Erik Cambria
*Temasek Laboratories, National University of Singapore*

Bebo White
*SLAC National Accelerator Laboratory, Stanford University*

Tariq S. Durrani
*Royal Society of Edinburgh*

Newton Howard
*Media Laboratory, Massachusetts Institute of Technology*

Guest Editorial

# Computational Intelligence for Natural Language Processing

The advent of the Social Web has provided people with new content-sharing services that allow them to create and share their own contents, ideas, and opinions, in a time- and cost-efficient way, with virtually millions of other people connected to the World Wide Web. This huge amount of information, however, is mainly unstructured (because it is specifically produced for human consumption) and hence not directly machine-processible. The automatic analysis of text involves a deep understanding of natural language by machines, a reality from which we are still very far off.

Hitherto, online information retrieval, aggregation, and processing have mainly been based on algorithms relying on the textual representation of web pages. Such algorithms are very good at retrieving texts, splitting them into parts, checking the spelling and counting the number of words. But when it comes to interpreting sentences and extracting meaningful information, their capabilities are known to be very limited. Word-based algorithms are limited by the fact that they can process only the information that they can 'see'. As human text processors, we do not have such limitations as every word we see activates a cascade of semantically related concepts, relevant episodes, and sensory experiences, all of which enable the completion of complex tasks (such as word-sense disambiguation, textual entailment, and semantic role labeling) in a quick and effortless way.

Computational models attempt to bridge such a cognitive gap by emulating the way the human brain processes natural language, e.g., by leveraging on semantic features that are not explicitly expressed in text. Computational models are useful both for scientific purposes (such as exploring the nature of linguistic communication), as well as for practical purposes (such as enabling effective human-machine communication).

Traditional research disciplines do not have the tools to completely address the problem of how language comprehension and production work. Even if you combine all the approaches, a comprehensive theory would be too complex to be studied using traditional methods. But we may be able to realize such complex theories as computer programs and then test them by observing how well they perform. By seeing where they fail, we can incrementally improve them. Computational models may provide very specific predictions about human behavior that can then be explored by the psycholinguist. By continuing in this process, we may eventually acquire a deep understanding of how human language processing occurs. To realize such a dream will take the combined efforts of forward-thinking psycholinguists, neuroscientists, anthropologists, philosophers, and computer scientists.

In this light, this Special Issue focuses on the introduction, presentation, and discussion of novel semantic approaches to natural language processing. The main motivation for the Special Issue, in particular, is to go beyond a mere word-level analysis of text and propose new computational-intelligence approaches that allow a more efficient passage from (unstructured) textual information to (structured) machine-processible data.

For this Special Issue, we received 43 valid submissions, of which only 18 were short-listed and eventually only 2 accepted for publication. The first article, "Frame-Based Detection of Opinion Holders and Topics: A Model and a Tool", by Aldo Gangemi, Valentina Presutti, and Diego Reforgiato, presents Sentilo, a model and a tool to perform holder and topic detection in opinion sentences. Sentilo implements an approach based on the neo-Davidsonian assumption that events/situations are the primary entities for contextualizing sentiment, which makes it able to distinguish holders, main topics, and sub-topics of an opinion. It uses a heuristic graph-mining approach that relies on FRED, a machine reader for the Semantic Web leveraging natural-language-processing and knowledge-representation components jointly with cognitively-inspired frames.

The second article, "A Probabilistic Generative Model for Mining Cybercriminal Networks from Online Social Media", by Raymond Lau, Yunqing Xia, and Yunming Ye, proposes the design, development, and evaluation of a novel

*(continued on page 63)*

*Digital Object Identifier 10.1109/MCI.2013.2291686*
*Date of publication: 14 January 2014*

## References


[1] G. Stegmayer, M. Gerard, and D. H. Milone, "Data mining over biological datasets: An integrated approach based on computational intelligence," *IEEE Comput. Intell. Mag.*, vol. 7, no. 4, pp. 22–34, Nov. 2012.

[2] O. G. Martinsen, S. Clausen, J. B. Nysaether, and S. Grimnes, "Utilizing characteristic electrical properties of the epidermal skin layers to detect fake fingers in biometric fingerprint systems—A pilot study," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 5, pp. 891–894, May 2007.

[3] F. Scotti and V. Piuri, "Adaptive reflection detection and location in iris biometric images by using computational intelligence techniques," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 7, pp. 1825–1833, July 2010.

[4] Q. Tao and R. Veldhuis, "Biometric authentication system on mobile personal device," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 4, pp. 763–773, Apr. 2010.

[5] M. G. Tsipouras, D. I. Fotiadis, and D. Sideris, "Arrhythmia classification using the RR-interval duration signal," *Comput. Cardiol.*, vol. 29, pp. 485–488, Sept. 2002.

[6] P. de Chazal, B. G. Celler, and R. B. Reilly, "Using wavelet coefficients for classification of the electrocardiogram," in *Proc. 22nd Annu. Int. Conf. IEEE Engineering Medicine Biology Society*, 2000, pp. 64–67.

[7] R. Almeida, S. Gouveia, A. P. Rocha, E. Pueyo, J. P. Martinez, and P. Laguna, "QT variability and HRV interactions in ECG: Quantification and reliability," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 7, pp. 1317–1329, July 2006.

[8] A. Kikuchi, T. Shimizu, A. Hayashi, T. Horikoshi, N. Unno, S. Kozuma, and Y. Taketani, "Nonlinear analyses of heart rate variability in normal and growth-restricted fetuses," *Early Human Develop.*, vol. 82, no. 4, pp. 217–226, Apr. 2006.

[9] D. Hang, S. Crozier, and S. Wilson, "A new heart rate variability analsis method by means of quantifying the variation of nonlinear dynamic patterns," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 9, pp. 1590–1597, Sept. 2007.

[10] A. Babloyantz and A. Destexhe, "Is the normal heart a periodic oscillator?," *Biol. Cybern.*, vol. 58, no. 3, pp. 203–211, 1988.

[11] O. Fojt and J. Holick, "Applying nonlinear dynamic to ECG signal processing," *IEEE Eng. Med. Biol. Mag.*, vol. 17, no. 2, pp. 96–101, Apr. 1998.

[12] S. Lee, T. L. Kang, H. Y. Quan, and X. Tian, "Analysis of HRV based on correlation dimension and largest Lyapunov exponent," *J. Biomed. Eng. Res.*, vol. 28, pp. 188–192, 2009.

[13] K. T. Lai and K. L. Chant, "Real-time classification of electrocardiogram base on fractal and correlation analyses," in *Proc. 20th Annu. Int. Conf. IEEE Engineering Medicine Biology Society*, 1998, pp. 119–122.

[14] C. Raab, J. Kurths, A. Schirdewan, and N. Wessel, "Normalized correlation dimension for heart rate variability analysis," *Biomedizinische Technik*, vol. 51, no. 4, pp. 229–232, Oct. 2006.

[15] B. Anuradha and V. C. Veera Reddy, "ANN for classification of cardiac arrhythmia," *ARPN J. Eng. Applied Sci.*, vol. 3, no. 3, pp. 1–6, June 2008.

[16] R. R. Sarvestani, R. Boostani, and M. Roopaei, "VT and VF classification using trajectory analysis," *Nonlinear Analysis: Theory, Methods Applicat.*, vol. 71, no. 12, pp. e55–e61, Dec. 2009.

[17] M. Roopaei, R. Boostani, R. R. Sarvestani, M. A. Taghavi, and Z. Azimifar, "Chaotic based reconstructed phase space features for detecting ventricular fibrillation," *Biomed. Signal Process. Control*, vol. 5, no. 4, pp. 318–327, Oct. 2010.

[18] A. S. Al-Fahoum and A. M. Qasaimeh, "ECG arrhythmia classification using simple reconstructed phase space approach," *Comput. Cardiol.*, vol. 33, pp. 757–760, Sept. 2006.

[19] Q. Jiao, Y. X. Guo, W. F. Cao, and Z. G. Zhang, "Short time ECG signal anaylsis based on the reconstruction of phase space," *Chinese J. Med. Instrum.*, vol. 32, no. 4, pp. 257–264, July 2008.

[20] H. I. Eisen, "Surgical ventricular reconstruction for heart failure," *New England J. Med.*, vol. 360, pp. 1781–1784, Apr. 2009.

[21] S. K. Singla and A. Sharma, "ECG as biometric in the automated world," *Int. J. Comput. Sci. Commun.*, vol. 1, no. 2, pp. 281–283, 2010.

[22] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG analysis: A new approach in human identification," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 3, pp. 808–812, June 2001.

[23] T. W. Shen, W. J. Tompkins, and Y. H. Hu, "One-lead ECG for identity verification," in *Proc. Second Joint 24th Annu. Conf. Engineering Medicine Biology/Annu. Fall Meeting Biomedical Engineering Society (EMBS/BMES)*, Texas, 2002, pp. 62–63.

[24] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold, "ECG to identify individuals," *Pattern Recognit.*, vol. 38, no. 1, pp. 133–142, Jan. 2005.

[25] H. Silva, H. Gamboa, and A. Fred, "One lead ECG based personal identification with feature subspace ensembles," in *Proc. Int. Conf. Machine Learning Data Mining Pattern Recognition*, Leipzig, Germany, 2007, pp. 770–783.

[26] F. Agrafioti and D. Hatzinakos, "ECG based recognition using second order statistics," in *Proc. Communication Networks Services Research Conf.*, 2008, pp. 82–87.

[27] Y. Wang, F. Agrafioti, D. Hatzinakos, and K. N. Plataniotis, "Analysis of human electrocardiogram for biometric recognition," *EURASIP J. Adv. Signal Process.*, 2008.

[28] A. D. C. Chan, M. M. Hamdy, A. Badre, and V. Badee, "Wavelet distance measure for person identification using electrocardiograms," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 2, pp. 248–253, Feb. 2008.

[29] I. Khalil and F. Sufi, "Legendre polynomials based biometric authentication using QRS complex of ECG," in *Proc. Int. Conf. Intelligent Sensors, Sensor Networks Information Processing*, 2008, pp. 297–302.

[30] S. Z. Fatemian and D. Hatzinakos, "A new ECG feature extractor for biometrics recognition," in *Proc. Int. Conf. Digital Signal Processing*, 2009, pp. 1–6.

[31] C. C. Chiu, C. M. Chuang, and C. Y. Hsu, "A novel personal identity verification approach using a discrete wavelet transform of the ECG signal," *Int. J. Wavelets, Multiresolution Inform. Process*, vol. 7, pp. 341–355, 2009.

[32] J. L. C. Loong, K. S. Subari, R. Besar, and M. K. Abdullah, "A new approach to ECG biometric systems: A comparative study between LPC and WPD systems," *World Acad. Sci., Eng. Technol.*, vol. 68, no. 44, pp. 759–764, 2010.

[33] D. P. Coutinho, A. L. N. Fred, and A. T. Figueiredo, "One-lead ECG-based personal identification using Ziv-Merhav cross parsing," in *Proc. Int. Conf. Pattern Recognition*, 2010, pp. 3858–3861.

[34] [Online]. Available: http://bcclib.ee.nchu.edu.tw/wp-content/uploads/2010/11/12-Lead_ECG_System.pdf

[35] W. Einthoven, G. Fahr, and A. de Waart, "Über die richtung und die manifeste grösse der potentialschwankungen im mennschlichen herzen und über den einfluss der herzlage auf die form des elektrokardiogramms," *Pflügers Arch Ges Physiol.*, vol. 150, nos. 6–8, pp. 275–315, 1913.

[36] W. Einthoven, G. Fahr, and A. de Waart, "On the direction and manifest size of the variations of potential in the human heart and on the influence of the position of the heart on the form of the electrocardiogram," *Amer. Heart J.*, vol. 40, no. 2, pp. 163–211, Aug. 1950.

[37] C. T. Chiang, "Contact type pulse measurement device," U.S. Patent 6 945 940 B1, 2005.

[38] F. Takens, "Detecting strange attractors in turbulence," in *Lecture Notes in Mathematics*, vol. 898, 1989, pp. 245–262.

[39] J. P. Huke, "Embedding nonlinear dynamical systems: A guide to Takens' theorem," Manchester Inst. Math. Sci., Univ. Manchester, Manchester, U.K., Tech. Rep., 2006.

[40] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica A*, vol. 9, no. 1–2, pp. 189–208, Oct. 1983.

[41] M. Sano and Y. Sawada, "Measurement of the Lyapunov spectrum from a chaotic time series," *Phys. Rev. Lett.*, vol. 55, no. 10, pp. 1082–1085, Sept. 1985.

[42] [Online]. Available: http://bcclib.ee.nchu.edu.tw/wp-content/uploads/file/The%20Confusion%20Matrix%20and%20the%20Threshold%20of%20Outlier.pdf

[43] M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural Network Design*. Boston, MA: PWS Publishing Co., 1995.

[44] C. Y. Chang, Y. C. Hong, P. C. Chung, and C. H. Tseng, "A neural network for thyroid segmentation and volume estimation in CT images," *IEEE Comput. Intell. Mag.*, vol. 6, no. 4, pp. 43–55, Nov. 2011.

[45] D. Guo and Y. Zhang "Novel recurrent neural network for time-varying problems solving," *IEEE Comput. Intell. Mag.*, vol. 7, no. 4, pp. 61–65, Nov. 2012.

[46] R. R. Hashemi, L. A. Blanc, C. T. Rucks, and A. Rajaratnam, "A hybrid intelligent system for predicting bank holding structures," *European J. Operational Res.*, vol. 109, pp. 390–402, Sept. 1998.


## Guest Editorial

*(continued from page 19)*

weakly-supervised cybercriminal network mining method to facilitate cybercrime forensics. There has been a rapid growth in the number of cybercrimes that cause tremendous financial loss to organizations. Recent studies reveal that cybercriminals tend to collaborate or even transact cyberattack tools via the 'dark markets' established in online social media. The method proposed by the authors is underpinned by a probabilistic generative model enhanced by a novel context-sensitive Gibbs sampling algorithm. Evaluated based on two social media corpora, their experimental results reveal that the proposed method significantly outperforms the best supervised baseline method.

These two articles are a solid and varied representation of some of the exciting challenges and solutions emerging in this field. We hope that you enjoy the special issue and that this research fosters future innovations. Finally, we would like to thank the current and previous editors in chief, namely Hisao Ishibuchi and Kay Chen Tan, for their precious help and guidance in laying out the contents of this Special Issue and the 40 reviewers that not only helped with the decision process but contributed with excellent reviews to make this issue special.