



# UNIVERSIDAD DE GRANADA

## Procesamiento del Lenguaje Natural basado en Lógica Difusa y su aplicación al Reconocimiento de la Voz

---

Inteligencia Computacional



Máster universitario en Ingeniería Informática

### Autor

Pablo Valenzuela Álvarez ([pvalenzuela@correo.ugr.es](mailto:pvalenzuela@correo.ugr.es))

# Índice

<b>1. Introducción</b>	<b>3</b>
<b>2. Lógica difusa en sistemas de NLP</b>	<b>4</b>
2.1. Problemas presentados	4
2.2. Sistema de lógica difusa	4
<b>3. Formas de adquisición de conocimiento</b>	<b>6</b>
3.1. Adquisición automática de tipos semánticos	6
3.2. Adquisición automática de tipos sintácticos y tipos de expresiones	6
<b>4. Resultados experimentales</b>	<b>9</b>
<b>5. Conclusiones del estudio</b>	<b>11</b>
<b>6. Trabajos relacionados</b>	<b>12</b>
6.1. Fuzzy approach for sentiment analysis (2017) [4]	12
6.2. Fuzzy logic applied to opinion mining (2021) [5]	12
<b>7. Referencias</b>	<b>13</b>

# 1. Introducción

La complejidad del lenguaje natural ha hecho que la gente busque diferentes técnicas de *soft computing* para su análisis. Entre los enfoques estadísticos, de conexiones u otras aproximaciones, la lógica difusa ha emergido como una alternativa efectiva para el reconocimiento del lenguaje natural.

Muchos de los fenómenos ocurridos en el lenguaje natural pueden describirse mediante matemáticas y la lógica difusa, estableciendo reglas que ayuden a reducir las dificultades en el análisis del habla.

La lógica difusa ha demostrado tener éxito al describir los significados de las palabras en relación a estos fenómenos externos mediante la utilización de descriptores difusos en los sistemas de control. Los términos “*caliente*” o “*frío*” pueden actuar como etiquetas para conjuntos difusos, a los cuales se pueden asignar grados de pertenencia según las lecturas de las temperaturas.

Es destacada también, su aplicación en la búsqueda en bases de datos impulsada por el lenguaje natural, donde la semántica de las palabras es expresada mediante funciones de membresía para las claves de búsqueda.

En esta presentación vamos a abordar el tratamiento interno del lenguaje en un sistema difuso. También revisaremos los principales problemas relacionados con el reconocimiento de la voz y se propondrán soluciones, mostrando los resultados obtenidos de los estudios experimentales realizados.

## 2. Lógica difusa en sistemas de NLP

### 2.1. Problemas presentados

Aplicando el sistema de reconocimiento a dominios restringidos, y limitando el tamaño del vocabulario y las construcciones semánticas, podemos encontrarnos con ciertos problemas:

- **Palabras fuera del vocabulario:** Un usuario puede pronunciar palabras que no están incluidas en el léxico del sistema. Por ejemplo, en un dominio de viajes aéreos, que contiene sólo ciudades estadounidenses, un usuario podría preguntar por ciudades de otro país.
- **Error del reconocedor de voz:** Pueden darse casos de detección de palabras erróneas por parte del reconocedor de voz. *Por ejemplo*, un caso de test: “*I need to get from Philadelphia Pennsylvania to Boston Massachusetts*”, el reconocedor identificó: “*I need to get from Philadelphia and stopping at to Boston Massachusetts*”.
- **Estructuras flexibles:** El usuario puede usar expresiones que la gramática del sistema no entienda. Mientras que el sistema pueda reconocer una estructura parecida a esta: “*a flight from Atlanta to Boston*”, un usuario puede decir “*a flight originating in Atlanta, destination Boston*”. Esta expresión puede resultar nueva para el sistema y no reconocerla bien.
- **Disfluencia o disritmia:** Un inicio falso de conversación, reformulaciones, palabras repetidas o mal pronunciadas, pausas, etc. Todos estos ejemplos hacen que el sistema confunda las relaciones semánticas.

### 2.2. Sistema de lógica difusa

El sistema de lógica difusa diseñado tiene dos principales tareas:

1. Evaluar si una palabra reconocida es semánticamente apropiada respecto a toda la frase reconocida.
2. Aplicar un conjunto de reglas de inferencia difusa para la predicción de posibles palabras que pueden faltar dada la palabra reconocida correctamente.

La semántica difusa usada es una tri-tupla  $\mathcal{L}(V,F,R)$  donde:

- $V$  es un conjunto de vocabulario lingüístico.
- $F$  es un conjunto de características del vocabulario.
- $R$  es un conjunto de reglas de inferencia.

Un miembro de  $V$  es llamado palabra, y tiene asociada el siguiente vector:

$$palabra(f_1=v_1, \dots, f_n=v_n)$$

Es decir, una palabra está representada por un vector de características de  $F$  y de valores simbólicos.

Un miembro de  $F$  es un símbolo que representa una característica. En este sistema, se usa el siguiente conjunto de características: {SYNTACTIC-TYPE, SEMANTIC-TYPE, PHRASE-TYPE CONTEXT-WORDS}.

El conjunto  $R$  consiste en dos tipos de reglas de inferencia:

- La regla de evaluación identifica el contexto de una palabra y evalúa si está bien reconocida.
- La regla de predicción predice posibles palabras faltantes basadas en la palabra correctamente evaluada.

Ambas reglas introducen información en el vector de características de una palabra de la siguiente forma: Para cada palabra  $X$ , si su vector de características contiene el contexto  $Y$ , se denota como la co-ocurrencia  $X * Y$ , y significa que cuando la palabra  $X$  es usada, la palabra  $Y$  puede aparecer.

Las reglas de evaluación son representadas así:

$$\begin{aligned} \text{E-rule} : X * Y \wedge X \wedge Y &\rightarrow \text{inc}^E(X) \\ \text{E-rule} : X * Y \wedge X \wedge \text{not}(Y) &\rightarrow \text{dec}^E(X) \end{aligned}$$

$Y$  significan que, cuando  $X$  es reconocida, y  $Y$  co-ocurre con  $X$ , se incrementa o decrementa la evaluación de  $X$  dependiendo de si  $Y$  es encontrada en el reconocimiento.

A su vez, las reglas de predicción son representadas así:

$$\begin{aligned} \text{P-rule: eval}(X) > \gamma \wedge X * Y &\rightarrow \text{inc}^P(Y) \\ \text{P-rule: eval}(X) < \gamma \wedge X * Y &\rightarrow \text{dec}^P(Y) \end{aligned}$$

$Y$  significan que, si la evaluación de  $X$  supera un umbral  $\gamma$ , se incrementa el peso en la predicción de  $Y$ , si  $Y$  está en el vector de co-ocurrencia  $X * Y$ . Si no supera el umbral, este valor disminuye.

### 3. Formas de adquisición de conocimiento

Un sistema con una lógica robusta requiere grandes cantidades de conocimiento, ya sea adquirido a través de conocimiento experto o de sistemas automáticos de adquisición de conocimiento.

#### 3.1. Adquisición automática de tipos semánticos

Dado un léxico general como por ejemplo WordNet [2], es posible clasificar las palabras recibidas de otro corpus en grupos semánticos del dominio de ese corpus. En este estudio se utilizó un método sencillo basado en un corpus dado y la utilización WordNet para el descubrimiento de tipos semánticos interesantes. La idea básica consiste en los siguientes puntos:

- WordNet posee una gran estructura jerárquica de palabras, divididas en varios niveles de abstracción según agrupación semántica.
- En un dominio restringido (por ejemplo, viajes), es preferible una categorización semántica “plana”, así podríamos tener uno o dos significados para cada palabra.
- Se usa un tercer nivel de abstracción para la agrupación semántica. Este nivel corresponde a categorías básicas como: “periodo de tiempo”, “ciudad”, “comida”, “movimiento”, etc.
- Se asume que para una categoría semántica interesante en un dominio, esta debería estar presente frecuentemente en el corpus. De esta manera se pueden encontrar casi todos los tipos semánticos de los tokens pertenecientes al corpus y luego seleccionar los que se repiten con alta frecuencia.
- El método del punto anterior hecho de forma automática, genera un léxico lleno de palabras con los tipos semánticos interesantes. Se experimentó con un conjunto ATIS corpus que contiene 152 mil tokens de palabras, 1571 tipos de palabras distintas.

#### 3.2. Adquisición automática de tipos sintácticos y tipos de expresiones

Aparte del tipo semántico, una palabra está también definida por su tipificación sintáctica y el tipo de expresión.

Para la tipificación sintáctica se usa un “etiquetador” basado en árboles entrenado en Penn Treebank (Dataset que contiene artículos del Wall Street Journal [3]). Después de un etiquetado automático, se aplican unas reglas a la secuencia de etiquetas extraídas que son parseadas en tres tipos de expresiones: *verb phrase*, *noun phrase* y *prepositional phrase*. El corpus ATIS usado contiene 13 mil frases que fueron parseadas y los resultados obtenidos mostraron un 97% de frases clasificadas correctamente.

Cada una de estas frases fue fragmentada en tokens, y cada token fue tipificada por el tipo semántico de su palabra “núcleo”. Esta palabra “núcleo” se decide siguiendo el siguiente criterio:

- Para cada *verb phrase*, el núcleo era el verbo.
- Para cada *prepositional phrase* y *noun phrase*, el último sustantivo era considerado como el núcleo.
- Para cada Be+adjective, el último adjetivo se considera el núcleo.

Siguiendo estas tareas, se obtuvo un total de:

- 66 tipos semánticos
- 12 tipos sintácticos: *noun*, *verb*, *be-verb*, *adjective*, *adverb*, *conjunction*, *interjection*, *auxiliary*, *numeral*, *pronoun*, *preposition* y *proper-noun*.
- 3 tipos de expresiones: *noun phrase*, *verb phrase* y *prepositional phrase*.

Los resultados obtenidos sobre los tipos de expresiones más frecuentes se muestran en la siguiente tabla:

Tipo de expresión	Frecuencia
City	12108
Flight	5774
Time	3574
Person	3371
Number	3344
Communicate	1948
Desire	1947
Leave	1665
Commercial-document	1421
Airline	1230
Arrival	839
Taxomic-group	805
Motion	788

Tabla 1. Resultados de la adquisición de tipos de expresiones.

Finalmente se construyeron las definiciones de las palabras basándose en los tipos sintácticos, semánticos y de expresión obtenidos con anterioridad. Y para cada palabra se construyó un vector de palabras-contexto (context-word) en el que se tienen en cuenta:

1. Las palabras en el lado izquierdo de la expresión.
2. Las palabras en el lado derecho de la expresión.
3. Las palabras en la expresión vecina de la izquierda.
4. Las palabras en la expresión vecina de la derecha

Una expresión vecina es diferenciada por el tipo de expresión sintáctica.

Entonces el vector se estructuró de la siguiente manera:

Word Vector:

- Semantic-Type
- Syntactic-Type
- Phrase-Type.
- Left-words-in-phrase
- Right-words-in-phrase
- Words-in-left-neighbor-phrase
- Words-in-right-neighbor-phrase

Y la fórmula de incremento en la evaluación es:

$$\Delta \epsilon = \frac{W_p (X_p - \frac{N_p}{\beta})}{\sum_{p=1}^n N_p}$$

Siendo:

- $X_p$ : número de palabras contextuales (context words) reconocidas.
- $N_p$ : número total de palabras contextuales (context words).
- $\beta$ : es un parámetro de control entre 1 y  $N_p$ .
- $W_p$ : peso del tipo de expresión p en la palabra evaluada.



## 4. Resultados experimentales

Como se ha descrito en las secciones anteriores, se ha implementado un sistema no supervisado para predecir las palabras reconocidas erróneamente por el reconocedor de voz de forma satisfactoria.

Los datos básicos usados consisten en el léxico semántico de WordNet y del corpus ATIS.

Las tareas sobre las que se experimentaron fueron la evaluación del reconocimiento de palabras y posibles palabras faltantes, para los que se usaron entradas de voz con errores (un total de 1967 frases). La precisión en el reconocimiento de palabras fué de un 81,9%, mientras que la precisión en el reconocimiento de las frases fué de un 35,13%.

En cuanto a las reglas de inferencia, dependiendo de un umbral, el sistema puede predecir más o menos las palabras faltantes. Se calculó como de bien las palabras predichas por el sistema pueden cubrir las palabras en las frases de referencia, contrastando con un ratio basado en la cantidad de palabras predichas y el tamaño total del modelo original. El indicador de rendimiento del sistema era el de conseguir un cubrimiento máximo reduciendo al mínimo este ratio.

La siguiente tabla muestra el rendimiento logrado con un umbral bajo ( $\gamma < 0.2$ ).

Fuzzy Logic-Based Word Prediction Results 1				
Uncovered words	Coverage percent	Average model size	Model reduction	Stop words
196 w.	99.11%	840 w.	34.5%	0 w.
286 w.	98.72%	615 w.	52.1%	43 w.
300 w.	98.65%	599 w.	53.3%	291 w.

Tabla 2. Resultados de la predicción de palabras testeados con un umbral de evaluación bajo. Test corpus = 22.26k words. Recognition accuracy = 81.90%. Original language model size = 1282 words.

Ante estos resultados, se vió que no afectaba mucho agrandar la lista de palabras vacías al porcentaje de cubrimiento, pero si lograba reducir el tamaño modelo. Sabiendo esto, se probó a aumentar el umbral.

Fuzzy Logic-Based Word Prediction Results 2				
Uncovered words	Coverage percent	Average model size	Model reduction	Stop words
321 w.	98.6%	580 w.	54.7%	291 w.
343 w.	98.45%	569 w.	55.5%	291 w.
376 w.	98.31%	555 w.	56.6%	291 w.
433 w.	98.05.%	538 w.	58%	291 w.
681 w.	96.94%	482 w.	62.3%	291 w.
881 w.	96.04%	401 w.	68.6%	291 w.

Tabla 3. Resultados de la predicción de palabras testeados con un umbral de evaluación alto.

Los resultados de la tabla 3 muestran que el sistema diseñado puede predecir hasta 96% de las palabras verdaderamente dichas por el usuario en un modelo de lenguaje dinámico de  $\frac{1}{3}$  del tamaño del modelo original.

## 5. Conclusiones del estudio

En este estudio se han delineado los principios y construcciones a seguir en un sistema dedicado al procesamiento del lenguaje natural (NLP) basado en lógica difusa. También se ha demostrado que para lograr un procesamiento robusto del lenguaje natural, es necesario seguir una serie de estrategias “soft computing”.

Para futuras investigaciones se sugieren explorar los siguientes caminos:

1. Métodos de adquisición de conocimiento más efectivos.
2. Incorporar información en el proceso de reconocimiento de las frases, ya que en muchas ocasiones el uso de la oración en un contexto del discurso viene determinado por las interpretaciones de su semántica.
3. Incorporar información de niveles más bajos, por ejemplo, la pronunciación de las palabras.

## 6. Trabajos relacionados

### 6.1. Fuzzy approach for sentiment analysis (2017) [4]

Los análisis de sentimientos tienen como objetivo identificar la polaridad de un documento sirviéndose del procesamiento del lenguaje natural, análisis de textos y la lingüística computacional.

Este artículo propone un sistema basado en reglas difusas para el análisis de sentimientos, que pueda ofrecer salidas más precisas mediante el uso de grados de pertenencia difusa. Según el estudio, comparándose con algoritmos usados comúnmente (árboles de decisión, naive bayes), los resultados obtenidos basándose en este enfoque muestran un rendimiento ligeramente mayor.

### 6.2. Fuzzy logic applied to opinion mining (2021) [5]

Desde la aparición de la web 2.0, se ha producido un gran aumento de la información generada por los usuarios. Dentro de esta información es fácil encontrar textos subjetivos/opiniones.

Este estudio describe diversos usos de la lógica difusa en la minería de opiniones en torno a tres tareas principales: procesamiento de características, clasificación de reseñas y emociones, prestando especial atención a las aplicaciones de análisis de sentimientos que usan la lógica difusa para lograr los objetivos deseados.

## 7. Referencias

- [1] Karray, Fakhri, et al. "Fuzzy Logic-Based Natural Language Processing and Its Application to Speech Recognition." *ResearchGate*,  
[https://www.researchgate.net/publication/239824041\\_Fuzzy\\_Logic-Based\\_Natural\\_Language\\_Processing\\_and\\_Its\\_Application\\_to\\_Speech\\_Recognition](https://www.researchgate.net/publication/239824041_Fuzzy_Logic-Based_Natural_Language_Processing_and_Its_Application_to_Speech_Recognition).
- [2]. *WordNet*, <https://wordnet.princeton.edu/>.
- [3]. "Penn Treebank Dataset." *Papers With Code*,  
<https://paperswithcode.com/dataset/penn-treebank>.
- [4] Jefferson, Chriss, et al. "Fuzzy approach for sentiment analysis."  
<https://ieeexplore.ieee.org/abstract/document/8015577>.
- [5] Serrano-Guerrero, Jesus, et al. "Fuzzy logic applied to opinion mining: A review." *The Free Dictionary*,  
<https://www.sciencedirect.com/science/article/abs/pii/S0950705121002811>.