

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO

CENTRO UNIVERSITARIO UAEM TEXCOCO
INGENIERÍA EN COMPUTACIÓN



Analizadores semánticos en el procesamiento del lenguaje natural

Que para obtener el Título de
Ingeniero en Computación

Presenta

Jorge Abel Pineda Bibriescas

Director de tesis

M en ISC Irene Aguilar Juárez

Revisores de tesis

M en C Gerardo Rafael Valencia
Valencia

Dr. Joel Ayala de la Vega

Texcoco, Estado de México, a 28 de Noviembre de 2014

RESUMEN

Esta tesis está orientada a documentar el papel de los analizadores semánticos en el procesamiento del lenguaje natural, además también se documentará como ejemplo el desarrollo de un analizador semántico.

El procesamiento del Lenguaje Natural (PLN) es una parte esencial de la inteligencia artificial que investiga y formula mecanismos computacionales efectivos que faciliten la interrelación hombre/máquina y permitan una comunicación mucho más fluida y menos rígida que los lenguajes formales.

Todo sistema de PLN intenta simular un comportamiento lingüístico humano; para ello debe tomar conocimiento tanto de las estructuras propias del lenguaje, como del conocimiento general acerca del discurso.

Debido a la complejidad necesaria para lograr que la computadora “comprenda” el lenguaje humano es que se pretende realizar este trabajo que incluye un analizador el cual permita facilitar esta tarea.

La manera en la que el analizador semántico realiza el análisis es la siguiente:

1. A partir de leer texto en lenguaje natural hace una comparación con una ontología (red Semántica) ya que trabajar con redes semánticas facilitará su uso.
2. Si las palabras se relacionan semánticamente, el analizador lo indica o indica lo contrario
3. Las búsquedas se realizan usando sinónimos y por herencia

Para su funcionamiento el analizador requiere una base de conocimientos, esta base de conocimientos debe tener la estructura de triplete, la cual contendrá (Concepto, relación, valor).

El analizador semántico estará orientado a la computadora para evitar inconsistencias en la introducción de texto. El usuario final estará enfocado en crear, enriquecer su base de conocimientos ya que mientras más información tenga mayor es el alcance del analizador semántico, además de que siendo más grande la base de conocimientos será mayor el conocimiento que obtendrá la computadora y evitará que ocurran inconsistencias.

CONTENIDO

Resumen.....	i
Dedicatoria.....	ii
Agradecimientos	iii
Contenido.....	iv
Listado de Imágenes	1
Listado de Tablas	3
Introducción	4
Capítulo 1.....	6
1.1 Problemática.....	6
1.2 Justificación.....	6
1.3 Objetivos	7
1.4 Metodología	8
1.5 Hipótesis.....	9
1.6 Variables de Medición.	9
1.7 Estadísticos.....	10
Capítulo 2.....	11
Wólffram Alpha.	17
LSA.....	18
Stilus:	18
EdAS:.....	19
Healthline:.....	19
Towards semantic search and inference in electronic medical records: An approach using concept-based information retrieval:	20
Khresmoi – multilingual semantic search of medical texts and images	20
Towards a semantic medical Web: HealthCyberMap’s tool for building an RDF metadata base of health information resources based on the Qualified Dublin Core Metadata Set: ...	21
Making ontologies work for resolving redundancies across documents:	21
Árboles de decisión como herramienta en el diagnóstico médico:.....	22

Otros analizadores.	22
Capítulo 3 Procesamiento de Lenguaje Natural.....	23
3.1 Inteligencia artificial.	24
3.1.1 Perspectiva histórica: de la Inteligencia Artificial.	25
3.1.2 Sistemas basados en conocimientos.....	26
3.2 Introducción al Procesamiento del lenguaje Natural.	27
3.2.1 Historia del Procesamiento del lenguaje Natural (PLN).....	27
3.3 Procesamiento del lenguaje natural.	29
3.3.1 Recuperación de información.	29
3.3.2 Interfaces en Lenguaje Natural.	30
3.3.3 Traducción Automática.....	31
3.3.4 Enfoque de dependencias.....	32
3.3.5 Introducción a las etiquetas Eagles.	33
3.3.6 Redes Semánticas.....	33
3.4 El Lenguaje.	35
3.4.1 Lenguaje natural.....	35
3.4.2 Lenguaje Formal.	36
3.4.3 Lingüística del texto.....	36
3.4.4 Niveles del Lenguaje.....	36
3.4.5 Textos en lenguaje natural.	37
3.5 Ontologías	37
Capítulo 4 Analizadores Semánticos	39
4.1 Análisis léxico.	39
4.1.1 Utilidad del análisis léxico.....	39
4.1.2 Funcionamiento del Analizador léxico.	40
4.2 Análisis Morfológico.	41
4.2.1 Analizador Morfológico.....	42
4.3 Analizador Sintáctico.	42
4.3.1 Análisis Sintáctico.	42
4.3.2 Funciones del Analizador Sintáctico.	44
4.4 Análisis Semántico.	45
4.4.1 Interpretación guiada por la semántica.....	47

4.4.2	Proceso de Interpretación Semántica.....	48
4.4.3	Análisis Semántico: Modelo Estructural.	48
4.4.4	Procesamiento semántico.....	49
4.4.5	Proximidad Semántica.	50
4.4.6	Aspectos Semánticos.	50
4.4.7	Funciones del análisis semántico.....	51
4.4.8	Reglas semánticas.	52
4.5	Análisis pragmático.....	52
Capítulo 5	Desarrollo del Analizador Semántico	54
5.1	Análisis de Analizador semántico.....	54
5.2	Diseño.....	57
5.2.1	Diagramas	71
5.3	Ontologías y Consultas.....	74
5.4	Pruebas	81
5.4.1	Pruebas de escritorio realizadas a la ontología planta.	81
Capítulo 6	resultados.....	84
6.1.1	Pruebas en GUI del Analizador Semántico.	84
6.2	Análisis comparativo.....	89
Conclusiones	90
Referencias	91

LISTADO DE IMÁGENES

Imagen 1.	Distinción entre IA como ciencia e IA como ingeniería (Pajares & Santos, 2006)	25
Imagen 2.	Ejemplo: de cómo está constituida una etiqueta.	33
Imagen 3.	Representación Gráfica: Grafo Orientado etiquetado.....	34
Imagen 4.	Conexión analizador léxico-sintáctico.....	41
Imagen 5.	Funcionamiento del parse.....	45
Imagen 6.	Se muestra un esquema grafico sobre el análisis semántico bajo las anteriores interpretaciones.	49
Imagen 1.	Algoritmo general de Analizador Semántico.....	57
Imagen 2.	Algoritmo general del Analizador Semántico Parte 2.....	58
Imagen 3.	Algoritmo de Herencia del Analizador Semántico.	59
Imagen 4.	Ontología para el caso1.....	60
Imagen 5.	Algoritmo que explica el funcionamiento del caso1: “trata de hallar la relación tiene y el valor ramas en el concepto árbol”.	61
Imagen 6.	Ontología para el caso 2.....	62
Imagen 7.	Algoritmo que explica el funcionamiento del Caso 2: “trata de hallar la relación cuenta con y el valor ramas en el concepto árbol”.....	63
Imagen 8.	Ontología para el caso 3.....	64
Imagen 9.	Algoritmo que explica el funcionamiento del Caso 3.”trata de hallar la relación Cuenta con y el valor ramas en el concepto arbolillo”.....	65
Imagen 10.	Ontología para el Caso 4.....	66
Imagen 11.	Algoritmo que explica el funcionamiento del Caso 4. “trata de hallar la relación cuenta con y el valor retoño dentro del concepto arbolillo”.....	67
Imagen 12.	Ontología para el Caso 5.....	68
Imagen 13.	Algoritmo que explica el funcionamiento del Caso 5. “trata de hallar la relación tiene y el valor ramas con el concepto planta leñosa”.....	69
Imagen 14.	Ontología para el caso 6.....	70
Imagen 15.	Algoritmo que explica el funcionamiento del Caso 6. “trata de hallar la relación color y el valor amarillo con el concepto arboleda”.....	71

Imagen 16.	Diagrama de secuencia	72
Imagen 17.	Diagrama de Clases.....	73
Imagen 18.	Diagrama de casos de uso.....	73
Imagen 19.	Ontología de planta.....	74
Imagen 20.	Ontología de plantas casi terminada	74
Imagen 21.	Ontología de planta terminada	75
Imagen 22.	Aplicación de consultas a la ontología planta.....	75
Imagen 23.	Aplicación de consultas a la ontología planta.....	76
Imagen 24.	Ultima consulta realizada a la ontología “planta.owl”	76
Imagen 25.	Primera parte de la ontología arboles.....	77
Imagen 26.	Segunda parte de la ontología	78
Imagen 27.	Árbol de la ontología Arboles y sus atributos.....	79
Imagen 28.	Consulta realizada a la ontología arboles.....	79
Imagen 29.	Árbol de ontología donde una clase muestra sus atributos.....	80
Imagen 30.	Imagen de de la consulta realizada a la ontología.....	80
Imagen 31.	Resultados del Analizador Semántico de la ontología planta.....	81
Imagen 32.	Resultados del Analizador Semántico de la ontología planta.....	81
Imagen 33.	Resultados del Analizador Semántico de la ontología planta.....	81
Imagen 34.	Resultados del Analizador Semántico de la ontología planta.....	82
Imagen 35.	Resultado del analizador Semántico a la ontología arboles.....	82
Imagen 36.	Resultado del analizador Semántico a la ontología arboles.....	82
Imagen 37.	Resultado del analizador semántico aplicado a la ontología arboles.....	83
Imagen 38.	Resultado del analizador semántico aplicado a la ontología arboles.....	83
Imagen 39.	Resultado del analizador semántico aplicado a la ontología arboles.....	83
Imagen 40.	Imagen de portada del Analizador Semántico.....	84
Imagen 41.	Imagen de la GUI del Analizador Semántico.....	85
Imagen 42.	Búsqueda de la Ontología en el Analizador Semántico.....	85
Imagen 43.	Ontología encontrada por el Analizador Semántico.....	86
Imagen 44.	Datos ingresados por el usuario (Tripleta).....	86
Imagen 45.	Respuesta del Analizador Semántico.....	87
Imagen 46.	Imagen de Lectura de la Ontología.....	87
Imagen 47.	Datos de la Ontología Seleccionada.....	88

LISTADO DE TABLAS

Tabla 1: Estadística de variables de comparación aplicada a ontologías de prueba.....	10
Tabla 2. Niveles de Conocimiento en el procesamiento del Lenguaje Natural.	28
Tabla 3: Comparativa entre Wólfam alpha y el analizador semántico, en ambos analizadores se realizaron las mismas pruebas; en las mismas ontologías.	89

INTRODUCCIÓN

Esta tesis documenta la importancia de los analizadores semánticos en el procesamiento del lenguaje natural. Durante toda la historia de la humanidad el conocimiento, en su mayor parte se comunica, se guarda y se maneja en la forma del lenguaje natural, griego, inglés, español, etc... En la época actual no se da ninguna excepción: el conocimiento sigue existiendo y creándose en la forma de documentos, libros, artículos aunque estos se guardan en forma electrónica o sea digital, el gran avance es que ahora contamos con las computadoras que pueden ser una ayuda enorme en el procesamiento de este conocimiento.

Sin embargo lo que es para nosotros los seres humanos no lo es para las computadoras, son archivos, unas secuencias de caracteres y nada más. Una computadora puede copiar tal archivo, respaldarlo, transmitirlo y borrarlo.

Para lo que se desarrollará un analizador semántico el cual ayuda a evitar las inconsistencias en una computadora, por medio de la introducción de textos en lenguaje natural los cuales son útiles en la comprensión humana, pero no para una computadora; por eso se pretende realizar este analizador.

El analizador funcionará de la siguiente forma: el cliente o usuario se encargara de crear la ontología nutrirla hacerla grande y que este en forma de tripleta; lo que realizara la computadora es la búsqueda de la tripleta ingresada por el cliente o usuario de tal manera que busque en la ontología. Si la encuentra mandará un mensaje de **“Hay Relación Semántica”** de otro modo mandará un mensaje de **“No hay Relación Semántica”**; Mientras la base de conocimientos este más completa mejor funcionará el analizador semántico.

En la introducción de textos en lenguaje natural, estos se transformarán en ontologías (Redes Semánticas), las cuales tiene la función de proporcionar la información que se pretende analizar de tal manera; esta ontología puede ser de distintas áreas no solo se enfoca en una sola sino que abarca muchas como por ejemplo: una ontología de (Plantas, Animales, Computadoras, mares... etc.). La base de conocimientos deberá contener una cierta forma la cual constara de un (Concepto, Relación, Valor) a lo que llamaremos tripleta; esta tripleta servirá al momento de buscar palabras concretas de tal manera que mientras exista algún valor de la tripleta lo buscará en toda la ontología.

Puntos en el Desarrollo de la tesis.

- Se analizarán los problemas que se presentan en el procesamiento del lenguaje natural para comprender la importancia de un analizador semántico.
- Se analizarán las aplicaciones de analizadores semánticos en México.
- Se analizaran los analizadores semánticos, su historia, sus etapas de evolución.

- Se documentaran otros tipos de analizadores como es el sintáctico o el analizador morfológico para explicar las diferencias del analizador semántico.

Capítulo 1

1.1 Problemática

Actualmente es común que la gente no conozca del tema por lo que es necesario, difundir este tipo de trabajos para extender el conocimiento sobre inteligencia artificial, ya que es una rama que amplía el horizonte en la tecnología del software.

Uno de los principales problemas de no tener un analizador semántico se presenta en la búsqueda de redes semánticas, porque al utilizar un buscador de palabras claves; el buscador usa un analizador sintáctico el cual en la mayoría de los casos no satisface la búsqueda.

Otro problema que se presenta es la búsqueda de información sin inconsistencias por lo que cuesta más trabajo encontrar lo que se busca.

Otro factor es que las búsquedas semánticas no son auto corregibles, es decir, la computadora al realizar una búsqueda y al no contar con un analizador semántico no puede evitar inconsistencias de tal modo que no se auto corrige. Teniendo un analizador semántico durante las búsquedas, la computadora podrá corregir sus errores.

1.2 Justificación

Un beneficio importante de este trabajo es difundir conocimiento sobre la inteligencia artificial, de tal manera que no solo se vea de manera teórica su uso, sino que se vea un caso práctico y así poder enseñar a los alumnos del centro universitario UAEM Texcoco lo que es un analizador semántico.

Otro punto es explicar el manejo de otros analizadores que hagan búsquedas; uno de los analizadores en línea en búsqueda de documentos es Wólfram alpha, este analizador de textos trabaja de manera semántica sin embargo no resuelve efectivamente todas las búsquedas, por ejemplo: ante la pregunta “¿Cuántos muertos en total hubo en la segunda guerra mundial?”, este analizador solo manda todo lo relacionado a la segunda guerra mundial como quienes participaron, las fechas de ataque, las conquista etc. De tal manera que no indica la relación semántica.

El desarrollo del analizador semántico será útil para evitar la ambigüedad en la lectura de documentos, lo que se pretende hacer es que mediante un par de palabras claves se pueda reducir el tiempo en búsquedas semánticas.

Un beneficio de contar con esta herramienta es que la computadora se autocorrija de modo que si un usuario o cliente buscan un documento en lenguaje natural, la computadora sea capaz de entenderlo y así generar la búsqueda con autocorrección, es decir, que no busque de una manera lineal sino que genere otras maneras de búsqueda las cuales satisfagan la necesidad del cliente.

Otro beneficio de esta herramienta será la forma en la que los alumnos conozcan lo que es un analizador semántico, como se constituye, para qué sirve, además este trabajo puede ser una plataforma para los que les interese entrar al campo de la inteligencia artificial.

Un beneficio más al crear este analizador es que sirva como componente para un buscador en diferentes áreas de la investigación de tal forma que sea emigrado a otros campos de estudio.

Otro factor que beneficia la creación de un analizador semántico es la de conocer a fondo las diferencias en los distintos tipos de analizadores como los analizadores sintácticos, los semánticos y permite conocer las características de cada uno.

Una de las tareas más importantes de la creación de un analizador semántico es la búsqueda y análisis de la relación semántica la cual es la relación de palabras, las cuales funcionen para un mejor análisis del documento que se pretende analizar.

1.3 Objetivos

Objetivo General.

- Desarrollar un analizador semántico con lenguaje java para identificar la relación semántica en varias ontologías para sentar las bases para el análisis de textos en lenguaje natural.

Objetivos Particulares.

- Documentar la situación actual de los estudios sobre el Procesamiento de lenguaje Natural y los Analizadores Semánticos para difundir esta área de conocimiento.
- Recopilar los artículos científicos, memorias de congresos, libros de inteligencia artificial, actas y todo material documental sobre el Procesamiento del lenguaje natural.
- Modelar y codificar un analizador semántico
- Redactar e ingresar una base de conocimientos.

1.4 Metodología

La Metodología para el desarrollo de software es un conjunto de pasos y procedimientos que deben seguirse para desarrollar software. La metodología utilizada es la del modelo de cascada ya que esta metodología conlleva un orden el cual si no se sigue no se puede pasar a la siguiente tarea o función y para el desarrollo del analizador semántico este método es muy útil, ya que si no se tiene una relación continua no se puede seguir con las demás tareas.

Aplicando las etapas del modelo de cascada al analizador semántico.

Las etapas del modelo de cascada son las siguientes:

- Análisis de requisitos
- Diseño del Sistema
- Diseño del programa
- Codificación
- Pruebas
- Implantación
- Mantenimiento

Análisis de requisitos: en esta etapa se analizara lo que el cliente necesita o bien lo que el sistema realizará como es la búsqueda de conceptos dentro de una ontología y sus relaciones y valores. Se realizará la búsqueda en toda la ontología hasta encontrar el concepto en caso de existir eso es lo que se logrará.

Diseño del Sistema: en este paso es la descomposición del problema una vez identificándolo lo que se pretende es realizar módulos los cuales sea más fácil realizar las tareas por separado. Por ejemplo un módulo sería la búsqueda del concepto en la ontología otro modulo es una vez encontrado buscar sus relaciones, otro modulo es de esas relaciones realizar la búsqueda del valor por tanto dividir el problema de tal forma que aunque se separen uniendo todos los módulos realicen la misma tarea principal unidos el cual es la búsqueda de palabras las cuales tengan relación semántica.

Diseño del programa: En esta etapa se realizara el algoritmo que cumpla con la necesidad del análisis de requisitos, además que también sea una base fundamental para la codificación además de revisar las herramientas necesarias las cuales se puedan emplear para facilitar la codificación.

Codificación: Para este caso la codificación es donde se implementa el código fuente, el cual es el analizador semántico además en esta etapa se verificara en que plataforma es la que se utilizara para la realización del analizador; la plataforma la cual se utilizara será java ya que es un entorno el cual siento que es más favorable por el uso de la herencia, ya que mucho del analizador utiliza en su mayoría la herencia.

Mantenimiento: Es la etapa en la cual el software es utilizado de tal manera el cliente, no le cause problemas ya que puede que no cumpla con lo pedido o no funcione en esta etapa es

asegurar que el analizador busque y que no haga búsquedas raras las cuales no sirvan a lo que el usuario pide.

Además de la metodología de cascada también se utilizara el análisis documental el cual consta de la descripción o el análisis de un documento en sus partes esenciales y recuperación. De esta metodología solo se aplicarán los 5 primeros pasos pues aún no será posible una completa implementación del analizador.

1.5 Hipótesis

Con el uso del analizador semántico desarrollado se aumentará el reconocimiento de existencia de coincidencias semánticas en las búsquedas en ontologías, respecto al Wólffram alpha.

1.6 Variables de Medición.

- Tiempo de Respuesta
- Reconocimiento de existencia de coincidencias semánticas en las búsquedas en ontologías

1.7 Estadísticos

Tabla 1: Estadística de variables de comparación aplicada a ontologías de prueba

Ontología	Wólfam alpha		Analizador Semántico	
	Tiempo de Ejecución	Existencia de relación semántica	Tiempo de Ejecución	Existencia de relación semántica
Plantas				
Arboles				
Orquesta Sinfónica				
Animales				
Insectos				
Preparación de Pizzas				
Países				
Carros				
Plantas				
Arboles				
Orquesta Sinfónica				
Animales				
Insectos				
Preparación de Pizzas				
Países				
Carros				

Capítulo 2

Estado del Arte Analizadores Semánticos En el Mundo.

El procesamiento del lenguaje natural tiene varios científicos interesados en su desarrollo, como muestra a continuación se describen algunos de los proyectos más recientes sobre el área en el mundo los países que participan son (Estados Unidos, México, Brasil, España, Suda África).

Nombre del proyecto: “Uso de Marcos para Convertir Textos a Redes Semánticas y para la integración de las ontologías resultantes”.

Quien lo hace: Dr. Adolfo Guzmán Arenas y Alma Delia Cuevas Rasgado.

Objetivos.

- Utilizando pequeñas redes que representan marcos (“Frames”) de un documento de texto convertirlos en Ontologías.
- Esta Ontología común o unificada servirá para contestar preguntas no triviales.
- La creación de un analizador sintáctico-semántico el cual será guiado por marcos.
- El analizador trabajara con 1000 marcos de 1000 verbos (acciones) más frecuentes en el español con niveles de generalidad.
- Además se creara un integrador funcionara de tal manera funcione con marcos ya instanciados, lo que es igual a manejar hipergrafos.
- La ventaja principal de ese trabajo es que podrá integrar automáticamente información de varios documentos fuentes sobre cierto tema, produciendo una estructura (ontología unificada o común) que contiene el conocimiento extraído de los documentos textuales.

Elementos: Marcos (“Frames”), Ontologías, analizador sintáctico-semántico, integrador de marcos, documentos textuales.

Resultados: Gracias al Uso de Marcos ha facilitado la disminución de preguntas no triviales.

(Guzman, 2013)

Nombre del proyecto: ANÁLISIS Y CALIFICACIÓN AUTOMATIZADA DE TEXTOS EN ESPAÑOL (Proyecto Kalt-Léxico).

Quien lo hace: Agustín Tristán L.

Objetivos:

- Desarrollo de Software para la calificación y análisis de textos en español.
- Ser desarrollado con rutinas propuestas del año 2000, pero retomando un esquema de programación en línea.
- Desarrollar este esquema en lenguaje C++, en ambiente de Windows y que sea compatible con equipos que puedan correr el Windows 98.
- Crear analizadores de Vocabulario, calculadora de indicadores, analizador sintáctico y graficador.
- Crear analizadores gramaticales, Frecuencia y ubicación de palabras, calculadora de parámetros semánticos por familiaridad de vocabulario, con jugador de verbos, calculadores de parámetros sintácticos y calculadora de indicadores de comprensión lectora.

Elementos: lenguaje C++, Windows, Windows 98, analizadores de vocabulario, calculadora de indicadores, parámetros semánticos y sintácticos, analizadores gramaticales.

Resultados: El proyecto Kalt-Lexico Se encuentra en fase de desarrollo y se pretende incluir dentro de un proyecto integral de exámenes en línea. (**Tristan, 2007**)

Nombre del proyecto: DIANA-Construcciones: Buscando el conocimiento en los textos.

Quien lo hace: Paolo Rosso,

Objetivos:

- Es desarrollar herramientas de procesamiento del lenguaje y recursos lingüísticos acorde con el modelo teórico y adecuado al Procesamiento del Lenguaje Natural.
- Verificar entidades y eventos con argumentos implícitos.
- El tratamiento del lenguaje Subjetivo tal y como se produce en los medios de comunicación Social.
- El desarrollo de recursos de tecnología lingüística que están en la base de las aplicaciones basadas en el procesamiento del lenguaje, extracción de información, búsquedas de respuestas y recuperación de información.

Elementos: procesamiento del lenguaje, recursos lingüísticos, Procesamiento del Lenguaje Natural, lenguaje Subjetivo, tecnología lingüística.

Resultados: Este proyecto está finalizando pruebas en la universidad de Cataluña, en modo beta ya que se aplican correcciones del lenguaje. (Martinez, 2013).

Nombre del proyecto: Analizador Semántico: Mujer Olímpica 2012

Quien lo hace: Vega Caridad, Richard

Objetivos:

- Pretende Mostrar con plataformas tecnológicas como son Web 2.0 una búsqueda de noticias que traten sobre los logros de las deportistas femeninas españolas en disciplinas olímpicas y paraolímpicas.
- Que esta aplicación podría ser configurada para el seguimiento de eventos especiales como juegos olímpicos del 2012.
- Ser el primer buscador de noticias semántico en dicha materia.
- Lograr alcanzar una gran trayectoria de búsquedas enfocadas a las mujeres en las olimpiadas 2012 utilizando analizadores semánticos de tal manera lograr la recuperación de información.

Elementos: Recuperación de la información, Analizador Semántico, Algoritmos.

Resultados: Este proyecto está finalizando pruebas en la universidad de Cataluña, en modo beta ya que se aplican correcciones del lenguaje. (Vega & Richard, 2011).

Estado del Arte Analizadores Semánticos en Diferentes Plataformas de Programación.

Para poder analizar el procesamiento del lenguaje natural varios científicos están en su desarrollo, a continuación se muestran diferentes plataformas de programación la cual atacan al problema los lenguajes de programación que son mencionados son (BISON, C#, JavaScrip...etc.).

Nombre del Proyecto: JTLex un generador de analizadores Semánticos.

Quien lo hace: Bavera Francisco, Nordio D., Aguirre Jorge

Objetivos:

- crear un generador de analizadores semánticos y léxicos al contrario de otros generadores existentes siguiendo los esquemas de traducción
- se basa en un nuevo formalismo en expresiones regulares traductoras
- su diseño como la especificación de los procedimientos con que el usuario implementa la semántica asociada a los símbolos son Orientados a Objetos.
- El lenguaje de implementación de JTLex es Java.
- Crear un entorno de generación de procesadores de lenguajes.

Elementos: generador de analizadores, Semántica, léxico, formalismo, expresiones regulares, lenguaje de implementación, procesadores de lenguajes

Resultados: Está en fase de desarrollo el cual todavía no se generan pruebas de este lenguaje. (Bavera, Arroyo, & Aguirre, 2013)

Nombre del Proyecto: ANTL (Another Tool Language Recognition)

Quien lo hace: Verónica Soto

Objetivos:

- Crear una categoría de meta-programas, es decir un programa que escribe programas.
- Desarrollar una descripción formal de la gramática de un lenguaje
- Diseñar acciones escritas en un lenguaje de programación de tal forma se transforme en un traductor e intérprete.
- Proporcionar facilidades para la creación de estructuras intermedias de análisis
- Proveer mecanismos para recuperarse automáticamente de errores y realizar reportes de los mismos.
- Tener todo el código fuente disponible y preparado para su instalación bajo plataformas de Linux, Windows, Mac OS X.
- Que ANTLR sea un programa escrito en java

Elementos: meta-programas, descripción formal de la gramática, traductor, intérprete, Java

Resultados: Este proyecto ha sido lanzado desde el año 2011 por lo cual está dando resultados positivos pero no ha logrado lo que su creadora esperaba.

(Soto, 2011)

Nombre del Proyecto: Analizador Semántico de Programas en Java.

Quien lo hace: Andrea Armijos

Objetivos:

- Hacer árboles de derivación y compararlo con árboles de sintaxis abstracta
- Realizar e comprender el análisis de tipos
- Determinar el tipo estático de cada expresión del programa
- Generar los parsers generados automáticamente
- Utilizar el análisis semántico FKScript

Elementos: arboles de derivación, arboles de sintaxis abstracta,

Resultados: Este proyecto está en etapa primaria ya que se inicio en este año pero se espera grandes resultados.

(Armijos, 2014)

Nombre del Proyecto: Analizador Semántico en FKScript.

Quien lo hace: Salvador Gómez Oliver.

Objetivos:

- Comprobación de la existencia de variables y funciones
- Calculo de tipos en expresiones
- Chequeo de tipos en instrucciones
- Chequeo de tipos en expresiones
- Enriquecimiento de los nodos del árbol AST cuyos nodos únicamente contienen un tipo (atributo Type) y un valor (atributo Text)
- Crear simples identificadores y literales
- Comentar aspectos prácticos del desarrollo de la etapa de análisis semántico del compilador para nuestro lenguaje FKScript.

Elementos: variables, funciones, expresiones, árbol AST, identificador, análisis semántico, lenguaje FKScript.

Resultados: Este proyecto aumento su eficacia un 20% como se había planeado aunque todavía no se desglosa por completo.

(Salvador, 2005)

Nombre del Proyecto: Análisis semántico-traductor descendente-usando polimorfismo con C#

Quien lo hace: Francisco Ríos Acosta.

Objetivos:

- La implementación de un analizador semántico en C#
- Declaración de constantes.
- Declaración de variables, asignación, lectura y visualización de la gramática.
- Hacer un traductor el cual ayude a fortalecer esta herramienta
- Utilizar el polimorfismo que se propone el cual consiste un objeto determinado.
- Crear un análisis sintáctico descendiente no recursivo predictivo.

Elementos: analizador semántico, C#, constantes, variables, asignación, visualización de la gramática, traductor, polimorfismo, análisis sintáctico descendiente, recursividad.

Resultados: no se tienen porque está en fase de prueba.

(Ríos, 2009)

Nombre del proyecto: Generador de analizadores semánticos en BISON

Quien lo hace: Miguel Cervantes Saavedra.

Objetivos:

- Traducir la especificación de una gramática de contexto libre
- Programar en C que implemente un analizador
- Asociar código (Acciones semánticas) a las reglas de la gramática
- Ejecutar cada vez que se aplique la regla correspondiente
- Realizar un análisis ascendente salto-reducción
- Agrupar en tokens y no terminales según a las reglas de la gramática
- Utilizar una pila donde se acumula símbolos
- Eliminar y mantener en ella la (Reducción)

Elementos: gramática, contexto libre, analizador, acciones semánticas, reglas de la gramática, análisis ascendente salto reducción, tokens, símbolos, reducción.

Resultados: Este proyecto se suspendió por falta de ingresos económicos del creador esta (incompleto).

(Cervantes, 2008)

Nombre del proyecto: Un modelo conciso de la cohesión en el texto y coherencia en la comprensión

Quien lo hace: Max Louwerse

Objetivos:

- Crear diversos aspectos de cohesión y coherencia estudiados extensamente desde una variedad común.
- crear indicaciones textuales desde las cuales debería construir las representaciones coherentes.
- Representar un modelo que clasifica la cohesión y la coherencia la gramática.
- Utilizar e ilustrar ejemplos de un corpus lingüístico.
- Discutir varias aplicaciones, incluyendo experimentos de tiempo de lectura y de seguimiento visual.

Elementos: cohesión, coherencia, variedad común, indicaciones textuales, representaciones coherentes, modelo de cohesión, coherencia gramática, corpus lingüístico, seguimiento visual.

Resultados: Este artículo fue terminado y cumplió con lo correspondiente a su tema.

(Louwerse, 2004)

Wólfram Alpha.

Este buscador ha originado una gran expectativa desde su lanzamiento en mayo 2009 esto se debe a su funcionamiento es capaz de realizar cálculos matemáticos tal como lo haría cualquier persona entre sus capacidades de cálculo constan la resolución de polinomios integrales y demás operaciones esto ha llevado a la gran comunidad de internet a considerar la resolución entre buscadores semánticos y no semánticos. Su funcionamiento es muy particular las inferencias son básicas y son realizadas por Cyc, un proyecto de inteligencia artificial que intenta ensamblar una ontología comprensiva y una base de conocimientos general todo esto con el objetivo de realizar razonamientos de tipo humano otro factor importante es RI, un grupo de expertos de Wólfram Alpha ha procesado la información presente a la Web y la han almacenado en una base de datos desde la cual se extrae un informe con los datos más relevantes evitando mostrar enlaces Web este buscador elimina paginas intermediarias y presenta el resultado directo en la interfaz del buscador.

Si bien es cierto este buscador recupera la información desde una base de datos y realiza inferencias sin embargo es más un sistema de respuestas que un buscador semántico como tal con esto lo corrobora el propio Stephen Wólfram “creador del buscador ya que considera que su aplicación es más un repositorio de conocimientos que un buscador”. (Morales, 2014)

LSA.

En (Garcia, 2013) LSA, Análisis semántico Latente es un tipo de análisis computacional que, está basado es un algoritmo matemático, permite determinar y cuantificar la similitud de significado entre piezas textuales pertenecientes a un mismo dominio de conocimiento.

Debido a esta funcionalidad se le ha llegado a contribuir como los seres humanos aprendemos, determinamos y usamos el significado de las palabras, capaz de emular y explicar cómo los seres humanos aprendemos, determinamos y usamos el significado de las palabras y el conocimiento en general.

Tomar posición al respecto, más allá de involucrarse en la discusión ya clásica acerca de si un mecanismo no dotado de conciencia puede dar respuesta a una cuestión tan compleja, supone conocer y comprender cuál es el sistema operativo del LSA, razón por la cual, en este artículo se presenta una breve descripción de este sistema a partir de sus fundamentos matemáticos, de la relación entre estos y los supuestos semánticos subyacentes, y de dos de las aplicaciones que se han desarrollado sobre su base operativa.

Para el estudio de los procesos cognitivos se trata fundamentalmente de una revisión y discusión bibliográfica que tiene por objetivo agrupar información que se encuentra dispersa en otros trabajos sobre el tema difundirla en nuestro idioma, comentar y cuestionar; desde una perspectiva lingüística, la naturaleza semántica atribuida a este análisis.

Otro trabajo importante para medir la similitud (ellos le llaman *confusión*, un número entre 0 y 1) entre dos conceptos es la teoría de la confusión , que mide (cuantifica) la confusión que existe cuando, por ejemplo, pido un mexicano y me dan un oaxaqueño (la confusión es 0, puesto que un oaxaqueño *es* un mexicano), pido un oaxaqueño y me dan un mexicano (la confusión es un número cercano a cero, porque el error o diferencia es poco); si pido un oaxaqueño y me dan un colombiano, la confusión aumenta; si me dan un croata, es mayor, y si me dan una salchicha o una iguana, cuando lo que pedí era un oaxaqueño, la confusión es grande, cercana a 1.

Stilus:

En (Daedalus, 2014) se describe a Stilus que permite revisar la ortografía y el estilo de sus textos en varios idiomas, según sus preferencias lingüísticas y desde cualquier ordenador dotado de un navegador estándar o Microsoft Word. Stilus no es el típico corrector automático basado en un análisis superficial del texto.

Su gran base léxica y la sensibilidad contextual con que revisa los escritos hacen posible que identifique con precisión errores «reales» en todos los niveles de lengua: ortografía, gramática, orto tipografía y estilo. Además, Stilus no corrige automáticamente el texto, sino que «revisa» y «detecta» posibles errores y los acompaña de sugerencias de reemplazo, explicaciones didácticas y referencias bibliográficas, para que sea el usuario quien confirme la conveniencia o no de «corregir». Con Stilus, el humano nunca pierde el control de la corrección.

En (Daedalus, 2014) se plantea la Evaluación de las posibilidades presentadas por diversas estructuras de datos multidimensionales en el campo de las búsquedas asociativas. Estructuras y

esquemas de búsqueda para el almacenamiento y recuperación de cadenas de caracteres mediante distancias evaluadoras de su similitud.

Herramientas de ayuda a la elaboración de documentos.

- Análisis de textos.
- Tratamiento de información textual.
- Recuperación de información.
- Morfología del español.
- Desambiguación funcional.
- Sintaxis del español.
- Lingüística computacional.
- Procesamiento de lenguaje natural.

EdAS:

En (García M. A., 2013) se relaciona la formación de palabras e incluyendo su morfología además de utilizar distintos arboles de dependencias para el estudio semántico de las oraciones buscadas.

Se plantea EdAS como el único editor web que existe para realizar análisis sintáctico, semántico y morfológico. EdAS es una nueva herramienta informática para practicar sintaxis a nivel académico e ideal para pizarras digitales.

Está comprobado que las nuevas tecnologías fomentan y favorecen la atención de los alumnos. Por ello, este editor de análisis sintáctico es una manera de acercar la sintaxis a nuestros alumnos, de un modo distinto, mucho más visual y atractivo.

Se evalúa el desempeño en la clasificación de datos a partir de los resultados obtenidos en la aplicación de algoritmos basados en árboles de decisión.

Los algoritmos de clasificación que utilizamos para evaluar a los árboles son los siguientes: (ID3, J48, Naive Bayes). Se consideró a un conjunto de prueba (322 casos El experimento se hizo con el propósito de analizar qué conjunto de datos es más significativo, a partir del porcentaje de casos que clasifica correctamente.

Healthline:

Este experimento fue realizado en Departamento de Patología del Hospital “Royal Hallamshire” en Sheffield, Reino Unido. Para más información revisar la bibliografía (Cross & Cols, 2000) **Healthline.com** ofrece una serie de herramientas que permiten buscar recomendaciones médicas en base a síntomas o tratamientos indicados por el usuario.

Symptom Checker. Indica al usuario las posibles causas del síntoma proporcionado como entrada, dichas causas las ordena de las más a las menos comunes. También muestra los tratamientos relacionados con la condición indicada como entrada, o las condiciones en las cuales se aplica el tratamiento indicado como entrada.

La información publicada en el sitio es revisada por un grupo de científicos expertos en medicina, los cuales la editan y le realizan los ajustes necesarios para que pueda ser comprendida por los usuarios no expertos. La información presentada no puede ser utilizada para reemplazar a un médico.

Towards semantic search and inference in electronic medical records: An approach using concept-based information retrieval:

En este artículo se presenta una propuesta para la búsqueda en registros médicos electrónicos, la cual es basada en concordancia de conceptos en vez de concordancia entre palabras claves. La metodología utilizada transforma las consultas y los documentos médicos de su formato original (lenguaje natural) a una ontología llamada SNOMED-CT. La ontología SNOMED-CT permite modelar el conocimiento médico el cual incluye: Procedimientos, Organismos, Estructura corporal, Farmacéutica.

Los conceptos están organizados en una jerarquía heredada que permite definir relaciones con otros conceptos. SNOMED-CT tiene aproximadamente 390,000 conceptos y 1, 400,00 relaciones.

Los resultados obtenidos muestran que la búsqueda basada en conceptos genera mejores resultados que la basada en palabras, sin embargo estos resultados dependen completamente de la correcta extracción de conceptos obtenidos de manera automática con la herramienta.

MetaMap (herramienta de procesamiento de lenguaje natural desarrollada por la biblioteca nacional de medicina de los Estados Unidos de América). Los errores en la extracción de conceptos impactan fuertemente al rendimiento de la herramienta. Es importante mejorar MetaMap para obtener mejores resultados. (Koopman, 2007)

Khresmoi – multilingual semantic search of medical texts and images

Es un proyecto que lleva dos años en desarrollo, el cual pretende crear un sistema de búsqueda multilingüe y multimodal para documentos e información relacionada. Uno de los problemas más grandes que debe enfrentar este sistema es la búsqueda sobre grandes cantidades de información médica, incluyendo:

- 1) Información médica general disponible en internet, Imágenes de radiología en 2D y 3D almacenadas en archivos de hospitales.
- 2)) Imágenes de resonancia magnética y tomografía por computadora.

El sistema permite la consulta tomando como entrada el texto en múltiples lenguajes proporcionado por el usuario, así como consultas basadas en imágenes. El sistema está dirigido principalmente a tres usuarios: público en general, médicos y radiólogos. El prototipo desarrollado hasta el momento permite realizar una búsqueda mediante una consulta ingresada en una caja de texto y los resultados son documentos relacionados con la consulta ordenada, de acuerdo a varios criterios (no definidos en el artículo) y a varios filtros establecidos (tampoco definidos).

(Cervantes R. , 2002)

Towards a semantic medical Web: HealthCyberMap's tool for building an RDF metadata base of health information resources based on the Qualified Dublin Core Metadata Set:

Propone estructurar los recursos de información médica disponibles en Internet en una forma novedosa que permita mejorar la recuperación de información y la navegación sobre esta.

Para lograr este objetivo se recolecta una colección de metadatos de recursos apropiados de forma no ambigua, lo cual permite conservar la semántica. En este artículo se modeló una ontología mediante un conjunto de metadatos Dublin Core (DC) con elementos extras para indicar la calidad y procedencia del recurso. El modelado de esta ontología se realizó en Protégé-2000 y mediante un plug-in disponible en esta herramienta llamado UMLS tab es posible realizar búsquedas sobre conceptos, y la información recuperada es agregada al conjunto de metadatos DC.

Los metadatos resultantes pueden ser utilizados con un motor de búsqueda e inferencia, así como integrarles una interfaz de navegación textual o visual. Con estos elementos finalmente se podría construir un portal médico con características de web semántica.

La explotación de los metadatos generados en este artículo quedan pendientes y sólo se menciona como podrán ser utilizados más no lo han realizado hasta la fecha. (Jr, 1996)

Making ontologies work for resolving redundancies across documents:

En (Haciendo ontologías para resolver las redundancias en los documentos) tienen como objetivo desarrollar un sistema para identificar conceptualmente documentos similares identificando los fragmentos que se superponen, es decir las redundancias.

Con la finalidad de cumplir con su objetivo, se desarrolló un prototipo de un sistema, mediante el uso de una ontología. Dicha representación se obtiene analizando un texto en lenguaje natural, del cual se obtienen estructuras semánticas - sintácticas de una frase y se construyen representaciones conceptuales con base a una ontología.

Cabe destacar que un texto en lenguaje natural presenta descripciones de eventos relacionados, es por ello que la ontología debe considerar sucesos, entidades y relaciones para comparar eventos descritos en diferentes niveles de especificidad. Por ejemplo: Un evento de limpieza presenta funciones para el agente que hace la limpieza, el objeto que limpia, el instrumento para llevar a cabo la limpieza y la suciedad que va a quitar.

Su enfoque inicia con una ontología básica y añade descripciones mejorando el desempeño del sistema, el proceso de concordancia se inicia mediante la búsqueda de correspondencia entre los conceptos del nivel medio, que a su vez puede tener correspondencia con los conceptos del nivel bajo; haciendo una comparación de grano fino. Un ejemplo donde se deben determinar las descripciones es: “hace que el cable sea más flexible” vs “hizo que el cable es demasiado rígido”, pero, ¿cómo determinar que son hechos similares? La ontología propuesta considera dimensiones tales como: alto/bajo para la talla; profundo/superficial para la profundidad y frío/caliente para la temperatura. Asimismo, se considera dentro de la ontología la polaridad implícita en los adjetivos multidimensionales por ejemplo “alta polaridad equivale a menor rigidez”.

El prototipo propuesto sólo ha sido probado por un conjunto de 15 pares de categorías similares pero es la base para definir una ontología que maneje conocimiento a gran escala. (O., y otros, 2003)

Árboles de decisión como herramienta en el diagnóstico médico:

En este artículo se evalúa el desempeño de clasificación de datos médicos a partir de los resultados obtenidos en la aplicación de algoritmos basados en árboles de decisión.

Para llevar a cabo este trabajo de investigación se manejaron dos bases de datos (Véase en la (Cross & Cols, 2000) que contienen información acerca de tumores (malignos o benignos) para la detección de cáncer de mama.

Los algoritmos de clasificación que se utilizan para evaluar a los árboles son los siguientes: (ID3, J48, Naive Bayes).

Se consideró a un conjunto de prueba (322 casos). El experimento se hizo con el propósito de analizar qué conjunto de datos es más significativo, a partir del porcentaje de casos que clasifica correctamente. (Jr, 1996)

Este experimento fue realizado en Departamento de Patología del Hospital “Royal Hallamshire” en Sheffield, Reino Unido. Para más información revise (Erandi, y otros, 2013).

Otros analizadores.

Son algunos que se vieron y que se utilizaron como antecesores para comparaciones ya que también son analizadores muy semejantes al analizador semántico se pretende obtener o realizar. El principal es wólfam alpha, hay otros que trabajan casi con el mismo principio pero no se mencionan como el Hakia y Lucene-SIREN que se explicaran brevemente a continuación.

Capítulo 3

PROCESAMIENTO DE LENGUAJE NATURAL

En este Capítulo se observan algunas definiciones teóricas que se deben conocer y se describen los antecedentes sobre el analizador semántico.

Los antecedentes del analizador semántico que hoy en día se utilizan, dependen de un análisis léxico ya que se requiere del reconocimiento de caracteres de entrada para formar componentes y así poder identificarlos, clasificarlos y poderlos ordenar; los componentes que integran el analizador semántico son otros analizadores los cuales son (El Analizador Léxico Grafico, El Analizador Sintáctico y el Analizador Morfológico), el primer componente del Analizador Semántico es el Analizador Léxico Grafico, el cual se encarga de eliminar espacios en blancos, tabuladores, saltos de línea y reconocer símbolos los cuales son válidos o inválidos entre otras cosas.

Como se menciona el analizador léxico grafico es una parte fundamental del Analizador Semántico ya que sin él no habría un lenguaje o vocabulario que analizar. Un ejemplo práctico es cuando tenemos un celular que se prende y se apaga con el reconocimiento de las palabras de prender y apagar; si se le añade otra palabra no la procesará como apágate o enciéndete no lo manejará por que no están explícitas en su lenguaje o vocabulario, de tal forma que si al celular por medio de programación se agranda su lenguaje o vocabulario, con esto ya aceptará las palabras de enciéndete y apágate, entonces podrá realizar las funciones mencionadas.

Otro ejemplo del analizador léxico grafico es cuando uno intenta agregar vocabulario como los sinónimos. Como humanos entendemos una gran cantidad de palabras ya sea estudiando, viajando o por los medios de comunicación.

Los componentes del analizador léxico grafico son:

- Leer caracteres de entrada
- Generar una secuencia de componentes léxicos
- Tokens
- La eliminación de comentarios
- Símbolos de puntuación

3.1 Inteligencia artificial.

Para (Pajares & Santos, 2006) El concepto de Inteligencia Artificial es el siguiente: “La inteligencia Artificial, IA, se puede definir como el estudio de las facultades mentales a través del uso de modelos computacionales”. Otra definición muy extendida es la de Hayes que considero que la inteligencia Artificial es “la implementación de razonamientos inteligentes mediante técnicas propias de la Computación”.

Pero para (Gomez, 1990) Es un área la cual abarca muchas disciplinas las cuales son la informática, la lógica y la filosofía estudia la creación y diseño de entidades capaces de razonar por sí mismas utilizando como paradigma la inteligencia humana. “Es la Ciencia e ingenio de hacer maquinas inteligentes, especialmente programas de cómputo inteligentes”.

Una de las aportaciones de la inteligencia artificial es el desarrollo de sistemas inteligentes que pueden igualar la capacidad mental del hombre y relacionar reglas de sintaxis del lenguaje hablado y escrito sobre la base de la experiencia, para luego hacer juicios acerca de un problema, cuya solución se logra con mejores juicios y más rápidamente que el ser humano.

Caracterizar los sistemas inteligentes incluye las tareas de:

- Reconocerlos y distinguirlos de los sistemas convencionales
- Clarificar el papel de la inteligencia en estos sistemas
- Identificar los problemas para los cuales la Inteligencia Artificial es la única o la mejor solución.

La inteligencia artificial como ciencia es para (Palma & Morales, 2008) La perspectiva científica de la IA busca una teoría computable del conocimiento humano. Es decir una teoría en la que sus modelos formales puedan ejecutarse en un sistema de cálculo y tener el mismo carácter predictivo que tienen, las ecuaciones de Maxwell en el electromagnetismo.

Siempre ha habido un fuerte debate asociado a las relaciones entre la IA y la cognición con las analogías mente-programa, cerebro-hardware y las consiguientes posturas a favor y en contra de la equivalencia entre pensar y computar.

La inteligencia artificial como ingeniería es para (Palma & Morales, 2008) la rama aplicada de la IA, conocida Ingeniería del conocimiento, el conocimiento es ahora de observación, modelado formalización y transformación por procesos de su mismo nivel o de meta niveles superiores. (Aprendizaje)

En la siguiente imagen se muestra las diferencias de la inteligencia artificial como ciencia y la inteligencia artificial como ingeniería. Ver la imagen 1.

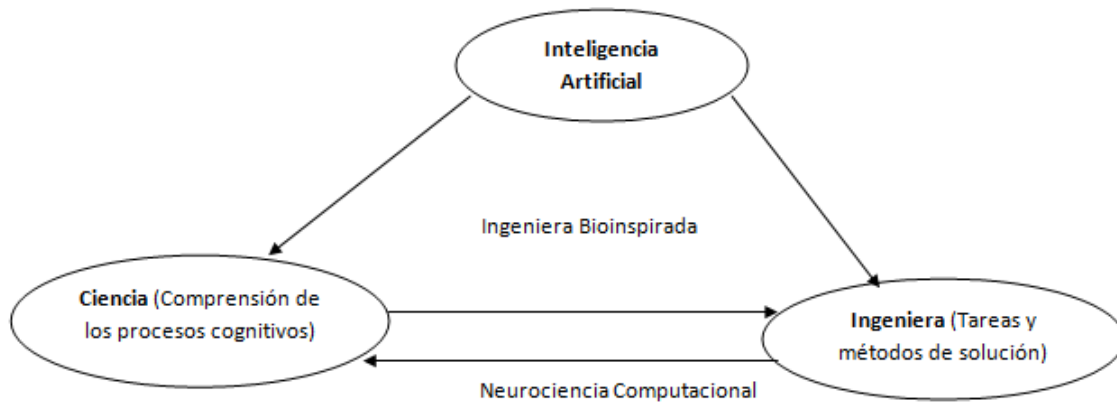


Imagen 1. *Distinción entre IA como ciencia e IA como ingeniería (Pajares & Santos, 2006)*

3.1.1 Perspectiva histórica: de la Inteligencia Artificial.

Pasaron 50 años después de acuñar el término de Inteligencia Artificial, lo poco que se ha investigado en sus fundamentos y en metodología, y la excesiva prisa mostrada en desarrollar aplicaciones. Tras estos algunos años en 1956, se acuña el término de Inteligencia Artificial y se abandona el conexionismo y el interés por el análisis del comportamiento y de los mecanismo soporte y se hace énfasis en el paradigma simbólico, que parte de descripciones de alto nivel para intentar dotar de inteligencia artificial a un programa, que siempre se supone que va a tener un usuario humano por lo que no hay que preocuparse del cuerpo del “agente” soporte del cálculo. (Pajares & Santos, 2006)

El optimismo inicial de la Inteligencia Artificial estuvo basado en limitar su dominio de interés a los micromundos formales (“mundo de los bloques”) y a los sistemas capaces de ser descritos de forma completa con métodos lógicos, mediante procedimientos generales de búsqueda. (Eslava, 2003)

Aunque por el paradigma simbólico (1956-1986) hubo un gran número relevante de trabajos basados en los principios de la cibernética, es en 1986 [Rumelhart y otros, 1986] cuando renace con fuerza la aproximación neuronal al problema de la inteligencia.

Para (Palma & Morales, 2008) En la década de los 90 [Arkin, 1998; Brooks, 1991] sale a la luz el interés de la aproximación situada, cerrando el lazo histórico que comenzó en 1943 con el reconocimiento de la necesidad de usar los tres paradigmas tanto en la perspectiva teórica como en la práctica.

De hecho tal debate, tal como ya hemos mencionado, el debate más importante en la IA no ha estado relacionado con la ingeniería, si no con la potencial validez del paradigma simbólico para explicar los procesos cognitivos.

Para (Palma & Morales, 2008) Los tres movimientos dignos de interés en el aspecto metodológico más allá de las propuestas iniciales de Leibniz, Turing y McCulloch son

1. El énfasis en la reutilización de componentes de modelado, a través del desarrollo de bibliotecas de tareas y métodos
2. El desarrollo de ontologías y servidores de terminología unificada
3. El trabajo de Newell, Marr, Maturana y Varela para especificar un marco de niveles y dominios de descripción de un cálculo que permite especificar de forma inequívoca que parte del conocimiento termina residiendo en la arquitectura de un computador y cual permanece en el dominio del observador externo y por consiguiente todavía no es computable.

3.1.2 Sistemas basados en conocimientos.

Un sistema basado en conocimiento (SBC) es una aplicación informática en la que aparece representado, como estructura de información procesable, el conocimiento necesario para resolver un determinado tipo de problemas separadamente del procedimiento para resolverlos.

Un sistema experto (SE) incorpora una significativa cantidad de conocimientos expertos acerca de un dominio bien delimitado y restringido, y puede usar técnica de razonamiento simbólico para resolver problemas en dicho dominio. Las características fundamentales del tipo de conocimiento involucrado es que es subjetivo, incompleto, incierto y sujeto a cambios.

Por lo tanto, un sistema experto es básicamente un programa basado en conocimiento que incorpora de forma computacional el conocimiento de un experto humano para realizar una tarea X en un dominio Y, de forma que para esa tarea, la eficiencia que subyace del programa es comparable a la que muestra el experto humano. Es esencial que el dominio Y sea lo más preciso posible y limitado en opciones, y que la tarea X pueda definirse de forma clara y razonable inequívoca.

Una de las características más significativas de los sistemas basados en reglas, que facilita la adquisición y mantenimiento del conocimiento, es la separación que establece entre el propio conocimiento, relativo al problema que se pretende resolver, y los procedimientos generales que lo manipulan, que son siempre los mismo; esta separación no es clara en los sistemas tradicionales en los cuales los programas contienen conocimiento sobre el problema. (Pajares & Santos, 2006)

Un sistema experto o sistema basado en conocimiento se puede definir como “Sistema que resuelve problemas utilizando una representación simbólica del conocimiento [Jackson 86]”.

Características importantes son:

- Representación explícita del conocimiento

- Capacidad de razonamiento independiente de la aplicación específica
- Capacidad de explicar sus conclusiones y el proceso de razonamiento
- Alto rendimiento en un dominio específico
- Uso de heurísticas contra el modelo matemáticos
- Uso de inferencia simbólica contra algoritmo numérico.

Algunas de esas propiedades se deben a la separación entre:

1. Conocimiento específico del problema (Base de Conocimiento)
2. Metodología para solucionar el problema (Maquina de inferencia)

(Eduardo, 2009)

3.2 Introducción al Procesamiento del lenguaje Natural.

El PLN es una parte esencial de la Inteligencia Artificial que investiga y formula mecanismos computacionales efectivos que faciliten la interrelación hombre/máquina, además intenta simular un comportamiento lingüístico humano. Para ello debe tomar conocimiento tanto de las estructuras propias del lenguaje, como el conocimiento general acerca del universo de discurso.

Una persona que participa en un diálogo sabe cómo combinar las palabras para formar una oración, conoce los significados de las mismas, sabe cómo estos afectan al significado global de la oración y posee un conocimiento del mundo en general que le permite participar en la conversación. El procesamiento del lenguaje natural se encarga de mantener constante comunicación tanto con las computadoras y los seres humanos de tal forma que allí una comunicación.

3.2.1 Historia del Procesamiento del lenguaje Natural (PLN).

Disciplina que inicia en 1940 y sus aplicaciones comenzaron a aparecer 2 años más tarde donde tuvo particular interés en la traducción automática. A principios la traducción consistía simplemente en el remplazo de palabras pero esto degeneró en problemas de sintaxis y semántica además de la falta de una estructura oracional de algunas lenguas. En 1970 la influencia de la inteligencia artificial tuvo interés ya no en la traducción del texto si no que optó por la realización de una presentación del significado, en esta época se construyó Elisa el primer sistema de preguntas y respuestas basado en el lenguaje natural, parte de los 70 y comenzó en los años 80 aparecen gramáticas orientadas al tratamiento computacional de la información. En esta época tuvo gran crecimiento la programación lógica como los lenguajes de programación a principio de los años 80 en vez de buscar nuevos lenguajes de programación se involucraron con lo que ya existían con esto hubo ese aumento. En la actualidad se utilizan técnicas estadísticas y de representación de conocimiento basadas en la **inteligencia artificial** que buscan mejorar los resultados de consultas de texto.

El PLN se define como “el reconocimiento y utilización de la información expresada en el lenguaje humano a través del uso de sistemas informáticos” para esto PLN hace uso de varias disciplinas del saber humano entre ellas: filosofía, matemática y psicología el tratamiento del lenguaje natural comprende según (Sosa, 1997). 4 fases:

- **Morfológico:** etapa temprana que se encarga de verificar que las unidades que forman una palabra pertenezcan al léxico del que se está tratando.
- **Sintáctico:** en esta etapa es la encargada de realizar si todas las combinaciones de los componentes de una oración son gramaticalmente correctas.
- **Semántico:** es la fase encargada de la representación del significado.
- **Pragmático:** encarga de añadir información en base al contexto en el que se está trabajando.

Una de las aplicaciones del procesamiento del lenguaje natural (PLN) está basada en el texto así el PLN como “Puede ser visto como una forma de abstraer textos que estamos procesando en una representación interna más concreta que nos facilite su manejo dentro de una aplicación o a la hora de comprender su significado”. (Sosa, 1997).

Tabla 2. Niveles de Conocimiento en el procesamiento del Lenguaje Natural.

Nivel	Declarativo (qué)	Procedural (cómo)
Fónico	Sonidos que hablamos	Formar Fonemas
Morfológico (léxico)	Unidades de palabra, Palabras	Formar Palabras , Derivar unidades de significado
Sintáctico	Roles Estructurales de palabras (o colección de palabras)	Formar oraciones
Semántico	Significado independiente del contexto	Derivar significado de oraciones
Discurso	Roles Estructurales de oraciones (o colección de oraciones)	Formar diálogos
Pragmático	Significado dependiente del contexto	Derivar significado de oraciones relativo al discurso circundante

Algunos de los problemas existentes en el PLN son:

1. Ambigüedades léxicas
2. Ambigüedad estructural
3. Relaciones Gramaticales

Un ejemplo donde se presentan estos problemas es: “Ana dijo que fue de noche cuando se torció el pie” donde fue se puede considerar como dos verbos de ser o de ir (ambigüedad léxica); y cuando (de noche) se torció el pie (relaciones gramaticales).

Como se puede apreciar varios problemas pueden existir en un mismo texto, los cuales pueden ser solucionados mediante un Analizador Semántico.

3.3 Procesamiento del lenguaje natural.

El recurso más importante que posee la raza humana es el conocimiento, o sea información, en la época actual de información, el manejo eficiente de este conocimiento depende el uso de todos los demás recursos naturales, industriales y humanos.

El esquema general de la mayoría de los sistemas y métodos que involucran el procedimiento de lenguaje es el siguiente:

- Primero, el texto no se procesa directamente si no se transforma en una representación formal que preserva sus características relevantes para la tarea o el método específico (por ejemplo un conjunto de cadenas de letras, una tabla de base de datos, un conjunto de predicados lógicos, etc.).
- Luego, el programa principal manipula esta representación, transformándola según la tarea, buscando en ella las subestructuras necesarias.
- Finalmente, si es necesario los cambios hechos a la representación formal (o la respuesta generada en esta forma) se transforma en el lenguaje natural.
- Entre las tareas principales del procesamiento de lenguaje natural se puede mencionar (Recuperación de información, Interfaces en el lenguaje natural y traducción automática). ((CIC-IPN), 2009)

3.3.1 Recuperación de información.

La aplicación del procesamiento de lenguaje natural más obvia y quizá más importante en el momento actual es la búsqueda de información (se llama también recuperación de información). Por un lado, en Internet y en las bibliotecas digitales se contiene una cantidad enorme de conocimiento que puede dar respuestas a muchísimas preguntas que tenemos. Por otro lado, hay tanta información que no sirve porque ya no se puede encontrarla. Hoy en día la pregunta ya no es “¿si se sabe cómo...?” sino “ciertamente se sabe, pero ¿dónde está esta información?”.

Técnicamente, rara vez se trata de decidir cuáles documentos (así se llaman archivos o textos en la recuperación de información) son relevantes para la petición del usuario y cuáles no. Usualmente, una cantidad enorme de documentos se puede considerar como relevantes en cierto grado, siendo unos más relevantes y otros menos. Entonces, la tarea se entiende cómo medir el grado de esta relevancia para proporcionar al usuario primero el documento más relevante; si no le sirvió, el segundo más relevante, etc.

El problema más difícil de la recuperación de información es, sin embargo, no de índole técnica sino psicológica: entender cuál es la necesidad real del usuario, para qué formula su pregunta.

Este problema se complica ya que no existe un lenguaje formal en el cual el usuario podría formular claramente su necesidad. (La dirección más prometedora de resolver este problema es, nuevamente, el uso de lenguaje natural.)

Las técnicas más usadas actualmente para la recuperación de información involucran la búsqueda por palabras clave: se buscan los archivos que contengan las palabras que el usuario teclee. Es decir, la representación formal usada es el conjunto de las cadenas de letras (palabras), usualmente junto con sus frecuencias en el texto (en número de ocurrencias). La claridad matemática de la tarea causó mucho avance en la teoría de estos métodos. Las ideas más usadas son los modelos probabilísticos y los procedimientos iterativos e interactivos: tratar de adivinar qué necesita el usuario preguntándole cuáles documentos le sirven).

Sin embargo, los métodos que involucran sólo las palabras (como cadenas de letras) pero no el sentido del texto son muy limitados en su capacidad de satisfacer la necesidad informática del usuario, es decir, de hallar la respuesta a la pregunta que tiene en mente. Se puede mejorar mucho aplicado las siguientes operaciones, desde las más sencillas hasta más complejas:

- Coincidencia de las formas morfológicas de palabras: buscando *pensar*, encontrar *piénsalo*. Este problema es bastante simple a resolver en el lenguaje inglés, al cual se dedica la mayor parte de investigación en el mundo. Sin embargo, para el español se convierte a un problema moderadamente serio, debido a la gran variedad de las formas de palabras en español.
- Coincidencia de los sinónimos, conceptos más generales y más específicos: buscando *cerdo*, encontrar *puerco*, *mascota*, *animal*, etc. Este problema prácticamente no depende de lenguaje (es tanto importante para el inglés como para el español), aunque los diccionarios que se usan sí son específicos para cada lenguaje.
- Tomar en cuenta las relaciones entre las palabras en la petición del usuario y en el documento: buscando *estudio de planes*, rechazar como no relevante *planes de estudio*. Para lograr este grado de calidad, se necesita reconocer (automáticamente) la estructura del texto y representarla en forma que permita la comparación necesaria, por ejemplo, en la forma de los grafos conceptuales. ((CIC-IPN), 2009)

3.3.2 Interfaces en Lenguaje Natural.

Las computadoras están entrando en todos los aspectos de nuestra vida cotidiana: en las oficinas, en las tiendas, en las escuelas, en los servicios públicos. Sin embargo, la gran mayoría del pueblo no tiene la preparación adecuada para usarlas ya que se requiere saber cómo se usan. Y nunca va a tener esta preparación, por el simple hecho: económicamente, más conviene que las máquinas se adapten al modo de la comunicación de las personas y no que todas las personas (las cuales sólo en el mundo hispanohablante son 400 millones), generación tras generación, aprendan cómo usar las máquinas –por ejemplo, aprendan el SQL para formular con precisión sus preguntas.

De esto surge la idea ya muy conocida de las películas de falta ciencia: las personas debemos poder hablar con las máquinas (o sea, los robots) como hablaríamos con nuestros sirvientes, dándoles órdenes en la forma cotidiana y escuchando sus respuestas.

Hablando de darles órdenes, no se trata de pronunciar los comandos especiales que de otra manera escogeríamos del menú: *abrir, edición, copiar, guardar, salir* (de modo similar a como se habla al perro pastor). En lugar de esto, se trata de hablar a la máquina como hablaríamos al ser humano.

Un tipo específico de las interfaces en lenguaje natural son preguntas complejas a una base de datos. Como un ejemplo, podemos mencionar el sistema TRAINS desarrollado en la Universidad de Rochester en los EE.UU. por James Allen. Este sistema vende los boletos de tren. El cliente – que puede ser cualquier persona sin ningún conocimiento sobre las máquinas– llama por teléfono a la estación de trenes y formula su necesidad: tengo que ir mañana en la tarde a Nueva York. El programa –sin siquiera darse cuenta el cliente de que habla con una máquina y no con una persona– trata de entender su pregunta e internamente la traduce a SQL, para ejecutar la búsqueda en su base de datos. Después, el programa conduce (por teléfono) el diálogo con el usuario, explicándole los precios y las condiciones, escuchando sus preguntas o comentarios sobre cuál boleto más le conviene, etc. Finalmente, si llegan al acuerdo, le reserva el boleto. Todo eso, nuevamente enfatizamos, no requiere del cliente ningún conocimiento previo sobre el manejo de los programas sino sólo el manejo natural de lenguaje que cotidianamente usa para hablar con otras personas.

El problema más importante de este tipo de aplicaciones es que –a diferencia de las aplicaciones en la recuperación de información– se requiere entender exactamente la intención del usuario, ya que el costo de error puede ser muy alto. Realmente, si el robot entiende incorrectamente el comando, puede hacer alguna acción destructiva o peligrosa. Si se malentiende la pregunta a la base de datos, la información proporcionada resultará incorrecta, lo que puede causar consecuencias graves.

Entonces, las interfaces en lenguaje natural en muchos casos requieren de las representaciones de información más detalladas y complejas, así como del análisis lingüístico más preciso y completo. ((CIC-IPN), 2009)

3.3.3 Traducción Automática.

Históricamente, el sueño de la traducción automática motivó las primeras investigaciones en lingüística computacional. A la primera vista, la traducción parece ser un trabajo bastante mecánico y aburrido, que puede fácilmente hacer la máquina: sustituir las palabras en un lenguaje con sus equivalentes en el otro.

Sin embargo, con los avances en los programas de traducción se hizo cada vez más obvio que esta tarea no es tan simple. Esto se debe, en una parte, a las diferencias entre los lenguajes las cuales varían desde muy obvias (por ejemplo, el orden de palabras es diferente en inglés y en español) hasta más sutiles (el uso de diferentes expresiones y diferente estilo).

El esquema general de prácticamente cualquier traductor automático es (en acuerdo con el esquema expuesto en la introducción a este artículo) el siguiente:

- El texto en el lenguaje fuente se transforma a una representación intermedia,

- De ser necesario, se hacen algunos cambios a esta representación,
- Luego, esta representación intermedia se transforma al texto en el lenguaje final.

En algunos sistemas, al texto generado con este esquema también se aplican algunos ajustes previstas por las heurísticas de traducción.

De la calidad esperada de traducción y de la proximidad de los dos lenguajes depende qué tan profundas son las transformaciones entre las dos representaciones, es decir, que tan diferente es la representación intermedia del texto en el lenguaje humano.

En algunos casos (se puede mencionar los traductores entre los lenguajes tan parecidos como español y catalán, portugués y gallego, etc.) basta con la representación morfológica: el análisis y generación de las palabras fuera del contexto. Por ejemplo, la palabra española *hijas* se analiza como HIJO-femenino-plural, se transforma (usando una tabla de correspondencias) a la representación FILHA-femenino-plural, de la cual se genera la palabra portuguesa *filhas* (aquí, HIJA y FILHA son claves de acceso a la base de datos que contiene las propiedades de las palabras en los lenguajes correspondientes).

En otros casos, cuando hay diferencias estructurales más profundos entre los lenguajes (que es el caso de casi cualquier par de idiomas, incluidos español-inglés, español-francés, etc.), se usan como la representación intermedia (que en este caso se llama interlingua) las estructuras formales de predicados lógicos o sus equivalentes, por ejemplo, redes semánticas (véase más abajo). Esta representación independiente de lenguaje es en realidad el sentido del mensaje que comunica el texto. Es decir, la transformación del texto a esta representación formal es comprensión del texto, y la transformación en el sentido contrario es generación: teniendo una idea, decirlo en el lenguaje humano.

De esta discusión queda completamente claro que en el caso de la traducción automática (de tipo interlingua) es aún mucho más importante que el programa comprenda el texto perfectamente y, por otro lado, pueda correctamente verbalizar el sentido dado. Cualquier error en comprensión del texto fuente causaría la traducción incorrecta, lo que, dependiendo de la situación del uso del texto traducido, puede resultar en consecuencias graves.

Entonces, este tipo de los sistemas de traducción requiere de toda la fuerza de la ciencia de lingüística computacional, los métodos más precisos y completos del análisis del texto y representación de su contenido.

((CIC-IPN), 2009)

3.3.4 Enfoque de dependencias.

En este enfoque las dependencias se establecen entre pares de palabras donde una es principal o lectora y la otra subordinada a (o dependiente de), la primera palabra de la oración tiene una palabra propia lectora la oración entera se ve como la estructura jerárquica de diferencia de

niveles o sea como un árbol de dependencias y la única palabra que no está subordinada a otra se la raíz del árbol.

La motivación de muchas dependencias sintácticas es el sentido de las palabras ya que sin este sentido fallaría al encontrar el significado de las palabras. (Calderon, 2001)

3.3.5 Introducción a las etiquetas Eagles.

Con el analizador morfológico para el castellano utiliza un conjunto de etiquetas para representar la información morfológica de las palabras este conjunto de etiquetas se basa en las etiquetas propuestas por el grupo de Eagles.

Para la anotación morfosintáctica de lesiones y corpus para todas las lenguas europeas así que está previsto que recojan los accidentes gramaticales existentes en las lenguas europeas es por esto que dependiendo de la lengua hay atributos que pueden no especificarse. Si un atributo no se especifica significa que o bien expresa un tipo de información que no existe en la lengua o que la información no se considera relevante la especificación del atributo se marca con el 0. (Ruben, 2011)

A continuación se muestra la imagen de cómo se clasifica una etiqueta, ver la imagen 2 a continuación.

ETIQUETAS			
Posición	Atributo	Valor	Código
Columna 1	Columna 2	Columna 3	Columna 4

Imagen 2. Ejemplo: de cómo está constituida una etiqueta.

3.3.6 Redes Semánticas.

La red semántica o esquema de representación en red es una forma de representación de conocimiento lingüístico en la que los conceptos y sus interrelaciones se representan mediante un grafo en caso de que no existan ciclos estas redes pueden ser visualizadas como arboles las redes semánticas son usadas entre otras cosas para representar mapas conceptuales y mentales.

En un grafo o red semántica los elementos semánticos se representan por nodos dos elementos semánticos entre los que se admiten se le da la relación semántica que representa la red estarán unidos mediante una línea, flecha o enlace. (Allan, 2010)

(Palma & Morales, 2008) Las redes semánticas (o redes conceptuales) son un formalismo o paradigma de representación de conocimiento basado en relación; es decir; este formalismo centra su atención en las reglas relacionadas que aparecen entre los conceptos o entidades de un dominio. Básicamente una red semántica [Quillian, 1968] es un grafo orientado formado por nodos etiquetados, que representan conceptos e instancias, arcos unidireccionales etiquetados, que representan conceptos o instancias.

En la siguiente imagen se muestra los elementos del grafo etiquetado y sus diferentes interpretaciones, del lenguaje natural y forma; Ver la imagen 3 a continuación.

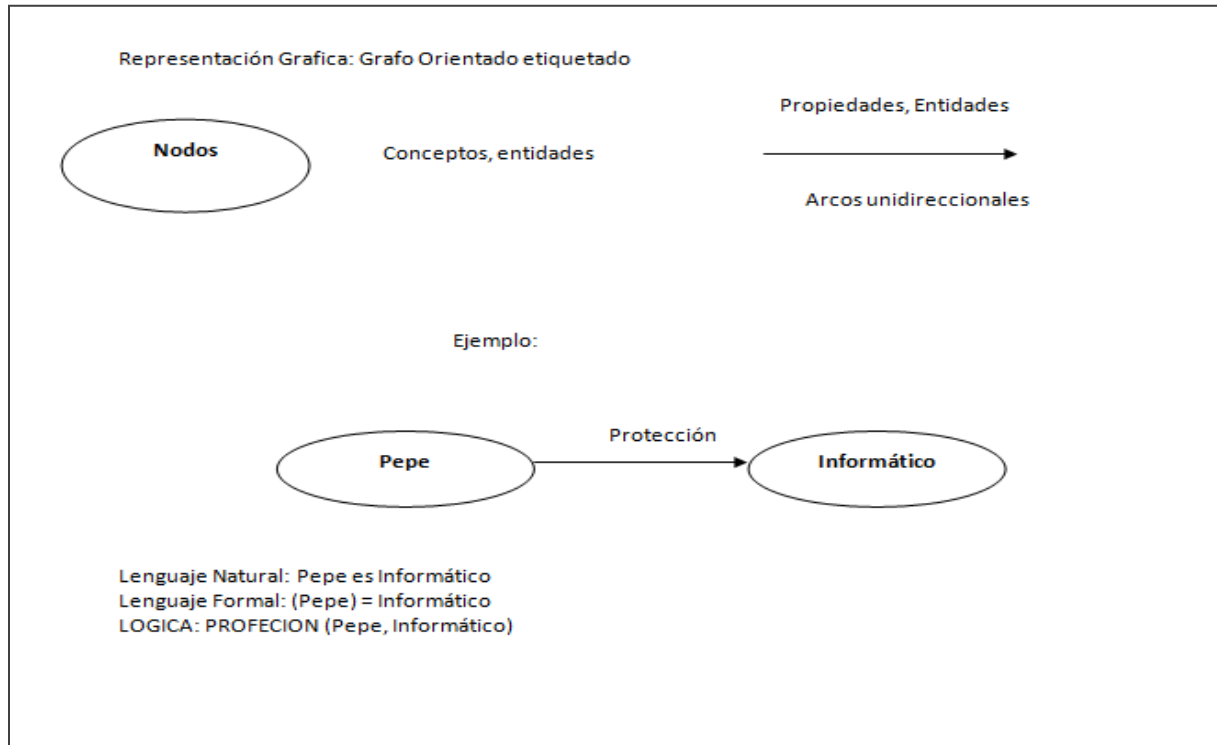


Imagen 3. Representación Gráfica: Grafo Orientado etiquetado

Conceptos Básicos en redes semánticas.

Es una red de representación del conocimiento mediante nodos (elementos del conocimiento o conceptos) y ramas o arcos (relaciones entre nodos, estas pueden ser de herencia o de descripción), fueron propuestas por [Quillian y Collins] en 1968 su concepción se basa en la asociación de conocimientos que realizan la mente humana. (Rodríguez C. , 2012).

Para (Pajares & Santos, 2006) las redes semánticas se utilizan para definir el significado de un concepto mediante su relación con otros conceptos. Para representarlas se suele utilizar una estructura de grafos de datos donde los conceptos se representan en los nodos y los enlaces definen las relaciones entre ellos. Los enlaces están etiquetados en lenguaje natural y especifican la naturaleza de esas relaciones.

También guardan una gran semejanza con la representación y la terminología de la programación orientada a objetos. Fue una técnica pionera en el modelado de objetos propuesta por [Quillian en 1968], recogiendo trabajos suyos previos de principios de los años 60.

El mecanismo por el que una red semántica permite inferir conocimiento se denomina **propagación de la activación** muestra como dos conceptos o nodos están relacionadas, aunque puede no ser evidente, tanto los marcos como las redes semánticas son paradigmas muy cercanos a la lógica de predicados. [Russell y Norving] propusieron en 1995 un mecanismo para transformar el conocimiento de ambas representaciones en expresiones de la lógica de primer

orden. La principal diferencia entre estas representaciones del conocimiento radica en su sintaxis, y en unas se basan en representaciones gráficas del conocimiento y otras en representaciones formales.

3.4 El Lenguaje.

Un lenguaje se considera como un conjunto frases y oraciones generalmente infinito y que se forma mediante combinaciones de palabras definidas en un diccionario previamente establecido. Estas combinaciones deben estructurarse correctamente es decir respetar un conjunto de reglas sintácticas que además deben tener sentido en un contexto dado a lo que se denomina semántica; A lo largo de la historia el ser humano ha utilizado el lenguaje para transmitir sus conocimientos, sentimientos, emociones, sensaciones, comunicarse con el resto de los humanos, ya sea de manera oral, gráfica escrita o por señas.

Cuando hablamos de lenguajes se pueden diferenciar dos clases muy bien definidas los lenguajes naturales (español, ruso, inglés, etc.) y los lenguajes formales (lenguajes de programación, lógica matemática, etc.), la diferencia entre estos lenguajes es que el lenguaje natural se centra en la comunicación entre seres humanos y lenguaje formal es creado por los seres humanos con propósitos alternos a la comunicación además de que este lenguaje formal se basa en resolver problemas. (Avilar, 1977)

3.4.1 Lenguaje natural.

El lenguaje natural fue desarrollado y organizado a partir de la experiencia humana; Podríamos definir el lenguaje natural como el medio principal para la comunicación humana con respecto a nuestro mundo el lenguaje nos permite designar las cosas reales y razonar acerca de ellas, así como también crear significados.

Una propiedad única de los lenguajes naturales es la polisemia, es decir la posibilidad de que una palabra en una oración tenga diversos significados incrementando así la riqueza semántica de una palabra, es por ello que a este tipo de lenguajes se les consideran expresivos y pueden ser utilizados para razonar situaciones complejas. (Gamallo & Ponche, 2012)

Podemos resumir que los lenguajes naturales se distinguen por las siguientes propiedades:

- Han sido desarrollados por enriquecimiento progresivo antes de cualquier intento de formación de una teoría.
- Existe dificultad o imposibilidad de una formalización completa
- Polisemia.

3.4.2 Lenguaje Formal.

Un lenguaje formal es un lenguaje artificial compuesto por símbolos y fórmulas que tiene como objetivo fundamental formalizar la programación de computadoras o representar simbólicamente el conocimiento científico. Las palabras y oraciones en un lenguaje formal están perfectamente definidas, una palabra mantiene el mismo significado prescindiendo del contexto o su uso. Los lenguajes formales son exentos de cualquier componente semántico fuera de sus operadores y relaciones y es gracias a esta ausencia de significado especial que los lenguajes formales pueden ser usados para moldear una teoría de la mecánica de la ingeniería eléctrica, en la lingüística u otra naturaleza, esto equivale a decir que durante la concepción de lenguajes formales toda la ambigüedad es eliminada. En resumen las características de los lenguajes formales son las siguientes: (Garuda, 2002)

- Se desarrollan de una teoría preestablecida
- Tienen componente semántico mínimo
- Posibilidad de incrementar el componente semántico de acuerdo con la teoría a formalizar
- La sintaxis produce oraciones no ambiguas
- Los números tienen un rol importante
- Poseen una completa formalización y por esto potencialmente posibilitan la construcción computacional.

3.4.3 Lingüística del texto.

La lingüística del texto ofreció en las últimas décadas del siglo XX un instrumental teórico-analítico de indudable valor para el conocimiento de las expresiones lingüísticas de dimensión textual por medio de las cuales se lleva a cabo la comunicación y que responden a la unidad del texto.

La lingüística del texto supuso una innovación radical no solo en los estudios lingüísticos, sino también en los estudios literarios, al proporcionar los fundamentos y líneas generales para el establecimiento de modelos de análisis y explicación de los textos de la lengua común y de los textos de lengua literaria a partir del reconocimiento y de la consideración del texto como unidad lingüística, partiendo del hecho de que la comunicación lingüística se hace por medio de textos. (Penas, 2009).

3.4.4 Niveles del Lenguaje.

La lingüística general comprende 5 niveles principales para análisis de estructura del lenguaje que son: (Catedra, 1987)

- a) Nivel Fonológico: Es el estudio de los fonemas de una lengua, los fonemas son unidades teóricas básicas postuladas para estudiar el nivel fónico-fonológico de la lengua humana.
- b) Nivel Morfológico: trata sobre la estructura de las palabras y las leyes para formas nuevas palabras a partir de unidades de significado más pequeñas llamadas morfemas.
- c) Nivel sintáctico: trata sobre como las palabras pueden unirse para construir oraciones y cuál es la función que cada palabra realiza en esa oración.
- d) Nivel Semántico: trata del significado de las palabras y de cómo se unen para dar significado a una oración.
- e) Nivel Pragmático: estudia la disciplina que toma en consideración de los factores extralingüísticos que determinan el uso del lenguaje que son (Emisor, Destinatario e Intención).

3.4.5 Textos en lenguaje natural.

Un texto en lenguaje natural es un lenguaje que usan las personas, como el español o el inglés, para diferenciarlo de los lenguajes artificiales que usan las computadoras, como el fortran o java. Los textos de uso común usados en un campo determinado, están escritos en un lenguaje natural, de forma tal que no se cuenta con un lenguaje controlado ya que el significado de una palabra puede depender por las características semánticas de otras palabras que se encuentran en el mismo texto. Por ejemplo, la palabra “caballo” puede significar cuadrúpedo, un instrumento de gimnasia, una pieza de ajedrez, por citar algunos. (Fernandez, 2004)

3.5 Ontologías

Con el avance tecnológico y científico de la humanidad; han brotado nuevos conceptos que deben relacionarse con otros para denotar un significado, la representación conceptual del conocimiento da relevancia ya que provee las deducciones a partir de un concepto dado.

La representación del conocimiento se puede realizar por medio de ontologías. *“Existen diversas definiciones de ontologías, la más difundida y aceptada es: Una conceptualización es una visión abstracta y simplificada del mundo que deseamos representar con algún propósito. Cada base de conocimientos, sistema basado en conocimiento o a gente a nivel del conocimiento está comprometido a cierta conceptualización, explícita o implícitamente. Una ontología es una especificación explícita de una conceptualización. El término se toma de la filosofía, donde una ontología es una explicación sistemática de la Existencia. Para los sistemas de Inteligencia Artificial, lo que “existe” es lo que puede ser representado.”* (Lamarca, 2013)

Una ontología almacena conceptos que están relacionados con sus significados, pudiendo así resolver problemas con la semántica (Lamarca, 2013) que una ontología es:

- a) Explícita por que define los conceptos, propiedades, relaciones, funciones, axiomas y restricciones que la componen;
- b) Formal porque es legible e interpretable por las maquinas.
- c) Una conceptualización porque es un modelo abstracto y una vista simplificada de

- las entidades que la representa;
- d) Compartida por que ha habido un consenso previo sobre la información que ha sido acordado por un grupo de expertos.

Para (Palma & Morales, 2008) una ontología es aquella rama del saber que trata sobre la esencia de las cosas que pertenece a través de los cambios. Uno de los frutos de la ontología ha sido las ontologías, para un filósofo una ontología es “un sistema particular de categorías sistematizado cierta visión del mundo” [Guarino, 1998] un ejemplo de ontología es la de [Lowe, 2006], que distingue entre universales (de un ordenador) y particulares (este ordenador).

Existen diferentes formalismos de representación de conocimientos para formalizar (e implementar) ontologías y cada uno de ellos tiene distintos componentes que pueden ser utilizados en estas tareas de formalización e implementación, sin embargo, dichos formalismos comparten un conjunto mínimo de componentes, a saber:

- **Clases:** representan conceptos tomados en su sentido más amplio en el dominio de los viajes por ejemplo: los conceptos que tenemos son: lugares, alojamientos, ciudades y medios de transporte. En la ontología, las clases están normalmente organizadas en taxonomías por medio de las cuales se pueden aplicar los mecanismos de herencia.
- **Relaciones:** que representan un tipo de asociaciones entre los conceptos del dominio y se definen formalmente como cualquier subconjunto de producto de n conjuntos es decir: $R \subset C_1 \times C_2 \times \dots \times C_n$. Las ontologías normalmente contienen relaciones binarias, cuyo primer argumento es conocido como rango.
- **Axiomas:** Según [Gruber, 1992] sirven para modernizar afirmaciones que son siempre ciertas. En las ontologías se utilizan generalmente para representar conocimientos que no se pueden definir formalmente a través de otros componentes; además sirven para verificar la consistencia de la ontología misma o la consistencia de los conocimientos almacenados en una base de conocimientos.
- **Instancias:** que se utilizan para representar elementos o individuos en una ontología.

El objeto de estudio de la ciencia de la ontología es el estudio de las categorías que existen en un dominio, el resultado de este estudio es a lo que denominamos ontología. Una ontología es un catálogo de los tipos de cosas que asumimos que existen en un dominio D desde la perspectiva de alguien que usa un lenguaje L con el propósito de hablar de D . (Bejar, 2007)

Los elementos de una ontología representan predicados, constantes, conceptos y relaciones pertenecientes a un lenguaje L cuando se usa para comunicar información sobre D , una ontología es pues un vocabulario.

Desarrollo de una ontología.

- Definir las clases que forman el dominio
- Organizar las clases en una jerarquía taxonómica
- Definir las propiedades de cada clase e indicar restricciones de sus valores
- Asignar valores a las propiedades para crear instancia

Capítulo 4

ANALIZADORES SEMÁNTICOS

En este capítulo se observara conceptos relacionados al analizador semántico los cuales son necesarios para conocer su estructura básica. Ya que el analizador semántico contiene otros analizadores como por ejemplo (léxico, morfológico y sintáctico).

Por otra parte este capítulo abarcará algunos aspectos teóricos del análisis semántico, sintáctico y léxico, además de revisar algunas funciones de los analizadores.

4.1 Análisis léxico.

Esta fase consiste en leer el texto del código fuente carácter a carácter e ir generando los tokens (caracteres relacionados entre sí), estos tokens constituyen la entrada para el siguiente proceso de análisis (análisis sintáctico). El agrupamiento de caracteres en tokens depende del lenguaje que vallamos a compilar; es decir, un lenguaje generalmente agrupara caracteres en tokens de otro lenguaje.

Los tokens pueden ser de dos tipos; cadenas específicas como palabras reservadas, puntos y comas, etc., y no específicas, como identificadores, constantes y etiquetas. La diferencia entre ambos tipos de tokens radica en si ya son conocidos previamente o no, por lo tanto y resumiendo el “analizador léxico lee los caracteres que componen el texto del programa fuente y suministra tokens al analizador sintáctico”.

El analizador léxico ira ignorando las partes no esenciales para la siguiente fase, como pueden ser espacios en blanco, los comentarios, etc., es decir, realiza la función de preprocesador en cierta medida. Los tokens se consideran entidades con dos partes: su tipo y su valor o lexema; algunos casos, no hay tipo (como en las palabras reservadas). Esto quiere decir que por ejemplo, si tenemos una variable una variable con nombre “x”, su tipo es identificador y su lexema es x.

(Ruiz, 2010)

4.1.1 Utilidad del análisis léxico.

El analizador léxico va leyendo del fichero de entrada los caracteres secuencialmente y los va agrupando, según las directivas programadas, en tokens con un significado conocido por el programador, los tokens serán la entrada del analizador sintáctico.

Podríamos pensar que el papel del analizador léxico podría ser asumido por el analizador sintáctico, pero el diferenciar dos entradas hace posible aplicar técnicas específicas para cada una

de ellas. El analizador léxico hace las funciones, a la vez de preprocesador ya que se encarga de eliminar los caracteres innecesarios para el proceso de compilación. Por ejemplo, elimina los espacios en blanco que hay entre palabra y palabra (siempre que no sean necesarios, como por ejemplo dentro de cadenas de texto).

Resumiendo, las funciones del analizador léxico son:

- Agrupar caracteres según categorías establecidas por la especificación del lenguaje fuente
- Rechazar texto con caracteres ilegales o agrupados según un criterio no especificado

(Ruiz, 2010)

4.1.2 Funcionamiento del Analizador léxico.

La relación entre el analizador léxico y el sintáctico es una relación de maestro-esclavo, el sintáctico demanda al léxico que lea el siguiente lexema y el léxico lee los caracteres necesarios del fichero de entrada hasta que consiguen completar un lexema, que será entregado a su maestro (el analizador sintáctico). (Ruiz, 2010)

El funcionamiento del analizador léxico es:

- Construir tokens validos a partir de la lectura carácter a carácter del fichero de entrada
- Pasar tokens validos al analizador sintáctico
- Gestionar el manejo del fichero de entrada
- Ignorar los espacios en blanco, comentarios y demás caracteres o tokens innecesarios para el proceso de análisis
- Avisar de los errores para el proceso de análisis
- Avisar de los errores encontrados en esta fase

En la siguiente imagen se observa el flujo de información entre las acciones del analizador léxico, este manda y recibe componentes del analizador sintáctico, véase en la imagen 4.

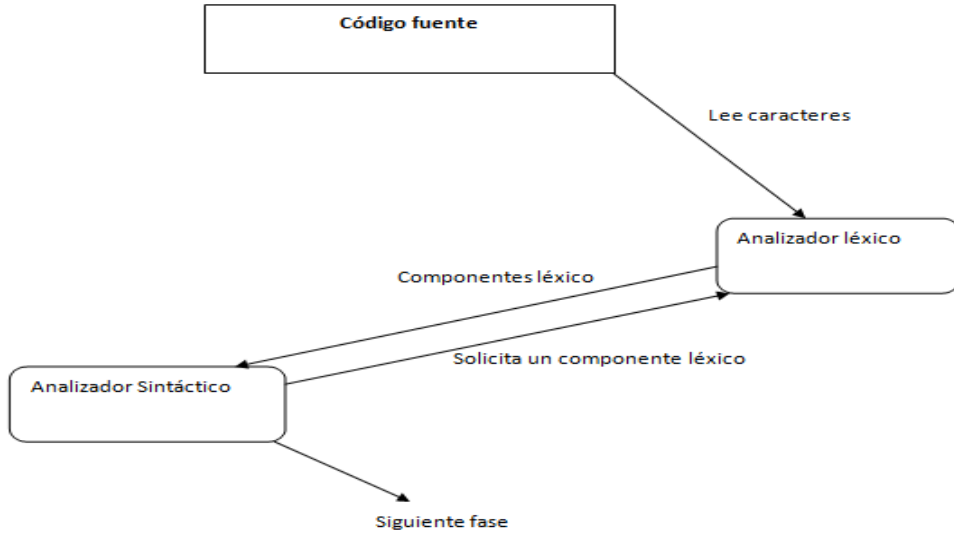


Imagen 4. *Conexión analizador léxico-sintáctico*

4.2 Análisis Morfológico.

Su función consiste en detectar la relación que se establece entre las unidades mínimas que forman una palabra, como puede ser el reconocimiento de sufijos o prefijos. Este nivel de análisis mantiene una estrecha relación con el léxico.

El léxico es el conjunto de información sobre cada palabra que el sistema utiliza para el procesamiento las palabras que forman parte del diccionario están representadas por una entrada léxica y en caso de que esta tenga más de un significado o diferentes categorías gramaticales, tendrá asignada diferentes entradas.

En el léxico se incluye la información morfológica, la categoría gramatical, irregularidades sintácticas y representación del significado.

Normalmente el léxico solo contiene la raíz de las palabras con formas regulares siendo el analizador morfológico el que se encarga de determinar si el género número o flexión que componen el resto de la palabras son adecuados.

4.2.1 Analizador Morfológico.

Constituyen el punto de partida de cualquier otro proceso, tanto como semántico y sintáctico en muchas aplicaciones son el único proceso que se aplica a textos por eso es muy útil ya que el analizador semántico reconoce por medio de texto, uno de los puntos más importantes de este analizador es relacionar las variantes flexivas de una palabra con su lema correspondiente y en asociar a cada forma correspondiente información morfosintáctica.

Sus características y sus atributos de las palabras nos sirven para que cada palabra tenga un significado real y no uno que no esté dentro de lo que se desea buscar algo sin sentido para eso sirve este tipo de analizador como en otros lenguajes podría tener diferentes significados por eso es muy necesario este tipo de analizador.

Uno de los principales problemas de este tipo de análisis son: el tratar con el análisis morfológico son las alternancias de los morfemas tanto léxicos como gramaticales y las reglas de combinación de raíces y sufijos flexivos. (Sosa, Procesamiento del Lenguaje Natural, 1997)

El primer problema de este analizador los morfemas pueden tener su raíz o morfema igual un ejemplo sería el (pod, pued, pud) ya que sus raíces son muy similares. A este problema es difícil manejarlo ya que se necesita reglas gramaticales para poder resolverlas, por otro lado también es muy difícil pero también son necesarias estas reglas para poder corregir el problema de este analizador. (Martinez R. , 2014)

4.3 Analizador Sintáctico.

Este analizador está enfocado a la gramática, esto une al analizador semántico de muchas maneras, una de ellas permite definir un lenguaje en este caso el español, esto es además de realizar reglas que nos permitan generar o producir cadenas de un mismo lenguaje.

Por otro lado el lenguaje natural será analizado por el analizador semántico por eso es importante y muy útil para este analizador conocer las reglas gramaticales de una oración que contiene un sujeto, verbo o acción y el complemento o predicado; para esta tesis trabaja de manera similar ya que se usa una tripleta la cual consta de (“**Concepto** que actúa como sujeto”, “**relación** actúa como verbo” y” **valor** actúa como predicado”). (Perez, 2013)

4.3.1 Análisis Sintáctico.

Tiene como función etiquetar cada uno de los componentes sintácticos que aparecen en la oración y analizar como las palabras se combinan para formar construcciones gramaticalmente correctas. El resultado de este proceso consiste en generar la estructura correspondiente a las categorías sintácticas formadas por cada una de las unidades léxicas que aparecen en la oración.

Las gramáticas, tal como se muestran en la siguiente figura, están formadas por un conjunto de reglas cuales son: (Sosa, Procesamiento del lenguaje Natural, 1997)

O --> SN, SV
SN --> Det, N
SN --> Nombre Propio
SV --> V, SN
SV --> V
SP --> Preposición, SN
SN = sintagma nominal
SV = sintagma verbal
Det = determinante

El analizador léxico (Ruiz, 2010) tiene como entrada el código fuente en forma de una sucesión de caracteres, el analizador sintáctico tiene como entrada los lexemas que le suministra el analizador léxico y su función es comprobar que están ordenados de forma correcta (dependiendo del lenguaje que queramos procesar) los dos analizadores suelen trabajar unidos e incluso el léxico suele ser una subrutina del sintáctico.

Al analizador sintáctico se le suele llamar parse, el parse genera de manera teórica un árbol sintáctico, este árbol se puede ver como una estructura jerárquica que para su construcción utiliza reglas recursivas. La estructuración de este árbol hace posible diferenciar entre aplicar unos operadores antes de otros en la evaluación de expresiones. (Ruiz, 2010)

4.3.1.1 Aspectos Sintácticos de la Oración.

“El proceso de análisis consiste en asignar a cada oración de un texto su estructura sintáctica” el análisis sintáctico automático se logra por medio de un programa informático que toma una oración como entrada y le asigna su estructura sintáctica. (Aguilar J. , 2006)

Para describir formalmente la estructura sintáctica de una oración existen dos enfoques principalmente: constituyentes y dependencias los dos enfoques emplean árboles para representar la estructura sintáctica de una oración pero se diferencian por el significado de los nodos y sus relaciones en el árbol. (Benot, 1987)

Aunque las palabras parecen seguir un orden lineal o secuencial dentro de la oración en realidad su esquema organizativo es más complejo, así mientras se contenga un orden se podrá conocer lo que funciona; para así evitar errores continuos. (Dorta, 2007)

4.3.1.2 Interpretación guiada por la Sintaxis.

El análisis semántico de la oración se guía mediante el proceso de análisis sintáctico. La construcción de la forma lógica se basa en la propiedad de composicionalidad. (Pajares & Santos, 2006) .

(Alicante, 2007-2008) Obtiene la forma lógica asociada a la oración durante el análisis sintáctico (o a partir del árbol de análisis) mediante un proceso composicional. El principio de composicionalidad es “El significado de cada constituyente oracional debe derivarse a partir del significado de sus constituyentes y de la regla sintáctica aplicada”.

4.3.2 Funciones del Analizador Sintáctico.

(Ruiz, 2010) La función principal del analizador sintáctico (AS o párise) es comprobar que los tokens que le suministra el analizador léxico (AL o léxer) van ordenados según la especificación de la gramática del lenguaje a compilar. Y si no es así, dar los mensajes de error adecuados, pero continuar funcionando sin detenerse, hasta que se llegue al final del fichero de entrada.

Es esencial que el proceso de análisis no se detenga al primer error encontrado, ya que así podrá informar al usuario en un solo informe de todos los errores generados. Aparte de esta función principal, el parse es la unidad que guía todo el proceso, o casi todo, de la compilación. Esto es así porque por un lado va solicitando al léxer los tokens y al mismo tiempo va dirigiendo el proceso de análisis semántico y generación de código intermedio.

Por lo que muchas veces se le llama al proceso de análisis semántico y generación de código intermedio, traducción dirigida por la sintaxis generalmente, los analizadores sintácticos obtienen un árbol teórico que permite expresar el orden de los lexemas según van apareciendo.

Ese árbol debe ser el modelo donde salga el análisis semántico, pero lo normal es que si utilizamos el método de la traducción dirigida por la sintaxis, no llegaríamos a obtener todos los errores. La generación del árbol ya que el parse realizara las acciones semánticas, es agregar los métodos para realizar pruebas las cuales avisarán de errores; los cuales podrán ser corregidos o afrontados.

Para generar parse, podemos utilizar dos técnicas, o bien lo hacemos a mano (es difícil aunque eficiente en su funcionamiento) o mediante herramientas que lo generan automáticamente (es menos eficiente pero más de implementar).

Para que aun así funcione, debemos especificar el lenguaje que debe poder leer, para especificar este lenguaje, debemos representarlo con unas reglas únicas y bien formadas de manera que el parse funcione de una manera bien definida. Es decir, lenguaje debe ser formal (tener unas reglas bien definidas). A estas reglas se le llama gramática. (Ruiz, 2010).

En la siguiente imagen se explica el comportamiento de la tabla de símbolos y las acciones del programa fuente, lexer y párise además de cómo influyen al árbol de análisis sintáctico, véase en la figura 5.

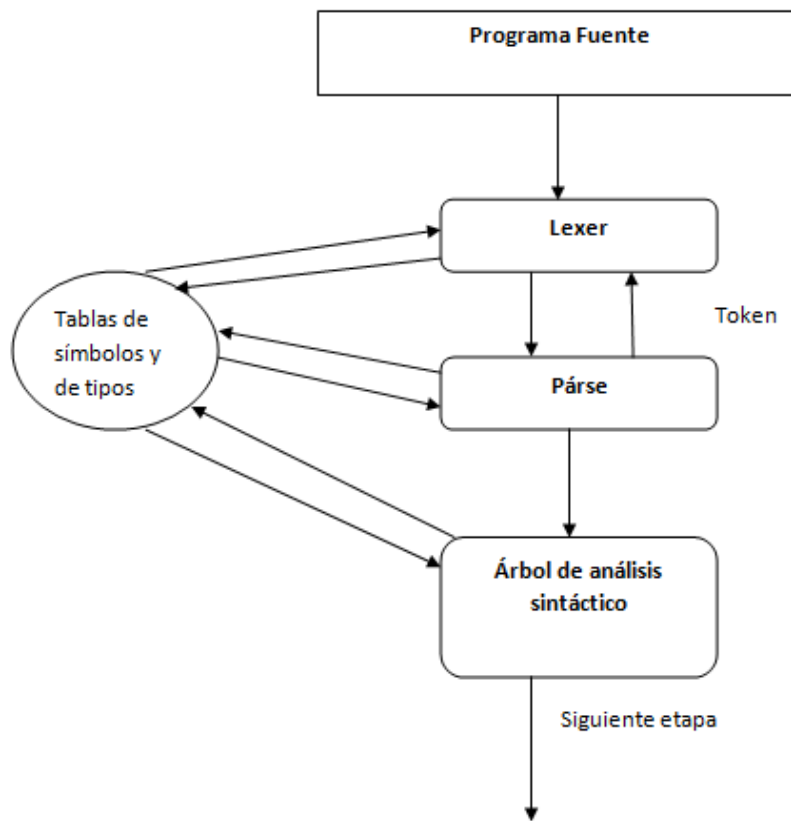


Imagen 5. *Funcionamiento del parse*

4.4 Análisis Semántico.

Es la extensión del análisis sintáctico para la comprensión de la redes semánticas y la base de conocimientos tanto para ver si tiene sentido y que este dirigida al lenguaje natural ya que es a eso a lo que se enfoca, para decirlo en otras palabras es darle significado asociado a las estructuras formales del lenguaje.

Este analizador computa información adicional necesaria para el procesamiento de un lenguaje una vez que se halla tomando bien un análisis completo y exacto del analizador sintáctico y genere resultados verídicos su objetivo primordial o principal del analizador semántico es que el programa analizado satisfaga las reglas requeridas por la especificación del lenguaje para garantizar su correcta ejecución.

Ventajas del análisis semántico

- La descripción formal de la semántica de lenguajes de programación es la descripción rigurosa del significado o comportamiento de programas, el lenguaje de programación además de verificar propiedades de relación con pruebas de corrección o información relacionada con su ejecución.
- Ser utilizados como base para la implementación de procesadores de lenguaje.
- Facilitar la comprensión de los lenguajes por parte del programador y como mecanismo de comunicación entre el diseñador del lenguaje e implementador, la especificación semántica de un lenguaje como documento de referencia aclara el comportamiento del lenguaje y sus diversas construcciones.
- Revela posibles ambigüedades existentes.
- Crear implementaciones de procesadores de lenguajes o en documentos descriptivos del procesamiento del lenguaje natural.

Como se ha mencionado el analizador semántico detecta la validez semántica de las sentencias aceptadas por el analizador sintáctico además se entiende que la semántica como el conjunto de reglas que satisfagan el significado de cualquier sentencia sintácticamente correcta y descrita en un determinado lenguaje. (Venegas, 2003)

En muchas aplicaciones del PLN los objetivos del análisis apuntan hacia el procesamiento del significado. En los últimos años las técnicas de procesamiento sintáctico han experimentado avances significativos, resolviendo los problemas fundamentales.

Sin embargo, las técnicas de representación del significado no han obtenido los resultados deseados, y numerosas cuestiones continúan sin encontrar soluciones satisfactorias.

Definir qué es el significado no es una tarea sencilla, y puede dar lugar a diversas interpretaciones. A efectos funcionales, para facilitar el procesamiento, la modularidad es una de las propiedades más deseables. Haciendo uso de esta concepción modular es posible distinguir entre significado independiente y significado dependiente del contexto.

El primero, tratado por la semántica, hace referencia al significado que las palabras tienen por sí mismas sin considerar el significado adquirido según el uso en una determinada circunstancia. La semántica, por tanto, hace referencia a las condiciones de verdad de la frase, ignorando la influencia del contexto o las intenciones del hablante. Por otra parte, el componente significativo de una frase asociado a las circunstancias en que ésta se da, es estudiado por la pragmática y conocido como significado dependiente del contexto.

Atendiendo al desarrollo en el proceso de interpretación semántica, es posible optar entre múltiples pautas para su organización, tal como se determinan en los siguientes párrafos.

En referencia a la estructura semántica que se va a generar, puede interesarnos que exista una simetría respecto a la estructura sintáctica, o por el contrario que no se dé tal correspondencia entre ellas. En el primer caso, a partir del árbol generado por el análisis sintáctico se genera una estructura arbórea con las mismas características, sobre la cual se realizará el análisis semántico. En el segundo caso, en la estructura generada por la sintaxis se produce un curso de transformaciones sobre las cuales se genera la representación semántica.

Cada una de las dos opciones anteriores puede implementarse de forma secuencial o paralela. En la interpretación secuencial, después de haber finalizado la fase de análisis sintáctico, se genera el análisis semántico. En cambio, desde un procedimiento en paralelo, el proceso de análisis semántico no necesita esperar a que el analizador sintáctico haya acabado toda su tarea, sino que puede ir realizando el análisis de cada constituyente cuando éste ha sido tratado en el proceso sintáctico.

Finalmente en combinación con cada una de las opciones anteriores, podemos escoger un modelo en el que exista una correspondencia entre reglas sintácticas y semánticas o, contrariamente, podemos optar por un modelo que no cumpla tal requisito. En caso afirmativo, para cada regla sintáctica existirá una regla semántica correspondiente.

El significado es representado por formalismos conocidos por el nombre de (*knowledge representation*).

El léxico proporciona el componente semántico de cada palabra en un formalismo concreto, y el analizador semántico lo procesa para obtener una representación del significado de la frase.

(Sosa, Procesamiento del Lenguaje Natural, 1997)

Para (Ruiz, 2010) el análisis semántico consiste en verificar que los tipos en las expresiones son posibles, que las funciones devuelven tipos adecuados, etcétera. Aunque algunas de las acciones no tendremos que realizarlas debido a que nuestro lenguaje es muy simple y no lo requiere.

Lo primero que debemos hacer es gestionar los símbolos de nuestro lenguaje, los ámbitos y los tipos, para ello, se utilizan unas estructuras de datos que son: la tabla de tipos y la pila de ámbitos. Hay muchas maneras de gestionar estas estructuras, pero lo normal es que haya una tabla de símbolos y otra de tipos para cada ámbito (un ámbito contiene las variables locales, los tipos locales, etc.). Cada función o procedimiento tiene su propio ámbito y cuando necesitamos un símbolo o tipo dentro del ámbito, lo primero es buscar y cuando necesitamos un símbolo o tipo dentro del ámbito, lo primero es buscar si está declarado en el mismo, y si no lo está, buscar a ver si lo está en un ámbito que lo englobe.

4.4.1 Interpretación guiada por la semántica.

Dado que finalmente se desea obtener es el significado de la oración, una aproximación apropiada consistirá en utilizar restricciones sintácticas concretas. En esta clase de analizadores se utilizan estructuras denominadas marcos (frames), cada uno de los cuales permite especificar acciones y tipo de objetos. Se trata de una forma de representación del conocimiento un marco se corresponde con una plantilla con una serie de entradas denominadas ranuras o slots para conservar instancias de acciones. (Rastier, 2001) Un algoritmo guiado por la semántica determina la acción básica descrita en la oración. Esta generalmente viene dada por el verbo principal así que el primer consiste en encontrar dentro de la oración los elementos para rellenar esos huecos. (Rodríguez, 2004) Un ejemplo de (Pajares & Santos, 2006), en la oración “el portero abre la puerta” en el verbo abre es el determine la acción principal: ABRIR que corresponderá a un marco con dos huecos: el Agente cuyo tipo sería portero y el Objeto que puede ser por ejemplo del tipo traspasar según alguna ontología predefinida sería puerta. El marco quedaría como sigue, dado sentido a la oración:

PASAR: Agente: portero, Objeto: puerta.

4.4.2 Proceso de Interpretación Semántica.

(Rodríguez, 2004) Es interpretación “Es el proceso de emparejar una oración con su forma lógica”. Interpretación aproximaciones: la información semántica completa el análisis sintáctico. Interpretación guiada por la semántica: la información sintáctica solo aparece como restricciones limitadas

(Pajares & Santos, 2006) Una vez definido un lenguaje de la forma lógica para la representación del significado, basado en el cálculo de predicados, se deben traducir las oraciones en Lenguaje Natural elegido. Este proceso de traducción se denomina *interpretación semántica* y básicamente consiste en obtener la forma lógica asociada e la oración a partir del árbol de análisis.

La interpretación semántica se debe realizar mediante un proceso composicional, o lo que es lo mismo el significado de cada constituyentes de la oración debe derivarse solamente a partir del significado de sus subconstituyentes. La composición semántica viene a suponer que la semántica de cualquier frase es una composición de las semánticas de sus subfrases.

4.4.3 Análisis Semántico: Modelo Estructural.

(Pajares & Santos, 2006) Menciona que “el análisis sintáctico se utilizaba los modelos más apropiados, ahora es necesario proceder de la misma forma, para ellos se desarrolla un modelo estructural” Allí era la gramática la que generaba estructuras complejas a partir de otras más simples. Siguiendo la misma estrategia, ahora se define las unidades básicas y a partir de ella se crean estructuras más complejas para llegar al significado de las oraciones

Para conseguir este objetivo se hace uso del principio de composicionalidad “el significado de una estructura o expresión compleja está en función de los diferentes significados de sus partes y de las reglas mediante las que se combinan” (Pajares & Santos, 2006). Para (Rodríguez, 2004) existen dos métodos relativos al proceso de interpretación semántica.

- La interpretación guiada por la sintaxis
- La interpretación guiada por la semántica.

En la siguiente imagen se muestra las acciones que lleva el análisis sintáctico para convertirse en un análisis semántico, como se muestra en la imagen 6.

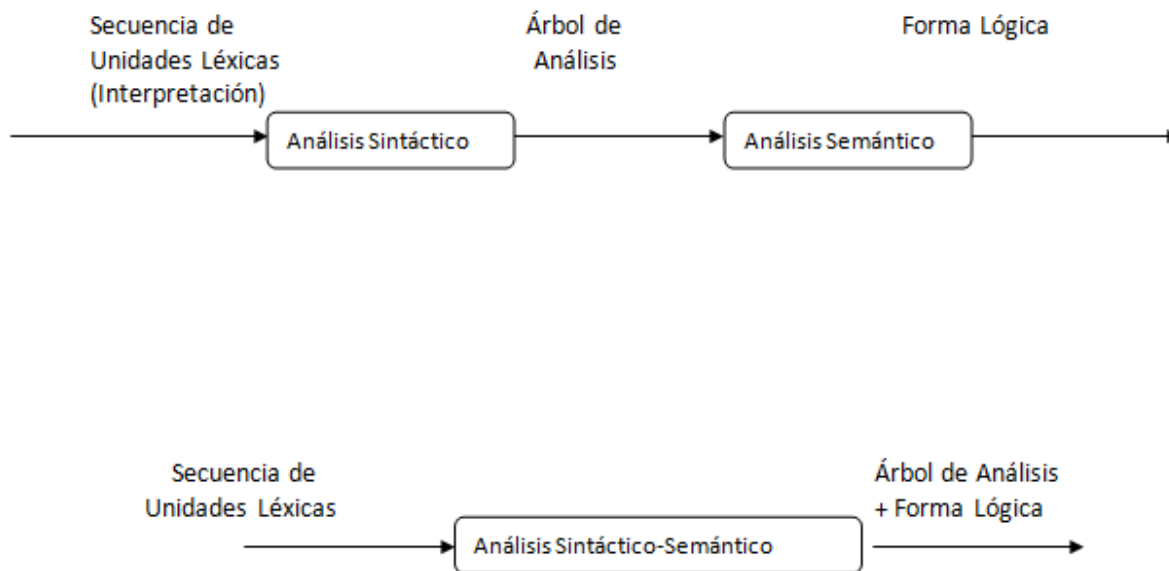


Imagen 6. Se muestra un esquema grafico sobre el análisis semántico bajo las anteriores interpretaciones.

4.4.4 Procesamiento semántico.

El Procesamiento semántico analiza un enunciado el cual está en lenguaje natural ya que se trata de encontrar su relación semántica, incluye representar la estructura léxica (que se lleva a cabo mediante un análisis sintáctico o gramatical), descomponerlo en elementos estructurales y comprobar su validez semántica; es decir, procesa el significado lo suficientemente para averiguar (entender) su significado (semántica), pero ¿Cuántas interpretaciones, puede tener un enunciado? Pueden ser diversas dependiendo de lo que cada persona pueda conocer sobre ello.

Para no caer en ambigüedades en las interpretaciones, el procesamiento semántico debe considerar.

1. Análisis sintáctico, que está relacionado con la estructura del lenguaje por lo que este evalúa las relaciones de concordancia entre los diferentes elementos estructurales.
2. Interpretación semántica, que verifica la concordancia entre sujeto o agente, verbo o acción, complemento directo o pasivo, más los adjetivos y adverbios que la oración que pueda tener. De ahí se obtiene el significado de toda la oración, y éste a su vez le asigna un significado a cada palabra de ella.

El procesamiento semántico tiene como entrada un enunciado y que a partir de una representación o estructuración de la información obtiene, mediante el análisis sintáctico, los datos que requiere el analizador semántico para buscar la relación semántica.

Un enunciado analizado nos permite obtener el significado exacto de una palabra, así si se desea extraer conocimientos acerca del procesamiento del lenguaje natural, se tendrá que realizar el procesamiento semántico del enunciado en cuestión para que su significado sea el mismo para cualquier persona que requiera saber sobre de ello.

Para poder Obtener el significado es importante trabajar con la estructura semántica, la cual se compone de relaciones, conceptos, partes y miembros. (Cervantes A. , 2011)

4.4.5 Proximidad Semántica.

Este modelo está relacionado con el conocimiento semántico que requiere para desambiguar oraciones porque sus diversas estructuras sintácticas son perfectamente posibles, o para enlazar frases circunstanciales que al no estar directamente enlazados con el mismo sentido del lexema; el lector requieren un método conectado con la semántica de contexto.

Así que el conocimiento que describe es una clase de conocimiento semántico del contexto que se trata de reconocer, las palabras que están relacionadas es decir que están “más cercanas” semánticamente o que son “semánticamente compatibles”. Por ejemplo en la frase conocida Veo un gato con un telescopio no es claro si telescopio está relacionado con ver o con gato. La información semántica permite decidir que telescopio está más próximo semánticamente de ver y no de la de gato.

La idea del empleo de la red semántica es la siguiente por ejemplo consideremos las frases: Me gusta beber licores con menta y me gusta beber licores con mis amigos. En ambas frases la clase semántica del sustantivo final ayuda a resolver la ambigüedad es decir con que parte de la frase están enlazadas las frases preposicionales con menta y con mis amigos ni menta ni amigos son palabras ambiguas pero no amigos está más cercana semánticamente a beber que a licores y menta está más cercana a licor que beber de esta forma se desambiguan los enlaces. (Justo, 2008)

4.4.6 Aspectos Semánticos.

La relación de concordancia no agota las relaciones que se establecen entre el sujeto y el predicado. Se vinculan además por relaciones del significado, las relaciones de significado no son exclusivas de la relación que se establece entre sujeto y predicado también se dan entre los distintos constituyentes del predicado.

La salida teórica de la fase de análisis semántico sería un árbol semántico que consiste en un árbol sintáctico en el que cada una de sus ramas ha adquirido el significado que debe tener, así como sus reglas semánticas las cuales ayudan a evitar ambigüedades.

Las reglas semánticas: son el conjunto de normas y especificaciones que definen al lenguaje de programación y están dadas por la sintaxis del lenguaje las reglas semánticas asignan un significado lógico a ciertas expresiones definidas en la sintaxis del lenguaje. La evaluación de las

reglas semánticas definen los valores de los atributos en los nodos del árbol de análisis sintáctico están dentro de la cadena de entrada.

Una regla semántica también puede tener efectos colaterales los cuales afectan el significado de las palabras para lo que se tiene que tomar precauciones para no entrar en ambigüedades.

Bueno con esto se dio una pequeña explicación de lo que se trata un analizador semántico el cual sirve generando arboles los cuales para poder estar bien deben cumplir con los anteriores analizadores, pero más el analizador sintáctico ya que es muy fundamental para su creación.

Otra parte importante son los aspectos del significado, sentido o interpretación de un determinado elemento, símbolo, palabra, expresión o representación formal. La semántica puede estudiarse desde diferentes puntos de vista:

- Semántica lingüística: trata de la codificación y decodificación de los contenidos semánticos en las estructuras lingüísticas.
- Semántica lógica: desarrolla una serie de problemas lógicos de significado, estudia la relación entre el signo lingüístico y la realidad.
- Semántica en ciencias cognitivas: intenta explicar porque nos comunicamos y cuál es el mecanismo psíquico que se establece entre hablante y oyente. (Gomez E. , 2010)

4.4.7 Funciones del análisis semántico.

El analizador semántico detecta la validez semántica de las sentencias aceptadas por el analizador sintáctico. El analizador semántico suele trabajar simultáneamente al analizador sintáctico y en estrecha cooperación se entiende por semántica como el conjunto de reglas que especifican el significado de cualquier sentencia sintácticamente correcta y escrita en un determinado lenguaje.

Las rutinas semánticas deben realizar la evaluación de los atributos de las gramáticas siguiendo las reglas semánticas asociadas a cada producción de la gramática. El análisis sintáctico es la fase en la que se trata de determinar el tipo de los resultados intermedios, comprobar que los argumentos que tiene un operador pertenecen al conjunto de los operadores posibles y si son compatibles entre sí, etc. En definitiva, comprobará que el significado de lo que se va leyendo es válido.

La salida “teórica” de la fase de análisis semántico sería un árbol semántico. Consiste en un árbol sintáctico en el que cada una de sus ramas ha adquirido el significado que debe tener. En el caso de los operadores polimorfismo (un único símbolo con varios significados), el análisis semántico determina cual es el aplicable.

Se compone de un conjunto de rutinas independientes, llamadas por los analizadores morfológicos y sintácticos. El análisis semántico utiliza como entrada el árbol sintáctico detectado por el análisis sintáctico para comprobar restricciones de tipo y otras limitaciones

semánticas y preparadas la generación de código. Las rutinas semánticas suelen hacer uso de una pila (la pila semántica) que contiene la información semántica asociada a los operandos en forma de registro semánticos.

4.4.8 Reglas semánticas.

Son el conjunto de normas y especificaciones que definen al lenguaje de programación y están dadas por la sintaxis del lenguaje. Las reglas semánticas asignan un significado lógico a ciertas expresiones definidas en la sintaxis del lenguaje.

La evaluación de las reglas semánticas definen los valores de los atributos en los nodos en los nodos del árbol de análisis sintáctico para la cadena de entrada. Una regla semántica también puede tener efectos colaterales, por ejemplo imprimir un valor o actualizar una variable global.

(Gustavo, 2010)

Para (Computacionales, 2010) Un lenguaje está formado por reglas semánticas que proporcionan el significado a una sentencia o instrucciones del lenguaje. Las reglas semánticas dadas en función de los atributos de los demás símbolos que componen la regla.

La evaluación de las reglas semánticas puede generar código, guardar información en una tabla de símbolos, la traducción de la cadena de componentes léxicos es el resultado obtenido al evaluar las reglas semánticas.

La evaluación de las reglas semánticas definen los valores de los atributos en los nodos del árbol de análisis sintáctico para la cadena de entrada. Una regla semántica puede tener colaterales, por ejemplo, imprimir un valor o actualizar una variable global.

4.5 Análisis pragmático.

Añade información adicional al análisis del significado de la frase en función del contexto donde aparece. Se trata de uno de los niveles de análisis más complejos, la finalidad del cual es incorporar al análisis semántico la aportación significativa que pueden hacer los participantes, la evolución del discurso o información presupuesta.

En general la pragmática incluye aspectos del conocimiento conceptual del mundo que van más allá de las condiciones reales literales de cada oración. Este conocimiento lo tienen en cuenta los hablantes cuando se comunican mediante una lengua les sirve para comprender mucha información sobreentendida pero no expresada explícitamente en las oraciones. Mientras la sintaxis y semántica estudia las oraciones, la pragmática estudia “las acciones del discurso” y las situaciones en las cuales el lenguaje es usado.

Muchas palabras y oraciones pueden ser ambiguas y tener más de un significado, su significado puede ser falso o producir implicaciones falsas el significado depende de los principios que usan las personas cuando hablan (por ejemplo ser relevantes y hacer énfasis en las oraciones

verdaderas). En este sentido la pragmática tiene dos conceptos importantes: la implicación y la presuposición de las oraciones. La implicación de una oración comprende la información que no es parte de su significado, pero que debe ser inferida por un oyente razonable. Las presuposiciones de una oración son las cosas que deben ser verdaderas para la oración sea verdadera o falsa.

Es decir, en base las presuposiciones (conocimiento verdadero de un dominio) las personas interpretan las oraciones y derivan conocimiento (implicaciones) que pueden ser o no verdaderos. El conocimiento lingüístico se incorporó al sistema de PLN a partir de los años sesenta, y se convierte en uno de sus componentes importantes.

A partir de ese momento se definió una área de conocimiento llamada lingüística computacional (CL Computational Linguistics) apoyada por la asociación para la lingüística computacional (ACL for Computational Linguistics). (Perez A. , 2006)

Capítulo 5

DESARROLLO DEL ANALIZADOR SEMÁNTICO

En este capítulo observaremos el desarrollo del analizador semántico, desde su análisis, diseño, compilación y pruebas además de ver como se utilizó la metodología de cascada para su realización.

5.1 Análisis de Analizador semántico.

Requerimientos del Analizador Semántico.

El analizador semántico necesita tener.

- Ontología (la ontología puede ser de cualquier área o tema de investigación, lo importantes es hacer una ontología completa)
- Que el programa reciba una tripleta del usuario (Concepto, Relación, Valor)
- Que la ontología cuente con una gran variedad de conceptos, ya que sin conceptos no podría realizar búsquedas semánticas.
- Que el concepto de la tripleta se encuentre en la ontología para que realice la búsqueda.
- Que cualquier concepto cuente con sinónimos para realizar búsquedas más detalladas.
- Que existan grupos de conceptos.
- Que cada concepto tenga un concepto padre para lo cual se pueda buscar relaciones semánticas en conceptos padres.
- Que el programa reciba la tripleta o sinónimos de la tripleta.

Función Analiza Semántica.

La función Analiza Semántica, Inicia con la búsqueda de la ontología (Cualquier ontología) a la cual se analizará el algoritmo, una vez encontrada la ontología el usuario mandará una tripleta la cual consta de (Concepto, Relación y Valor) una vez ingresados.

El algoritmo toma los valores de la tripleta ingresada; realizará la búsqueda en toda la ontología del **Concepto** si existe el concepto, analizará sus relaciones hasta encontrar la **Relación** ingresada en la tripleta. Si se encuentra buscará el **Valor**; cuando la tripleta ingresada se

encuentra entonces mandará un mensaje de “**Hay Relación Semántica**”. Es importante mencionar que existe relación semántica cuando los tres elementos son aplicados en la realidad, o son comúnmente usados por las personas, lo que no ocurre con (Polinomio, Cuenta con, Ruedas) porque el concepto polinomio es abstracto mientras que las Ruedas son concretas y no es usado regularmente por la gente.

En caso de que el **Concepto** no exista su nombre tal como en la tripleta se pide entonces realizará una función llamada Sinónimos la cual ejecuta la opción de buscar los sinónimos del concepto, si uno de ellos concuerda con el nombre del **Concepto** introducido en la tripleta entonces seguirá con la búsqueda de la **Relación** y el **Valor**.

La función Sinónimos no solo aplica a los conceptos si no que a las **Relaciones** y **Valores** de tal modo que la búsqueda no se vea afectada siempre y cuando exista el nombre de la relación y el valor en la ontología.

Otra parte la cual maneja el analizador Semántico es la función filtro que busca por grupos de tal forma si un grupo determinado de miembros pertenecen a un grupo también realizara esa búsqueda como, por ejemplo si hablamos de un grupo de plantas el analizador buscará en todo el grupo de plantas sus relaciones y valores; pero si no son miembros y son partes de un concepto no realizará la búsqueda. Para resumirlo solo realizara la búsqueda en miembros (member) no en partes (part) de un concepto.

La última función de herencia trata de buscar la **Relación** y **Valor** por los padres (subset) del **Concepto**; funciona así si el **Concepto** introducido en la tripleta existe pero no tiene la **Relación** y el **Valor** introducido no existen en ese Concepto entonces ahí es donde se realiza esta función la cual busca a sus padres de tal forma en el padre busca la Relación y el Valor, si existe entonces manda el mensaje de “Hay Relación Semántica”, en caso de no estar ira haciendo la búsqueda hacia arriba hasta llegar al nodo padre de la ontología si esta la relación y el valor en el nodo padre entonces manda el mensaje de “Hay Relación Semántica” sino entonces mandara el mensaje de “No Hay Relación Semántica”; En caso de que el Concepto no tenga padres entonces mandara un mensaje de “Este concepto no tiene Padres”.

Aplicando el Modelo de Cascada al Analizador Semántico.

Para el desarrollo del analizador semántico se utilizó el modelo de cascada ya que fue la metodología que más se ajustó para su realización. El modelo de cascada consiste en seguir una serie de pasos los cuales te lleven a la realización del producto (Software), una característica del modelo de cascada es que tiene un orden el cual no se puede romper.

Los pasos del modelo de cascada son: (Análisis de Requisitos, Diseño del programa, Codificación y Pruebas)

Análisis de Requisitos: En esta etapa se examinaron los requerimientos necesarios para el analizador semántico que son: (contar con una Ontología, utilizar una tripleta y ver diferentes maneras de búsqueda como buscar por sinónimos, herencia y por grupos).

Diseño del programa: Para esta etapa se utilizó unos diagramas que muestran el funcionamiento completo del analizador semántico con el fin de obtener una búsqueda más completa y exacta; observe las Imágenes 1, 2 y 3 que a continuación se muestran.

Codificación: Para esta etapa se observa la codificación del analizador semántico, las funciones que lleva a cabo, como se ingresa la tripleta y como realiza la búsqueda de la tripleta en la ontología.

Pruebas: En esta última etapa se observa cómo se aplican ejemplos y como el analizador semántico los resuelve.

5.2 Diseño

Analizador Semántico Función (Sinónimos parte1).

En la siguiente imagen se muestra el diagrama de flujo de la primera etapa de búsqueda desde la carga de la ontología hasta las diferentes opciones que tiene para diferentes casos; ver la imagen 1 a continuación.

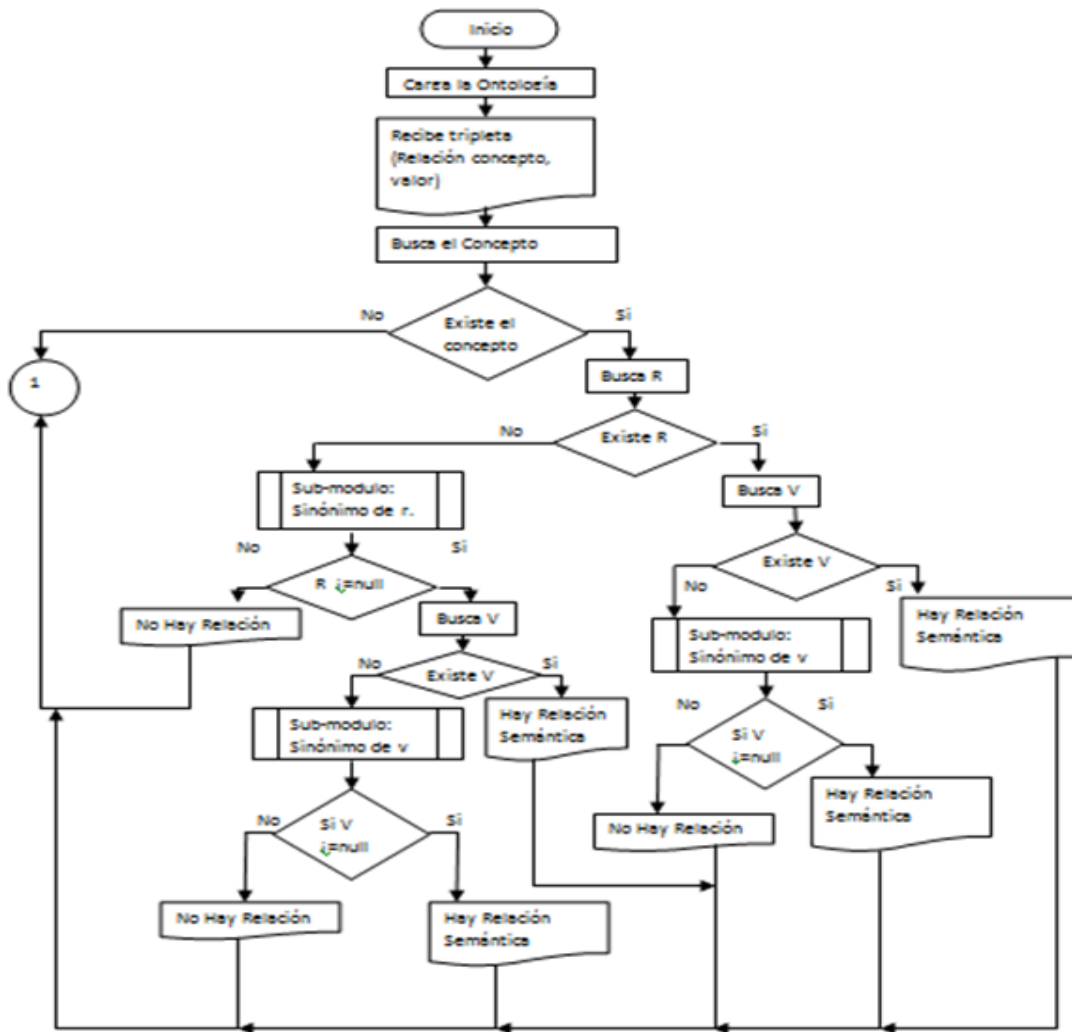


Imagen 1. Algoritmo general de Analizador Semántico

Analizador Semántico Función (Sinónimos parte2).

En la siguiente imagen se muestra la continuación del primer diagrama de flujo, como al igual que el otro se puede observar las diferentes opciones de búsqueda que puede realizar; hasta terminar; ver la imagen 2 a continuación.

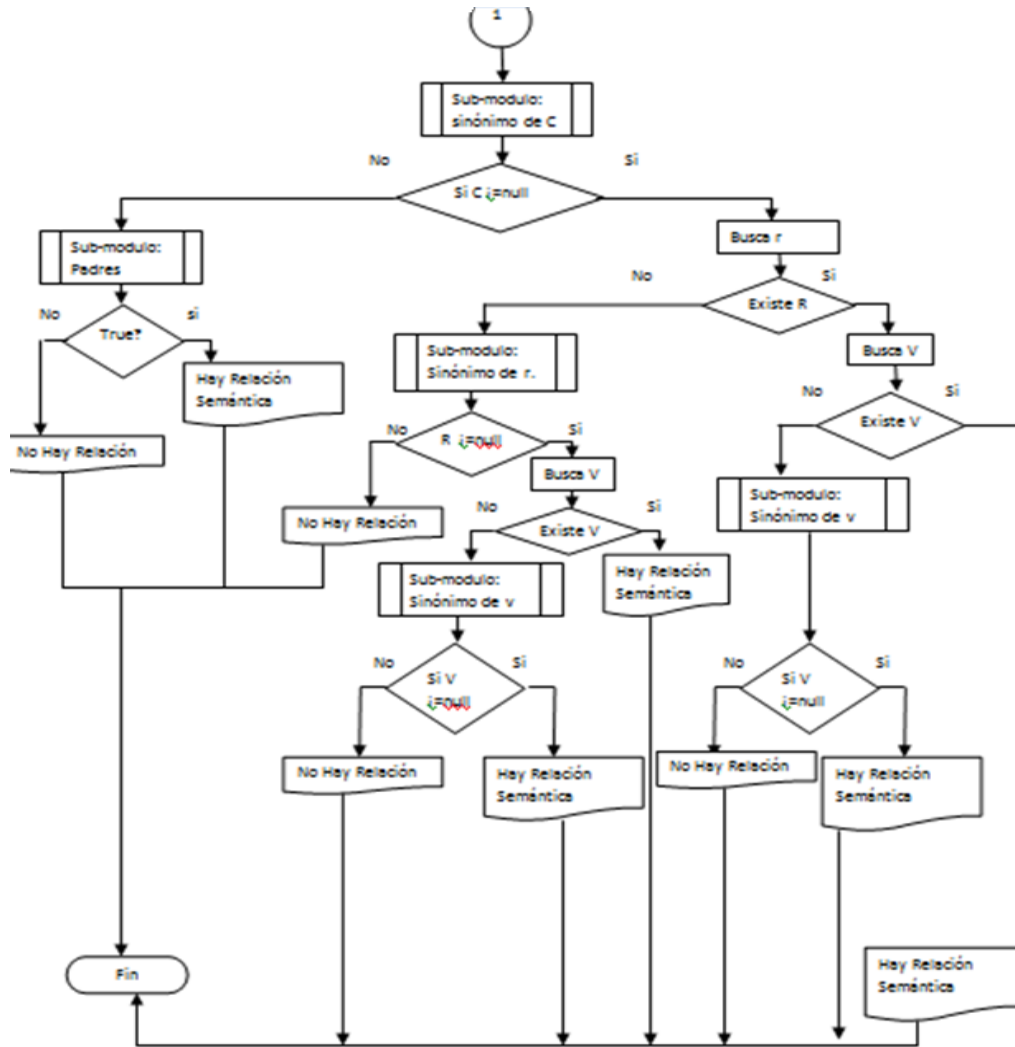


Imagen 2. Algoritmo general del Analizador Semántico Parte 2

Analizador Semántico Función (Herencia).

En la siguiente imagen se muestra una nueva función en la cual se realizará la búsqueda por herencia; ver la imagen 3 a continuación.

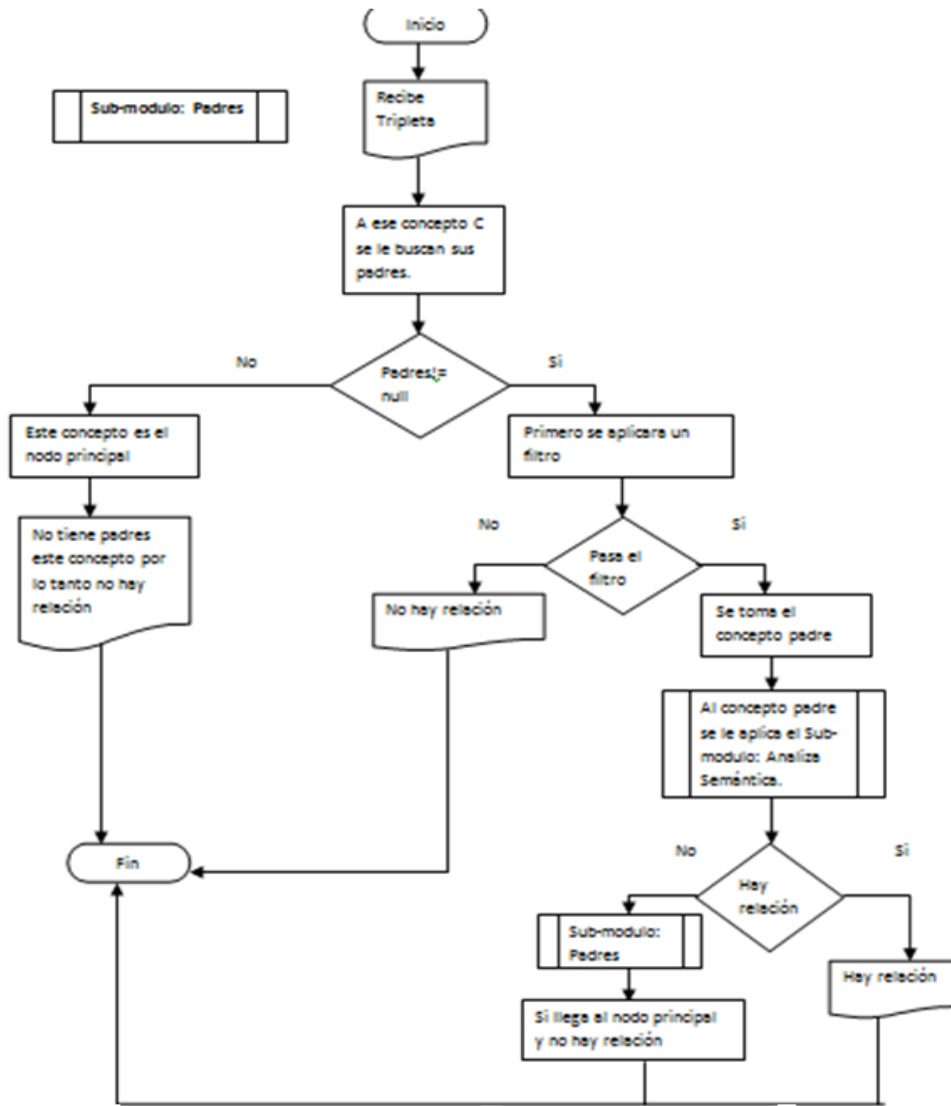


Imagen 3. Algoritmo de Herencia del Analizador Semántico.

Caso 1: Este ejemplo “Hallar la relación tiene y el valor ramas en el concepto árbol”

Concepto= Árbol; Relación= Tiene; Valor= Ramas.

Explicación: Acerca del funcionamiento del algoritmo Analizador semántico: se considera la siguiente tripleta: (Concepto, Relación, Valor) por ejemplo: (Árbol, Tiene, Ramas).

Lo primero que realiza el algoritmo es la búsqueda de la ontología llamada: **OntoPrueba.ont** una vez encontrada, toma los valores de la tripleta ejemplo. Realizará la búsqueda en toda ontología del concepto: **Árbol**, una vez encontrado analizará sus relaciones hasta encontrar la relación llamada: **Tiene**, si se encuentra buscará el valor: **Ramas**, cuando la tripleta ejemplo ingresada se encuentra en la ontología entonces manda un mensaje de “**Hay relación semántica**”.

Es importante mencionar que existe relación semántica cuando los tres elementos en nuestro ejemplo (Carro, Cuenta con, Ruedas) son aplicados en la realidad, o son comúnmente usados por las personas, así como se muestra en la Imagen 4 a continuación.

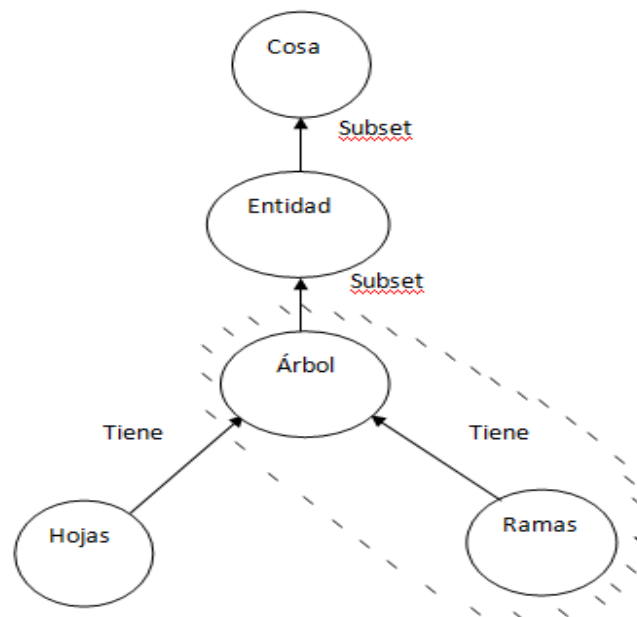


Imagen 4. *Ontología para el caso1.*

Explicación: El Algoritmo inicia con la búsqueda de la **Ontología** (como por ejemplo: OntoPrueba.ont) una vez encontrada la **Ontología**; recibe la tripleta del usuario (**Concepto=C, Relación=R, Valor=V**), una vez ingresada la tripleta realiza la búsqueda del **Concepto** en la **Ontología**, Si existe realiza la siguiente función la cual es buscar la **Relación**, si existe entonces busca el **Valor** y si existe en la **Ontología** manda un mensaje de “**Hay Relación Semántica**”; como se observa a continuación en la imagen 5.

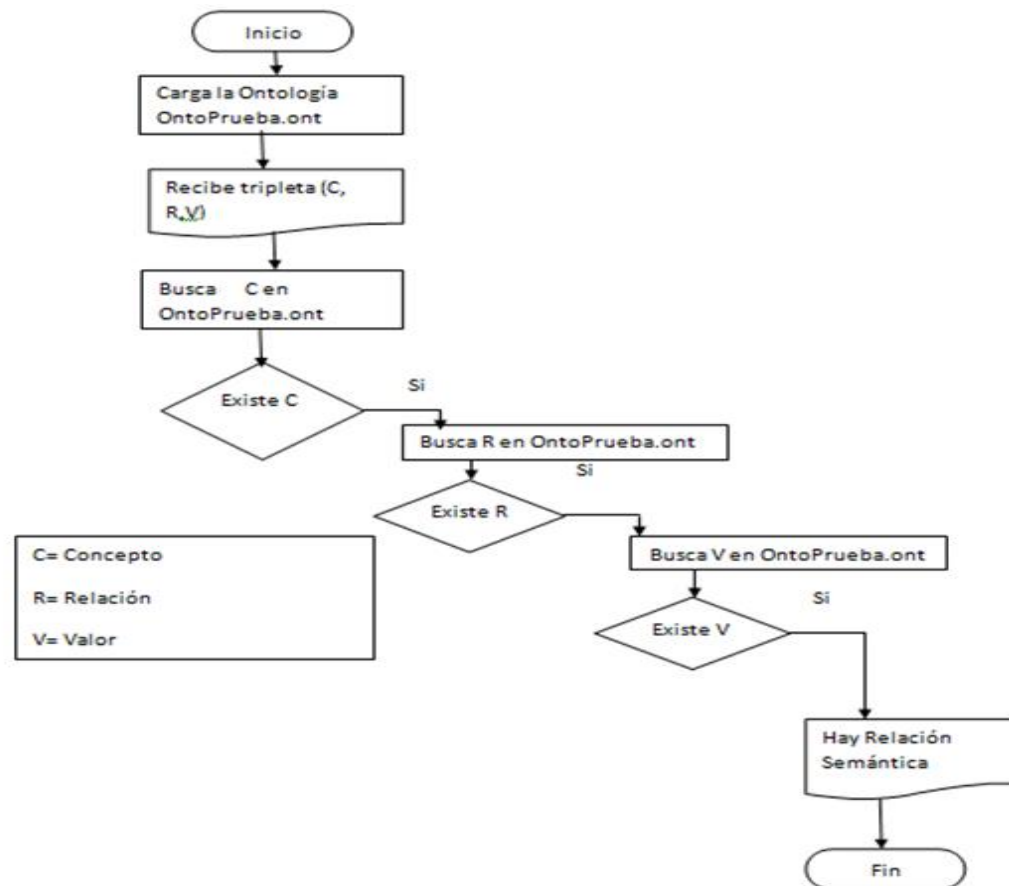


Imagen 5. Algoritmo que explica el funcionamiento del caso1: “trata de hallar la relación tiene y el valor ramas en el concepto árbol”.

Caso 2: Este ejemplo “Hallar la relación cuenta con y el valor ramas en el concepto árbol”

Concepto= **Árbol**; Relación= **Cuenta con**; Valor= **Ramas**.

Explicación: Para este caso la tripleta utilizada es (**Árbol, Cuenta con, Ramas**), la ontología en este y en los siguientes casos se llamará: **OntoPrueba.ont**, buscará el concepto: **Árbol**; si lo encuentra buscará en sus relaciones, la relación llamada: **Cuenta con** al no encontrarlo realizará la búsqueda por palabra en la ontología; si la encuentra comprobará a que concepto pertenece, si se llama: **Tiene**; entonces es su sinónimo; a esto se le llama búsqueda por sinónimos.

Después buscará el valor: **Ramas**, si se encuentra entonces manda un mensaje de “**Hay Relación Semántica**”; así como se muestra en la Imagen 6 a continuación.

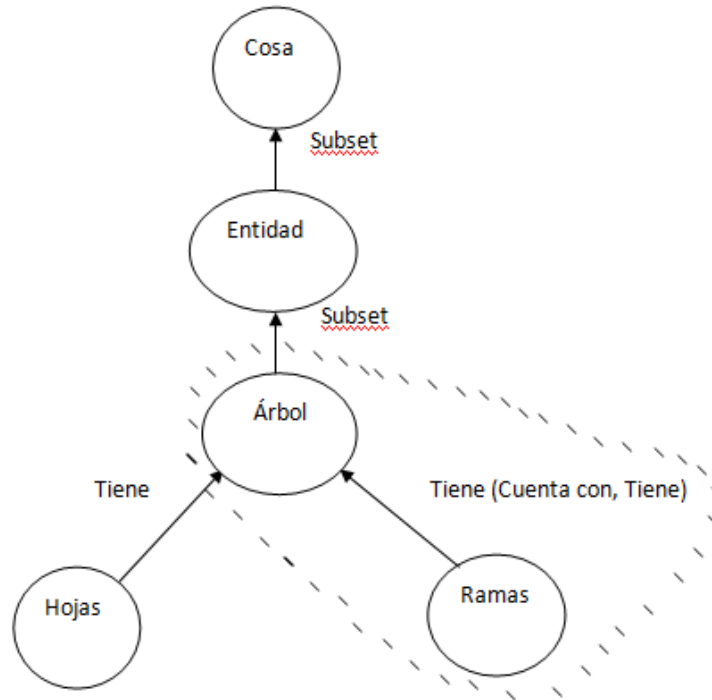


Imagen 6. *Ontología para el caso 2.*

Explicación: El Algoritmo inicia con la búsqueda de la **Ontología** (como por ejemplo: OntoPrueba.ont) una vez encontrada la **Ontología**; recibe la tripleta del usuario (**Concepto=C, Relación=R, Valor=V**), una vez ingresada la tripleta realiza la búsqueda del **Concepto** en la **Ontología**, Si existe realiza la siguiente función la cual es buscar la **Relación** sino existe en la **Ontología** realiza la función **Sinónimos** la cual es buscar un sinónimo de la **Relación**, si este sinónimo es igual al nombre de la **Relación** ingresado en la tripleta, Entonces realizará la búsqueda del Valor y si este se encuentra en la ontología entonces mandará un mensaje de “**Hay Relación Semántica**”; como se observa a continuación en la imagen 7.

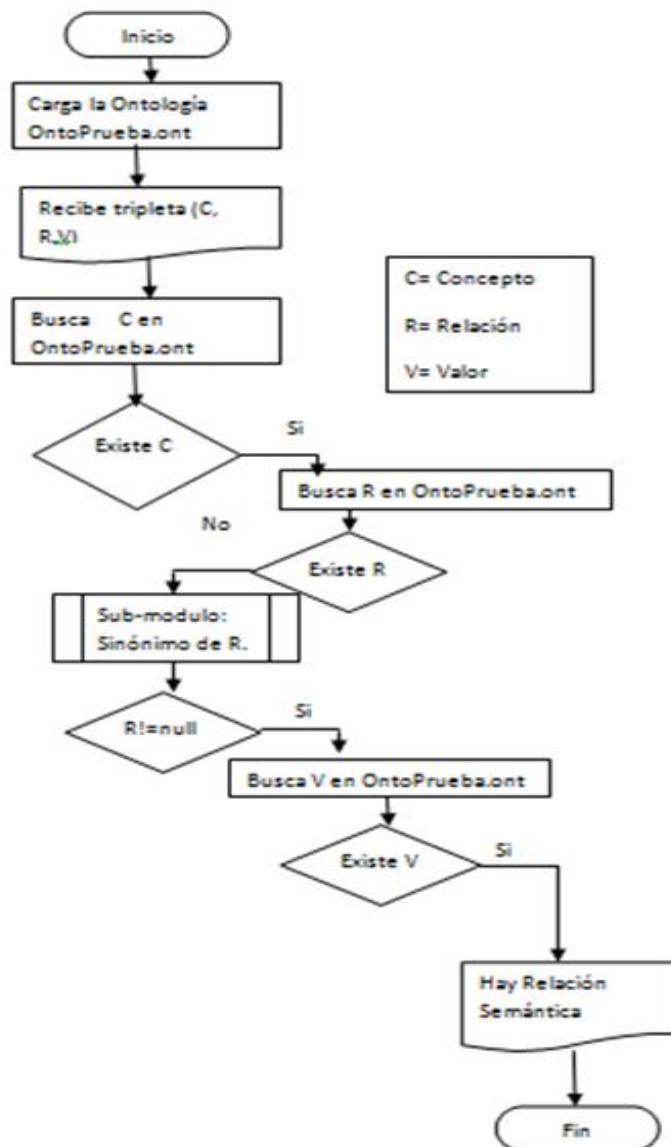


Imagen 7. Algoritmo que explica el funcionamiento del Caso 2: “trata de hallar la relación cuenta con y el valor ramas en el concepto árbol”

Caso 3: Este ejemplo “Hallar la relación Cuenta con y el valor ramas en el concepto arbolillo”

Concepto= Arbolillo; Relación= Cuenta con; Valor= Ramas.

Explicación: Bueno como en los casos anteriores y posteriores se utilizará una tripleta, para este caso la tripleta utilizada es (Arbolillo, Cuenta con, Ramas), en la ontología: **OntoPrueba.ont**, Buscará el concepto: **Arbolillo** si no lo encuentra realizará la búsqueda por sinónimos; si encuentra a que concepto pertenece, y este se llama: **Árbol**; buscará en sus relaciones, la relación llamada: **Cuenta con**; si no la encuentra realizará la búsqueda por sinónimos; si lo encuentra verificara a que concepto pertenece, si ese concepto se llama: **Tiene**, buscará el valor:

Ramas ,si se encuentra entonces manda un mensaje de “**Hay Relación Semántica**”; así como se muestra en la Imagen 8 a continuación.

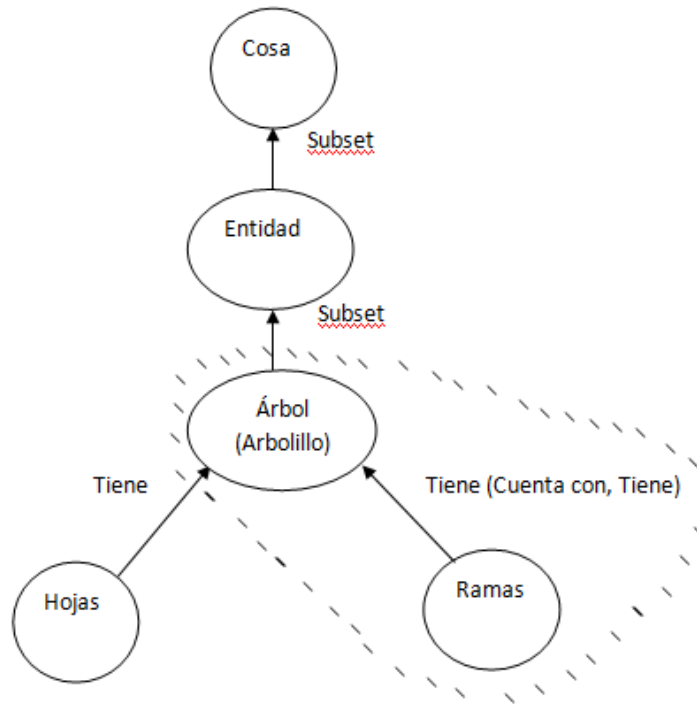


Imagen 8. *Ontología para el caso 3.*

Explicación: El Algoritmo inicia con la búsqueda de la **Ontología** (como por ejemplo: OntoPrueba.ont) una vez encontrada la **Ontología**; recibe la tripleta del usuario (**Concepto=C, Relación=R, Valor=V**), una vez ingresada la tripleta realiza la búsqueda del **Concepto** en la **Ontología**, Sino existe realiza la función Sinónimos la cual consta de buscar sinónimos del concepto si uno de los sinónimos del concepto es igual al ingresado al de la tripleta realiza la siguiente función la cual es buscar la **Relación** sino existe en la **Ontología** realiza la función **Sinónimos** la cual es buscar un sinónimo de la **Relación**, si este sinónimo es igual al nombre de la **Relación** ingresado en la tripleta, entonces realizará la búsqueda del Valor y si este se encuentra en la ontología entonces mandara un mensaje de “**Hay Relación Semántica**”; como se observa a continuación en la imagen 9.

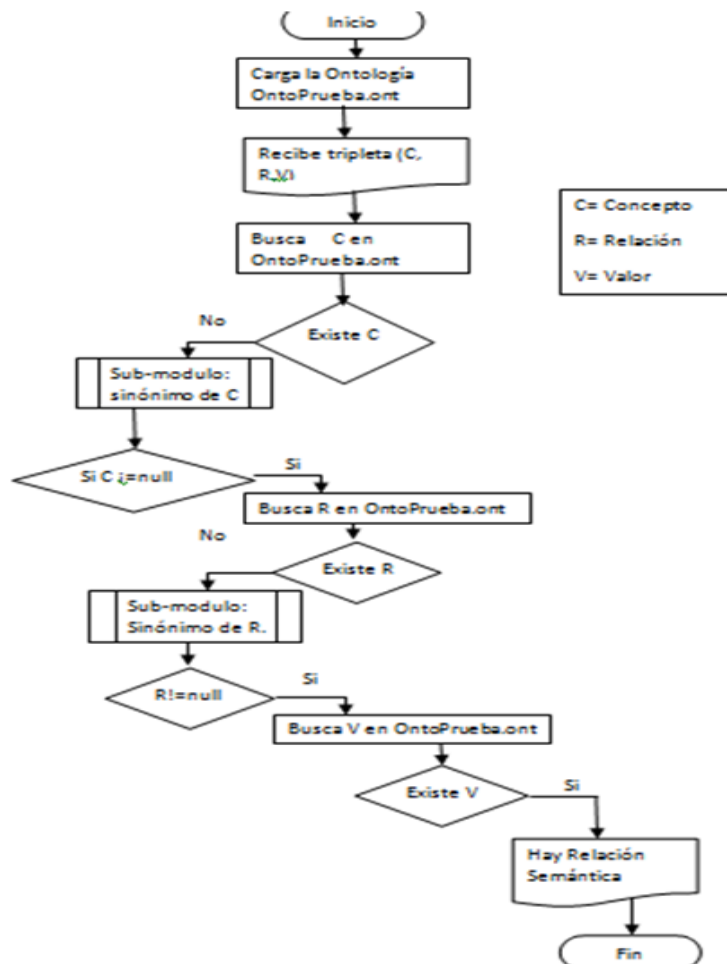


Imagen 9. Algoritmo que explica el funcionamiento del Caso 3. "trata de hallar la relación Cuenta con y el valor ramas en el concepto arbolillo"

Caso 4: Este ejemplo "Hallar la relación cuenta con y el valor retoño con el concepto arbolillo"

Concepto= Arbolillo; Relación= Cuenta con; Valor= Retoño.

Explicación: para este caso la tripleta será (Arbolillo, Cuenta con, Retoño) en la ontología: **OntoPrueba.ont**, Buscará el concepto: **Arbolillo** si no lo encuentra realizará la búsqueda por sinónimos; si lo encuentra verificará a que concepto pertenece, si se llama: **Árbol**, entonces buscará en sus relaciones la relación llamada: **Cuenta con** si no la encuentra realizará la búsqueda por sinónimos hasta encontrar el concepto al que pertenece, si lo encuentra y este se llama: **Tiene**, buscará el valor: **Retoño**; si no lo encuentra buscará por sinónimos; si lo encuentra verificará a que concepto pertenece, si este se llama: **Ramas**; entonces mandara un mensaje de "**Hay relación Semántica**"; así como se muestra en la Imagen 10 a continuación.

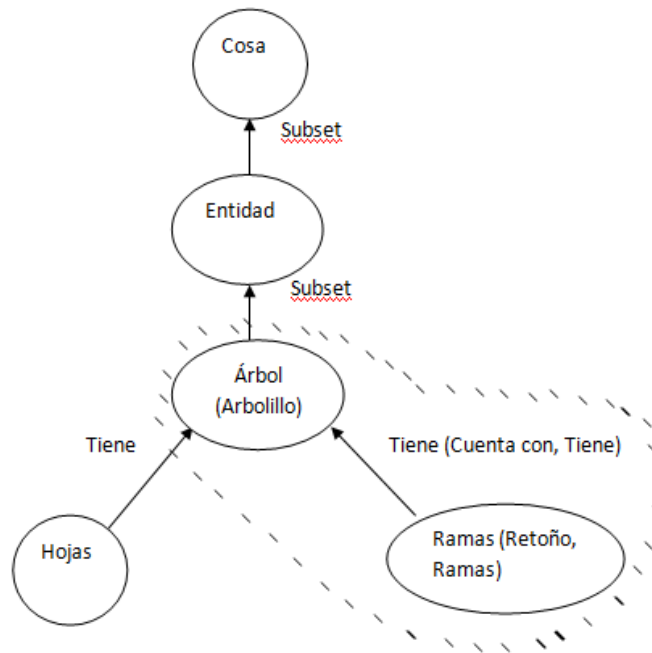


Imagen 10. *Ontología para el Caso 4.*

Explicación: El Algoritmo inicia con la búsqueda de la **Ontología** (como por ejemplo: OntoPrueba.ont) una vez encontrada la **Ontología**; recibe la tripleta del usuario (**Concepto=C, Relación=R, Valor=V**), una vez ingresada la tripleta realiza la búsqueda del **Concepto** en la **Ontología**, Sino existe realiza la función **Sinónimos** la cual consta de buscar sinónimos del **Concepto** si uno de los sinónimos del **Concepto** es igual al ingresado al de la tripleta realiza la siguiente función la cual es buscar la **Relación** sino existe en la **Ontología**; entonces realiza la función **Sinónimos** la cual es buscar un sinónimo de la **Relación**.

Si este sinónimo es igual al nombre de la **Relación** ingresado en la tripleta, entonces ejecutara la siguiente función de buscar el **Valor** sino lo encuentra realizara la función **Sinónimos**, y si un **Sinónimo** del **Valor** es igual al ingresado al de la tripleta entonces mandara un mensaje de “Hay Relación Semántica”; como se observa a continuación en la imagen 11.

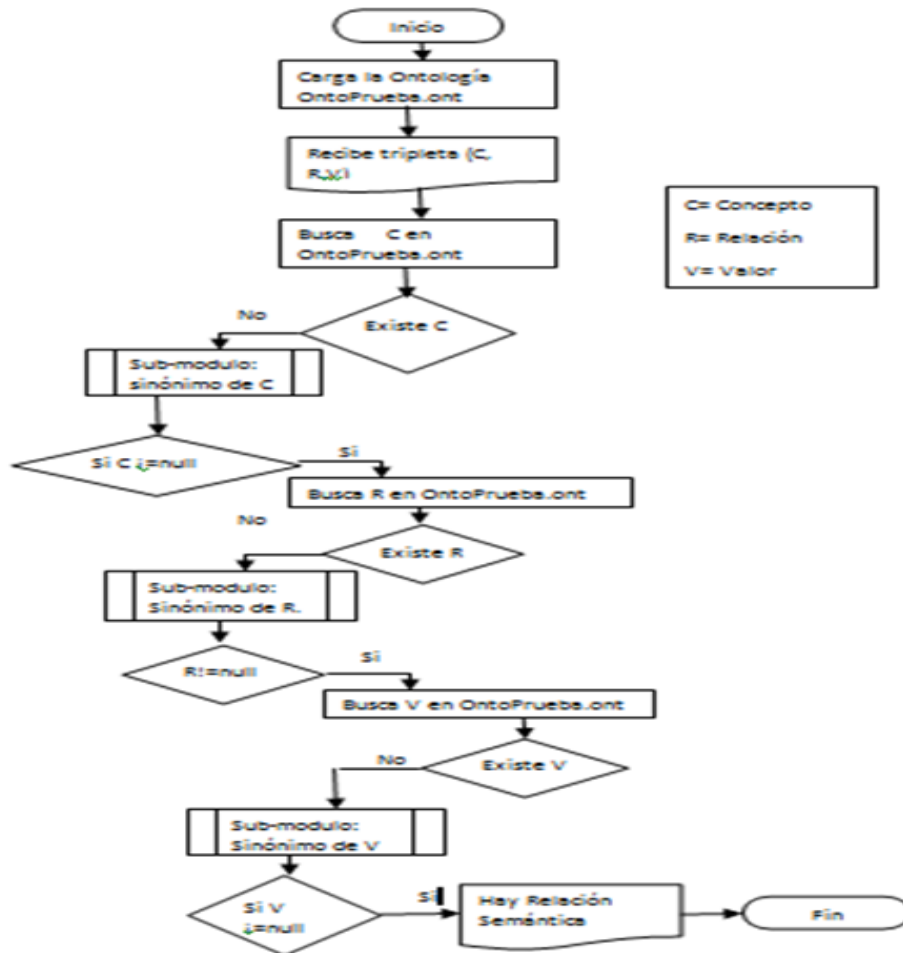


Imagen 11. Algoritmo que explica el funcionamiento del Caso 4. “trata de hallar la relación cuenta con y el valor retoño dentro del concepto arbolillo”

Caso 5: Este ejemplo “Hallar la relación tiene y el valor ramas con el concepto planta leñosa”.

Concepto= Planta leñosa; Relación= Tiene; Valor= Ramas.

Explicación: En este Ejemplo se analizará la tripleta (planta leñosa, tiene, ramas) en la ontología llamada **OntoPrueba.ont**, Buscará el concepto: **Planta leñosa** si no lo encuentra realizará la búsqueda de sinónimos, en caso de no encontrarlo entonces realizará otra búsqueda la cual consiste en la herencia; si lo encuentra por herencia pero no encuentra sus relaciones entonces; del concepto buscado irá buscando sus padres de tal forma buscará en sus relaciones la relación: **Tiene** si lo encuentra buscará el valor: **Ramas**; si lo encuentra en uno de sus padres del concepto buscado entonces mandará un mensaje de “**Hay relación Semántica**”, si no lo encuentra en ninguno de sus padres entonces mandará un mensaje de “**No Hay relación Semántica**”; así como se muestra en la Imagen 12 a continuación.

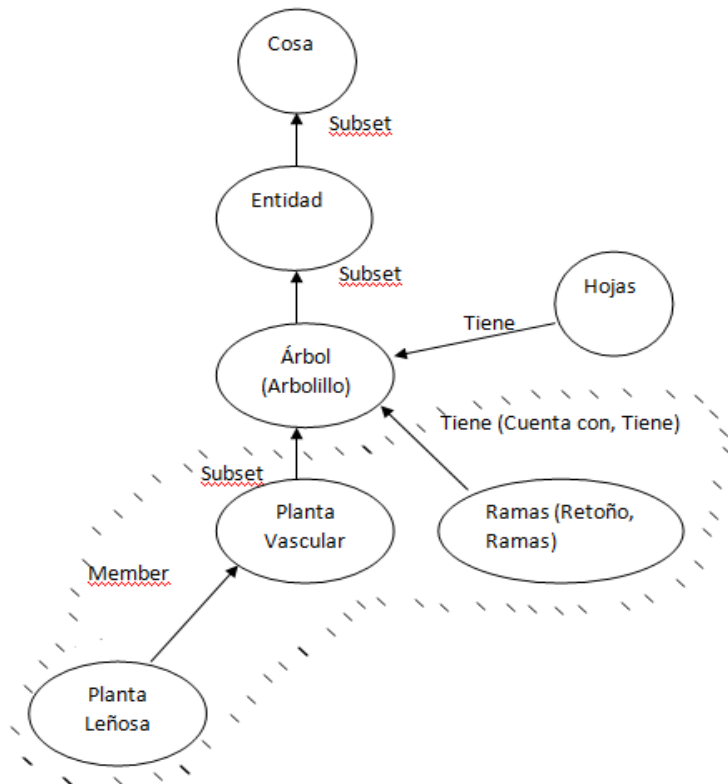


Imagen 12. *Ontología para el Caso 5.*

Explicación: El Algoritmo inicia con la búsqueda de la **Ontología** (como por ejemplo: OntoPrueba.ont) una vez encontrada la **Ontología**; recibe la tripleta del usuario (**Concepto=C, Relación=R, Valor=V**), una vez ingresada la tripleta realiza la búsqueda del **Concepto** en la **Ontología**, Sino existe realiza la función **Sinónimos** la cual consta de buscar sinónimos del **Concepto** pero si no lo encuentra por la función **Sinónimos** entonces realiza la siguiente función de **Padres** (Función Herencia) la cual es buscar sus padres del **Concepto** si tiene padres el **Concepto** entonces se realiza la búsqueda de la **Relación** y el **Valor** en el **Concepto** padre si son iguales a los ingresados en la tripleta entonces mandara un mensaje de “**Hay Relación Semántica**”; como se observa a continuación en la imagen 13.

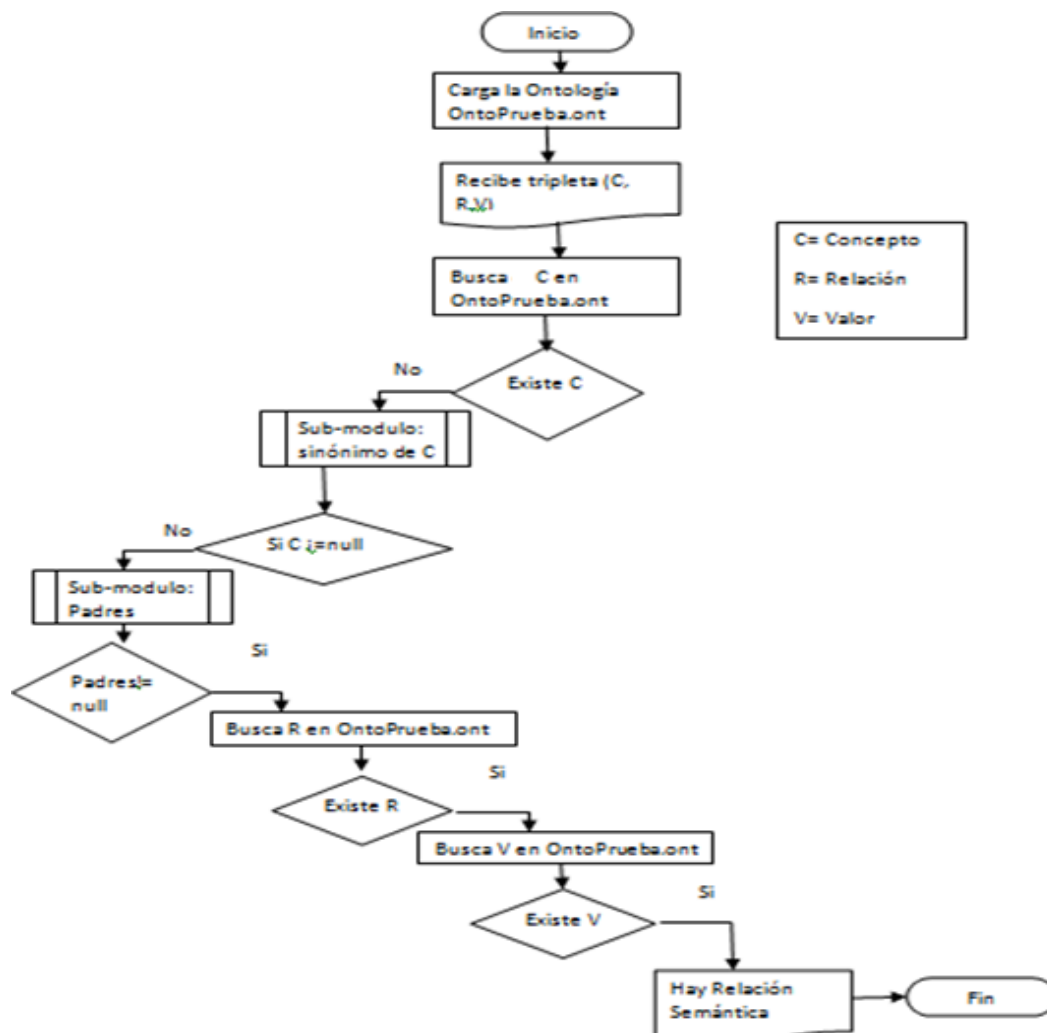


Imagen 13. Algoritmo que explica el funcionamiento del Caso 5. “trata de hallar la relación tiene y el valor ramas con el concepto planta leñosa”

Caso 6: Este ejemplo “Hallar la relación color y el valor amarillo con el concepto Arboleda”

Concepto= Arboleda; Relación= Color; Valor= Amarillo.

Explicación: Para este Ejemplo se utilizará la siguiente tripleta (Arboleda, Color, Amarillo) en la ontología como en todos los casos **OntoPrueba.ont**, en la cual buscara el concepto: **Arboleda** si no lo encuentra realizará la búsqueda de sinónimos, si no lo encuentra entonces buscará por herencia de tal forma si tampoco lo encuentra entonces mandara un mensaje de que “No hay relación Semántica”, ya que si un elemento de la tripleta no está entonces no podrá realizar la siguiente búsqueda; así como se muestra en la Imagen 14 a continuación.

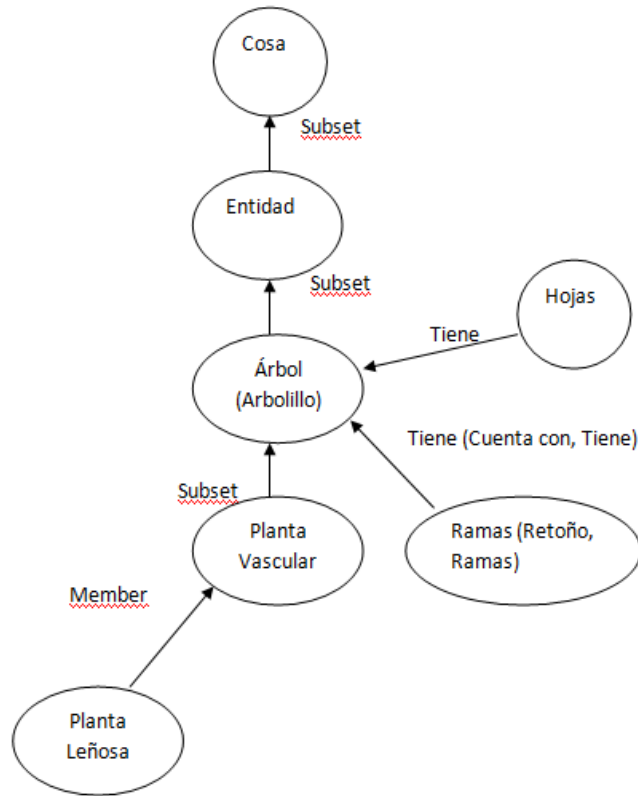


Imagen 14. *Ontología para el caso 6.*

Explicación: El Algoritmo inicia con la búsqueda de la **Ontología** (como por ejemplo: OntoPrueba.ont) una vez encontrada la **Ontología**; recibe la tripleta del usuario (**Concepto=C, Relación=R, Valor=V**), una vez ingresada la tripleta realiza la búsqueda del **Concepto** en la **Ontología**, Sino existe realiza la función **Sinónimos** la cual consta de buscar sinónimos del **Concepto** pero si no lo encuentra por la función **Sinónimos** pero si no existe entonces manda un mensaje de “**No Hay Relación Semántica**”; como se observa a continuación en la imagen 15.

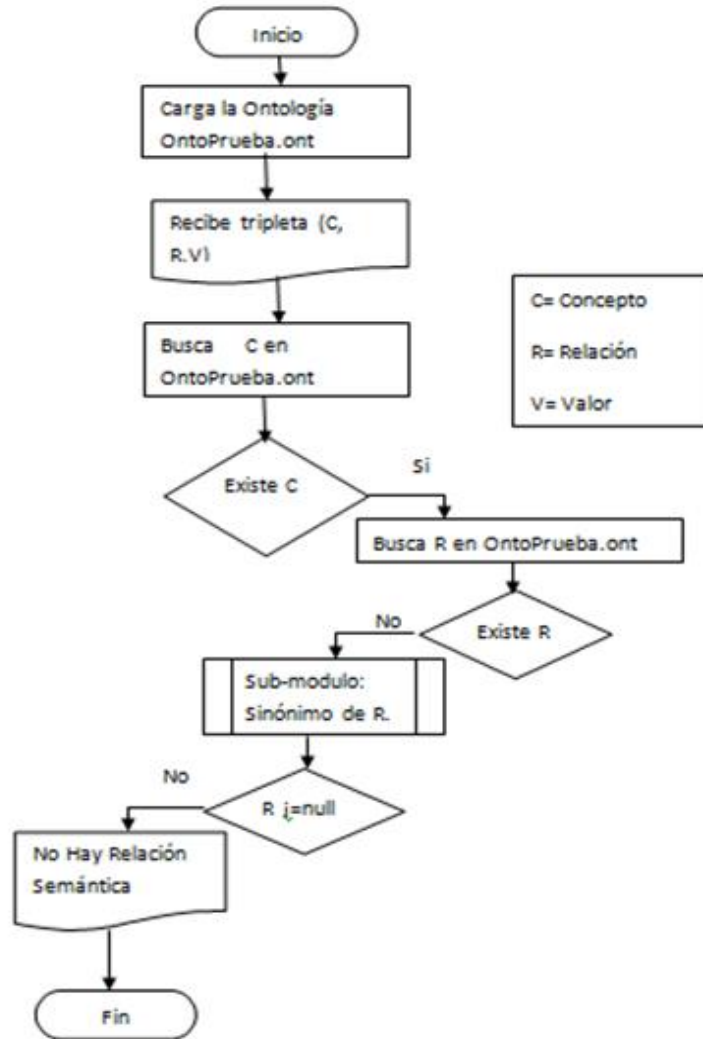


Imagen 15. Algoritmo que explica el funcionamiento del Caso 6. “trata de hallar la relación color y el valor amarillo con el concepto arboleda”

5.2.1 Diagramas

En esta parte se observará los siguientes diagramas los cuales son (Secuencias, Clases y de Caso de Uso), con el fin de observar un poco más a fondo el funcionamiento del analizador semántico.

En la imagen 16 se muestra los diagramas de Secuencias.

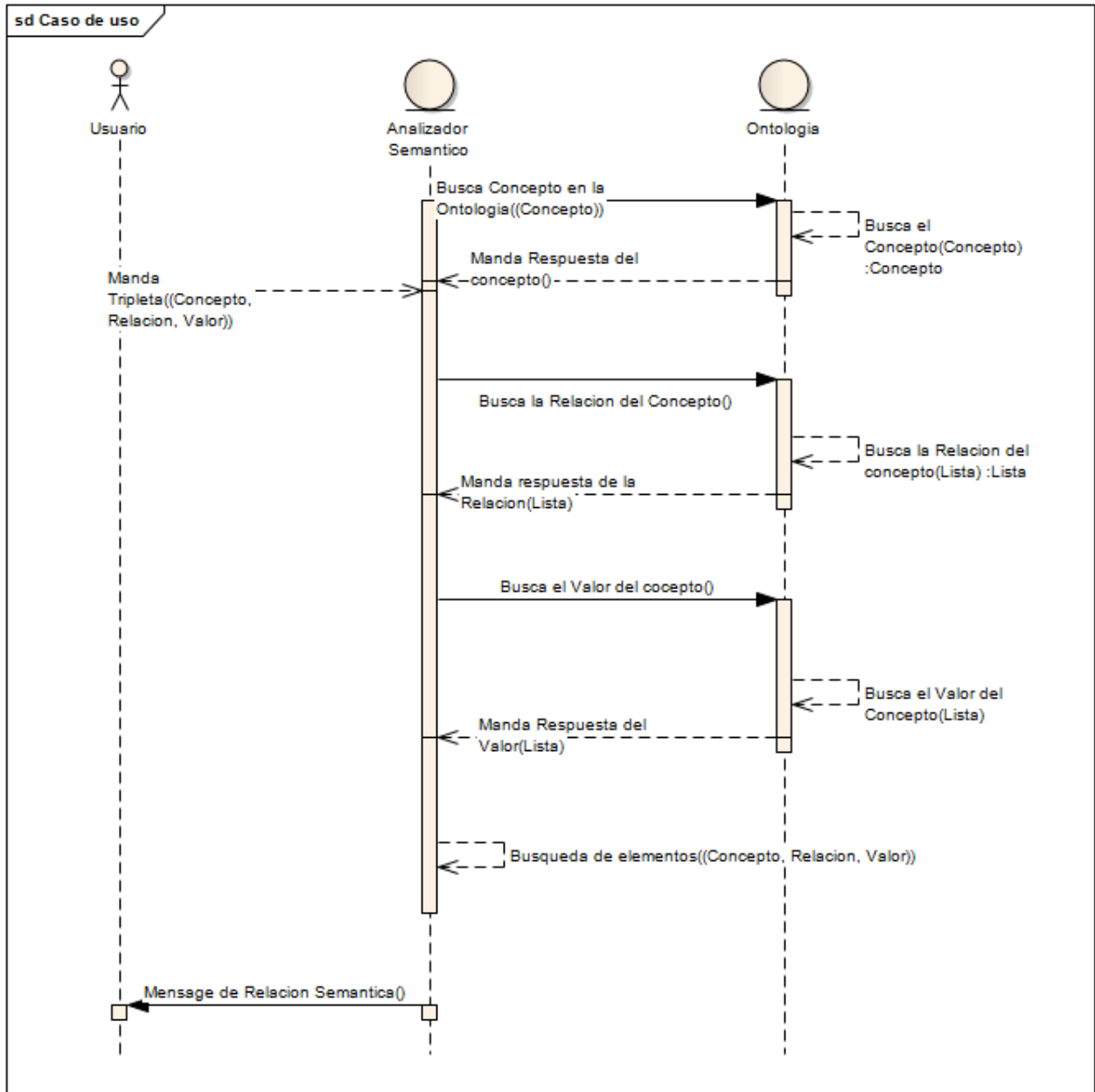


Imagen 16. Diagrama de secuencia

En la siguiente imagen se muestra el Diagrama de Clases. 17

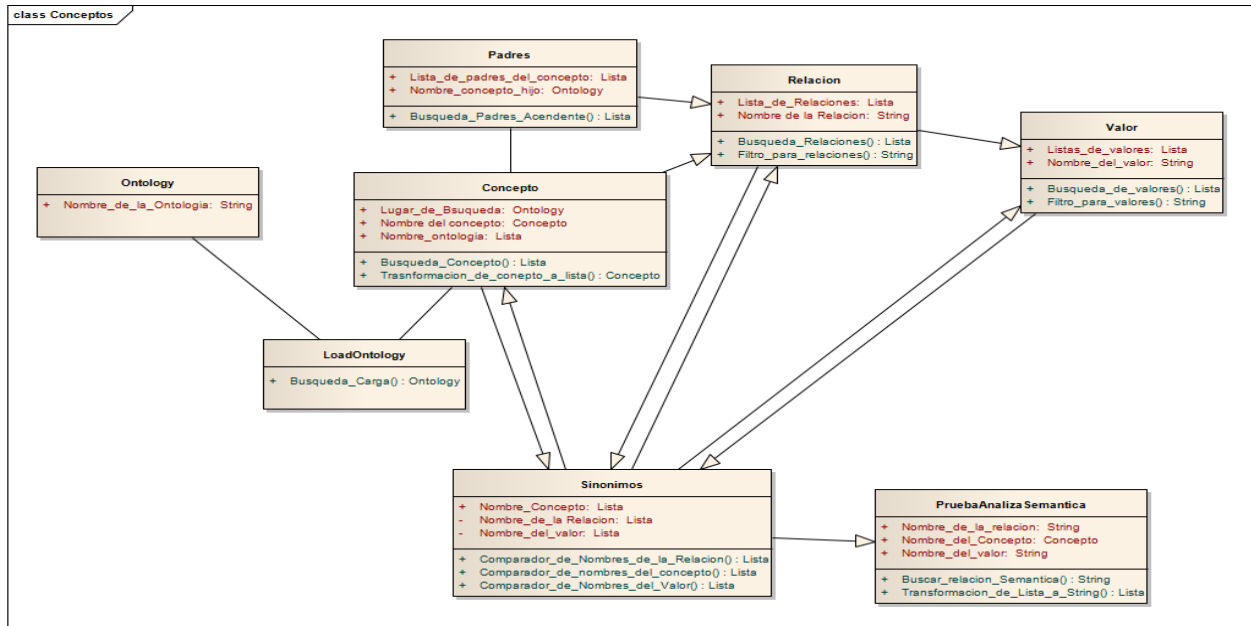


Imagen 17. Diagrama de Clases.

En la siguiente imagen se muestra el Diagrama de Casos de Uso. 18

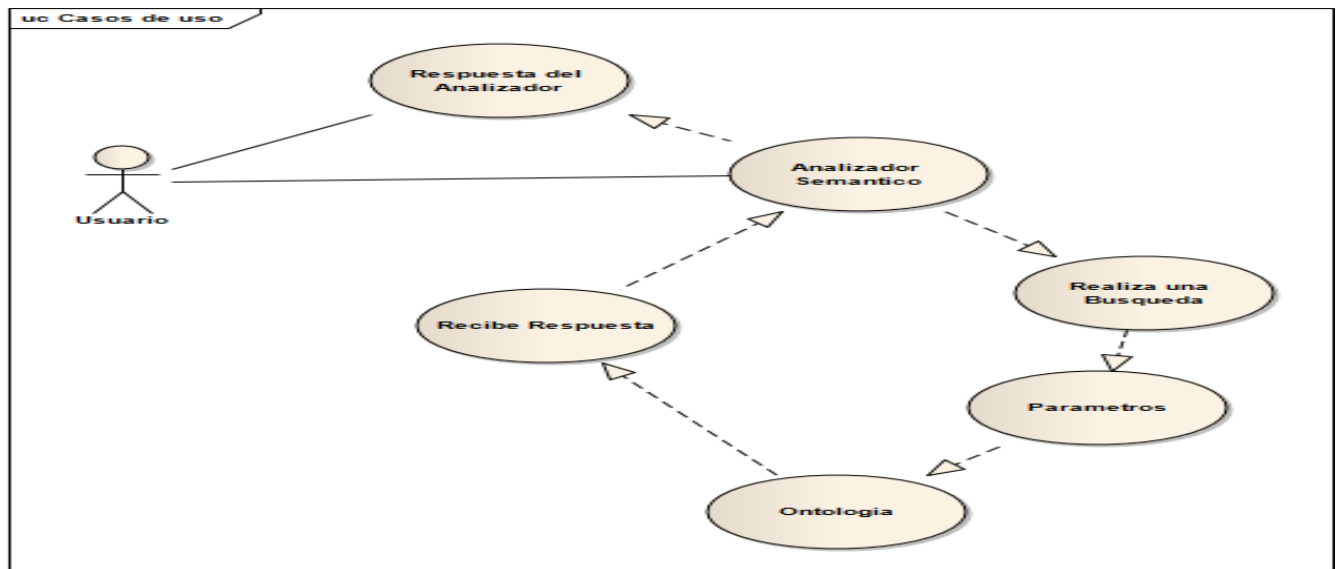


Imagen 18. Diagrama de casos de uso.

5.3 Ontologías y Consultas.

Ontologías: En esta parte se mostrarán las ontologías que se utilizaron, para demostrar que si cumple el analizador semántico mostrando relación semántica. La primera ontología a la que se le realizaron pruebas se llama “planta.owl”, a la cual se verá el árbol semántico con el que está estructurado. En la imagen 19.

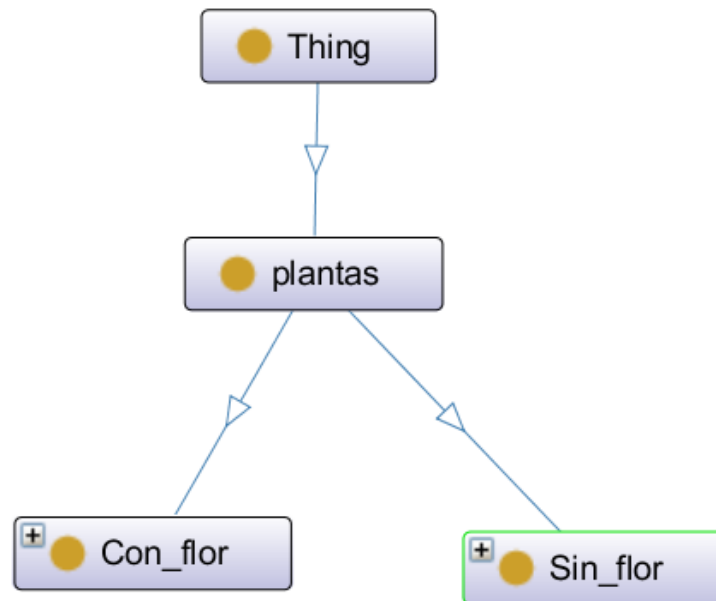


Imagen 19. *Ontología de planta.*

En la siguiente imagen se muestra cómo se va completando la red semántica de la misma ontología “planta.owl”, como lo muestra la imagen a continuación 20.

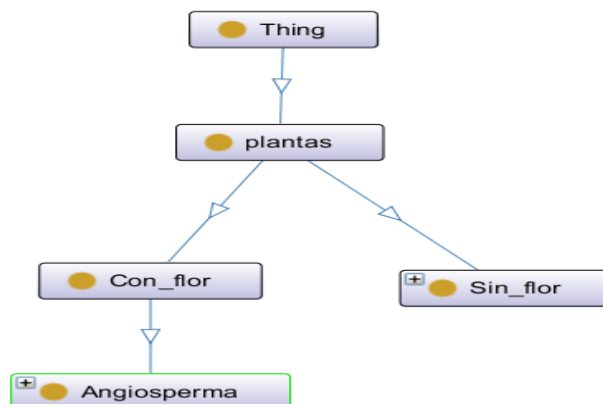


Imagen 20. *Ontología de plantas casi terminada*

En la siguiente imagen se observa la ontología ya terminada, como se muestra en la imagen 21.

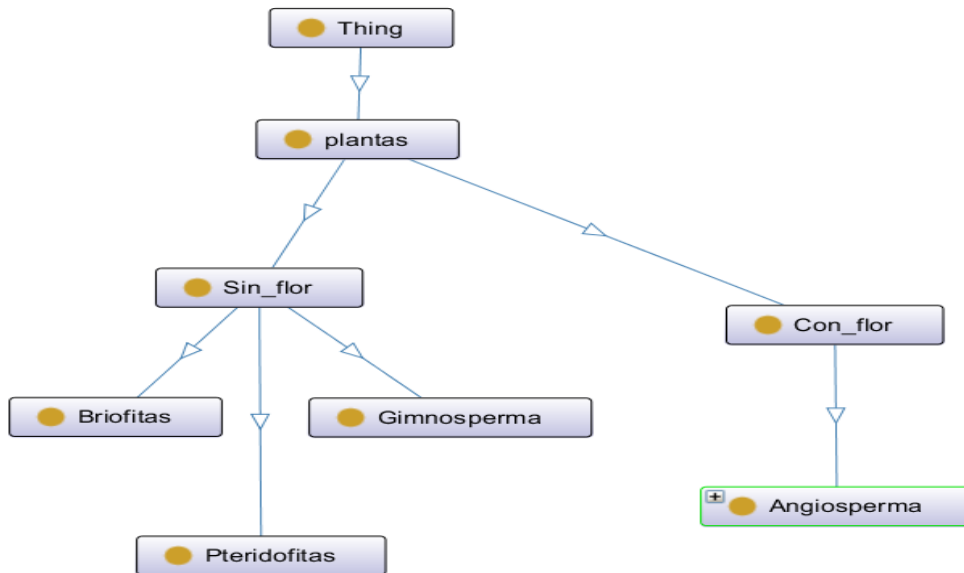


Imagen 21. *Ontología de planta terminada*

Consultas.

En este apartado se muestra algunas consultas en DL query realizadas a la ontología “planta.owl”; para ver si está bien diseñada en el árbol semántico, como se muestran a continuación en la imagen 22.

DL query:

Query (class expression)

plantas and Con_flor

Execute Add to ontology

Query results

Sub classes (2)

- Angiosperma
- Nothing

Imagen 22. *Aplicación de consultas a la ontología planta*

En la imagen se muestra otra consulta realizada a la misma ontología planta 23.

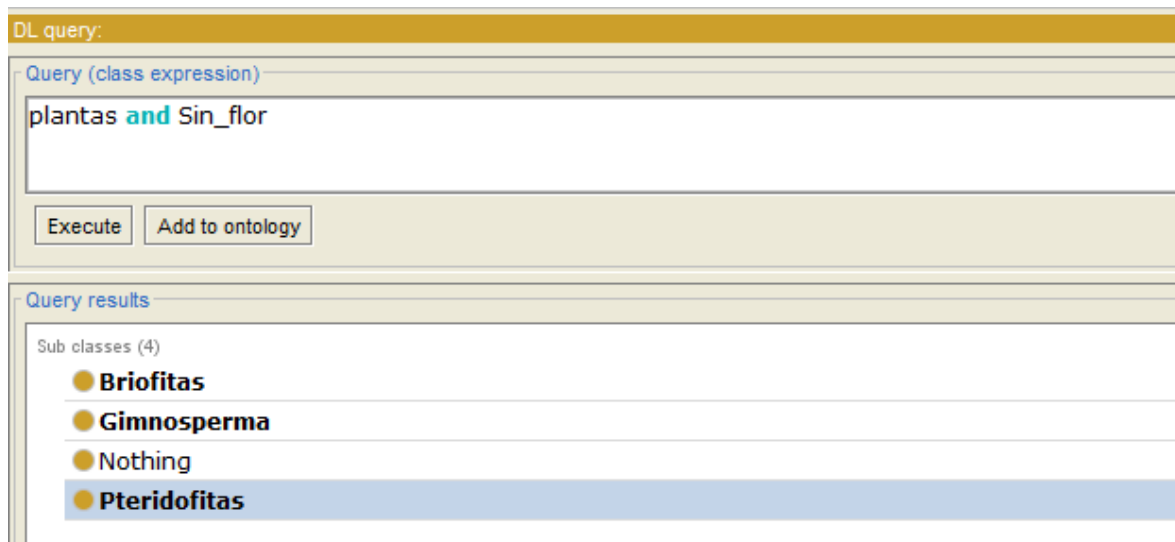


Imagen 23. Aplicación de consultas a la ontología planta

En la imagen 24 se muestra la última consulta realizada a la ontología planta como se muestra a continuación.

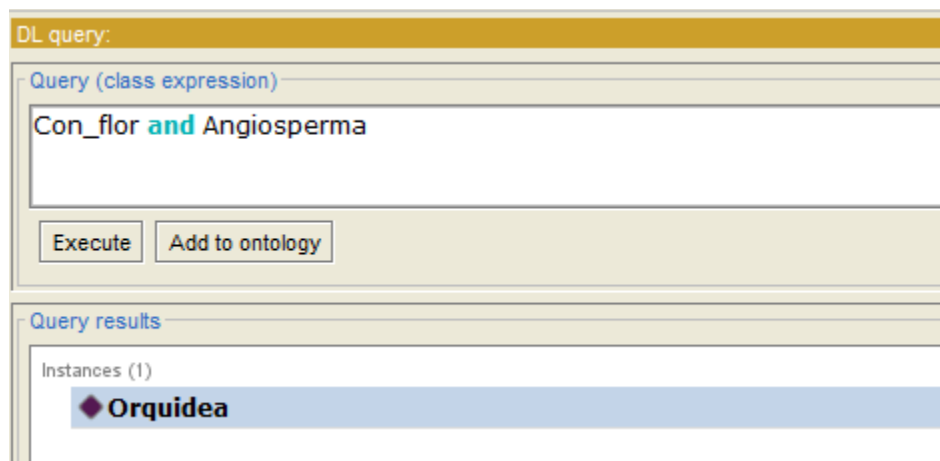


Imagen 24. Última consulta realizada a la ontología "planta.owl"

La siguiente ontología es de los diferentes tipos de arboles a esta ontología tiene el nombre de "arboles.owl", a la cual también se le realizaron consultas para verificar semántica. En la imagen 25 se muestra una parte de la ontología arboles.

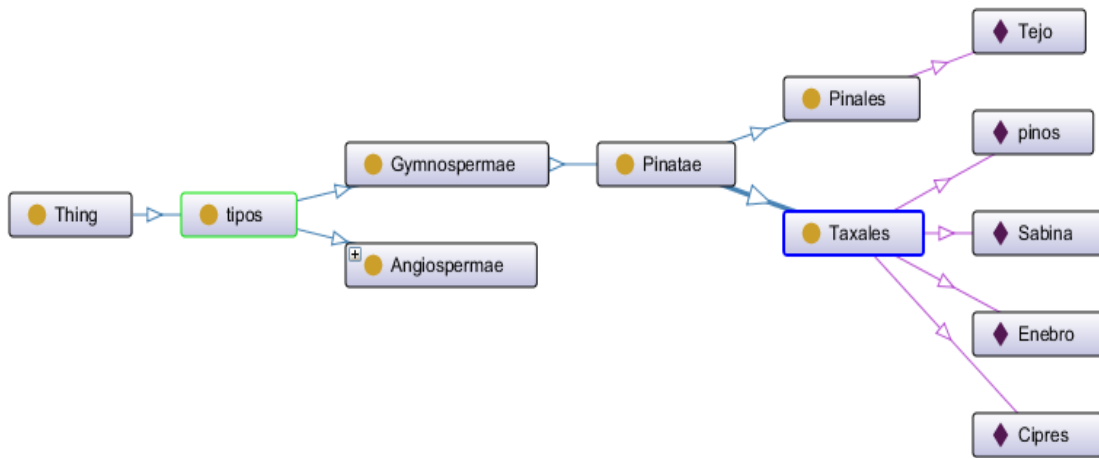


Imagen 25. *Primera parte de la ontología arboles.*

En la imagen 26 se muestra la otra parte de la ontología.

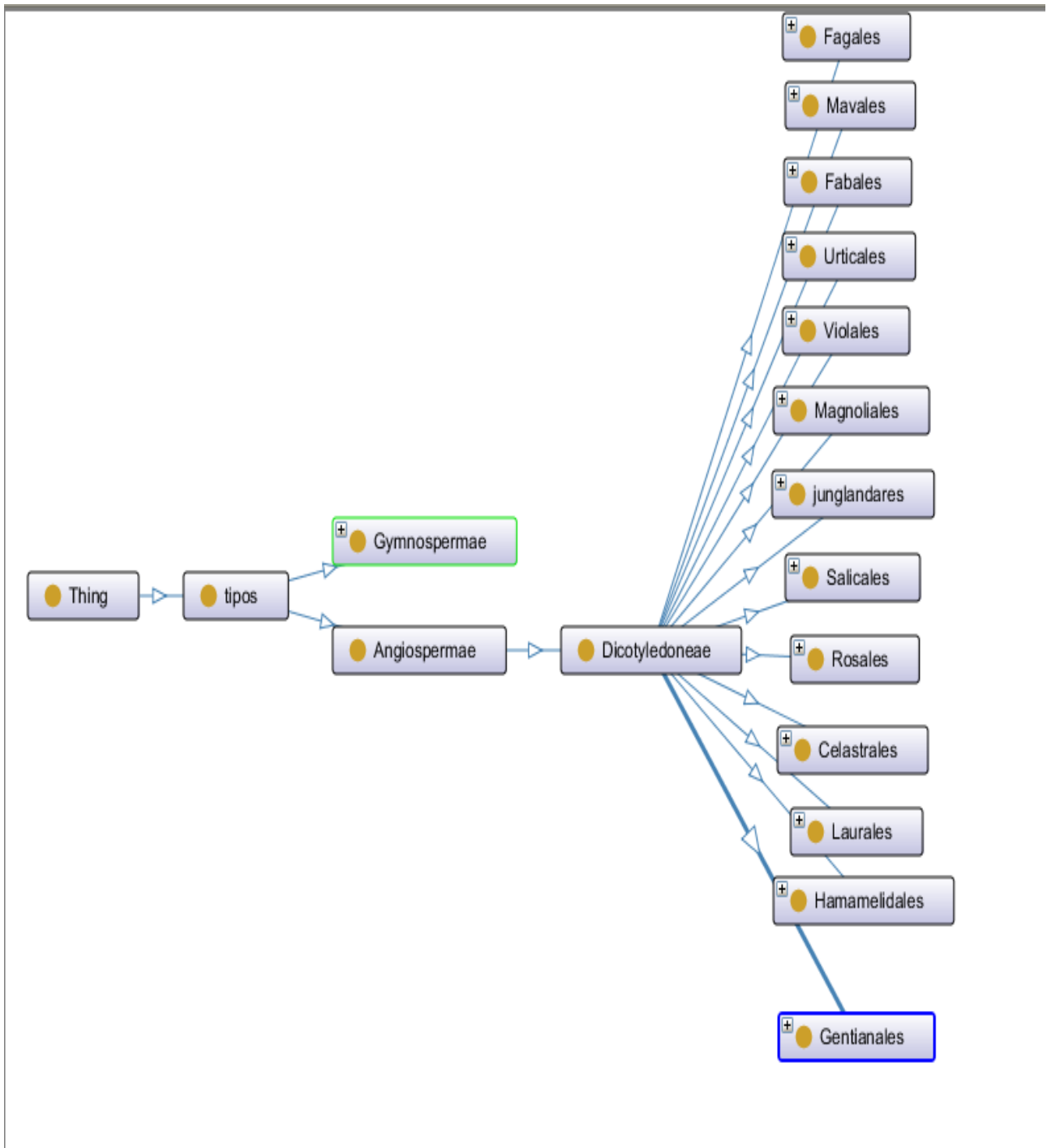


Imagen 26. Segunda parte de la ontología

En la imagen 27 se muestra algunos atributos que tiene el árbol para observar cómo está constituido.

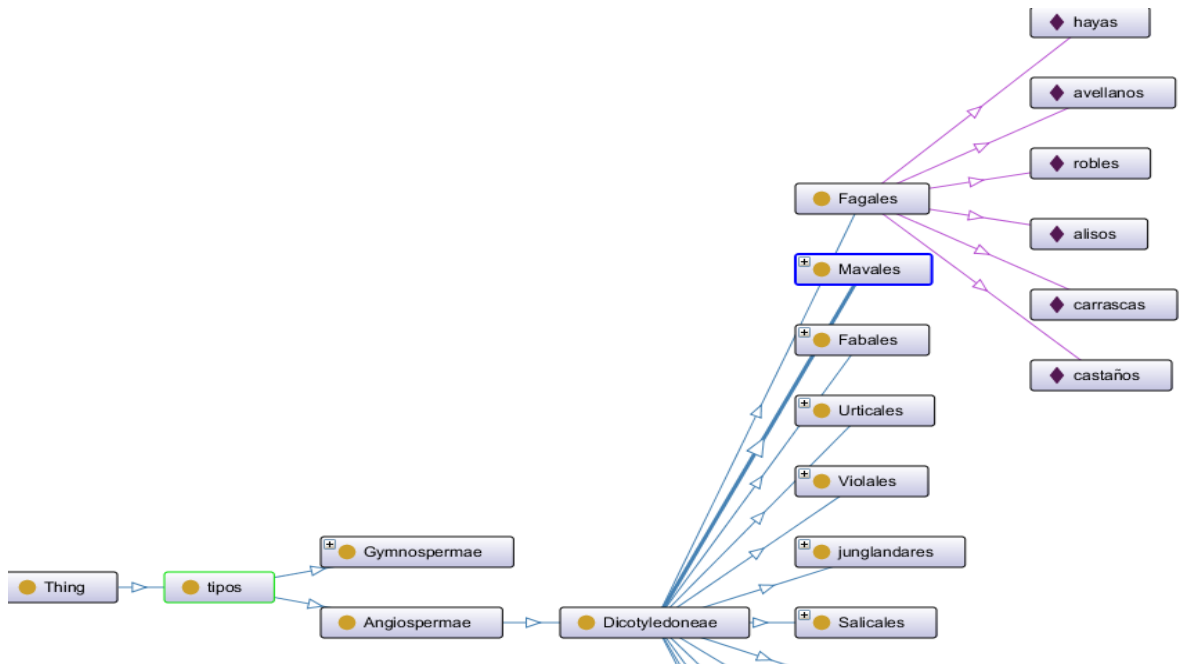


Imagen 27. *Árbol de la ontología Arboles y sus atributos.*

En la imagen 28 se muestra la consulta a la ontología en DL query para demostrar que está bien el árbol.

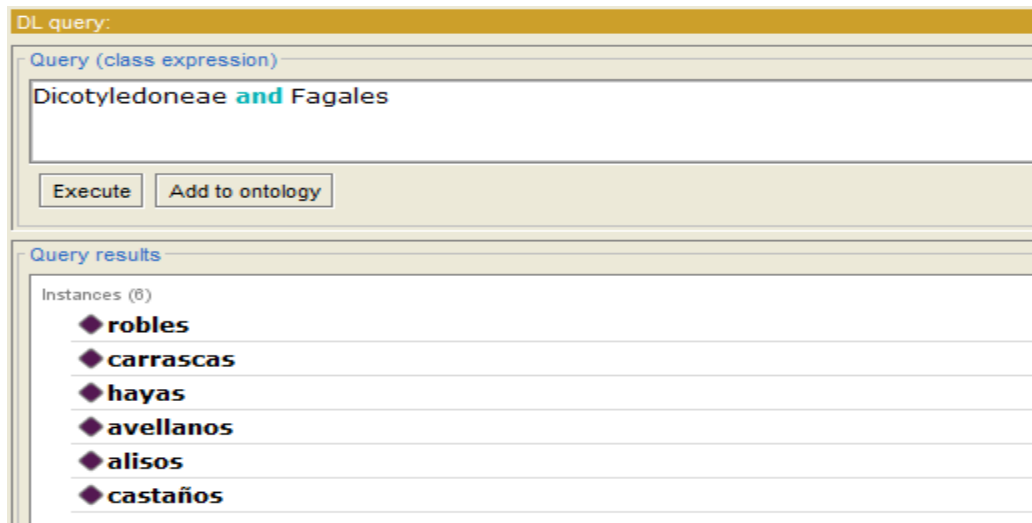


Imagen 28. *Consulta realizada a la ontología arboles.*

A continuación se muestran la imagen 29, donde se observa el árbol semántico.

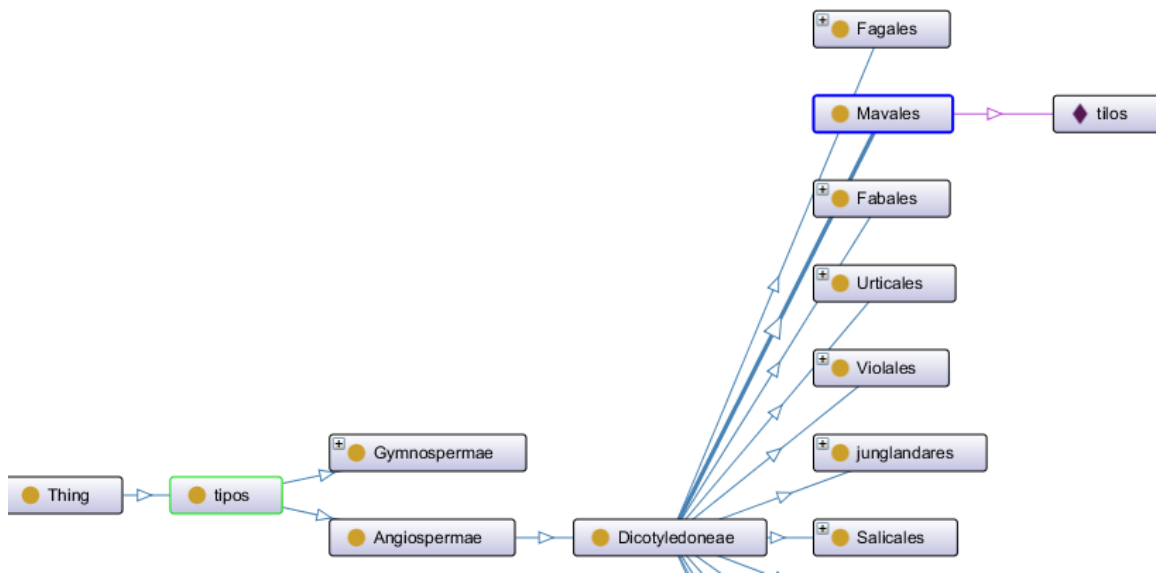


Imagen 29. *Árbol de ontología donde una clase muestra sus atributos.*

En la siguiente imagen 30 se muestra la consulta realizada para verificar el árbol semántico.

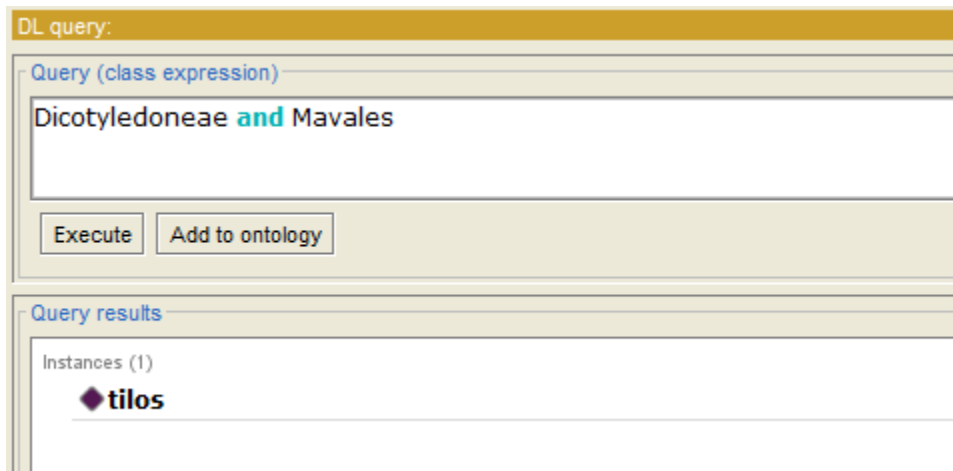


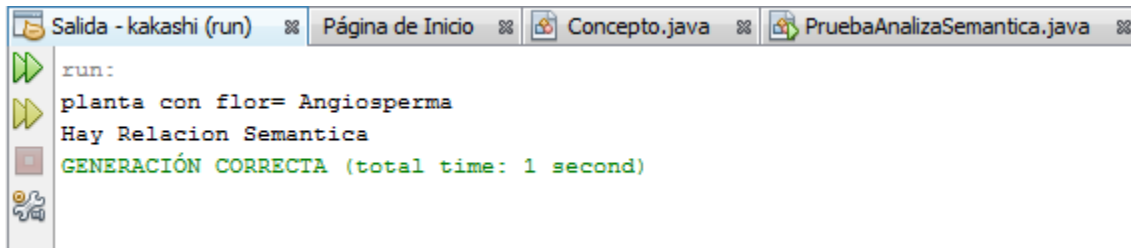
Imagen 30. *Imagen de de la consulta realizada a la ontología.*

5.4 Pruebas

En este apartado se muestran la semántica que tiene el analizador semántico ya terminado, mostrando la terminación final del programa.

5.4.1 Pruebas de escritorio realizadas a la ontología planta.

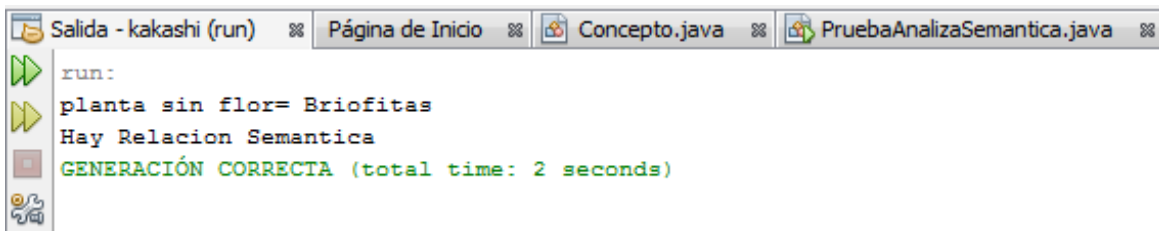
En la imagen 31 se muestra la prueba del analizador semántico funcionando de igual forma como se mostro en la ontología,



```
Salida - kakashi (run)  Pagina de Inicio  Concepto.java  PruebaAnalizaSemantica.java
run:
planta con flor= Angiosperma
Hay Relacion Semantica
GENERACIÓN CORRECTA (total time: 1 second)
```

Imagen 31. Resultados del Analizador Semántico de la ontología planta.

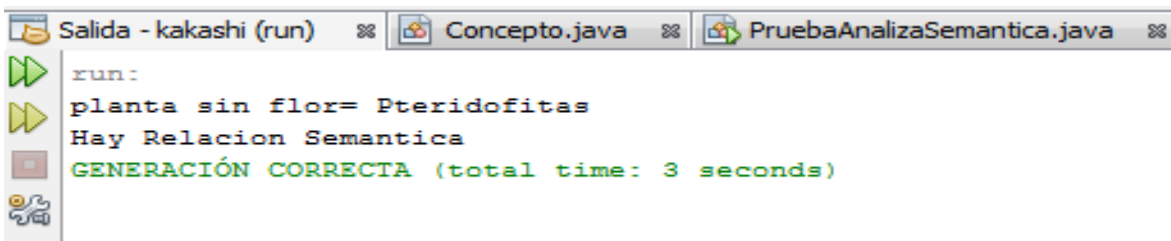
En la imagen 32 se muestra otro ejemplo aplicado a la ontología planta.



```
Salida - kakashi (run)  Pagina de Inicio  Concepto.java  PruebaAnalizaSemantica.java
run:
planta sin flor= Briofitas
Hay Relacion Semantica
GENERACIÓN CORRECTA (total time: 2 seconds)
```

Imagen 32. Resultados del Analizador Semántico de la ontología planta

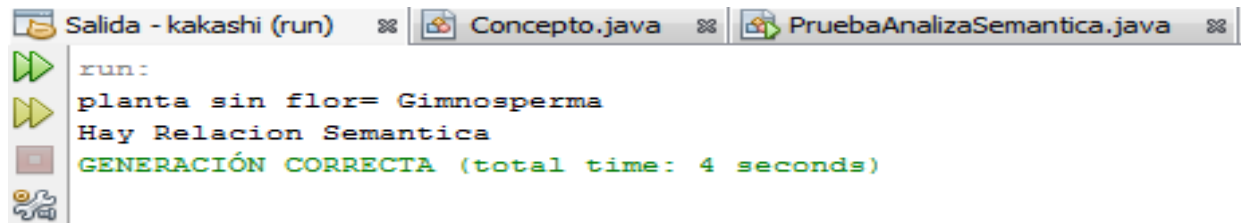
En la siguiente imagen se muestra otro ejemplo aplicado a la ontología planta. 33



```
Salida - kakashi (run)  Concepto.java  PruebaAnalizaSemantica.java
run:
planta sin flor= Pteridofitas
Hay Relacion Semantica
GENERACIÓN CORRECTA (total time: 3 seconds)
```

Imagen 33. Resultados del Analizador Semántico de la ontología planta

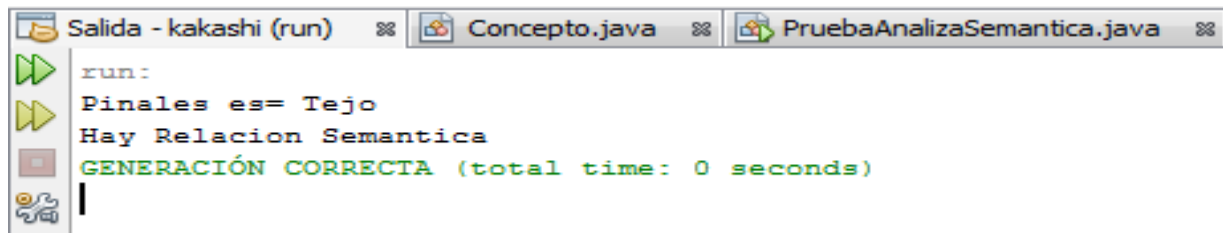
En la imagen 34 se muestra otro ejemplo aplicado al analizador semántico de la ontología planta.



```
Salida - kakashi (run)  Concepto.java  PruebaAnalizaSemantica.java
run:
planta sin flor= Gimnosperma
Hay Relacion Semantica
GENERACIÓN CORRECTA (total time: 4 seconds)
```

Imagen 34. Resultados del Analizador Semántico de la ontología planta

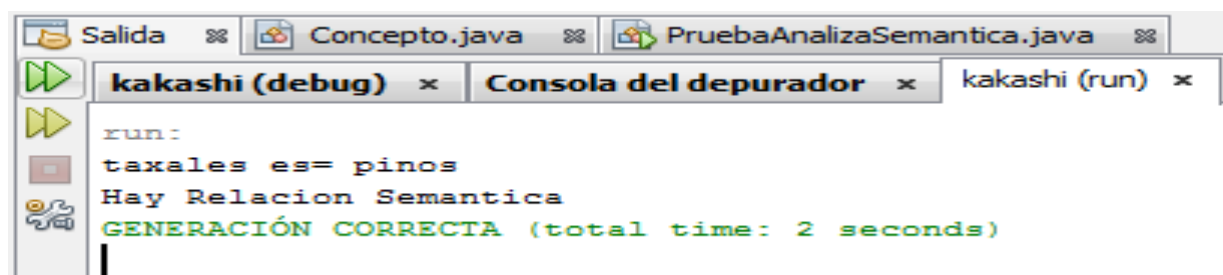
En la imagen 35 se muestra el resultado de la consulta realizada a la ontología arboles en el analizador semántico.



```
Salida - kakashi (run)  Concepto.java  PruebaAnalizaSemantica.java
run:
Pinales es= Tejo
Hay Relacion Semantica
GENERACIÓN CORRECTA (total time: 0 seconds)
```

Imagen 35. Resultado del analizador Semántico a la ontología arboles.

En la imagen 36 se muestra otro resultado de la consulta realizada a la misma ontología en el analizador semántico.



```
Salida  Concepto.java  PruebaAnalizaSemantica.java
kakashi (debug) x  Consola del depurador x  kakashi (run) x
run:
taxales es= pinos
Hay Relacion Semantica
GENERACIÓN CORRECTA (total time: 2 seconds)
```

Imagen 36. Resultado del analizador Semántico a la ontología arboles.

En la siguiente imagen se muestra el resultado del analizador semántico aplicado a la ontología árboles. 37

```
run:
taxales es= sabina
Hay Relacion Semantica
GENERACIÓN CORRECTA (total time: 2 seconds)
```

Imagen 37. Resultado del analizador semántico aplicado a la ontología arboles.

En la imagen 38 se muestra el resultado aplicado del analizador semántico aplicado a la ontología arboles.

```
run:
taxales es= enebro
Hay Relacion Semantica
GENERACIÓN CORRECTA (total time: 3 seconds)
```

Imagen 38. Resultado del analizador semántico aplicado a la ontología arboles.

En la imagen 39 se muestra el resultado aplicado del analizador semántico aplicado anteriormente en las consultas de la ontología arboles.

```
run:
taxales es= cipres
Hay Relacion Semantica
GENERACIÓN CORRECTA (total time: 2 seconds)
```

Imagen 39. Resultado del analizador semántico aplicado a la ontología arboles.

Capítulo 6

RESULTADOS.

6.1.1 Pruebas en GUI del Analizador Semántico.

En la imagen 40 se muestra la pantalla de inicio el cual ayuda a iniciar el programa del analizador semántico.

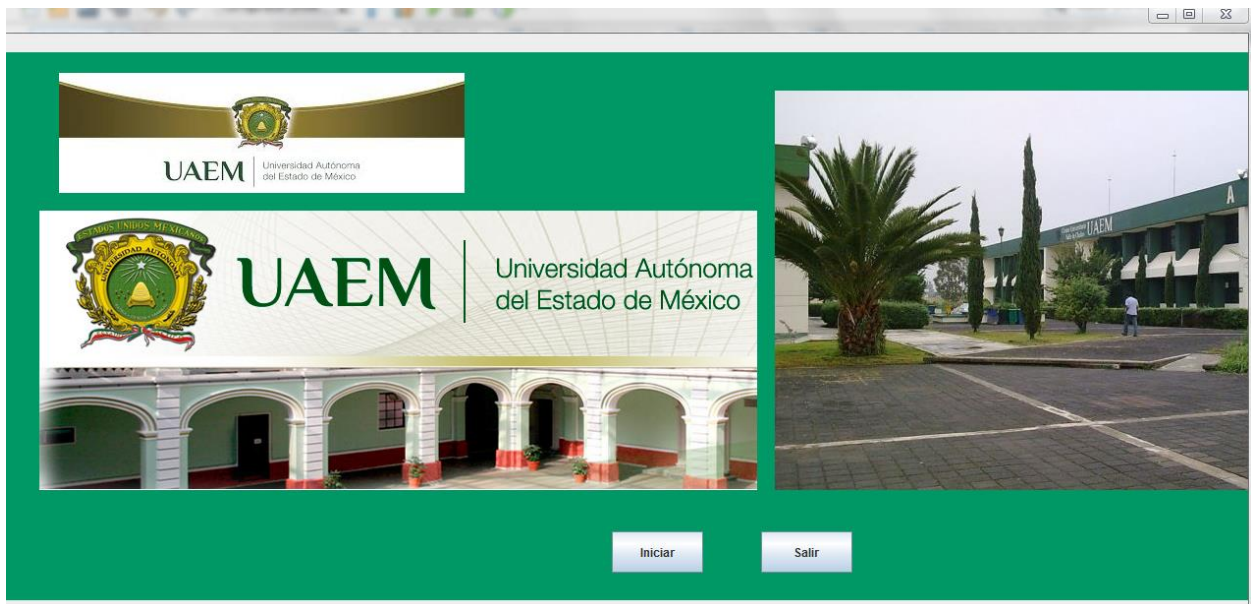


Imagen 40. *Imagen de portada del Analizador Semántico.*

En la imagen 41 se muestra la GUI del analizador semántico y los componentes que tiene.

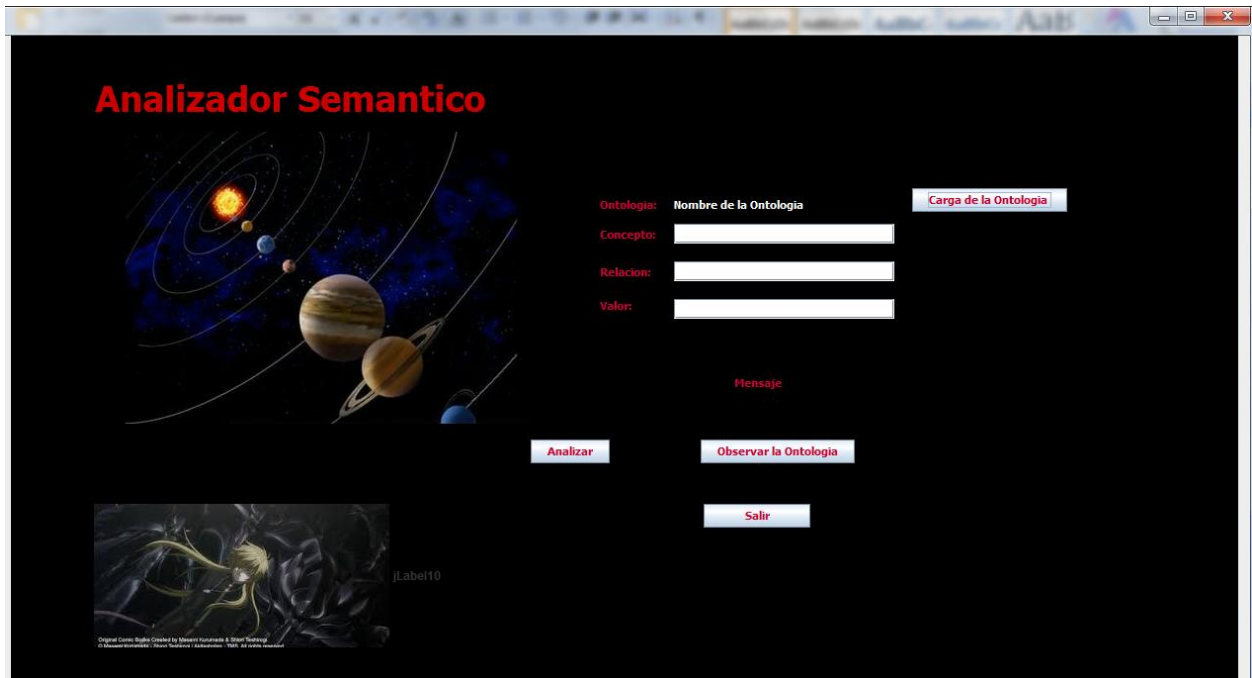


Imagen 41. Imagen de la GUI del Analizador Semántico.

En la imagen 42 se muestra el analizador semántico, mostrando la ontología que se desea utilizar.

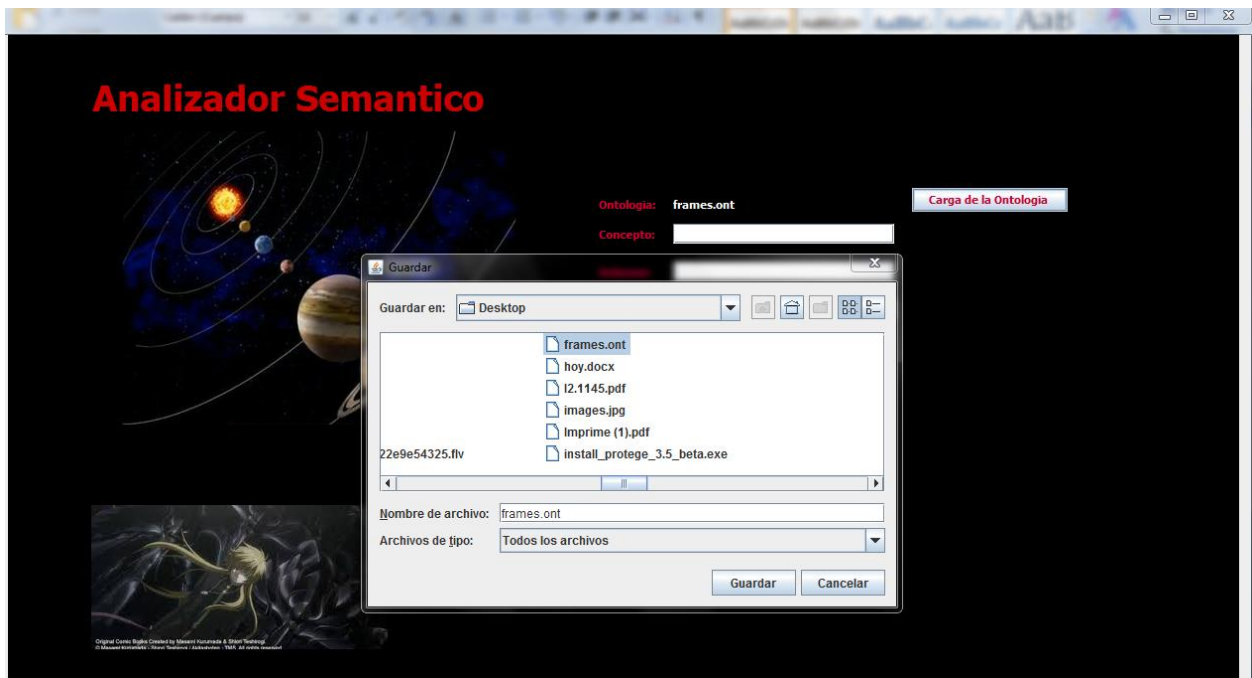


Imagen 42. Búsqueda de la Ontología en el Analizador Semántico.

En la imagen 43 se muestra la ontología seleccionada por el usuario.

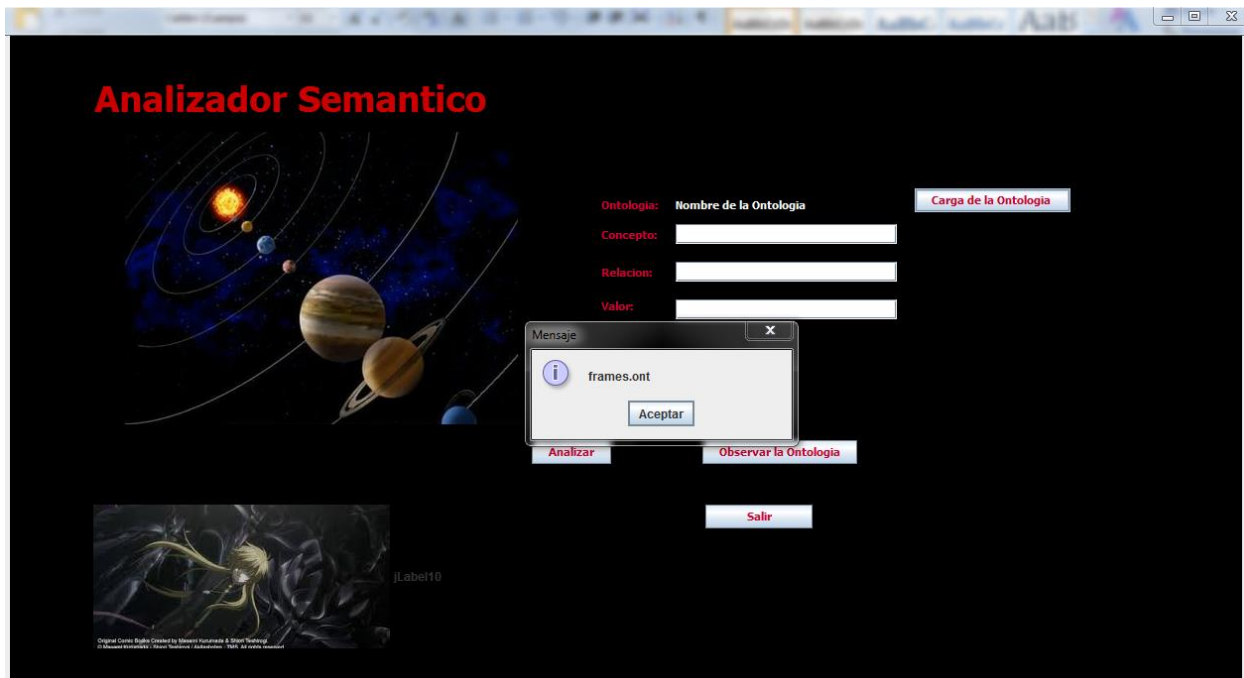


Imagen 43. *Ontología encontrada por el Analizador Semántico.*

En la imagen 44 se muestra la triplete ingresada por el usuario.

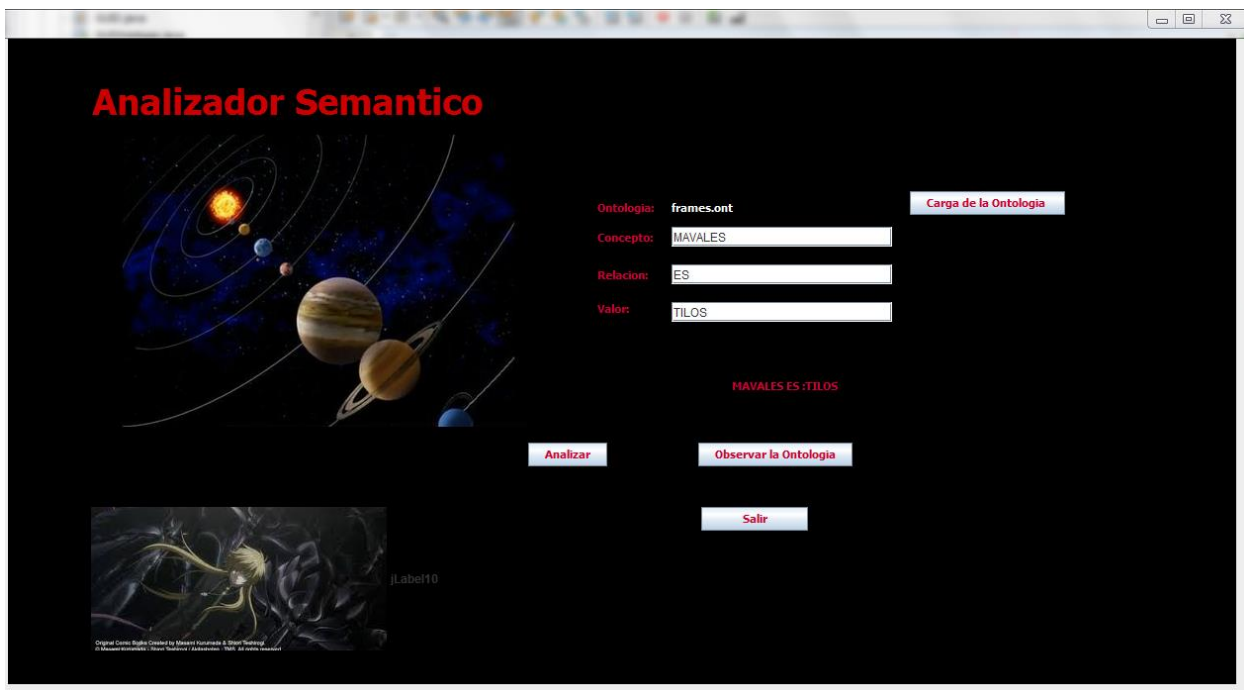


Imagen 44. *Datos ingresados por el usuario (Triplete).*

En la imagen 45 se muestra el resultado del analizador.

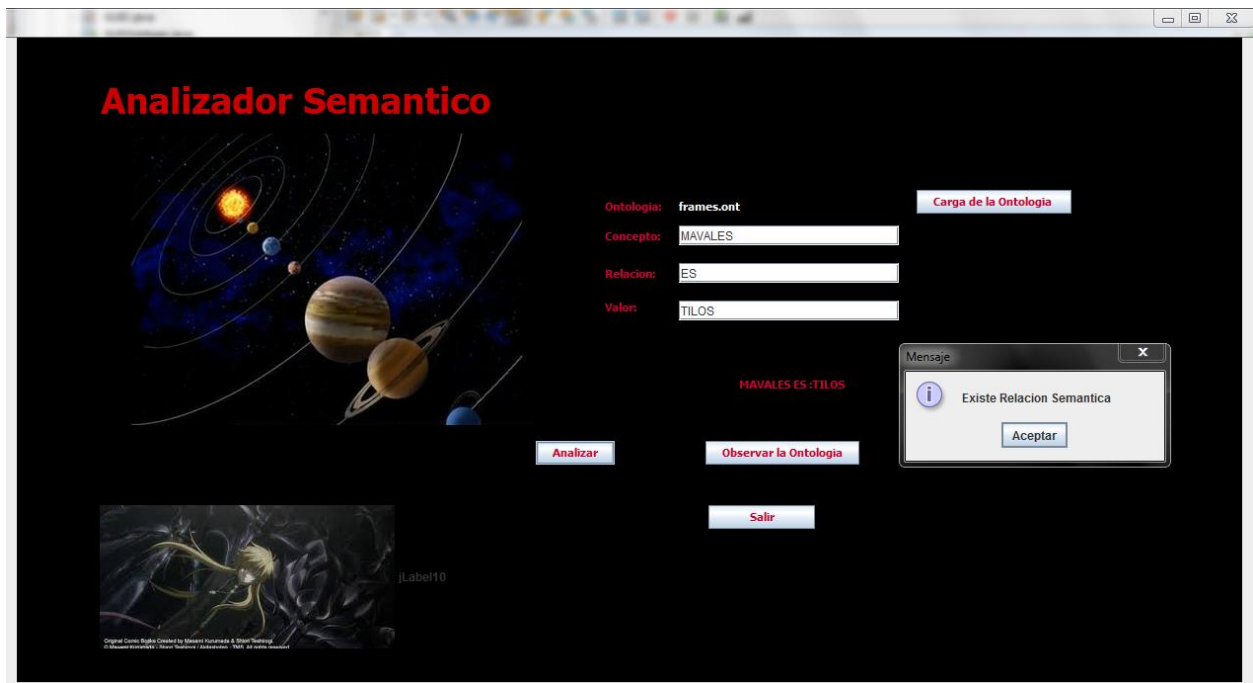


Imagen 45. Respuesta del Analizador Semántico.

En la imagen 46 se muestra la antepenúltima ventana la cual muestra como está constituida la ontología.

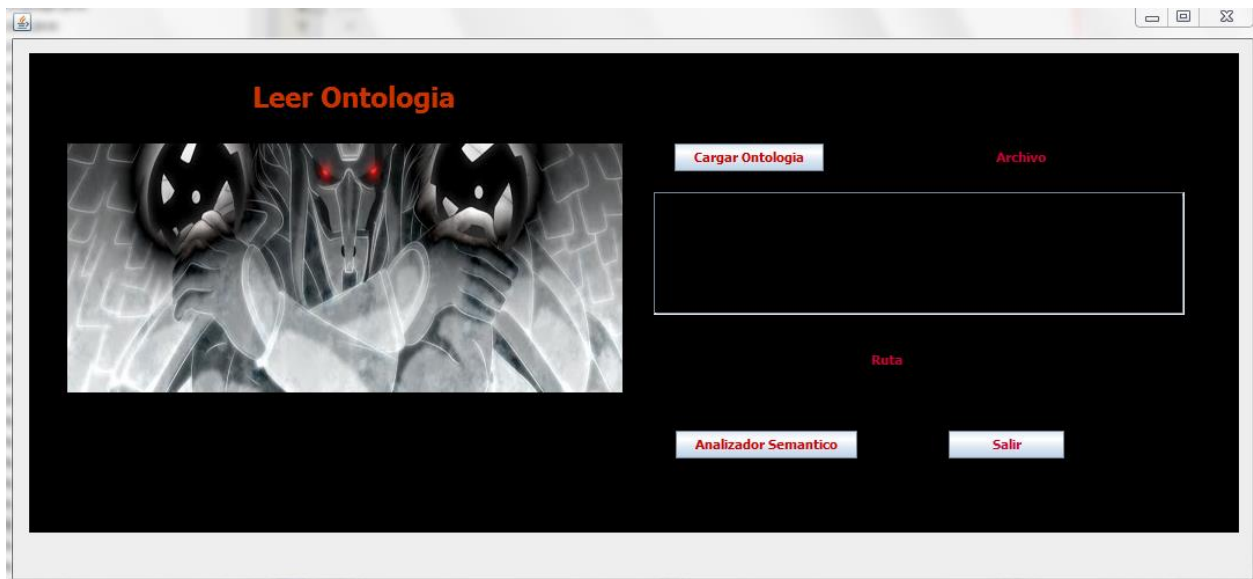


Imagen 46. Imagen de Lectura de la Ontología.

En la imagen 47 se muestra el contenido de la ontología utilizada.

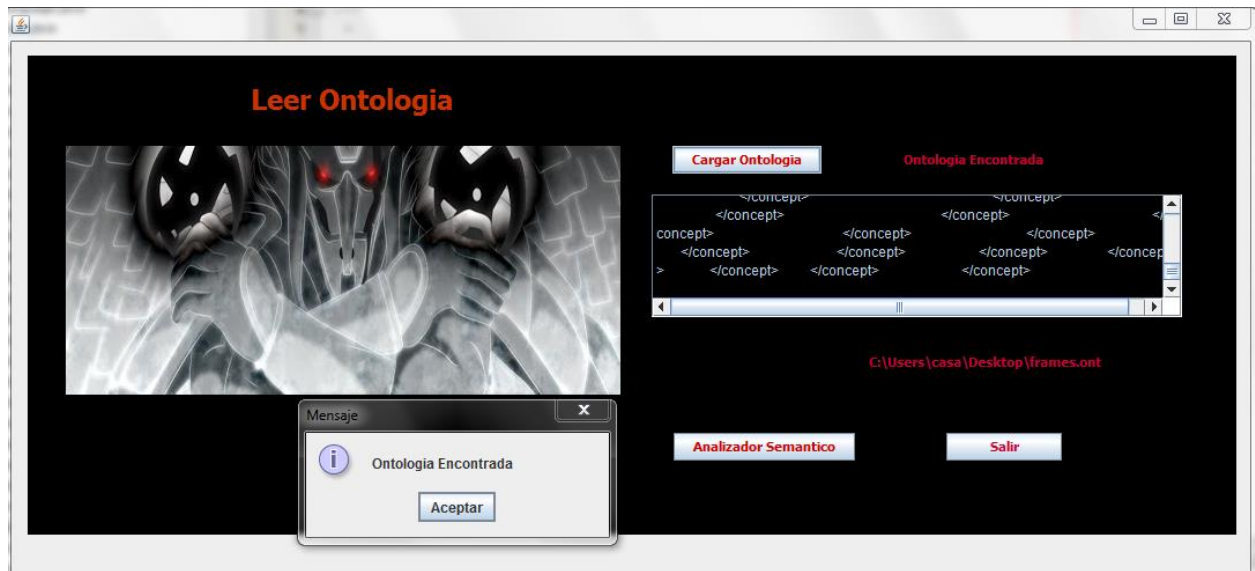


Imagen 47. Datos de la Ontología Seleccionada.

6.2 Análisis comparativo

Diseño de pruebas: A continuación se muestra una tabla en la cual se realizaron pruebas a los dos analizadores como el analizador semántico como a Wólfram alpha utilizando las misma ontologías, para ambos casos se utilizaron las misma oraciones semánticas; Wólfram alpha es un analizador con una base de conocimientos ya establecida por eso no es posible que acepte algunas oraciones, por eso es que en algunos casos no hubo relación semántica por otro lado el analizador se puede hacer más grande añadiéndole conceptos nuevos de tal manera tubo mas aciertos que el Wólfram alpha como se ve en la tabla a continuación.

Tabla 3: Comparativa entre Wólfram alpha y el analizador semántico, en ambos analizadores se realizaron las mismas pruebas; en las mismas ontologías.

Ontología	Analizador Semántico		Wólfram alpha	
	Tiempo de Ejecución	Existencia de relación semántica	Tiempo de Ejecución	Existencia de relación semántica
Plantas	0-3 segundos	Si	0-5 segundos	Si
Arboles	1-4 segundos	Si	0-7 segundos	Si
Orquesta Sinfónica	0-2 segundos	Si	0-8 segundos	Si
Animales	0-3 segundos	Si	0-2 segundos	No
Insectos	1-3 segundos	Si	0-3 segundos	No
Preparación de Pizzas	0-2 segundos	Si	0-8 segundos	Si
Países	0-3 segundos	Si	0-7 segundos	Si
Carros	0-2 segundos	Si	0-6 segundos	Si
Felinos	1-4 segundos	Si	0-6 segundos	Si
Frutas	0-3 segundos	Si	0-7 segundos	Si
Verduras	0-2 segundos	Si	0-5 segundos	Si
Meses	0-2 segundos	Si	0-8 segundos	Si
Futbol	1-3 segundos	Si	0-8 segundos	Si
Transporte	1-4 segundos	Si	0-7 segundos	Si
Personas	0-3 segundos	Si	0-8 segundos	Si
Ejercicios	1-4 segundos	Si	0-3 segundos	No

CONCLUSIONES

Los resultados del analizador semántico fueron satisfactorios ya que cumplió con los requerimientos para los que fue creado, realizó búsquedas semánticas exitosas lo que completaría al analizador semántico sería ser implementado a la Web para ser búsquedas más completas y útiles pero por la falta de tiempo ya no se pudo llevar a cabo.

Los beneficios de contar con un analizador semántico es la búsqueda de relaciones semánticas en ontologías lo cual es útil, si se tiene una gran cantidad de datos en una ontología y si se desea encontrar un concepto, con sus atributos o información este analizador puede proporcionarla a demás de que es dinámica su base de conocimientos.

En un futuro este proyecto se puede aplicar a la Web, además de poder implementarse no solo a una base de conocimientos sino a una base de datos de cualquier tipo; al analizador semántico se le puede crear un contestador de preguntas inteligentes la cual por medio de palabras clave active el analizador semántico y este mande una respuesta basándose en la ontología o base de datos además de proporcionar información de lo que se busca.

Wólfam alpha es un analizador muy interesante porque es un analizador semántico que trabaja con base de datos que funciona bien en ocasiones pero como no se incrementado el conocimiento de su base de datos entonces no realiza las búsquedas con conceptos desconocidos y también no realiza búsquedas de conceptos por sinónimos del concepto, lo bueno es que es una herramienta que trabaja en la web además de que proporciona información de lo que se busca si se busca bien, esto hace referencia a si se busca con las palabras clave este analizador proporcionara información de lo que se busca.

Una similitud entre Wólfam alpha y el analizador semántico es que trabaja con palabras claves, estas palabras sirven para hacer búsquedas más rápidas y en algunos mas practicas, una de las diferencias de estos analizadores es que uno realiza búsquedas por otros nombres y por herencia pero el otro no, hay varias similitudes como contradicciones pero con los resultados obtenidos se puede concretar que ambos son útiles y funcionan.

REFERENCIAS

- (CIC-IPN), (-U.-C. (30 de Octubre de 2009). *Asociacion Mexicana para el procesamieto del lenguaje Natural*. Obtenido de Asociacion Mexicana para el procesamieto del lenguaje Natural: <http://www.cicling.org/ampln/index-tmp.htm>
- (Mayo de 2014). Recuperado el Mayo de 2014, de <http://www.hardwarecult.com/foro/index.php?topic=2070.0>
- (Abril de 2014). Recuperado el Abril de 2014, de <http://cursocompiladoresuaeh.blogspot.mx/2010/11/unidad-vi-analisis-semantico.html>
- (Mayo de 2014). Recuperado el Mayo de 2014, de <http://www.saber.cic.ipn.mx/cake/SABERsvn/trunk/Repositorios/webVerArchivo/569/1>
- (2014). Recuperado el 2014, de <http://www.fedora-es.com/manuales/redes/tanenbaum.pdf>
- (2014). Recuperado el 2014, de <http://www.gelbukh.com/thesis/Sulema%20Torres%20Ramos%20-%20MSc.pdf>
- ABUD Figueroa. (2005). MECSE: Conjunto de Métricas para evaluar Software educativo. (IPN, Ed.) *UPIICSA*, V(XIII), 39. Obtenido de <http://www.revistaupiicsa.20m.com/Emilia/RevSepDic05/Antonieta.pdf>
- actividad-2-analizador-lxico-sintctico-y-semntico*. (Mayo de 2014). Recuperado el Mayo de 2014, de *actividad-2-analizador-lxico-sintctico-y-semntico*: http://www.slideshare.net/maryr_/actividad-2-analizador-lxico-sintctico-y-semntico
- Aguilar, A. M. (3 de Marzo de 2004). *Teoria De Chomsky*. Recuperado el 3 de Septiembre de 2014, de *Teoria De Chomsky*: http://www.csub.edu/~tfernandez_ulloa/spanishlinguistics/chomsky%20y%20la%20gramatica%20generativa.pdf
- AGUILAR, C. J., Muñoz, A., & Zechinelli, M. (s.f.). Administración de contenidos digitales mediante objetos de aprendizaje. Recuperado el agosto de 2010, de <http://bibliotecadigital.conevyt.org.mx/colecciones/documentos/somece/16.pdf>
- Aguilar, J. (11 de Octubre de 2006). *Analisis sintactico automatico*. Obtenido de *Analisis sintactico automatico*: <http://www.gedlc.ulpgc.es/investigacion/desambigua/morfosintactico.htm>
- Alicante, U. d. (2007-2008). *Interpretacion Semantica*. Obtenido de *Interpretacion Semantica*: http://rua.ua.es/dspace/bitstream/10045/3874/1/ILN_Tema4.pdf

- Allan, V. (26 de Mayo de 2010). *Redes Semanticas*. Obtenido de Redes Semanticas:
<http://es.slideshare.net/victorallan/redes-semanticas-y-metaforas>
- Analisis de Documentos*. (2 de Septiembre de 2014). Recuperado el 2 de Septiembre de 2014, de Analisis de Documentos:
https://www.ulpgc.es/descargadirecta.php?codigo_archivo=7614
- Analisis Semantico*. (Mayo de 2014). Recuperado el Mayo de 2014, de Analisis Semantico:
<http://www.giaa.inf.uc3m.es/docencia/II/PL1/analisisSemantico05-06.pdf>
- Analizadores*. (Abril de 2014). Recuperado el Abril de 2014, de Analizadores:
<http://books.google.com.mx/books?id=YTPHHZY0SkC&pg=PA18&lpg=PA18&dq=que+es+un+analizadores+morfologicos&source=bl&ots=aX80zW6zDx&sig=tecEl7sN3xSGdukPjr2TawwhUHA&hl=es&sa=X&ei=qUd2U-P6Gs2bqAaogIH4Dw&ved=0CFQQ6AEwBw#v=onepage&q=que%20es%20un%20analizadores>
- Analizadores*. (Abril de 2014). Recuperado el Abril de 2014, de Analizadores:
<http://dspace.utpl.edu.ec/bitstream/123456789/1610/3/005X1214.pdf>
- Armijos, A. (7 de Febrero de 2014). *Analizador Semántico de Programas en Java*. Obtenido de Analizador Semántico de Programas en Java: <http://prezi.com/nuqvwxhhjyoy/analisis-semantic-de-programas-escritos-en-java/>
- Arteta, I. (3 de Marzo de 2013). *Modelo de Cascada*. Recuperado el 31 de Agosto de 2014, de Modelo de Cascada: <http://modelo-cascada.blogspot.mx/>
- Avilar, D. (31 de Agosto de 1977). *La lengua*. Obtenido de La lengua:
<http://es.wikipedia.org/wiki/Lenguaje>
- Base de Conocimientos*. (Mayo de 2014). Recuperado el Mayo de 2014, de Base de Conocimientos: http://es.wikipedia.org/wiki/Base_de_conocimiento
- Bavera, F., Arroyo, M., & Aguirre, J. (2013 de Octubre de 2013). *JTLex un generador de analizadores Semanticos*. Obtenido de JTLex un generador de analizadores Semanticos:
http://biblioteca.universia.net/html_bura/ficha/params/title/jtlex-generador-analizadores-lexicos-traductores/id/55880399.html
- Bejar, J. (2007). *Ontologias*. Obtenido de Ontologias:
<http://www.lsi.upc.edu/~bejar/ia/transpas/teoria/3-RC3-Ontologias.pdf>
- Benot, E. (Mayo de 1987). *Enfoque sintactico de la oracion*. Obtenido de Enfoque sintactico de la oracion: <https://www.overdrive.com/media/359766/sobre-las-partes-de-la-oracion-y-el-enfoque-sintactico-funcional-de-la-gramatic>
- Calderon, J. F. (2001). *Enfoque de dependencias*. Obtenido de Enfoque de dependencias:
<http://www.eumed.net/libros-gratis/2008b/409/La%20Teoria%20de%20la%20Dependencia.htm>

- Catedra, J. (12 de Junio de 1987). *Los Niveles del lenguaje*. Obtenido de Los Niveles del lenguaje: <http://roble.pntic.mec.es/msanto1/lengua/niveles.htm>
- Cepa, D. (2013). *Niveles de Lenguaje*. Recuperado el 27 de Abril de 2014, de Niveles de Lenguaje: <http://es.slideshare.net/rafikylopez/niveles-de-uso-del-lenguaje-29557744>
- Cervantes, A. (febrero de 2011). *Procesamiento semantico*. Obtenido de Procesamiento semantico: <http://www.buenastareas.com/ensayos/Procesamiento-Sem%C3%A1ntico/1625950.html>
- Cervantes, M. (26 de Enero de 2008). *Generador de analizadores semánticos en BISON*. Obtenido de Generador de analizadores semánticos en BISON: <http://ccia.ei.uvigo.es/docencia/PL/bison.pdf>
- Cervantes, R. (1 de Julio de 2002). *Kheresmoi-multilingual Semantic Serach of medical text and imagenes*. Obtenido de Kheresmoi-multilingual Semantic Serach of medical text and imagenes: <http://healthcybermap.semanticweb.org/>
- CIEP (Coordinación de Innovación Educativa y Pregrado). (2008). *Universidad de Guadalajara*. Recuperado el diciembre de 2012, de <http://www.crea.udg.mx/crea.jsp>
- Computacionales, L. S. (17 de Noviembre de 2010). *Reglas Semanticas*. Obtenido de Reglas Semanticas: <http://commpi.blogspot.mx/2010/11/reglas-semanticas.html>
- Cross, S., & Cols. (2000). Decision Support Technologies Are Appropriate for the Cytodiagnosis of Cancer? En *Artificial Intelligence Techniques in Breast* (págs. 265-295). World Sientific.
- Daedalus. (20 de Marzo de 2014). *Analizador Morfosintactico*. Obtenido de Analizador Morfosintactico: http://www.mystilus.com/Analizador_morfosintactico
- Dorta, G. (2007). *Aspectos sintacticos*. Obtenido de Aspectos sintacticos: http://www.scielo.org.ve/scielo.php?pid=S0798-97092007000200001&script=sci_arttext
- Eduardo, M. (29 de Agosto de 2009). *Sistemas Basados en Conocimientos*. Obtenido de Sistemas Basados en Conocimientos: <http://ccc.inaoep.mx/~emorales/Cursos/Representa/node14.html>
- Erandi, R., Martinez, B., Cruz, N., Acosta, H. G., Rebate, I., Suarez, M., & Pavon, P. (2013). Arboles de desicion como herramienta en el diagnostico medico. *Revista medica de la universidad veracruzana*, 19-24.
- Eslava, M. (28 de Junio de 2003). *Inteligencia Artificial y Sistemas Experto*. Obtenido de Inteligencia Artificial y Sistemas Experto: <http://es.slideshare.net/melieslava/inteligencia-artificial-y-sistemas-expertos-237929>
- Estructura de la investigacion Cientifica*. (Mayo de 2014). Recuperado el Mayo de 2014, de Estructura de la investigacion Cinetifica:

<https://sites.google.com/site/conocimientocspina/estructura-de-la-investigacion-cientifica/observacion-estructurada-o-formalizada>

Estructura de un programa en Java. (Febrero de 2014). Recuperado el Febrero de 2014, de Estructura de un programa en Java: <http://milibretajava.blogspot.mx/2012/07/estructura-de-un-programa-en-java.html>

Fernandez, A. (Marzo de 2004). *Procesamiento de textos en lenguaje Natural*. Obtenido de Procesamiento de textos en lenguaje Natural: <http://uvadoc.uva.es/handle/10324/2943>

Freeling. (Diciembre de 2013). Recuperado el Diciembre de 2013, de Freeling: http://translate.googleusercontent.com/translate_c?depth=1&hl=es&prev=/search%3Fq%3Dfreeling%26es_sm%3D93%26biw%3D1366%26bih%3D667&rurl=translate.google.com.mx&sl=en&u=http://nlp.lsi.upc.edu/freeling/doc/userman/html/node3.html&usg=ALkJrhYUiBXE84J6gctmeX

Gamallo, P., & Ponche, C. (2012). *Lenguaje Natural*. Obtenido de Lenguaje Natural: <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/issue/view/208>

Garcia, A. (16 de Octubre de 2013). *Análisis Sintáctico*. Obtenido de Análisis Sintáctico: http://angarmegia.com/analisis_gramatical.htm

Garcia, M. A. (23 de Noviembre de 2013). *Editor de análisis*. Obtenido de Editor de análisis: <http://www.analisisintactico.com/>

Garuda, M. (Febrero de 2002). *Lenguaje Formal*. Obtenido de Lenguaje Formal: <http://www.icarito.cl/enciclopedia/articulo/primer-ciclo-basico/lenguaje-y-comunicacion/comunicacion-oral/2010/03/50-8919-9-lenguaje-formal-e-informal.shtml>

Gomez, E. (Marzo de 2010). *Aspectos semánticos*. Obtenido de Aspectos semánticos: <http://www.redalyc.org/articulo.oa?id=34702801>

Gomez, R. (1990). *Inteligencia Artificial ¿Hacia donde nos lleva?* Obtenido de Inteligencia Artificial ¿Hacia donde nos lleva?: <http://www.comoves.unam.mx/numeros/articulo/2/la-inteligencia-artificial-hacia-donde-nos-lleva>

Gonzales, J. (2007). *Lenguaje Formal*. Recuperado el 15 de Abril de 2014, de Lenguaje Formal: http://187.174.253.10/Biblionetica/diccionario/apenjkml/lenguaje_formal.htm

Gustavo, G. (2010). *Análisis Semántico*. Obtenido de Análisis Semántico: <http://cursocompiladoresuaeh.blogspot.mx/2010/11/unidad-vi-analisis-semantico.html>

Guzman, A. (2013). *Uso de Marcos para convertir redes Semánticas*. Recuperado el 2 de Septiembre de 2014, de Uso de Marcos para convertir redes Semánticas: http://www.tamps.cinvestav.mx/seminario_2011_4

Herencia. (Febrero de 2014). Recuperado el Febrero de 2014, de Herencia: <http://todoenjava.blogspot.mx/2012/01/herencia.html>

- Herencia*. (Abril de 2014). Recuperado el Abril de 2014, de Herencia:
http://es.wikipedia.org/wiki/Herencia_m%C3%BAltiple
- Inteligencia Artificial*. (Mayo de 2014). Recuperado el Mayo de 2014, de Inteligencia Artificial:
<https://www.google.com.mx/webhp?sourceid=chrome-instant&ion=1&espv=2&ie=UTF-8#q=inteligencia%20artificial>
- Investigacion 2*. (Abril de 2014). Recuperado el Abril de 2014, de Investigacion 2:
<http://www.psicol.unam.mx/Investigacion2/pdf/METO2F.pdf>
- Java*. (mayo de 2014). Recuperado el mayo de 2014, de Java:
https://www.java.com/es/download/faq/whatis_java.xml
- Java Introduccion*. (Febrero de 2014). Recuperado el Febrero de 2014, de Java Introduccion:
http://www.java.com/es/download/whatis_java.jsp
- Jr, Q. (1996). *Leraning Decision Tree Classifiers*. Computing Surveys.
- Justo, F. (2008). *Proximidad Semantica*. Obtenido de Proximidad Semantica:
<http://www.hispanoteca.eu/Foro-preguntas/ARCHIVO-Foro/hom%C3%B3nimos-hom%C3%B3fonos-par%C3%B3nimos.htm>
- Koopman, B. (2007). *Towards Semantic Serach and Interence in electronic medical records: an approach using concept-base information*. Obtenido de Towards Semantic Serach and interence in electronic medical records: an approach using concept-base informaction:
<http://dx.doi.org/10.4066/AMJ.2012.1362>
- Lamarca, J. (8 de Diciembre de 2013). *Ontologias*. Obtenido de Ontologias:
<http://www.hipertexto.info/documentos/ontologias.htm>
- Lenguaje*. (2013-20014). Recuperado el 27 de Abril de 2014, de Lenguaje:
<http://www.significados.info/lenguaje/>
- Lenguaje Natural*. (2006). Recuperado el 15 de Abril de 2014, de Lenguaje Natural:
<https://www.inbenta.com/es/quienes-somos/lenguaje-natural>
- Listas*. (Abril de 2014). Recuperado el Abril de 2014, de Listas:
<http://sebastianpuente.blogspot.mx/2012/04/listas-en-java.html>
- Louwerse, M. (18 de Agosto de 2004). *Un modelo conciso de la cohesión en el texto y coherencia en la comprensión*. Obtenido de Un modelo conciso de la cohesión en el texto y coherencia en la comprensión: http://www.scielo.cl/scielo.php?pid=S0718-09342004005600004&script=sci_arttext
- Martinez, A. (2013). *Proyecto Diana*. Recuperado el 3 de Septiembre de 2014, de Proyecto Diana: <http://clic.ub.edu/es/proyectos>
- Martinez, R. (2014). *Analizador morfológico*. Obtenido de Analizador morfológico:
<http://www.aulafacil.com/cursos/16857/primaria/lenguaje-primaria/lengua-sexto-primaria-11-anos/analisis-morfologico>

- MC GREAL, R. (s.f.). The need for Open Educational Resources for ubiquitous learning. Recuperado el septiembre de 2012, de <http://auspace.athabasca.ca/bitstream/2149/3169/1/PEREL%20JanFINAL.pdf>
- Metodo Experimental*. (abril de 2014). Recuperado el abril de 2014, de Metodo Experimentla: <http://www.monografias.com/trabajos11/methods/methods.shtml#EXPERIM>
- Metodo-Experimental*. (2014). Recuperado el 2014, de Metodo-Experimental: <http://www.alegsa.com.ar/Dic/debugger.php>
- Metodologia de Cascada*. (s.f.). Recuperado el 31 de Agosto de 2014, de Metodologia de Cascada: http://www.ecured.cu/index.php/Metodolog%C3%ADa_de_cascada
- Metodos*. (2014). Recuperado el 2014, de Metodos: <http://puntocomnoesunlenguaje.blogspot.mx/2012/04/metodos.html>
- Metodos y tecnicas de Investigacion*. (Abril de 2014). Recuperado el Abril de 2014, de Metodos y tecnicas de Investigacion: <http://www.gestiopolis.com/economia/metodos-y-tecnicas-de-investigacion.htm>
- MORALES Gamboa, R. (junio de 2011). Modelo de objetos de aprendizaje para la producción y gestión de contenidos educativos. *Ingeniare*, 19(1). doi:10.4067/S0718-33052011000100001
- Morales, L. (2014). *Wolfram alpha*. Obtenido de Wolfram alpha: <http://www.wolframalpha.com/examples/>
- O., J., G., D., Reinhard, S., Crouch, R., Paiya, V., & Vanden, M. (2003). "Haciendo Ontologias". 26.
- Observacion*. (2014). Recuperado el 2014, de Observacion: <http://es.wikipedia.org/wiki/Observaci%C3%B3n>
- Observacion*. (Abril de 2014). Recuperado el Abril de 2014, de <http://www.ecured.cu/index.php/Observaci%C3%B3n>
- Ontologias*. (abril de 2014). Recuperado el abril de 2014, de Ontologias: <http://es.wikipedia.org/wiki/Ontolog%C3%ADa>
- Ortiz, D., & Pernia, A. L. (2006). *Procesamiento del Lenguaje Natural*. Recuperado el 3 de Septiembre de 2014, de Procesamineto del Lenguaje Natural: http://www.elprofesionaldelainformacion.com/contenidos/1997/enero/procesamiento_del_lenguaje_natural_revisin_del_estado_actual_bases_tericas_y_aplicaciones_parte_i.html
- Pajares, G., & Santos, M. (2006). En *Inteligencia Artificial e Ingenieria del Conocimiento* (págs. 4,5). MADRID.
- Palma, J., & Morales, R. (2008). Inteligencia Artificial. En J. Palma, & R. Morales, *Inteligencia Artificial Metodos, tecnicas y aplicaciones* (pág. 7). India.

- Penas, M. A. (2009). *La linguística del texto*. Obtenido de La linguística del texto:
http://www.academia.edu/2292610/T.Albaladejo._La_linguistica_del_texto_y_el_analisis_interdiscursivo_en_la_literatura_comparada
- Peralta, M. (Octubre de 2014). *Sistema de información*. Obtenido de Sistema de información:
<http://www.monografias.com/trabajos7/sisinf/sisinf.shtml>
- Perez, A. (Julio de 2006). *Linguística Informativa*. Obtenido de Linguística Informativa:
http://www.saber.ula.ve/bitstream/123456789/13157/1/hc_propuestatesis.pdf
- Perez, A. (11 de Mayo de 2013). *Analizador Sintáctico*. Obtenido de Analizador Sintáctico:
<http://es.slideshare.net/alexisdario/analizador-sintactico-20951453>
- Pérez, C. (2002). *Definición de ontología como especificación de conocimiento*. Recuperado el 27 de Abril de 2014, de Definición de ontología como especificación de conocimiento:
<http://elies.rediris.es/elies18/531.html>
- R. J. (20 de Junio de 2011). *Etiquetas Eagles*. Obtenido de Etiquetas Eagles:
<http://www.geekfactory.mx/tutoriales/tips-para-eagle-esquemáticos-más-limpios-con-labels/>
- Rastier, F. (2001). *Semántica Interpretativa*. Obtenido de Semántica Interpretativa:
<http://books.google.com.mx/books?id=IaDg41b4gQQC&pg=PA324&lpg=PA324&dq=Interpretaci%C3%B3n+guiada+por+la+sem%C3%A1ntica.&source=bl&ots=12Bgdieffi&sig=Mv2IIZqOM751MxgMGE5VbWAeoEw&hl=es-419&sa=X&ei=3pE3VPafDcWOyASH0YCwCQ&ved=0CCwQ6AEwAg#v=onepage&q=Interpre>
- Redes Semánticas*. (2002). Recuperado el 27 de Abril de 2014, de Redes Semánticas:
http://ce.azc.uam.mx/profesores/clc/03_docencia/posgrado/i_artificial/RedesSemánticas.pdf
- Redes Semánticas*. (2014). Recuperado el 2014, de Redes Semánticas:
<http://toalablog.wordpress.com/2012/10/17/redes-semánticas-2/>
- Redes SEMánticas*. (Abril de 2014). Recuperado el Abril de 2014, de Redes SEMánticas:
<http://di002.edv.uniovi.es/~ortin/publications/semántico.pdf>
- Ríos, F. (23 de Septiembre de 2009). *Análisis semántico-traductor descendente-usando polimorfismo con C#*. Obtenido de Análisis semántico-traductor descendente-usando polimorfismo con C#: <http://es.scribd.com/doc/103521215/Analisis-Semántico-Traductor-Descendente>
- Rodríguez, C. (28 de Mayo de 2012). *Redes Semánticas*. Obtenido de Redes Semánticas:
<http://es.slideshare.net/gglloo-gloglogloglo/redes-semánticas-13110503>
- Rodríguez, J. M. (2004). *ESTRUCTURA SEMÁNTICA Y ANÁLISIS IDEOLÓGICO*. Obtenido de ESTRUCTURA SEMÁNTICA Y ANÁLISIS IDEOLÓGICO:
<http://www.vinv.ucr.ac.cr/latindex/filologia-30-2/10-Rodríguez.pdf>

- Ruiz, J. (2010). Compiladores. En J. Ruiz, *Compiladores teoria e implementacion* (pág. 7). alfaomega.
- Salvador, O. (12 de Abril de 2005). *Analizador Semántico en FKScript*. Obtenido de Analizador Semántico en FKScript: http://www.sgoliver.net/blog/?page_id=115
- SEP, Secretaría de Educación Pública. (2009). Telecurso Introducción al uso de los multimedia en la educación / Lección 2. México.
- Sistemas de Informacion*. (Febrero de 2014). Recuperado el Febrero de 2014, de Sistemas de Informacion: http://es.wikipedia.org/wiki/Sistema_de_informaci%C3%B3n
- Sosa, E. (Enero de 1997). *Procesamiento del lenguaje Natural*. Obtenido de Procesamiento del lenguaje Natural: http://www.elprofesionaldelainformacion.com/contenidos/1997/enero/procesamiento_del_lenguaje_natural_revisin_del_estado_actual_bases_tericas_y_aplicaciones_parte_i.html
- Sosa, E. (Enero de 1997). *Procesamiento del Lenguaje Natural*. Obtenido de Procesamiento del Lenguaje Natural: http://www.elprofesionaldelainformacion.com/contenidos/1997/enero/procesamiento_del_lenguaje_natural_revisin_del_estado_actual_bases_tericas_y_aplicaciones_parte_i.html
- Soto, V. (3 de Febrero de 2011). *ANTL*. Obtenido de ANTL: <http://veritosoto.wordpress.com/category/uncategorized/page/2/>
- Tecnicas de Investigacion*. (Mayo de 2014). Recuperado el Mayo de 2014, de Tecnicas de Investigacion: <http://www.cvalle.edu.mx/PhpSystem/Tareas/adjuntos/tecnicasinv.pdf>
- Tipos de Analizadores*. (Abril de 2014). Recuperado el Abril de 2014, de Tipos de Analizadores: <http://www.giaa.inf.uc3m.es/docencia/II/PL1/analisisSemantico05-06.pdf>
- Tristan, A. (28 de Septiembre de 2007). *Analisis Automatizada de Textos*. Recuperado el 2 de Septiembre de 2014, de Analisis Automatizada de Textos: <http://www.ieia.com.mx/materialesreuniones/2aReunionRegional/0.ConferenciasMagnas/DrAgustinTristanLopez.pdf>
- Vega, C., & Richard. (28 de Julio de 2011). *Mujer Olimpica 2012*. Recuperado el 3 de Septiembre de 2014, de Mujer Olimpica 2012: <http://e-archivo.uc3m.es/handle/10016/13189>
- Venegas, R. (2003). *Analisis Semantico*. Obtenido de Analisis Semantico: http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-09342003005300008
- Ventajas y desventajas del metodo experimental*. (Abril de 2014). Recuperado el Abril de 2014, de Ventajas y desventajas del metodo experimental: <http://simulacion2011.blogspot.mx/2011/02/ventajas-y-desventajas-de-la.html>

