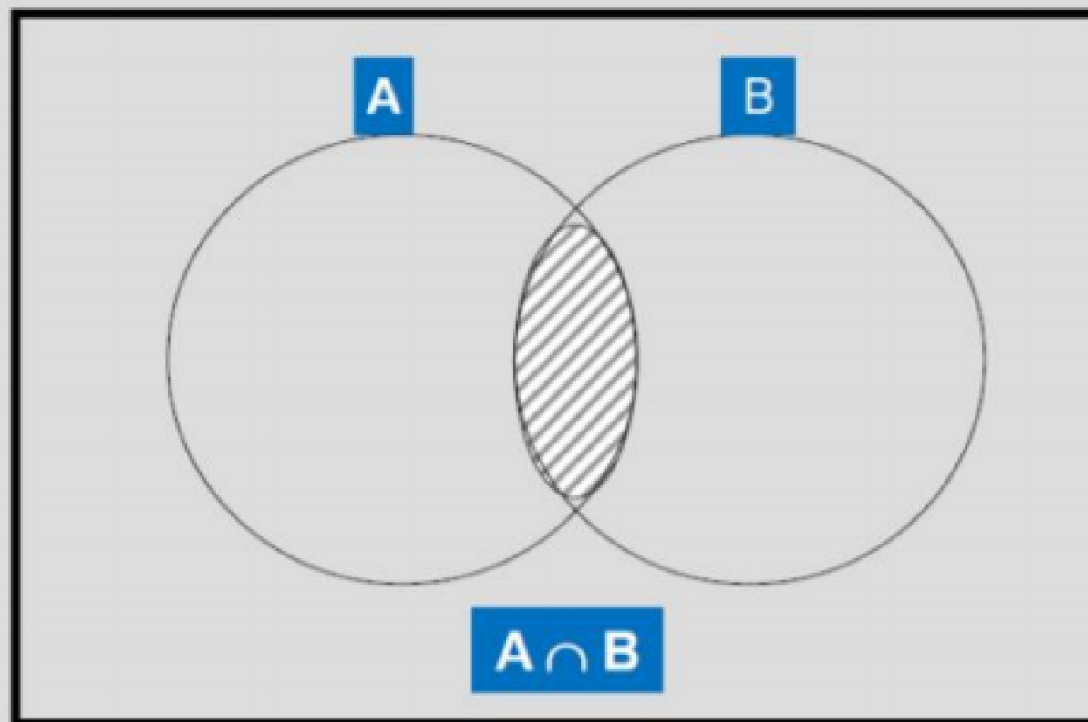
Y para complementar estas relaciones, vamos a añadir algunas operaciones conocidas, ¿las recuerdan?:

Unión de conjuntos: cuando dos conjuntos A y B comparten la totalidad de sus objetos se considera que establecen una unión, esto es $\{x \mid x \in A \text{ o bien } x \in B\}$.



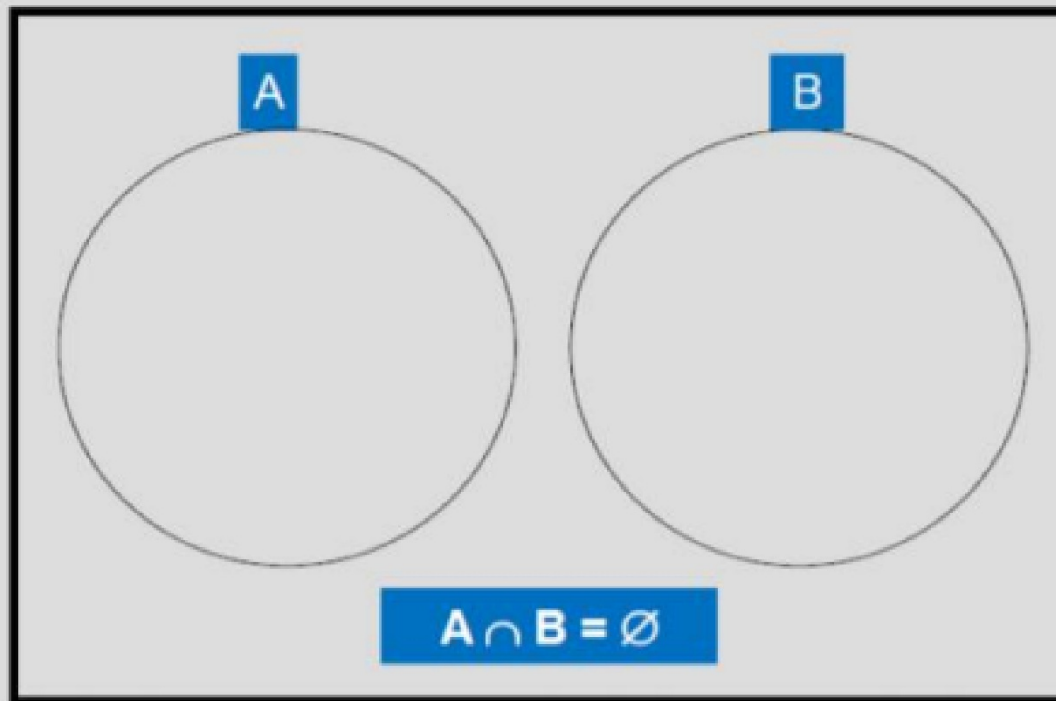
¡Hagamos un grupo! (5)

Intersección de conjuntos: dos conjuntos A y B comparten una parte de la totalidad de sus objetos, esto es $\{x \mid x \in A \text{ y } x \in B\}$.



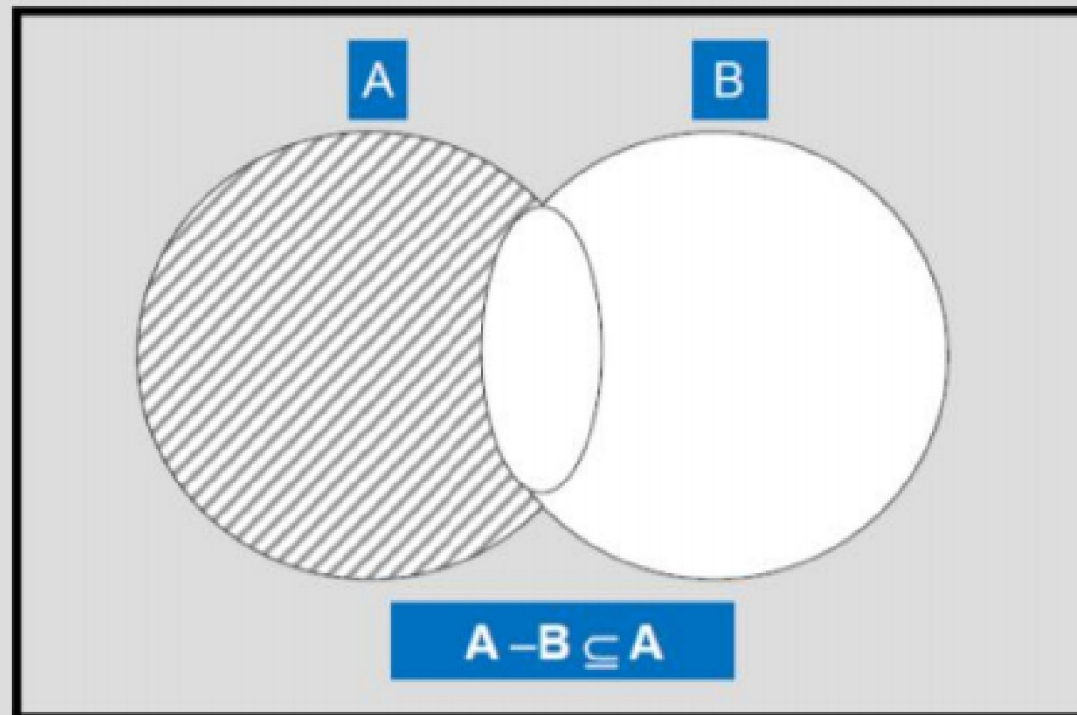
¡Hagamos un grupo! (6)

Conjuntos disjuntos: dos conjuntos A y B no tienen ningún objeto que compartir, esto es $\{A \cap B = \emptyset\}$.



¡Hagamos un grupo! (7)

Diferencia entre conjuntos: los objetos que aparecen en el conjunto A y que no aparecen en el conjunto B, es decir: $\{A - B \subseteq A\}$:



Experimento (1)

Vamos a crear una gramática que describa la composición de frases nominales en 2 lenguas diferentes: español e inglés. Esta gramática es de juguete, pues lo que queremos aprender es cómo aplicar la teoría de conjuntos para describir la composición de estas frases.

```
>>> sentence = [("the", "DT"), ("little", "JJ"), ("yellow", "JJ"), ❶  
... ("dog", "NN"), ("barked", "VBD"), ("at", "IN"), ("the", "DT"), ("cat", "NN")]  
  
>>> grammar = "NP: {<DT>?<JJ>*<NN>}" ❷  
  
>>> cp = nltk.RegexpParser(grammar) ❸  
>>> result = cp.parse(sentence) ❹  
>>> print(result) ❺  
(S  
  (NP the/DT little/JJ yellow/JJ dog/NN)  
  barked/VBD  
  at/IN  
  (NP the/DT cat/NN))  
>>> result.draw() ❻
```

Para empezar, veamos cómo se constituyen las frases nominales en estas lenguas conforme a los ejemplos de la siguiente lámina.

Experimento (2)

Vamos a partir de las siguientes hipótesis:

1. En inglés, una frase nominal sigue una secuencia Artículo + Adjetivo + Nombre: *The White House, The New English Grammar, A big fish, those intelligent students...*
2. En español, una frase nominal sigue una secuencia de Artículo + (Adjetivo) + Nombre + (Adjetivo), esto es, el adjetivo puede ir antes o después del nombre: *la prensa mexicana, el gran robo, mi nueva computadora negra...*

Experimento (3)

Supongamos que queremos enseñarle a alguien que no sabe ni inglés ni español que estas dos lenguas siguen un orden de palabras distinto en frase nominal. Una forma de hacerlo es por ensayo y error:

- **En inglés:** An important theory [es correcta] / Book this expensive [es incorrecta]
- **En español:** Un libro caro [es correcta]/ un buen libro [es correcta]/ un buen libro caro [¿suena bien?]/ Caro libro buen un [es incorrecta]

Experimento (4)

Si la secuencia que presentan el inglés y el español son regulares, otra vía para enseñar el orden de palabras es hacer reglas formales, por ejemplo:

Regla de construcción (RC)

1. Inglés: Art + (Adj*) + N
2. Español: Art + (Adj*) + N (Adj*)

Operadores (Op)

1. + = "Seguido de"
2. (...) = "Una palabra que puede aparecer o no"
3. * = "Se repite más de una vez"

Listas de palabras (LP)

1. Inglés: {The, a, this, dog, place, friend, brave, nice, intelligent}
2. Español: {El, un, ese, perro, lugar, amigo, fiero, buen, inteligente}

Experimento (5)

¿Qué frases son correctas, y que frases no lo son?:

FN1 = El perro fiero

FN2 = Un perro fiero

FN3 = Ese perro fiero

FN4 = El buen perro

FN5 = El buen perro fiero

FN6 = El fiero buen perro

FN7 = El perro buen inteligente fiero (¿?)

FN8 = El perro amigo buen inteligente fiero (¿?)

FN 9 = El perro amigo lugar buen inteligente fiero ¡Error!

Experimento (6)

¿Qué nos muestran nuestros resultados? Que hay secuencias de frases nominales que son adecuadas, y otras son erróneas, p.e.:

FN1 = Art + N + Adj

FN2 = Art + Adj + N

...

FN6 = Art + Adj + Adj + N

∈

Frase nominal en español
(FNE) =

Art + (Adj*) + N (Adj*)

FN7 = Art + N² + Adj⁴

FN8 = Art + N² + Adj⁵

FN9 = Art + N³ + Adj⁵

∉

Frase nominal en español
(FNE) =

Art + (Adj*) + N (Adj*)

Experimento (7)

¿Y qué otras relaciones podemos establecer? Veamos algunos casos:

1. Tanto las frases nominales en español e inglés son subgrupos de un grupo más general denominado **frase nominal**:

Frase nominal en español
(FNE) =

Art + (Adj*) + N + (Adj*)

Frase nominal en inglés
(FNI) =

Art + (Adj*) + N



Frase nominal (FN) =

Art + (Adj) + N

Experimento (8)

¿Y qué otras relaciones podemos establecer? Veamos algunos casos:

2. Por otro lado, podemos determinar si dos secuencias son equivalentes o no, esto es:

$$FN1 = Art + N + Adj = FN2 = Art + Adj + N$$

Si no es el caso, entonces lo representamos como:

$$FV = V + FN \neq FN = Art + (Adj) + N$$

Muchas gracias