

TÉCNICAS PARA EL DESARROLLO DE INTERFACES EN LENGUAJE NATURAL

J. Guadalupe Ramos Díaz¹, Adrián Núñez Vieyra², Rogelio Ferreira Escutia³,
Anastacio Antolino Hernández⁴, Heberto Ferreira Medina⁵

Eje1. La investigación en las Ciencias Básicas
Mesa 2. Ciencias de la Ingeniería y Tecnología

Palabras Clave: (Asistente virtual, chatbot, similitud de coseno, similitud semántica)

La capacidad creativa del ser humano ha permitido la invención de infinidad de máquinas que realizan tareas que facilitan la vida y la manipulación de nuestro entorno. La comunicación con las máquinas había sido siempre por medios manuales, mecánicos, eléctricos o electrónicos. Recientemente, se han dado avances muy significativos en el campo del procesamiento del lenguaje natural, una disciplina de la inteligencia artificial, que permite dar un paso contundente en la intención de lograr una comunicación verbal con las máquinas.

En este trabajo presentamos un repaso a la evolución de las interfaces en lenguaje natural, en particular contrastamos las herramientas para hacer búsquedas en texto completo (FULL TEXT) que soportan los manejadores de bases de datos, con respecto a los asistentes virtuales que se pueden construir con *frameworks* adhoc y, finalmente, con el tipo de asistentes susceptibles de desarrollo empleando las librerías de las inteligencias artificiales generales como ChatGPT de OpenAI. La comparación permite reconocer qué enfoque se debe de aplicar al momento de tener la necesidad de desarrollar un asistente virtual.

Introducción

Indudablemente uno de los campos del conocimiento que ha sufrido mayor evolución en los últimos años es el relacionado con la inteligencia artificial. El avance tanto en el hardware que permite mayor capacidad de procesamiento, como en el software que ofrece nuevas técnicas y recursos a disposición de los desarrolladores de tecnología hacen posible la construcción de sistemas inteligentes cada vez más sofisticados.

¹ Tecnológico Nacional de México, Campus Instituto Tecnológico de Morelia; jose.rd@morelia.tecnm.mx

² Tecnológico Nacional de México, Campus Instituto Tecnológico de Morelia; adrian.nv@morelia.tecnm.mx

³ Tecnológico Nacional de México, Campus Instituto Tecnológico de Morelia; rogelio.fe@morelia.tecnm.mx

⁴ Tecnológico Nacional de México, Campus Instituto Tecnológico de Morelia; anastacio.ah@morelia.tecnm.mx

⁵ Tecnológico Nacional de México, Campus Instituto Tecnológico de Morelia; heberto.fm@morelia.tecnm.mx

Una de las disciplinas de la inteligencia artificial, i.e., el procesamiento de lenguaje natural ha sorprendido al mundo entero con la irrupción de los *modelos avanzados de lenguaje* empleados por reconocidos chatbots, e.g., ChatGPT (OpenAI, 2023) y Google Bard (2023) que pueden responder en lenguaje natural a preguntas formuladas por humanos, también en lenguaje natural, en muy diversas temáticas.

La tecnología para construir sistemas capaces de realizar diálogo (Xu, et al., 2000) es añeja y se podía considerar madura a partir de los métodos para su creación (Bhagwat, 2018). En relación con ella se desplegaron esfuerzos gubernamentales (NSF, 2016) para enfocar los riesgos y retos de la aparición de módulos de software parlantes con capacidad de acción, que fueron conocidos como asistentes cognitivos inteligentes (NSF, 2016; Maier, et al., 2019).

No obstante, la aparición de los modelos avanzados de lenguaje ha hecho que los desarrolladores se replanteen la forma más conveniente para desarrollar chatbots que se puedan emplear como interfaces en lenguaje natural de sistemas de software y que recojan las virtudes avanzadas de los populares chatbots como Chatgpt (OpenAI, 2023).

En el presente trabajo revisamos algunas técnicas de construcción de interfaces en lenguaje natural, echamos un vistazo a la técnica formal básica para asociar texto, la similitud de coseno, enseguida mencionamos la búsqueda Full Text de los gestores de bases de datos, después nos centramos en las herramientas específicas para creación de sistemas de diálogo, e.g., Dialogflow (Dialogflow, 2023). Finalmente mostramos las ventajas que representa usar una librería como openai (OpenAI, 2023) en el desarrollo de interfaces en lenguaje natural y presentamos nuestras conclusiones.

Cálculos básicos de semejanza entre texto

Un ejercicio fundamental para iniciar la aproximación a la realización de interfaces en lenguaje natural puede realizarse a partir del cálculo de la semejanza de texto entre bloques de texto. Un usuario puede emitir una consulta q escrita en lenguaje natural. Se supone tener almacenado un conjunto de n bloques b de texto. El procedimiento para responder a la consulta planteada consiste en buscar el bloque de texto b que más se parezca a la consulta q , el cual será enviado como respuesta al usuario. La fórmula para calcular el “parecido” entre trozos de texto requiere convertir cada trozo a una representación vectorial y entonces aplicar la fórmula de similitud de coseno:

Similitud de coseno es una medida estándar entre bloques de texto escritos en lenguaje natural. Si $t1$ y $t2$ son dos bloques de texto, entonces su similitud viene dada por el producto punto de los vectores $V(t1)$ y $V(t2)$, dividido entre el producto de las longitudes euclidianas de $V(t1)$ y $V(t2)$. Formalmente:

$$sim(t1, t2) = \frac{\vec{V}(t1) \cdot \vec{V}(t2)}{|\vec{V}(t1)| |\vec{V}(t2)|}$$

Figura 1 Fórmula para cálculo de similitud

El valor devuelto por la similitud de coseno es un valor entre 0 y 1. Cuando el valor calculado sea más cercano a la unidad estaremos hablando de mayor similitud. Esta técnica elemental requiere que haya términos comunes en el texto para que el cálculo devuelva valores distintos de 0.

Adicionalmente a la similitud de coseno hay otras técnicas para el cálculo de semejanza, e.g., coeficiente de empalme simple, coeficiente Jaccard, distancia Manhattan y distancia euclidiana (Alvarez 2009), entre otras. Esta técnica formal permitió el desarrollo de los primeros chatbots.

Búsquedas Full Text en gestores de base de datos

Las búsquedas a texto completo son una herramienta con la que cuentan los gestores de bases de datos, por ejemplo, el bien conocido: MySQL. Una búsqueda a texto completo regresa cualquier bloque que contiene al menos un emparejamiento. Un emparejamiento ocurre cuando un bloque analizado contiene total o parcialmente los términos o palabras clave especificados en el texto de la consulta. Para realizar consultas a texto completo se requiere haber generado un índice FULLTEXT sobre el atributo de la tabla que se desea consultar, en la Fig. 2, el atributo "descripcion" cuenta con dicho índice.

Visible	Key	Type	Unique	Columns
<input checked="" type="checkbox"/>	PRIMARY	BTREE	YES	idLibros
<input checked="" type="checkbox"/>	idLibros_UNIQUE	BTREE	YES	idLibros
<input checked="" type="checkbox"/>	natural	FULLTEXT	NO	descripcion

Figura 2 Atributo "descripcion" con índice FULLTEXT

Una consulta que no encuentra la ocurrencia de los términos buscados en el atributo analizado de la tabla devuelve cero, o un valor aproximado a cero en caso de que haya pocas ocurrencias. En el ejemplo de la Fig. 3, se intenta hacer emparejamiento entre el contenido del campo "descripcion" y el texto "bien mal sombrero". Como resultado hay dos registros que hacen emparejamiento. El texto que arroja mayor similitud es el siguiente: "Este es un libro que trata de metáforas que se construyen alrededor del bien y del mal. Fausto es un personaje que representa al mal". Como puede observarse hay

dos términos que aparecen en el registro encontrado. Esta técnica se ha podido emplear en el diseño de interfaces, siempre y cuando haya términos comunes entre el texto de la consulta y el texto almacenado.

```

1 • SELECT *, MATCH (descripcion) AGAINST ('bien mal sombrero')
2     as similitud
3 FROM libros
4 WHERE MATCH (descripcion) AGAINST ('bien mal sombrero')
5 order by similitud asc;

```

editorial	titulo	descripcion	similitud
béry ilce	El Principito	El aviador cuenta que una vez, cuando era un n...	0.0000000018859283024
Porrua	Fausto	Este es un libro que trata de metáforas que se ...	0.2718571722507477

Figura 3 Consulta en lenguaje natural

Marcos de trabajo para el desarrollo de interfaces en lenguaje natural

Existe hoy en día una extensa oferta para el desarrollo de chatbots e incluso una versión más sofisticada, i.e., para la generación de asistentes cognitivos inteligentes (NSF, 2016). Algunas herramientas son, por ejemplo: Chatfuel, Microsoft Bot Framework, Wit.ai, Dialogflow, IBM Watson, Pandorabots, Botpress, Botkit, RASA Stack y ChatterBot. Estas herramientas permiten modelar el diálogo de diferente manera, por ejemplo, Dialogflow (2023) se basa en dos elementos principales:

Intents: Es el nombre técnico que Dialogflow da a las intenciones de comunicación, es decir, a las órdenes que un usuario dirige hacia un bot. Un intent es un modelo de orden en lenguaje natural que hay que construir. El intent puede embeber palabras clave (*keywords*) y sus variaciones léxicas, sinónimos y cantidades; en esta parte Dialogflow explota las características de *machine learning* de Google para reconocer automáticamente variaciones de acepciones de palabras.

Context: Los contextos son mecanismos que funcionan como estados previos obligatorios para que se pueda reconocer un intent (llamados context de entrada) y por otro lado, cuando un intent es reconocido puede llevar a cabo la definición de un context de salida, al terminar su tarea, que podría ser condición para que se activen otros intents, de esta manera un context puede verse como una estafeta que entrega un intent hacia otro intent.

Con estos dos elementos es posible desarrollar sistemas de diálogo con bifurcaciones complejas. Si bien, las intenciones de comunicación permiten variar la forma en que el usuario se dirige al bot, también es cierto que aún se percibe como un autómatá limitado.

Los marcos de trabajo para el desarrollo de asistentes virtuales se podrían considerar como tecnología madura y definitiva, sin embargo, la aparición de tecnologías emergentes en el campo de la inteligencia artificial planteó nuevos enfoques.

Inteligencias artificiales generales

La reciente irrupción de las inteligencias artificiales generales (AGI por sus siglas en inglés) tales como ChatGPT (OpenAI, 2023) o Bard (2023) han supuesto un nuevo hito por su renovada capacidad para procesar diálogo en lenguaje natural. Evidentemente para la industria y usuarios finales resulta muy atractivo incorporar las nuevas habilidades artificiales en el desarrollo de interfaces que prometen una experiencia de usuario superior a la provista por las tradicionales interfaces gráficas o por los sistemas de diálogo con funcionamiento similar al de un autómatas. Una característica particular que aportan las nuevas AGIs es la posibilidad de analizar semejanza semántica entre bloques de texto. Veamos por ejemplo las frases de la Fig. 4.

```
fraseF1 = "monarcas era un equipo de futbol"  
fraseF2 = "barcelona hizo epoca como conjunto futbolero"  
fraseDif = "ciencias computacionales es muy solicitada"
```

Figura 4 Frases para prueba

Se observa que la fraseF1 y la fraseF2 pertenecen al mismo contexto, lo que no ocurre con la tercera. Una vez que se obtienen los vectores asociados a cada frase se invoca el cálculo de similitud semántica a partir de la librería de OpenAI (Fig. 5).

```
similitud_F1vsF2 = cosine_similarity(vfraseF1,vfraseF2)  
similitud_F1vsFDif = cosine_similarity(vfraseF1,vfraseDif)  
similitud_F2vsFDif = cosine_similarity(vfraseF2,vfraseDif)
```

Figura 5 Invocación al cálculo de similitud entre frases

A continuación, en la Fig. 6, se presentan los resultados de semejanza entre las 3 frases.

```
Similitud de, "monarcas era un equipo de futbol" con "barcelona hizo epoca como conjunto futbolero" es: 0.8764746966737981  
Similitud de, "monarcas era un equipo de futbol" con "ciencias computacionales es muy solicitada" es: 0.742653621637637  
Similitud de, "barcelona hizo epoca como conjunto futbolero" con "ciencias computacionales es muy solicitada" es: 0.75017789283
```

Figura 6 Resultados de cálculo de similitud semántica

La similitud mayor se obtiene justamente entre la fraseF1 y fraseF2. Hay que observar que no hay palabras en común y sin embargo los resultados son los esperados. Esto se

debe a que los vectores que se obtuvieron consideran la técnica de “word embedding”. Esto significa que incluyen numéricamente la información de relación de aparición entre ellas en el lenguaje español. Esto se obtiene a partir del análisis de millones de documentos por OpenAI (2023). Esta característica permite concebir chatbots con habilidades semánticas que se suman al alto grado de capacidad de procesamiento de diálogos complejos.

Conclusiones

Se presentaron 4 técnicas que se emplearon a lo largo del proceso de evolución del desarrollo de interfaces en lenguaje natural. Creemos que en el futuro inmediato las interfaces de usuario sufrirán transformaciones radicales a partir de la influencia de las AGIs. Esto abre, nuevamente en las ciencias computacionales, una beta de trabajo científico e industrial que generará una revolución en la forma en la que las personas se comunican con las máquinas. Incluso, se puede llegar a pensar que podría haber comunicación verbal entre computadoras y/o dispositivos en un futuro corto. Por todo lo anterior, la presente revisión a las técnicas presentadas constituye un tema pertinente y relevante que se continuará abordando en futuras investigaciones.

Referencias

- Alvarez R, J.D. (2009), *Clasificación Automática de Textos usando Reducción de Clases basada en Prototipos*, Tesis de maestría, Instituto Nacional de Astrofísica, Óptica y Electrónica, INAOE, Asesores: Manuel Montes y Gómez y Luis Villaseñor Pineda.
- Bard (2023), *Google Bard*, <https://bard.google.com/>
- Bhagwat, V. A (2018) *Deep Learning for Chatbots*, Master's Projects. 630. https://scholarworks.sjsu.edu/etd_projects/630
- Brownlee J. (2017), *Deep Learning for Natural Language Processing, Develop Deep Learning Models for Natural Language in Python*. Machine Learning Mastery.
- DialogFlow (2023), *DialogFlow, Documentation*, <https://dialogflow.com/>
- Maier, T., Menold, J., McComb, C. (2019), *Towards an ontology of cognitive assistants*. Proceedings of the Design Society: International Conference on Engineering Design 1(1), 2637– 2646.
- NSF (2016) *Intelligent Cognitive Assistants Workshops*, Workshop Summary and Recommendations, ICA-1 Report 2016, May 8-10, 2016, Washington, D.C., USA, vol. 1. Joint Semiconductor Research Corporation and NSF.
- OpenAI (2023), *Creating safe AGI that benefits all of humanity*, <https://openai.com/>, Septiembre 2023.
- Xu, Wei & Rudnicky, A., (2000). *Language modeling for dialog system*. Sixth International Conference on Spoken Language Processing, ICSLP 2000 / INTERSPEECH , 118-121.