

Memoria de traducción morfológica

María Elena **Acevedo-Mosqueda**
Marco Antonio **Acevedo-Mosqueda**
Fabiola **Martínez-Zúñiga**
Federico **Felipe-Durán**

Instituto Politécnico Nacional,
Escuela Superior de Ingeniería Mecánica y Eléctrica.
Unidad Profesional "Adolfo López Mateos" Zacatenco,
Telecomunicaciones, Edif. Z-4, 3er Piso,
Col. Lindavista, Del. Gustavo A. Madero, CP 07738,
Ciudad de México.
MÉXICO.

correo electrónico (email): eacevedo@ipn.mx
macevedo@ipn.mx
fmartinezzu@ipn.mx
ffelipe100@hotmail.com

Recibido 10-11-2014, aceptado 14-03-2015.

Resumen

En este trabajo se implementa una memoria de traducción mediante herramientas de inteligencia artificial. La memoria de traducción se puede concebir como una tarea de reconocimiento de patrones. Este tipo de memoria se utiliza para implementar un diccionario técnico automatizado Español/Inglés-Inglés/Español, el cual es una herramienta muy útil para estudiantes de las ciencias de la computación. En particular, se utilizan las *memorias asociativas bidireccionales morfológicas*, en su modo autoasociativo, que muestran tener una recuperación correcta de todos los patrones aprendidos.

Palabras clave: inteligencia artificial, memorias asociativas, reconocimiento de patrones, memoria de traducción.

Abstract (Morphological Translation Memory)

In this work, a translation memory is implemented with Artificial Intelligence tools. The translation memory can be conceived as a pattern recognition task. This kind of memory

is used to implement a computing technical Spanish-English/English-Spanish dictionary that is a very useful tool for students of computer sciences. In particular, we use a Morphological Bidirectional Associative Memory in an autoassociative way which allows to recall all the learning patterns.

Key words: artificial intelligence, associative memories, pattern recognition, translation memory.

1. Introducción

Un sistema de memorias de traducción [1], como su nombre lo indica, memoriza las traducciones hechas por un traductor humano. Muchos de los sistemas de memorias de traducción consisten en una base de datos que almacenan el texto original junto con su traducción como segmentos de pares.

Actualmente, la mayoría de los traductores profesionales utilizan la memoria de traducción como base fundamental [2,3]. Los sistemas comerciales más conocidos son: *Across*, *Déjà Vu*, *MemoQ*, *MultiTrans*, *SDL Trados*, *Similis*, *Transit and Wordfast*.

Además de la memoria o archivo de traducción como elemento principal, un sistema de memoria de traducción [4] consiste en un arreglo de herramientas y funcionalidades que asisten al traductor humano, las cuales, usualmente incluyen: (a) un editor multilinguaje para leer los textos fuente y escribir las traducciones en todos los formatos de archivo relevantes que vienen que diferentes procesadores de texto; (b) un programa de administración de terminología para manejar el almacenamiento, la recuperación y la actualización de las bases de datos terminológicas; (c) un término automático para el reconocimiento de características para buscar automáticamente en la base de datos todos los términos que aparecen en los segmentos del texto fuente; (d) una herramienta de concordancia que permita a los usuarios recuperar todas las instancias de una cadena específica desde la memoria de traducción; (e) un rasgo estadístico para proveer la cantidad de texto que puede ser reusada de una memoria de traducción para traducir un nuevo documento; y (f) una herramienta de alineamiento para crear nuevas bases de datos para memorias de traducción desde documentos previamente traducidos que solamente están disponibles como archivos de texto fuente y objetivo separados,

mediante la comparación de un archivo fuente y su traducción, empatando los segmentos correspondientes y ligándolos como unidades de una memoria de traducción.

Un diccionario manejado por una computadora [5] es un diccionario en forma electrónica que puede ser almacenado en una base de datos y puede ser accesado mediante una aplicación de software. El diccionario puede manejar un solo lenguaje o puede ser multilingüe para poder manejar traducciones entre dos a más lenguajes o la combinación de ambos.

La forma general de una entrada de un diccionario bilingüe [6] es: $w_p^i \rightarrow v_1, v_2, \dots, v_k$, donde w_p^i es el i -ésimo significado de la palabra w con parte de la lengua p en el lenguaje fuente. Cada v_j es una traducción en el lenguaje objetivo para el significado w_p^i . Se asume que cada v_j , implícitamente, tiene la misma parte de la lengua p como w_p^i . También es importante notar que ningún significado se asocia implícitamente con v_j .

En este trabajo, no se tiene el problema de la desambiguación debido a que nuestro propósito es asociar cada palabra objetivo v_j con un solo significado w_p^i .

Por otra parte, la Inteligencia Artificial (IA) [7], definida ampliamente, tienen que ver con el comportamiento artificial de artefactos. El comportamiento inteligente implica percepción, razonamiento, aprendizaje, comunicación y actuación en ambientes complejos. Además, las memorias asociativas son una subárea de la IA y son parte relevante de las ciencias de la computación. Los modelos de memorias asociativas tienen una gran importancia y dinamismo en las actividades desarrolladas por varios grupos de investigación, específicamente, aquellos que investigan tópicos relacionados con la teoría y las aplicaciones de la clasificación y el reconocimiento de patrones.

2. Desarrollo

2.1. Memorias asociativas

La mente humana es asociativa [8], *i.e.*, los seres humanos tenemos una memoria asociativa debido a que mediante la asociación podemos recordar lo que hemos aprendido: objetos, seres humanos, conceptos, ideas abstractas, aún si los objetos están contaminados con ruido (caras con barba o anteojos o canciones incompletas).

Una Memoria Asociativa (MA) [9] es una estructura de contenido direccionable que mapea un conjunto de patrones de entrada a un conjunto de patrones de salida. Una estructura de contenido direccionable se refiere a la organización de una memoria en donde la memoria se accesa mediante su contenido, contrariamente a una dirección explícita en un sistema de memoria de una computadora tradicional. La recuperación de los datos se basa en un grado de similaridad entre el patrón de entrada y los patrones almacenados en la memoria.

Una memoria asociativa puede formularse como un sistema de entrada y salida, idea que se esquematiza en la Figura 1.

En este esquema, los patrones de entrada y salida están representados por vectores columna denotados por \mathbf{x} y \mathbf{y} , respectivamente. Cada uno de los patrones de entrada forma una asociación con el correspondiente patrón de salida, la cual es similar a la una pareja ordenada; por ejemplo, los patrones \mathbf{x} y \mathbf{y} del esquema anterior forman la asociación (\mathbf{x}, \mathbf{y}) .

Los patrones de entrada y salida se denotarán con las letras negrillas, \mathbf{x} y \mathbf{y} , agregándoles números naturales como superíndices para efectos de discriminación simbólica. Por ejemplo, a un patrón de entrada \mathbf{x}^1 le corresponderá el patrón de salida \mathbf{y}^1 , y ambos formarán la asociación $(\mathbf{x}^1, \mathbf{y}^1)$; del mismo modo, para un número entero positivo k específico, la asociación correspondiente será $(\mathbf{x}^k, \mathbf{y}^k)$.

La memoria asociativa \mathbf{M} se representa mediante una matriz, la cual se genera a partir de un conjunto finito de asociaciones conocidas de antemano: este es el *conjunto fundamental de aprendizaje*, o simplemente *conjunto fundamental*.

El conjunto fundamental se representa de la siguiente manera: $\{(\mathbf{x}^\mu, \mathbf{y}^\mu) \mid \mu = 1, 2, \dots, p\}$ donde p es un número entero positivo que representa la cardinalidad del conjunto fundamental.



Fig. 1. Una memoria asociativa se puede representar como un sistema de entrada y salida

A los patrones que conforman las asociaciones del conjunto fundamental se les llama *patrones fundamentales*. La naturaleza del conjunto fundamental proporciona un importante criterio para clasificar las memorias asociativas:

Una memoria es *autoasociativa* si se cumple que $\mathbf{x}^\mu = \mathbf{y}^\mu$ para todo μ elemento de $\{1, 2, \dots, p\}$, por lo que uno de los requisitos que se debe de cumplir es que $n = m$.

Una memoria *heteroasociativa* es aquella en donde existe μ elemento de $\{1, 2, \dots, p\}$ para el que se cumple que $\mathbf{x}^\mu \neq \mathbf{y}^\mu$. Nótese que puede haber memorias heteroasociativas con $n = m$.

En los problemas donde intervienen las memorias asociativas, se consideran dos fases importantes: la fase de aprendizaje, que es donde se genera la memoria asociativa a partir de las p asociaciones del conjunto fundamental, y la fase de recuperación que es donde la memoria asociativa opera sobre un patrón de entrada.

2.2. Memorias Asociativas Morfológicas (MAM)

El concepto de redes neuronales morfológicas fue propuesto por Davidson [10] y Ritter [11]. Ritter y Sussner puntualizaron que son un tipo novedoso de redes neuronales artificiales y que con capaces de resolver cualquier problema computacional convencional. En 1998, Ritter y sus colegas establecieron el concepto de Memorias Asociativas Morfológicas (MAM) [12]. Las MAM son esencialmente diferentes a las redes neuronales tradicionales tales como Lemmatrix [13], Correlograph [14], Linear Associator [15,16] y Hopfield [17]. Las reglas de estímulos de las neuronas, reglas asociativas, entre otras, difieren de las redes neuronales tradicionales. Se pueden procesar no sólo dos valores de patrones, sino que también patrones con valores reales tales como el procesamiento y reconocimiento de imágenes en escala de grises. Al mismo tiempo, una MAM tiene una capacidad ilimitada de almacenamiento, una velocidad de recuperación de un solo paso y gran tolerancia al ruido aditivo y sustractivo.

Los cálculos básicos que ocurren en una red morfológica están basados en la estructura algebraica $(R, \vee, \wedge, +)$, donde los símbolos \vee y \wedge denotan las operaciones binarias de máximo y mínimo, respectivamente.

Sean $(\mathbf{x}^1, \mathbf{y}^1), (\mathbf{x}^2, \mathbf{y}^2), \dots, (\mathbf{x}^p, \mathbf{y}^p)$ los p pares de vectores con $\mathbf{x}^k = (x_1^k, x_2^k, \dots, x_n^k) \in \mathbb{R}^n$ y $\mathbf{y}^k = (y_1^k, y_2^k, \dots, y_m^k) \in \mathbb{R}^m$ para $k = 1,$

$2, \dots, p$. Para un conjunto dado de asociaciones de patrones $\{(\mathbf{x}^k, \mathbf{y}^k) \mid k = 1, 2, \dots, p\}$ se definen un par de matrices de patrones asociados (X, Y) , donde $X = (\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^p)$ y $Y = (\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^p)$. Por tanto, X es de dimensión $n \times p$ con la ij -ésima entrada x_{ij}^k y Y tiene dimensión $m \times p$ con la ij -ésima entrada y_{ij}^k . Ya que $\mathbf{y}^k \nabla (-\mathbf{x}^k)^t = \mathbf{y}^k \Delta (-\mathbf{x}^k)^t$, la carga notacional se reduce denotando esta identidad morfológica por $\mathbf{y}^k \times (-\mathbf{x}^k)^t$. Con cada par de matrices (X, Y) asociamos dos memorias asociativas naturales M y W de dimensiones $m \times n$, las cuales están definidas por las ecuaciones (1) y (2):

$$M = \bigvee_{k=1}^p (\mathbf{y}^k \times (-\mathbf{x}^k)^t) \quad (1)$$

$$W = \bigwedge_{k=1}^p (\mathbf{y}^k \times (-\mathbf{x}^k)^t) \quad (2)$$

Estas dos ecuaciones describen la fase de entrenamiento de una memoria asociativa morfológica.

Las ecuaciones para la fase de recuperación se pueden observar en las ecuaciones (3) y (4).

$$\mathbf{y} = M \Delta \mathbf{x} \quad (3)$$

$$\mathbf{y} = W \nabla \mathbf{x} \quad (4)$$

En la siguiente subsección se presenta un ejemplo numérico de las fases de aprendizaje y recuperación de una memoria morfológica.

2.3. Ejemplo

Se presenta un ejemplo que permita explicar ambas fases: de entrenamiento y recuperación. Se utilizarán patrones similares a los usados en este trabajo. El valor de cada elemento de los vectores será $a \in \mathbb{Z}^+$ y $a = \{0, 1, \dots, 255\}$. Sean (x^p, x^p) los pares de patrones del conjunto fundamental con $p = \{1, 2, 3\}$. La cardinalidad de x es 3. Como se puede observar, la memoria será autoasociativa. A continuación se presentan los tres patrones que se utilizarán en este ejemplo.

$$x^1 = \begin{bmatrix} 72 \\ 79 \\ 76 \\ 65 \end{bmatrix} \quad x^2 = \begin{bmatrix} 104 \\ 111 \\ 118 \\ 122 \end{bmatrix} \quad x^3 = \begin{bmatrix} 86 \\ 87 \\ 88 \\ 89 \end{bmatrix}$$

Ahora, se calcula la primera asociación.

$$\mathbf{x}^1 \times (-\mathbf{x}^1)' = \begin{bmatrix} 72 \\ 79 \\ 76 \\ 65 \end{bmatrix} \times (-(72 \ 79 \ 76 \ 65)) =$$

$$= \begin{bmatrix} 72-72 & 72-79 & 72-76 & 72-65 \\ 79-72 & 79-79 & 79-76 & 79-65 \\ 76-72 & 76-79 & 76-76 & 76-65 \\ 65-72 & 65-79 & 65-76 & 65-65 \end{bmatrix}$$

$$\mathbf{x}^1 \times (-\mathbf{x}^1)' = \begin{bmatrix} 0 & -7 & -4 & 7 \\ 7 & 0 & 3 & 14 \\ 4 & -3 & -3 & 11 \\ -6 & -14 & -14 & 0 \end{bmatrix}$$

Se utiliza la ecuación (1) para obtener la memoria autoasociativa morfológica del tipo *máx*.

$$M = \begin{bmatrix} 0 & -7 & -4 & 7 \\ 7 & 0 & 3 & 14 \\ 4 & -3 & -3 & 11 \\ -6 & -14 & -14 & 0 \end{bmatrix} \mathbf{V} \begin{bmatrix} 0 & -7 & -14 & -18 \\ 7 & 0 & -7 & 11 \\ 14 & 7 & 0 & -4 \\ 18 & 11 & 4 & 0 \end{bmatrix} \mathbf{V}$$

$$\mathbf{V} \begin{bmatrix} 0 & -1 & -2 & -3 \\ 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 \\ 3 & 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & -2 & 7 \\ 7 & 0 & 3 & 14 \\ 14 & 7 & 0 & 11 \\ 18 & 11 & 4 & 0 \end{bmatrix}$$

La memoria *mín* se construye de manera similar a la memoria *máx* sólo que, en este caso, se aplica el operador mínimo. Entonces, la memoria obtenida es:

$$W = \begin{bmatrix} 0 & -7 & -14 & -18 \\ 1 & 0 & -7 & -2 \\ 2 & -3 & 0 & -4 \\ -6 & -14 & -11 & 0 \end{bmatrix}$$

Ahora, se le presenta el primer patrón a la memoria asociativa *máx* para ilustrar la fase de recuperación.

$$M\Delta\mathbf{x}^1 = \begin{bmatrix} 0 & -1 & -2 & 7 \\ 7 & 0 & 3 & 14 \\ 14 & 7 & 0 & 11 \\ 18 & 11 & 4 & 0 \end{bmatrix} \Delta \begin{bmatrix} 72 \\ 79 \\ 76 \\ 65 \end{bmatrix} =$$

$$= \begin{bmatrix} (0+72)\wedge(-1+79)\wedge(-2+76)\wedge(7+65) \\ (7+72)\wedge(0+79)\wedge(3+76)\wedge(14+65) \\ (14+72)\wedge(7+79)\wedge(0+76)\wedge(11+65) \\ (18+72)\wedge(11+79)\wedge(4+76)\wedge(0+65) \end{bmatrix} =$$

$$= \begin{bmatrix} 72\wedge 78\wedge 74\wedge 72 \\ 79\wedge 79\wedge 79\wedge 79 \\ 86\wedge 86\wedge 76\wedge 76 \\ 90\wedge 90\wedge 80\wedge 65 \end{bmatrix} = \begin{bmatrix} 72 \\ 79 \\ 76 \\ 65 \end{bmatrix}$$

Se puede observar que se recupera el patrón \mathbf{x}^1 . De hecho, dado que la memoria asociativa morfológica que se está manejando es el tipo autoasociativa, todos los patrones aprendidos serán recuperados. Esta es la característica principal de las memorias autoasociativas: no tienen factor de olvido.

2.4. Memoria Asociativa Morfológica Bidireccional (MAM)

El diagrama a bloques de la MAM Bidireccional se presenta en la Figura 2. Este modelo es muy parecido al de la Memoria Asociativa Bidireccional Alfa-Beta (BAM Alfa-Beta) [18], la única diferencia es que las memorias Alfa-Beta manejan números binarios, mientras que las morfológicas utilizan números reales.

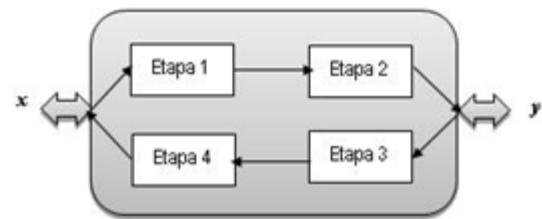


Fig. 2. Diagrama de bloques de la memoria asociativa bidireccional morfológica.

La Etapa 1 y 3 son similares en cuanto a su función, lo mismo sucede con las Etapas 2 y 4. Es por esto que sólo se describirá el sentido hacia adelante, es decir, se presentará el proceso de las Etapas 1 y 2.

En la fase de aprendizaje, la Etapa 1 es la encargada de aplicar una transformada vectorial de expansión dimensional con un vector similar al *one-hot* de la BAM Alfa-Beta, sólo que en este caso sustituimos el 1 por un valor de 255, como se muestra a continuación.

$$x^1 = \begin{bmatrix} x_1^1 \\ x_2^1 \\ \vdots \\ x_n^1 \\ 255_{(1)} \\ 0_{(2)} \\ \vdots \\ 0_{(p)} \end{bmatrix}, x^2 = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \\ 0_{(1)} \\ 255_{(2)} \\ \vdots \\ 0_{(p)} \end{bmatrix}, \dots, x^p = \begin{bmatrix} x_1^p \\ x_2^p \\ \vdots \\ x_n^p \\ 0_{(1)} \\ 0_{(2)} \\ \vdots \\ 255_{(p)} \end{bmatrix}$$

donde p es el número de patrones que se van a asociar. También se realiza la transformada vectorial de expansión dimensional con un vector similar al *zero-hot* pero utilizamos un -255 en vez del cero y ceros en lugar de los unos.

$$\bar{x}^1 = \begin{bmatrix} \bar{x}_1^1 \\ \bar{x}_2^1 \\ \vdots \\ \bar{x}_n^1 \\ -255_{(1)} \\ 0_{(2)} \\ \vdots \\ 0_{(p)} \end{bmatrix}, \bar{x}^2 = \begin{bmatrix} \bar{x}_1^2 \\ \bar{x}_2^2 \\ \vdots \\ \bar{x}_n^2 \\ 0_{(1)} \\ -255_{(2)} \\ \vdots \\ 0_{(p)} \end{bmatrix}, \dots, \bar{x}^p = \begin{bmatrix} \bar{x}_1^p \\ \bar{x}_2^p \\ \vdots \\ \bar{x}_n^p \\ 0_{(1)} \\ 0_{(2)} \\ \vdots \\ -255_{(p)} \end{bmatrix}$$

Con estos vectores transformados, se crean dos memorias autoasociativas morfológicas: *máx* y *mín*, respectivamente.

$$M = \bigvee_{k=1}^p (\mathbf{x}^k \times (-\bar{\mathbf{x}}^k)) \tag{1}$$

$$W = \bigwedge_{k=1}^p (\bar{\mathbf{x}}^k \times \mathbf{x}^k) \tag{2}$$

La Etapa 2 es igual que en la BAM Alfa-Beta, es decir, se crea el *Linear Associator* modificado con los patrones de salida y . En la fase de recuperación, la tarea de la Etapa 1 es la de recuperar los vectores similares al *one-hot* y *zero-hot*, de manera que estos vectores se le presenten al *Linear Associator* modificado para tratar de recuperar el patrón correspondiente. En la Figura 3, se puede observar la tarea de la Etapa 1.

La posición k en donde se encuentra el 255 o -255 nos indica a que columna de la matriz del *Linear Associator* tenemos que acceder para obtener el patrón de salida y correspondiente al patrón de entrada x .

3. Diseño

El diseño del diccionario se basó en la investigación de [19], por tanto, se decidió desarrollar un diccionario *click-on/key-in*.

El sistema propuesto está basado en el reconocimiento de patrones. Los patrones a reconocer serán las palabras. Por otra parte, el uso del concepto de memorias asociativas nos permite asociar palabras de diferentes lenguajes, para nuestro caso,

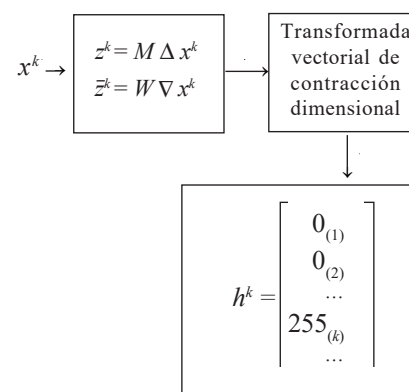


Fig. 3. Diagrama a bloques del proceso de la Etapa 1 en la fase de recuperación.

asociamos palabras en español con su correspondiente traducción en inglés. Por lo tanto, el primer paso es el reconocimiento de patrones. Entonces, cada una de las letras de cada palabra utilizada en el diccionario debe ser codificada con su correspondiente ASCII. El sistema, automáticamente, busca la palabra de mayor longitud y ajusta la longitud de las demás palabras agregándoles ceros. Una vez realizado esto, se procede a crear la memoria autoasociativa bidireccional. A todos los vectores de entrada que representan las palabras en español se les agrega un vector similar al vector *one-hot* para crear una memoria *máx*. También se le agrega a cada vector de entrada un vector similar al *zero-hot*. Cada vector *one-hot* y *zero-hot* tendrá 75 elementos ($p = 75$) debido a que el número total de palabras contenidas en el diccionario es de 75.

Se utilizan las ecuaciones (1) y (2) para implementar la fase de aprendizaje. En una matriz aparte, se guardan las palabras en inglés indicando con un índice la localización de cada palabra. Para la fase de recuperación, se presenta la palabra en español, se convierte a su correspondiente ASCII y se ajusta su longitud. Si se va a procesar con la memoria *máx*, se le agregan 75 elementos al final del vector con valor 255, si se va a procesar con la memoria *mín*, entonces se le agregan 75 elementos con valor -255. El procedimiento de la fase de recuperación es similar al de las memorias asociativas bidireccionales Alfa-Beta [18]. El resultado que se obtiene de la fase de recuperación es el índice que nos indica el lugar en donde se encuentra, dentro de la matriz, la palabra en inglés correspondiente a la palabra en español introducida.

Para crear las memorias autoasociativas morfológicas *máx* y *mín* para la traducción de inglés a español se sigue el mismo algoritmo descrito para la traducción Español-Inglés.

4. Resultados

El sistema traductor Español/Inglés-Inglés/Español, se implementó en una laptop Dell con un procesador Intel Core i3 utilizando el lenguaje de programación Matlab. En la Figura 2 se muestra un ejemplo de pantalla del sistema. Se puede observar que se tiene la opción de cambiar el lenguaje de traducción. Las listas desplegables marcadas con los números 1 y 3 permiten la elección de la palabra. El botón activa la traducción. También se tiene la posibilidad de introducir la palabra mediante el teclado.

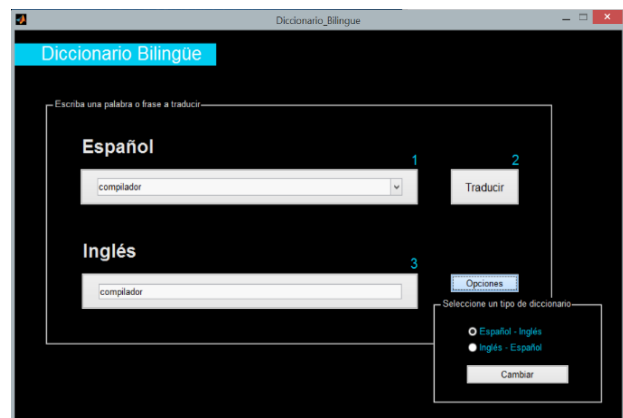


Fig. 4. Ejemplo de la pantalla del diccionario bilingüe asociativo.

Para probar la efectividad de recuperación (reconocimiento de patrones) de la memoria asociativa bidireccional morfológica, se crearon las memorias con diferentes números de palabras. Primero se utilizaron 10 palabras, luego 20, 30 y así sucesivamente hasta llegar a 100, después se incrementó de 100 en 100 el número de palabras hasta llegar a 500. En todos los casos, se recuperaron todos los patrones. En la Tabla 1 se muestran los resultados de recuperación del diccionario bilingüe asociativo.

Cabe mencionar que cuando cambia el número de patrones (palabras) el número de elementos que se agregan a los vectores originales también cambia; por tanto, cuando se le agregan más palabras al diccionario lo que está sucediendo es que se le está agregando más ruido, ya sea aditivo o sustractivo, a los patrones que se van a reconocer. Entonces va a llegar un punto en el que la cantidad de ruido sea demasiado grande y la memoria ya no sea capaz de reconocer los patrones. Esto depende del número de patrones y de la máxima longitud de las palabras introducidas. La máxima cantidad de ruido aditivo o sustractivo que soporta una memoria asociativa es alrededor de 97%. Por lo tanto, la cantidad de palabras (CP) que se pueden manejar para una longitud máxima de palabra fija, LM , y un vector *one-hot* o *zero-hot* con K elementos, siendo la longitud total del vector $CP + LM$, es

$$CP = \frac{(K + LM) * 97}{100}$$

Tabla 1. Porcentaje de recuperación del diccionario bilingüe asociativo dependiendo del número de palabras contenidas.

Núm. de palabras	% de recuperación
10	100
20	100
30	100
40	100
50	100
60	100
70	100
80	100
90	100
100	100
200	100
300	100
400	100
500	100

5. Conclusiones

Las memorias autoasociativas morfológicas no presentan factor de olvido por lo tanto, recuperan todos los patrones aprendidos. En este trabajo se utilizaron este tipo de memorias para implementar un diccionario técnico bilingüe basado en una memoria de traducción. El máximo de palabras que contiene este diccionario es de 500. Con esta cantidad de palabras el diccionario asociativo presentó un 100% en la recuperación, lo que quiere decir que su reconocimiento de patrones es correcto.

Por otra parte, las memorias asociativas morfológicas pueden manejar ruido aditivo o sustractivo con sus memorias *máx* y *mín*, respectivamente. Durante el proceso de recuperación, se le agrega estos tipos de ruido a los patrones que se van a reconocer. Mientras el ruido agregado no rebasa el 97%, la memoria será capaz de reconocer los patrones de manera correcta. Mientras más grande sea la longitud máxima de las palabras o frases que contenga el diccionario mayor será la cantidad de estas que pueden ser almacenadas.

Cabe resaltar que se espera que la MAM bidireccional tenga el mismo comportamiento que la BAM Alfa-Beta, es decir,

que no importando el número de patrones que se van a asociar, la recuperación siempre será correcta: todos los patrones aprendidos son recuperados.

Agradecimientos

Los autores agradecen el apoyo económico de SNI, SIP, EDI y EDD para la realización de esta investigación.

Referencias

- [1] Zerfass A. *Evaluating Translation Memory Systems*. <http://www.mt-archive.info/LREC-2002-Zerfass.pdf>.
- [2] Massion F. *Translation-Memory-Systeme im Vergleich*. Reutlingen Doculine, 2005.
- [3] Lagoudaki E. *Translation Memory Systems: Enlighthening user's perspective*. London: Imperial College, 2006.
- [4] Reinke U. State of the Art in Translation Memory Systems. *Translation: Computation, Corpora, Cognition*. Vol. 3(1), pp 27-48, 2013.
- [5] Deshpande A, A survey: structure of machine readable dictionary, *International Journal of Engineering and Innovative Technology*, Vol. 1(4), pp 27-30, 2012.
- [6] Flati T y Navigli R, *Journal of Artificial Intelligence Research*, Vol. 43, pp 135-171, 2012.
- [7] Nilson NJ, *Inteligencia Artificial: Una nueva síntesis*, McGraw-Hill 2000
- [8] Yáñez C, Acevedo ME, Gómez P, *Modelos asociativos Alfa-Beta*, Editorial CIDETEC-IPN, 2007.
- [9] Vyas P, Kacheria R, HDL based implementation of Palm associative memory. *Asian Journal of Computer Science and Information Technology*. Vol 3(8), pp 109-116, 2013.
- [10] Davidson JL, Hummer F. Morphology Neural Networks: An Introduction with applications. *Circuits Systems Signal Process*. Vol. 12(2), pp 177-210, 1993.
- [11] Ritter GX, Sussner P. An introduction to morphological neural networks. *Proceedings of the 13th International Conference on Pattern Recognition*. Vol. 4, pp 709-717, 1996.
- [12] Ritter GX, Sussner P, Díaz-de-León JL. Morphological Associative Memories. *IEEE Transactions on Neural Networks*. Vol. 9(2), pp 281-293, 1998.
- [13] Steinbuch, K., "Die Lernmatrix", *Kybernetik*, vol. 1, no. 1, pp. 36-45, 1961.
- [14] Willshaw, D., Buneman, O. & Longuet-Higgins, H., "Non-holographic associative memory", *Nature*, no. 222, pp. 960-962, 1969.

- [15] Anderson, J. A., "A simple neural network generating an interactive memory", *Mathematical Biosciences*, vol. 14, pp. 197-220, 1972.
- [16] Kohonen, T., "Correlation matrix memories", *IEEE Transactions on Computers*, C-21, vol. 4, pp. 353-359, 1972.
- [17] Hopfield, J.J., "Neural networks and physical systems with emergent collective computational abilities", *Proceedings of the National Academy of Sciences*, vol. 79, pp. 2554-2558, 1982.
- [18] Acevedo ME, *Memorias Asociativas Bidireccionales Alfa-Beta*, Tesis Doctoral, CIC-IPN, 2006
- [19] Liu TC, Fan M, Paas F, Effects of digital dictionary format on incidental acquisition of spelling knowledge and cognitive load during second language learning: Click-on vs key-in dictionaries, *Computers & Education*, Vol. 70, pp 9-20, 2014.

Latindex

Sistema Regional en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal

<http://www.latindex.unam.mx>