Artificial Neural Networks for the Analysis of Text across Cultures and Languages

Joe K. Wölfel

MERL

Hao Chen
Jang Hyun Kim
Monica Murero
Bhawna Sharma
Joseph Woelfel

University at Buffalo, State University of New York

Raymond Hsieh

California University at Pennsylvania

October 30, 2005

Abstract

Raw sensory data impinge on individual brains as a stream of ever changing impulses, but language transforms this inchoate stream into a coherent framework of discrete objects. Each language performs this "chunking" function in its own way.

Research on neural networks, real and artificial, cast new light on the interface between the incoherent continuous flow of sensory data and the coherent, categorical model formed by individuals and cultures. Artificial neural networks such as Catpac™ (Woelfel, 1993), made it possible to project textual data onto Galileo coordinates, which in turn made it possible to compare texts quantitatively across samples and time (Hao, 2004; Hsieh, 2004). Using procedures developed by Woelfel, et. al.,(1988); Hao, (2004); Hsieh, (2004),we project text in several languages onto Galileo coordinates and calculate precise quantitative comparisons among the underlying concepts in the texts as they are encoded into the various languages. So far, research has shown that the neural model can be used even across different languages. So far, however, such research has been restricted to languages that can be encoded in ASCII text.

In the present article, we describe Wölfpak, a Unicode based artificial neural network for the analysis of text in any language (Woelfel, et. al, 2005). We present examples in English, German, French, Hindi, Chinese and Korean. Although the current prototype version of Wölfpak does not allow the production of Galileo coordinates, this is a straightforward modification which will be made available in updated versions.

Beta versions of the Wölfpak prototype for Windows and Macintosh are available at jwoelfel@galileoco.com.

References

Foldy, J., and J. Woelfel, (1992). "Cognitive Processes as Damped Harmonic Oscillators, *Quality and Quantity*, **26**,

Pearson, E. S., (editor) (1948) *Karl Pearson's Early Statistical Papers*, Cambridge: Cambridge University Press.

Stevens, S.S. (1939). Psychology and the science of science. *Psychological Bulletin*, *36*, 221-263.

Torgerson, Warren S.,"*Multidimensional scaling: I. theory and method*," Psychometrika, vol. 17, no. 4, pp. 401--419, Dec 1952.

Woelfel, J., "Artificial neural networks in policy research: A current assessment," Journal of Communication, 43(1), 63-80. 1993.

Woelfel, J., R. Holmes and D. L. Kincaid, "Rotation to congruence for general riemann surfaces under theoretical constraints," in G. A.   Barnett and J. Woelfel, Readings in the Galileo System: Theory, Methods and Applications, Dubuque, IA, Kendall-Hunt, 1988. (pp. 219-234)

Woelfel, J., G.A. Barnett and R. Pruzek, "Rotation to simple processes: The effect of alternative rotation rules on observed patterns in time-ordered measurements," *Quality and Quantity*, 23:3-20 (1989).

Woelfel, J., and G. A. Barnett (1992) "Procedures for controlling reference frame effects in the measurement of multidimensional processes," *Quality and Quantity* **26:** 367-381.

Woelfel, Joe K.,Raymond Hsieh, Hao Chen, Jennie Hwang, Pauline Hope Cheong, Devon Rosen, Joseph Woelfel, "Wölfpak™: A Neural Network for Multilingual Text Analysis", Paper presented at the 25th annual meetings of the International Society for Network Analysis, Los Angeles, CA, February 16-20, 2005