

XV Jornadas de Investigación y Cuarto Encuentro de Investigadores en Psicología del Mercosur. Facultad de Psicología - Universidad de Buenos Aires, Buenos Aires, 2008.

Algunas ventajas del enfoque conexionista en cuanto al procesamiento de lenguaje natural.

Bruno, Mariano, Grieco, Gustavo y Saura, Silvana.

Cita:

Bruno, Mariano, Grieco, Gustavo y Saura, Silvana (2008). *Algunas ventajas del enfoque conexionista en cuanto al procesamiento de lenguaje natural*. XV Jornadas de Investigación y Cuarto Encuentro de Investigadores en Psicología del Mercosur. Facultad de Psicología - Universidad de Buenos Aires, Buenos Aires.

Dirección estable: <https://www.aacademica.org/000-032/9>

ARK: <https://n2t.net/ark:/13683/efue/00Q>

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

ALGUNAS VENTAJAS DEL ENFOQUE CONEXIONISTA EN CUANTO AL PROCESAMIENTO DE LENGUAJE NATURAL

Bruno, Mariano; Grieco, Gustavo; Saura, Silvana
Universidad Nacional de Rosario -IRICE-CONICET-. Argentina

RESUMEN

En el presente trabajo, se examinan algunas ideas fundamentales relativas al procesamiento de lenguaje natural. En primer lugar, se exploran los modelos de procesamiento de lenguaje natural basados en los supuestos chomskianos de gramáticas formales. En segundo lugar, se introducen las características generales de los modelos conexionistas. A continuación, se revisan los modelos de procesamiento de lenguaje natural basados en el modelo conexionista de SRN (simple recurrent networks). Posteriormente, se analizan algunas de las características más destacables de este modelo. Finalmente, se consideran algunas de las ventajas que posee la perspectiva conexionista por sobre los modelos de procesamiento de lenguaje natural clásicos.

Palabras clave

Procesamiento de lenguaje natural Gramáticas formales Conexionismo Redes de recurrencia simple

ABSTRACT

SOME ADVANTAGES OF THE CONNECTIONIST APPROACH REGARDING NATURAL LANGUAGE PROCESSING

In the present work, some main ideas regarding natural language processing are examined. Firstly, some models of natural language processing based on Chomsky's formal grammars approach are explored. Secondly, some general characteristics of the connectionist models are introduced. Later on, natural language processing models based on the Simple Recurrent Network connectionist model are explored. Next, the most remarkable features of these models are analysed. Finally, some of the advantages of the connectionist perspective in comparison to traditional natural language processing models are highlighted.

Key words

Natural language processing Formal grammars Connectionism Simple recurrent networks

1. INTRODUCCIÓN

El objetivo que persigue el presente trabajo radica en examinar algunas ideas fundamentales relativas al procesamiento de lenguaje natural provenientes del enfoque conexionista. En primer lugar, resulta necesario señalar que el procesamiento de lenguaje natural o NLP (*Natural Language Processing*) constituye, dentro de la ciencia cognitiva, un campo de investigación en donde convergen varias disciplinas: la inteligencia artificial, la lingüística, la psicología, la filosofía, la neurobiología y la antropología (Gardner, 1985). Dentro del campo específico de la inteligencia artificial, el objetivo del NLP es la creación de programas que puedan analizar, comprender y generar el lenguaje humano. Esto es, diseñar programas que ejecuten o simulen la comunicación humana.

Cabe destacar, que al momento de crear un programa de este tipo se debe tener en cuenta qué teoría o supuestos de base se

están utilizando para el diseño del mismo. Esto es, resulta necesario preguntarse qué concepto de lenguaje se utiliza y que arquitectura de procesamiento se considera que posee la mente humana. Más específicamente, una máquina capaz de comunicarse utilizando el lenguaje natural ¿Deberá aprender el lenguaje o se le “programará” explícitamente cómo deben combinarse las palabras?

En el presente trabajo se analizarán dos enfoques. Por una parte, se explorarán los programas que surgieron a partir de la implementación de los diversos enfoques de Noam Chomsky (1998, 1999).

En líneas generales, los enfoques chomskianos de adquisición del lenguaje natural postulan que los niños poseen las nociones de verbo, sustantivo y otras estructuras de lenguaje básicas al momento del nacimiento. Esto es, son innatas. Posteriormente, en términos simplificados, a partir del contacto con el entorno, los niños ajustan sus parámetros para poder dominar la gramática de su lengua nativa. Esta gramática formal es la que le permitirá al niño generar oraciones a partir de determinadas reglas explícitas y de su aplicación secuencial.

Por otra parte, el otro enfoque que se analizará en el presente trabajo es el conexionista. Este enfoque está basado en redes neuronales artificiales (RNA). Las RNA simulan el comportamiento de las neuronas en el cerebro, procesando información de entrada (*input*) y generando salidas (*output*).

Por último, se examinarán algunas de las ventajas que el enfoque conexionista posee frente a los modelos de procesamiento de lenguaje natural derivados de los enfoques chomskianos.

2. EL ENFOQUE DE LAS GRAMÁTICAS FORMALES

El primer enfoque presentado es el más tradicional y está basado en los enfoques chomskianos que, en mayor o menor medida, se estructuran en torno a la noción de gramáticas formales (Chomsky, 1999). Una gramática formal se compone de un conjunto bien definido de reglas y símbolos, y un procedimiento algorítmico secuencial que permite aplicar esas reglas a los símbolos. El ejemplo canónico de gramáticas formales es el de la oración bimembre verbal, donde primero va el sujeto, luego el predicado, que contiene el verbo, con posibles modificadores y objetos directos e indirectos. Con esta información y una lista de palabras previamente clasificadas como verbos, sujetos, etc, resulta evidente cómo reconocer o fabricar oraciones.

Desde el punto de vista computacional, los enfoques chomskianos sugieren el uso de categorías para estas gramáticas formales, dependiendo del tipo de algoritmo que puede reconocer o generar dichas gramáticas. Más aún, el problema de construir un programa que permita reconocer oraciones con gramáticas formales está muy bien estudiado y no resulta compleja su implementación.

La razón para presentar el procesamiento natural usando reglas formales reside en que dicha teoría está basada fuertemente en la metáfora del ordenador (Gardner, 1985; Johnson Laird, 1988 y Newell, 1980) cuya esencia consiste en considerar que la mente humana procesa información aplicando un algoritmo que utiliza reglas bien definidas.

3. EL ENFOQUE CONEXIONISTA

El enfoque conexionista está basado en redes neuronales artificiales (Rumelhart, Hinton y McClelland, 1986; Anderson, 1995; O'Reilly y Munakata, 2000). Las RNA son un conjunto de técnicas de aprendizaje automatizado que simulan (con mayor o menor grado de abstracción) el procesamiento cognitivo humano inspirado en la arquitectura de procesamiento de información que utiliza el cerebro.

Las redes neuronales artificiales están constituidas por unidades (similares a neuronas) y conexiones entre las mismas. Cada conexión posee un valor numérico que representa la fuerza con la que una unidad afecta a otra a la cuál está conectada. La información que se ingresa a la red, se propaga a través de sus unidades. Cabe mencionar que, la conexión entre dos unidades puede ser excitatoria o inhibitoria. Cuando esta conexión entre dos unidades es excitatoria y la primera se activa, la segunda

tiende a activarse también. Inversamente, cuando la conexión es inhibitoria, la segunda unidad tiende a inhibirse.

3.1. Modelo Conexionista Feedforward Multicapa

En un modelo conexionista *feedforward* (o de propagación de la activación hacia adelante) multicapa las unidades se ordenan en capas (Anderson, 1995). Una capa es un conjunto de unidades en donde la información se procesa en forma paralela. La propagación de la información que entra en la RNA se procesa, en primer lugar, en una capa llamada “capa de entrada”. Luego esta capa enviará información a una segunda capa de unidades llamada “capa oculta”. Por último, la capa oculta enviará la información a la denominada “capa de salida”. La “capa oculta” posee una influencia sobre el resultado que será de vital importancia para la RNA ya que en ella se formará una representación interna de los datos de entrada para generar una adecuada clasificación de los mismos.

Para poder procesar usando una RNA, hace falta reunir un conjunto de entradas y salidas de ejemplos correctos definidos en función de un problema particular. Comenzando con conexiones aleatorias, se “entrena” a la red, mostrándole las entradas junto con las salidas deseadas. Luego se compara la salida obtenida con la salida deseada con el fin de “recompensar” sus aciertos, y “castigar” sus equivocaciones (Roseblatt, 1962 y Rumelhart, Hinton y Williams, 1986).

3.2. Redes de Elman

Jeffrey L. Elman es un psicolingüista pionero en la aplicación exitosa de un tipo particular de RNA, denominada red de recurrencia simple (o SRN según sus siglas en inglés). Este tipo de RNA posee conexiones que retroalimentan las unidades, sólo en una capa en particular. Esta capa se denomina “capa contextual” y equivale a la memoria a corto plazo de la red. Esta sencilla modificación permite tratar el problema del lenguaje natural. Esto se debe a que, hasta ahora se han analizado las RNA en relación con problemas estáticos que no requerían procesamiento secuencial. En el reconocimiento de lenguaje natural, las letras se ordenan para formar palabras, las palabras forman frases, y las frases, textos.

En 1990, Elman introdujo una SRN que lleva su nombre, buscando la forma de plasmar el paso del tiempo en el aprendizaje del lenguaje. Para llevar a cabo su propósito, se basó en una RNA de tipo *feedforward* con tres capas de unidades: una de entrada, una oculta y otra de salida. Luego adicionó una capa más que denominó “capa contextual”. Esta nueva capa funciona copiando los valores de activación del instante previo de la capa oculta, para luego utilizarlos durante el procesamiento de la muestra actual. La capa contextual posee, en un primer momento, activaciones que no influyen en el resultado del procesamiento de la primera entrada.

De esta forma, la activación previa influye sobre la actual, capturando las propiedades específicas que poseen las secuencias de símbolos. Así, la RNA obtiene una memoria sencilla, en la cual el pasado inmediato posee gran influencia. En consecuencia, al obtener una memoria de corto plazo, se puede llegar a procesar secuencias de entidades, sean estos, hechos sucesivos en el tiempo o cadenas de palabras.

Las simulaciones de Elman se realizaron entrenando una SRN con oraciones muy simples pero semánticamente correctas (Elman, 1990; Elman y otros, 1999). El objetivo del entrenamiento consistió en que la red fuera capaz de predecir la próxima palabra que seguía en una oración incompleta. Esta es una característica humana y Elman, precisamente, buscaba generar un modelo conexionista que permitiera explicar dicha característica humana del procesamiento de la información.

Sin embargo, lo que resultó sorprendente, fue que, al hacer un análisis matemático del conjunto de valores de conexión de la capa oculta, se descubrió que había ciertas palabras eran “cercanas” entre sí.

Para entender este resultado de la simulación de Elman, hay que tener una idea de la “cercanía” entre entidades. En la vida cotidiana se sabe cuando dos objetos están cerca o lejos, simplemente midiendo la distancia física entre ellos.

Sin embargo, resulta posible definir otra noción de distancia. Un ejemplo sencillo sería, en la categorización de los animales, perro y lobo estarían semánticamente cercanos entre sí, pero distanciados de canario. Esta SRN pueden analizarse con una distancia especial que hace que los conceptos cercanos entre sí tiendan a confundirse en la predicción.

Utilizando esta noción especial de distancia, se procedió a analizar la capa oculta de la red de Elman en función de las palabras que componían las oraciones con las que se había entrenado a la misma, y se observó una característica sorprendente: los sustantivos y los verbos se encontraban relativamente lejos y los seres vivos estaban alejados en relación con los no vivos. Simultáneamente, los animales estaban distanciados de las personas. Con esta información se configuró un esquema que puso de manifiesto que la RNA estaba captando las categorías subyacentes a las oraciones simples a las cuales había sido expuesta. En cierto sentido, se puede afirmar que la red había capturado la semántica implícita a partir de los discursos con los cuales había sido entrenada (Bruno, 2005).

4. REFLEXIONES FINALES

Una de las primeras ventajas que se aprecian al describir el modelo conexionista es la manera en que una RNA se utiliza. Basta exponerla a oraciones bien formadas para que aproxime las categorías implícitas. Esta es una diferencia notable con los modelos basados en los enfoque chomskianos donde se necesitan, básicamente, reglas explícitas. Precisamente, la adquisición de categorías debe realizarse de forma manual, esto es, especificando las reglas para la categorización.

La idea de reglas explícitas en los modelos tradicionales conduce a descubrir una ventaja interesante de los modelos conexionistas aplicados al lenguaje natural, esta es, la alta tolerancia a fallos. Un error de tipeo en una palabra, un verbo mal conjugado o una oración "demasiado poética", no invalidará totalmente la respuesta de la red, sólo disminuirá su eficiencia. En cambio, una gramática formal probablemente fallará al querer aplicar una regla y podrá dar cómo resultado una oración incomprensible.

Otra importante ventaja de las RNA es su plausibilidad biológica. Aunque esta característica varía según el tipo de red, la representación del conocimiento y la forma de ajuste de los pesos, en general, el rendimiento de estas técnicas se está acercando rápidamente al de los humanos. Por el contrario, en los sistemas basados en reglas, la plausibilidad biológica está descartada, debido a que toda evidencia parece indicar que nuestros cerebros no funcionan simulando operaciones computacionales lógicas.

4.1. ¿Una Nueva Visión del Innatismo?

Las nuevas ideas aportadas por Elman y colaboradores (1999) permiten cambiar radicalmente el enfoque innatista en el aprendizaje del lenguaje natural.

Los enfoques chomskianos sobre la adquisición del lenguaje están fuertemente basados en la idea de que resulta imposible adquirir estructuras y conceptos primarios del lenguaje. En consecuencia, se los supone innatos y se reformula la teoría a partir de la introducción de la noción de parámetros para explicar la existencia heterogénea de idiomas. Sin embargo, a la luz de las simulaciones de Elman (1990) y la teoría de redes neuronales artificiales, las hipótesis chomskianas comienzan a presentar ciertas inconsistencias.

Una SRN expuesta a oraciones simples categoriza correctamente. La evidencia que surge de las simulaciones, junto con el mayor grado de plausibilidad biológica que poseen las RNA, conduce a pensar que lo realmente innato sería la capacidad de comprender un lenguaje a partir de la simple exposición al mismo y no mediante el uso de una gramática innata.

BIBLIOGRAFÍA

- ANDERSON, J.A. (1995). *An Introduction to Neural Networks*. Cambridge: The MIT Press.
- BRUNO, M. (2005). *Proposición de Modelos Formales de las Nociones Freudianas Ligadas al Concepto de Representante Psíquico Mediante las Concepciones Conexionistas Asociadas al Esquema de Representación Distribuida*. Tesis doctoral no publicada, Universidad Nacional de Rosario, Rosario.
- CHOMSKY, N. (1998). *El conocimiento del lenguaje*. Madrid: Alianza Editorial.
- CHOMSKY, N. (1999). *Estructuras sintácticas*. Buenos Aires: Siglo XXI.
- ELMAN, J.L.; BATES, E.A.; JOHNSON, M.H.; KARMILOFF-SMITH, A.; PARISI, D. y PLUMKETT, K. (1999). *Rethinking Innateness*. Cambridge: The MIT Press.
- ELMAN, J.L. (1990) Finding structure in time. *Cognitive Science*, 14(2):179-211.
- GARDNER, H. (1985). *La Nueva Ciencia de la Mente. Historia de la Revolución Cognitiva*. Buenos Aires: Editorial Paidós.
- NEWELL, A. (1980). Physical Symbol System. *Cognitive Science*, 4: 135-183.
- O'REILLY, R.C. y MUNAKATA, Y. (2000). *Computational Explorations in Cognitive Neuroscience*. Cambridge: The MIT Press.
- ROSEMBLATT, F. (1962). *Principles of Neurodynamics*. New York: Spartan.
- RUMELHART, D.E.; HINTON, G.E. y MCCLELLAND, J.L. (1986). A General Framework for Parallel Distributed Processing. En J. L. McClelland y D. E. Rumelhart (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol 1. Págs. 45-76. Cambridge: The MIT Press.
- RUMELHART, D.E.; HINTON, G.E. y WILLIAMS, R.J. (1986). Learning Internal Representations by Error Propagation. En J. L. McClelland y D. E. Rumelhart (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol 2. Págs. 319-362. Cambridge: The MIT Press.