**AGENTE INTELIGENTE CON**

**PROCESAMIENTO DE LENGUAJE NATURAL**

**Jerson Martínez, Julio Guadalupe**

***Facultad de Ingeniería Industrial y de Sistemas, Universidad Nacional de Ingeniería, Lima, Perú***

[***jcesar20112594d@gmail.com***](mailto:jcesar20112594d@gmail.com)

[***jersonrok@gmail.com***](mailto:jersonrok@gmail.com)

**2. DEFINICIÓN DE LA TAREA**

Existen numerosas aplicaciones industriales del reconocimiento del habla. Todas estas aplicaciones tienen algunas características comunes:

* La ***dificultad de la tarea***, debida al tamaño y tipo de los vocabularios y a la dependencia del hablante que usa el sistema.
* La ***necesidad de aprendizaje automático*** por parte del sistema, lo cual supone un gran coste computacional por un lado y la necesidad de obtener un buen número de muestras de entrenamiento.
* La necesidad de un grado mínimo de efectividad.
* La existencia de **entornos de trabajo ruidosos**

El sistema descrito en este artículo se ha aplicado a tareas de reconocimiento de voz. Para poder comunicar a un agente inteligente de manera remota usando el Lenguaje Natural. Para facilitar el reconocimiento de voz en cuanto a conjugación de verbos, singularizar y pluralizar sustantivos; hemos usado un módulo de conjugación de verbos [15].

**3. ESTRUCTURA DEL SISTEMA DE RECONOCIMIENTO DEL HABLA**

En la figura 1 se muestra la estructura general de un sistema de reconocimiento del habla continua típico, dentro del marco del reconocimiento de formas.

Los pasos básicos del procesamiento de la señal son los siguientes:

**1.- Análisis de características:** Se realiza un análisis frecuencial y/o temporal de la señal de voz para obtener vectores de observación o de características.

**RESUMEN:** *El presente trabajo tiene por objetivo hacer una revisión de las principales características de las técnicas de procesamiento del lenguaje natural, centrándose en su aplicación a la recuperación de información y áreas afines. El módulo de lenguaje natural traduce los comandos del operador a una representación en un lenguaje intermedio para, que posteriormente se ejecute el programa mediante un intérprete.*

**PALABRAS CLAVE:** Procesamiento del lenguaje natural, lenguajes, recuperación de la información, reconocimiento de voz.

**1.- INTRODUCCIÓN**

El instrumento que los seres humanos utilizamos para comunicar el conocimiento es el lenguaje natural. Actualmente, buena parte del saber humano se encuentra en forma digital en distintos tipos de colecciones de datos. Los volúmenes de información son inmensos, según la International Data Corporation [8], el mundo generó 1,8 zettabytes de información digital en 2011 y “en 2020 el mundo va a generar 50 veces [esa cantidad]” [13]. Las computadoras archivan esta información pero sin el Procesamiento de Lenguaje Natural (PLN) es difícil aprovecharla. El procesamiento del lenguaje involucra una transformación a una representación formal, manipula esta representación y por último, si es necesario, lleva los resultados nuevamente a lenguaje natural. Los campos de desarrollo de PLN incluyen la recuperación y extracción de información, traducción automática, sistemas de búsquedas de respuestas, generación de resúmenes automáticos, minería de datos, análisis de sentimientos, entre otras.

En este artículo se presenta la arquitectura de un sistema de lenguaje natural para teleoperar un agente inteligente, que realizará diversas tareas de servicio. El sistema incluye un módulo de reconocimiento de voz continua en tiempo real que se comunica a través de una red local con el controlador del robot. La interfaz de lenguaje natural está diseñada de manera que sea flexible para los distintos módulos del sistema completo, que incluye: un módulo de aprendizaje de nuevas tareas.

**4.- Análisis sintáctico:** En esta fase se añaden restricciones al sistema de reconocimiento de unidades de tal forma que las secuencias de unidades consideradas se correspondan con secuencias de unidades que formen palabras (que ya se consiguió en la decodificación léxica) y que estas palabras formen una secuencia válida según se especifica en una gramática de palabras.

**5.- Análisis semántico:** Esta fase añade nuevas restricciones a las posibles secuencias de reconocimiento.

La metodología utilizada en nuestro sistema de reconocimiento del habla continua es la basada en los modelos ocultos de Markov (Hidden Markov Models, HMM). [16] En la figura 2 se muestra un sistema típico de reconocimiento de palabras conectadas usando HMMs.

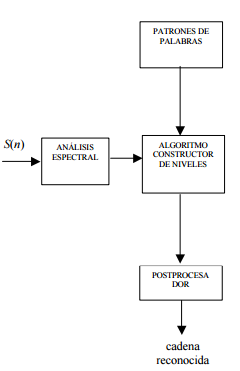


Figura 2. Sistema típico de reconocimiento de palabras conectadas usando modelos de Markov

Los pasos que se siguen son los siguientes:

***a.- Análisis espectral:***La señal de voz s(n) se convierte en un conjunto de vectores LPC [16] o en un conjunto de vectores de características. Esto define la secuencia de observación **O** de la cadena desconocida.

**2.- Sistema de reconocimiento de unidades:** Se debe realizar la elección de la unidad de reconocimiento adecuada para nuestra aplicación. En nuestro sistema se ha elegido como unidad la palabra, debido a que nuestro vocabulario no es demasiado grande y la tarea a la que se aplica el reconocimiento es restringida. Las palabras se modelan usando modelos ocultos de Markov [16].

**3.- Decodificación léxica:** En esta fase se añaden restricciones al sistema de reconocimiento de unidades de tal forma que las secuencias de unidades consideradas correspondan a secuencias de unidades de voz que existan en un diccionario de palabras.

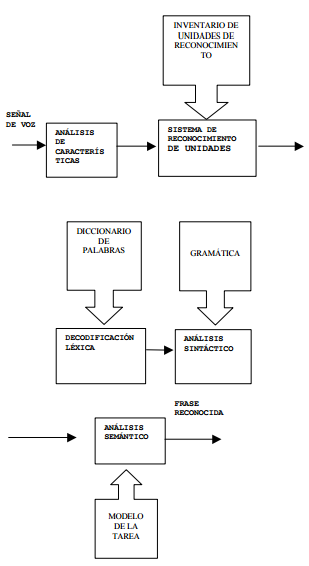


Figura 1. Estructura global de un sistema de reconocimiento del habla continua

***b.- Algoritmo de construcción de niveles:***Un nivel es una posición de palabra en una cadena. La secuencia de vectores de características de la cadena desconocida se compara con los HMMs de palabra única usando un algoritmo de puntuación de Viterbi [16]. La salida de este proceso es un conjunto de cadenas candidatas, generalmente de longitudes diferentes, ordenadas por puntuaciones.

***c.- Postprocesador:*** Sobre la cadena obtenida se realizan pruebas adicionales (por ejemplo, de duración), para eliminar candidatos no probables. El postprocesador elige la cadena más probable para las cadenas candidatas.

**4. Procesamiento del lenguaje natural:**

En esta sección se describe el esquema general de procesamiento del lenguaje natural usado en nuestra aplicación. Una vez que disponemos de la trascripción de la voz en un texto ASCII obtenida por el módulo de reconocimiento del habla, el módulo siguiente consiste en interpretar la frase y traducirla a un lenguaje comprensible por el robot.

Existen dos enfoques generales para realizar este procesamiento:

* Diseño
* Aprendizaje o adaptación

El ***diseño*** consiste en crear y programar una gramática específica para nuestra aplicación y realizar las tres fases típicas de análisis:

1. **Análisis léxico:** consiste en identificar cada unidad mínima de la frase.
2. **Análisis sintáctico:** consiste en identificar la estructura sintáctica de la frase, obteniendo un árbol sintáctico.
3. **Análisis semántico:** consiste en interpretar y traducir la frase.

Mediante las *técnicas de aprendizaje o adaptación*, el sistema se adapta a las restricciones sintácticas y semánticas de nuestra aplicación a partir de ejemplos [17]. Pueden usarse, por ejemplo, redes neuronales recursivas tipo **RAM** (*Recursive HeteroAssociative Memories*), que básicamente consisten en un codificador que transforma la frase en una representación analógica interna y un decodificador que transforma esa representación interna en una orden o programa para el robot. El funcionamiento del codificador y decodificador es lo que se debe aprender.

El análisis léxico y análisis sintáctico son las fases más algorítmicas del proceso. Existen varios algoritmos que se pueden utilizar, independientemente de la aplicación. El objetivo de estas dos fases es obtener una estructura en árbol que represente la sintaxis de la orden de entrada.

El análisis semántico es dependiente de la aplicación. En esta fase queremos obtener el significado de la frase para poder traducirla al lenguaje objeto. Por tanto, es la parte más crítica y menos algorítmica del diseño. En la figura 3 se muestra el esquema de un sistema típico de procesamiento del lenguaje natural.

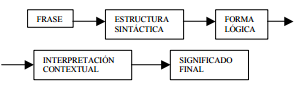


Figura 3. Esquema del procesamiento de lenguaje natural

En las fases de análisis sintáctico y semántico se intenta obtener una estructura con significado a partir de la sucesión de palabras recibidas del reconocedor de voz. En el caso de la interfaz con un robot, esta estructura puede ser representada directamente por la secuencia de primitivas para el robot en el lenguaje destino.

En la mayoría de sistemas de procesamiento del lenguaje natural, y en particular en la interfaz con un robot, es necesario realizar dos fases posteriores al análisis semántico: *la integración del discurso y el análisis pragmático.*

* En la fase de *integración del discurso* se obtienen las coordenadas de los objetos y localizaciones indicadas por el operador. Estas coordenadas se obtendrán finalmente consultando la base de datos del sistema.
* En la fase de *análisis pragmático* se traducen las instrucciones intermedias a las que se ha llegado a las instrucciones directamente entendibles por el robot (o las instrucciones definidas en el protocolo de comunicación con el robot).

**5.- Desarrollo del sistema de programación con lenguaje natural**

**5.1.- Arquitectura del sistema**

El sistema se compone de 2 módulos básicos que están interrelacionados: el módulo de lenguaje natural y el módulo de aprendizaje automático.

Estos 2 módulos se comunican a través de la información almacenada en una base de datos de tareas, donde el sistema de aprendizaje almacena las instrucciones de las tareas aprendidas

En la figura 4 se muestra la relación entre los módulos, donde L significa que el módulo accede a la base de datos para leer y E para escribir.

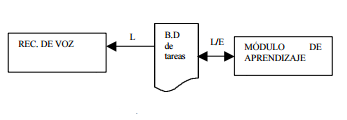


Figura 4. Arquitectura del Sistema

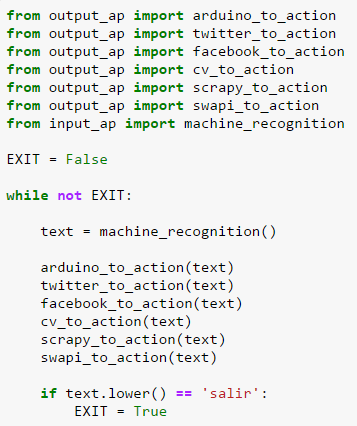
**5.2 Módulo de procesamiento del lenguaje**

La representación independiente del contexto que se ha elegido para la aplicación es un lenguaje intermedio que denominaremos lenguaje destino. Este lenguaje destino está formado por las primitivas básicas que puede ejecutar el agente, pero tienen un nivel de abstracción superior para que sea más sencillo el procesamiento del lenguaje. Los lenguajes de programación existentes para agentes son inadecuados como lenguajes intermedios, debido principalmente a dos razones:

1. Las funciones de movimiento y las representaciones del entorno obtenidas por los sensores típicos de un robot se parecen poco a las representaciones que un operador tiene en mente cuando emite un comando en lenguaje natural.
2. Las estructuras de programación típicas de estos lenguajes (secuencia, selección y repetición) son muy distintas a las estructuras que se obtienen después de analizar una frase en lenguaje natural.

El lenguaje intermedio usado en nuestro sistema se basa en procesos concurrentes, con la posibilidad de realimentación por voz, con el objetivo de facilitar el proceso de traducción al lenguaje directamente ejecutable por el agente. El siguiente código muestra los comandos que se emplea para las instrucciones y el núcleo que incluye los métodos actuadores:





**6. Conclusiones**

Se ha diseñado un sistema teleoperado que admite órdenes complejas usando un reconocedor de voz continua, con la estructura apropiada para su integración con un módulo de aprendizaje. La principal conclusión que obtenemos es que para el diseño de interfaces en lenguaje natural entre agentes y operador, es conveniente el uso de un lenguaje objeto adecuado para las estructuras obtenidas después del análisis del comando de entrada.

El trabajo futuro incluye: optimización de los algoritmos de análisis sintáctico y semántico.

**Agradecimientos**

Este trabajo ha sido realizado con los amplios conocimientos adquiridos en el curso de Sistemas Operativos. Por ello , nuestro agradecimientos al profesor.

**REFERECNIAS:**

**[8] J. F. Gantz and D. Reinsel, “Extracting value from chaos.” International Data Corporation. 2011. [Online]. Available:http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf**

**[13] G. S. Ingersoll, T. S. Morton, and A. L. Farris, “Taming text: how to find, organize, and manipulate It”, Manning Publications, December 28, 2012. [Also Online]. Available:** [**www.safaribooksonline.com**](http://www.safaribooksonline.com)

**[15]** [**http://www.clips.ua.ac.be/pages/pattern-es**](http://www.clips.ua.ac.be/pages/pattern-es)

**[16] Wang, X., Pols, L. “A preliminary study about robust speech recognition for a robotics application”, Institute of Phonetic Sciences, University of Amsterdam, Proceedings 21 (1997), 11—20.**

**[17]** Salerno, J.J. “An application of a recurrent neural model for parsing natural language”, International Journal on Artificial Intelligence Tools [Architectures, Languages, Algorithms], 6 Iss:3 p. 397 – 419. Sept 1997.