

Affective Computing: Sentiment, Metaphor and Terminology¹

Khurshid Ahmad,
Trinity College,
Dublin, IRELAND.

Abstract

Affective computing involves understanding and representing how subjective experience is articulated, and how this articulation impacts on others especially in the critical areas of finance, commerce, well-being and security. There is a considerable pride in affixing oneself as rational but that leaves little room for exuberant, contrarian, or *irrational* behavior and beings. I will look very briefly on limits of rationality, some of the developments in computer-based affect analysis and finish this paper by letting the reader know about my modest contribution to affective computing.

Bounded Rationality?

There is evidence of *bounded rationality* ([41]) in reactions to the news about financial markets appears to change the (numerical) prices and volumes of assets (e.g. shares, bonds, commodities) traded in the markets. Traders, for instance, keep on buying ever larger volumes of assets even when there is incontrovertible evidence that the prices are falling. Contrariwise, traders keep on selling ever larger volumes of assets when the market is rising. It has been suggested that there are limits to rational behavior at certain times, especially during the onset of a crisis. The rationalist approach is to discount the news altogether and focus on prices/volume: the problem here is that human beings do not always make rational choices ([42][40]); the stakeholders in a financial market, especially at the onset of boom or bust in the market, behave in an *exuberant* manner [39]. Humans have a propensity to choose radically different solutions to the same problem if the problem is expressed or framed differently ([30][31]). The images collateral to economic/financial news, for example, strikes and civil unrest, facial gestures of regulators, traders and investors, plays a key role in ‘framing’ news and blogs. Expert traders and regulators usually make judicious choices in aggregating linguistic, numerical and gestural

¹ Khurshid Ahmad (2011). Invited Keynote Speech: Affective Computing: Sentiment, Metaphor and Terminology. *First International Workshop on Semantic and Collaborative Technologies for the Web, Bucharest, Romania* (EU Project ERRIC, FP7-REGPOT-2010-1/264207). June 2011.

information. It has been argued that ‘early’ financial decision making involves fast processes akin to mental arithmetic ([47]) and visual enumeration [17]). The mood and emotive state of traders plays an important role in financial trading. The affect content of the stakeholders in the market appears to have an impact on prices and volumes of asset traded.

Bounded rationality appears to be a characteristic feature when we look at almost any walk of human endeavour where an exchange takes place: money exchanged for goods and services, politicians writing manifestos pleading for votes [35] and voters’ ‘emotive’ response [36], description of ‘other’ people [2] and so on. In any of these exchange processes we see contrarian behavior, impact of framing, legacy effects and plain bias/prejudice. In such exchanges one has to take into the mood and emotive state of the stakeholders. Thus, we see that the advertisement and marketing world is keen to quantify the affect content of the reaction of the customer/voter/tourist to goods and services on offer [44] .

The metaphorical use of language plays an important role in how we articulate our subjective experience especially our moods and emotions. An integrated study of how metaphors and emotive language, indicating the strength of emotion together with its polarity and action-orientation, are used in describing subjective experience on the one hand, and the relationship (if any) of the experience with observable facts, is critical for the development of affective computing and sentiment analysis [3].

Computational analysis of exuberant markets and the impact of news

‘Sentiment analysis is the task of identifying positive and negative opinions, emotions, and evaluations’ ([49]). Keynes had suggested that there are “animal spirits” in the market ([32]), and the contrariness and obfuscation of traders and regulators is caused by the spirits leading to *irrational exuberance* ([39]). The impact of news on financial markets can be substantial and, according to the Nobel Laureate Robert Engle ([23]), asymmetric: the arrival of ‘bad news’ has a longer lasting effect on prices, and particularly on volumes of shares traded, when compared to the arrival of ‘good news’. Economists typically use proxies for the news content – change in the values of currencies, bonds, or aggregated indices like the Dow Jones, NASDAQ.

News was traditionally interpreted as a causal variable in financial market models: readily quantifiable aspects of news as a proxy for the news itself and includes the timing of news arrival, the volume of news items and the type of news ([14]). News proxies were used with some effect to show that ‘negative’ news has much longer lasting impact than the positive news ([24][23]). This kind of sentiment analysis is almost always conducted post-hoc ([15][18]) but there are some exceptions where the researchers have attempted to predict the market ([33]). News content analysis, rather than the use of proxies, is becoming more important recently. Continual records of market analysts’ opinions on commercial news channels shows that positive news has short term (1 minute) impact on prices but the negative news impacts for 15 minutes

([16]). It has been argued that the use of optimistic language in press releases of a firm appears to increase the firm's future earnings, whilst pessimistic news has the opposite effect ([20]). Tetlock has shown that the negative affective component of news reports does have a longer lasting impact by analyzing opinion columns in financial news papers ([46][45]).

Dictionaries of Affect

One of the pioneers of political theory and communications in the early 20th century, Harold Lasswell ([34]), has used sentiment to convey the idea of an attitude permeated by feeling rather than the undirected feeling itself. This approach to analysing contents of political and economic documents – called content analysis – was given considerable fillip in the 1950's and 1960's by Philip Stone of Harvard University who created the so-called General Inquirer System ([43]) and a large digitised dictionary – the *GI Dictionary* also known as the *Harvard Dictionary of Affect*.

There are two conventional methods of creating affect dictionaries: First is to use a dictionary like the Harvard University's General Inquirer (GI) lexicon ([43]) and to rely on the judgment of the dictionary makers. We have looked at the structure and content of GI dictionary and compared it with other dictionaries of affect, including derivatives of the WordNet: we found that the whilst the major dictionaries of affect share less than 50% of the vocabulary, yet the sentiment computation is only marginally effected by the choice of dictionaries ([22][21]). Second, affect dictionaries can be created from a collection of texts –the so-called local grammar approach, which was pioneered by Zellig Harris [28] and his associates especially Maurice Gross [27], was adopted by us in the development of a sentiment extraction system: the system extracted statistically significant collocational patterns of candidate terms in a corpus of specialist texts ([6])– these patterns invariably comprise metaphorical words relating to change and affect ([4][12]) in close proximity with names of key objects, e.g. shares, currencies, and the names of key stakeholders/enterprises ([48]). The two methods are complementary, and the corpus based method can be used adaptively in order to cope with new words and changing sentiment lexicon.

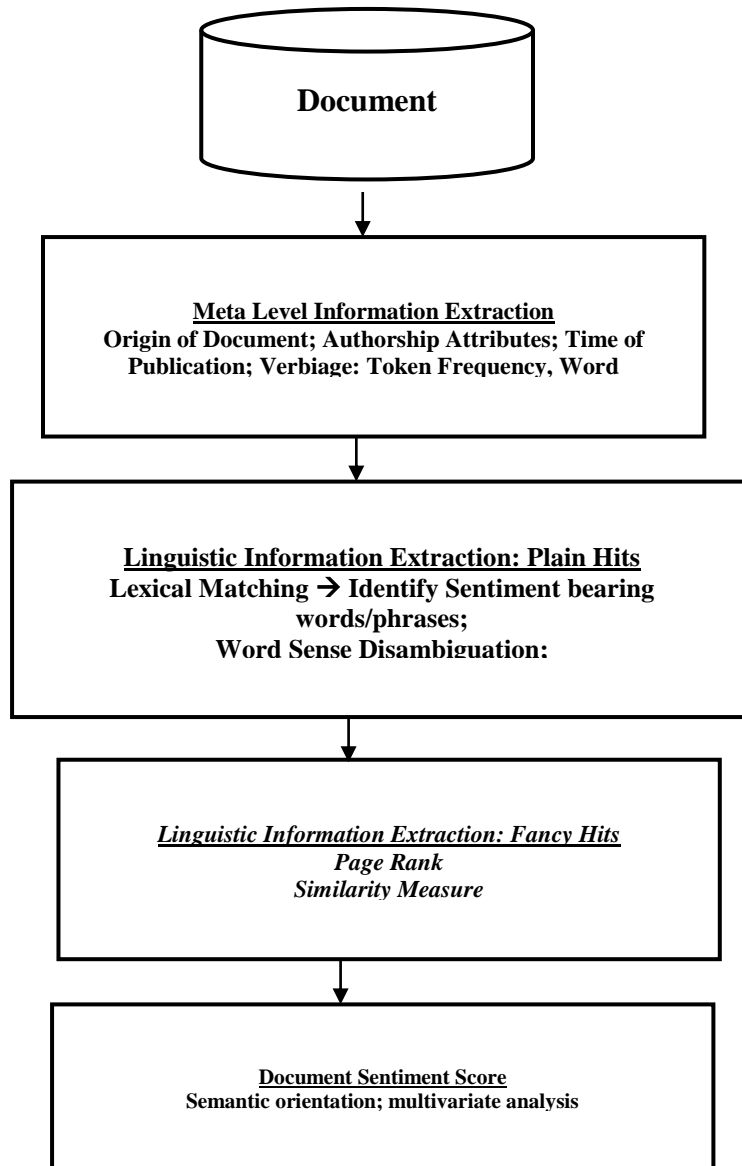
Machine learning algorithms are increasingly being used to identify sentiment bearing patterns automatically: typically, supervised learning algorithms are used, one uses a training corpus of texts comprising pre-tagged sentiment phrases and an external or exogenous variable, for instance prices of financial assets, and the algorithm learns to associate sentiment bearing phrases with changes in the direction of prices ([33],[25]).

Affect Analysis Systems

Typical affect analysis systems, especially those available commercially by major news vendors effectively manage a document stream. Each document in the stream is subjected to three different types of information extraction which leads to the

computation of a *sentiment score* usually related to various statistics based on the distribution of polarity of sentiment-bearing phrases in texts (see [1], [38] for details).

First, meta-level or document external information is extracted which may comprise time of publication, authorship attribution data, length of the document, originating and target sources. Second, information is extracted at least two levels of linguistic description – (i) lexical level where sentiment bearing words/phrases are extracted by using an affect dictionary and (ii) syntactic level of linguistic description by looking at syntactic categories of words/phrases in the document primarily for disambiguation purposes. Third, phase of linguistic information extraction relates to the authoritativeness of the document using Page Ranking and similarity measures. Finally, the system outputs a sentiment score which may be a sentiment word/phrase count or other a weighted average of the count (See figure below)



My contributions to affective computing

I have been working on sentiment analysis since 1997 and have focused on developing methods and techniques that will help in the extraction of affect – sentiment is a form of affect- from a collection of texts. The methods focused on the relationship between domain specific terms and affect words. The initial area of application was financial markets: The integration of sentiment index, extracted from full-text newswires in English, and stock exchange indices was reported by us ([10], [11]). Subsequently, the use of computational grids and cloud computing was demonstrated and evaluated for the purposes of automatic extraction of sentiment words and technical terms from a corpus of financial news (c. 800,000 newswires) ([7] [4] [26]).

The method developed by my colleagues was demonstrated to work on financial newswires in Arabic [12], Chinese ([9], [29]), and Urdu [13]. The work now embraces informal texts (blogs for example) [37], news reports about ideologically motivated groups [2]. Working in collaboration with the Trinity Business School, we were able to adapt the econometric notions of return– the rate of change of the price/volume of an asset- and volatility –the variance of the return- for financial and ideological sentiment analysis ([5]). Results have recently been published about the nature of the distribution of sentiment returns in German [37], North American, Japanese markets of the Irish markets [19].

I am currently studying how metaphorical language is conceived and used, how such complex psycho-linguistic constructs can be represented on computer systems and in different languages, and how the practice of affect analysis in fields as diverse as financial trading, brand management, product reviews. This study may help in formulating research questions in this critical area of how we the humans are stimulated by and respond to subjective judgments. Papers on this broad topic of affective computing have recently published as a collection comprising contributions from scholars metaphor and experimental psychology, artificial intelligence, computational linguistics and information extraction, translation studies, brand management and financial trading [3].

References

- [1] Abbassi, A., Chen, H., & Salem, A. (2008). Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums. *ACM Transactions on Information Systems*. Vol 26 (No.3). (Article 12).
- [2] Ahmad, K. (2008) Edderkoppspinn eller nettverk: News media and the use of polar words in emotive contexts. , *Synaps*, 21, 2008, p20 - 36
- [3] Ahmad, K. (2011) (Ed.), *Affective Computing and Sentiment Analysis: Metaphors, Emotions and Terminology*. Heidelberg, Springer Verlag, 200 pp.
- [4] Ahmad, K. & L. Gillam and D.Cheng. 2005 ‘Textual and Quantitative Analysis: Towards a new, e-mediated Social Science’. Proc. of the 1st Int. Conf. on e-Social Science (Manchester, July 2005).

- [5] Ahmad, K. (2008) The 'return' and 'volatility' of sentiments: An attempt to quantify the behaviour of the markets?. In Proc. of EMOT 2008: Sentiment Analysis: Emotion, Metaphor, Ontology and Terminology. Workshop (13th Language Resources and Evaluation Conf., 27 May 2008. Marrakesh, Morocco). (to appear in K. Ahmad (Ed.) 2011)
- [6] Ahmad, K. *Pragmatics of Specialist Terms and Terminology Management*. In (Ed.) Petra Steffens. Machine Translation and the Lexicon. 3rd Int. EAMT Workshop, Heidelberg, Germany, April 26-28, 1993.) Heidelberg (Germany): Springer. pp.51-76. (Lecture Notes on Artificial Intelligence, Vol. 898, Eds. J. G Carbonell & J. Siekmann).
- [7] Ahmad, K., & L. Gillam and D. Cheng. 2005: 'Society Grids'. In (Eds.) Simon Cox and David Walker. *Proc. UK e-Science All Hands Meeting 2005*. Swindon: EPSRC Sept 2005. pp 923-930.
- [8] Ahmad, K., Bale, T., & Casey, M. Connectionist Simulation of Quantification Skills. *Connection Science* Vol. 14 (No. 3). pp 165-201.
- [9] Ahmad, K., D. Cheng and Y. Almas, (2006) 'Multi-lingual Sentiment Analysis of Financial News Streams.' In (Eds.) Stefano Cozzini et al. Proc. 1st Int. Conf. on Grid in Finance (Palermo, February 2006) (http://pos.sissa.it/archive/conferences/026/001/GRID2006_001.pdf)
- [10] Ahmad, K., D. Cheng, T. Taskaya, S. Ahmad, L. Gillam, P. Manomaisupat, H. Traboulsi and A. Hippisley. (2006) 'The mood of the (financial) markets: In a corpus of words and of pictures.' In (Eds.) A. Wilson, D. Archer and P. Rayson. *Corpus linguistics around the world*. Amsterdam/New York: Rodopi, pp 18-32.
- [11] Ahmad, K., Taskaya Temizel, T., Cheng, D., Ahmad, S., Gillam, L., Manomaisupat, P., Traboulsi, H. and Casey, M. 2004 "Fundamental Data to Satisfy the Chartist." *The Technical Analyst*, vol. 1(3), pp. 36-39.
- [12] Almas, Y., & K. Ahmad. (2006) 'LoLo: A System based on Terminology for Multilingual Information Extraction'. In (Eds.) Mary Elaine Califf et al. *COLING ACL 2006: Workshop on Information Extraction beyond the Document*, Sydney, Australia. Association of Computational Linguistics. pp56 – 65. (<http://acl.ldc.upenn.edu/W/W06/W06-0207.pdf>)
- [13] Almas, Y., and K. Ahmad, (2007) A note on extracting 'sentiments' in financial news in English, Arabic & Urdu, The Second Workshop on Computational Approaches to Arabic Script-based Languages, Linguistic Society of America 2007 Linguistic Institute, Stanford University, Stanford, California., July 21-22, 2007, Linguistic Society of America, 2007, pp1 - 12
- [14] Antweiler, W., & Frank, M. Z. (2004) Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance*, pp 1259-1294, 2004.
- [15] Bauwens, L.; W.B. Omrane; and P. Giot. (2005). "News Announcements, Market Activity and Volatility in the Euro/Dollar Foreign Exchange Market." *Journal of International Money and Finance*. Vol. 24, pp 1108-1125.
- [16] Busse, J. and T. Clifton Green. Market efficiency in real-time. *Journal of Financial Economics*, Vol. 65, pp 437, 2002.
- [17] Casey, M. & Ahmad, K. 'A competitive neural model of small number detection'. *Neural Networks* Vol. 19 (No. 10), pp 1475 - 1489

- [18] Chang, Y., and Taylor, S.J. (2003) Information Arrivals and Intraday Exchange Rate Volatility. *J. Int. Financial Markets, Institutions and Money*, vol 13, pp85-112.
- [19] Daly, N., K. Ahmad, and C.Kearney. (2009) Correlating Market movements with consumer confidence and sentiments: A longitudinal study. *Proc. of Text Mining Workshop*, Leipzig, Germany.
- [20] Davis, A.K., Piger, J.M., Sedor, L.M.: *Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases*. Technical Report, Federal Reserve, Bank of St Louis, (2006)
- [21] Devitt, Ann and Khurshid Ahmad, Sentiment Analysis and the Use of Extrinsic Datasets in Evaluation, *Language Resources and Evaluation Conference (LREC 2008)*, Marrakech, Morocco, 28-30 May, 2008
- [22] Devitt, Ann., and Ahmad, Khurshid. 'Sentiment Polarity Identification in Financial News: A Cohesion-based Approach'. *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics. June 23–30, 2007, Prague, Czech Republic*. Stroudsburg, PA: Association for Computational Linguistics (ACL) pp 984-991
- [23] Engle, R.F (2003) Risk and Volatility: Econometric Models and Financial Practice," Nobel Prize in Economics documents 2003-4, Nobel Prize Committee, http://nobelprize.org/nobel_prizes/economics/laureates/2003/engle-lecture.pdf
- [24] Engle, R.F., & V. K. Ng, 1991. Measuring and Testing the Impact of News on Volatility, NBER Working Papers 3681, National Bureau of Economic Research, Inc.
- [25] Généreux, M.,Thierry Poibeau and Moshe Koppe Sentiment analysis using automatically labelled financial news. In (Ed). K. Ahmad (2011)
- [26] Gillam, L., K. Ahmad, G. Dear. 2005 : 'Grid-enabling Social Scientists: The FINGRID infrastructure.' *Proc. of the 1st Int. Conf. on e-Social Science* (Manchester, July 2005).
- [27] Gross, Maurice (1997). "The construction of local grammars," In (Eds.) E. Roche and Y. Schabès, *Finite-State Language Processing, Language, Speech, and Communication*, Cambridge, Mass.: MIT Press, 1997 , pp. 329-354.
- [28] Harris, Zellig. (1991) A Theory of Language and In-formation: A Mathematical Approach. Oxford: Clarendon Press.
- [29] Hippiseley, A., Cheng, D., & Ahmad, K. (2005) Head-Modifier Principle in Chinese. *Natural Language Engineering* Vol. 11 (2), pp 129-157.
- [30] Kahneman, D & Tversky, A. (1979) Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, Vol. 47 (No. 2) (Mar., 1979), pp. 263-292
- [31] Kahneman, D. A. (2002). Maps of Bounded Rationality: The [2002] Sveriges Riksbank Prize [Lecture] in Economic Sciences. (http://nobelprize.org/nobel_prizes/economics/laureates/2002/kahnemann-lecture.pdf visited 14 April 2009).
- [32] Keynes, John Maynard (1936) *The General Theory of Employment, Interest and Money*. London: Macmillan (reprinted 2007)
- [33] Koppel, Moshe, and Shtrimberg, Itai. (2004), "Good News or Bad News? Let the Market Decide" In *AAAI Spring Symposium on Exploring Attitude and Affect in Text* , Palo Alto CA, pp. 86-88.
- [34] Lasswell, Harold D. (1948). *Power and personality*. London:Chapman & Hall.

- [35] Namenwirth, Zvi., and Lasswell, Harold D. (1970) *The changing language of American values : a computer study of selected party platforms*. Beverly Hills (Calif.) :Sage Publications.
- [36] Neuman, W.R., Marcus, G.E., Crigler, A.N., and Mackuen, M. (2007). (Eds.) *The Affect Effect: Dynamics of emotion in political thinking and behaviour*. Chicago & London: The University of Chicago Press.
- [37] Remus. R., K. Ahmad, G. Heyer. (2009) Integrating Indicators: Data Mining [German] Financial News, Blogs, Market Movement and Consumer Surveys for Sentiment. *Proc. of Text Mining Workshop*, Leipzig, Germany.
- [38] Shanahan, J.G., Qu, Y., & Weibe. J. (2006). (Eds.) *Computing Attitude and Affect in Text: Theory and applications*. Dordrecht: Springer.
- [39] Shiller, R. (2005). *Irrational Exuberance* (2nd Edition). Princeton: Princeton University Press, 2005.
- [40] Simon, H. A. (1978). *Rational Decision-Making in Business Organization: The [1978] Sveriges Riksbank Prize [Lecture] in Economic Sciences* (http://nobelprize.org/nobel_prizes/economics/laureates/1978/simon-lecture.pdf, site visited 14 April 2009).
- [41] Simon, H. A., (1955). A behavioral model of rational choice, *Quarterly Journal of Economics*, Vol. 69, pp 99-118 (1955).
- [42] Simon, H. A., (1956) Rational choice and the structure of the environment, *Psychological Review*, Vol 63: 129-138 (1956). pp 241-268
- [43] Stone, Philip,J., Dunphy, D.C., Smith, Marshall, S.S., Ogilvie and associates. (1966). *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge/London: The MIT Press.
- [44] Teichert,T., Gerhard Heyer, Katja Schöntag and Patrick Mairif. Co-Word Analysis for Assessing Consumer Associations: A Case Study in Market Research. In (Ed.) Ahmad (2011)
- [45] Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008), More Than Words: Quantifying Language to Measure Firms' Fundamentals. *Journal of Finance*; Jun2008, Vol. 63 (No. 3), pp 1437-1467
- [46] Tetlock, Paul C. (2007). Giving Content to Investor Sentiment: The Role of Media in the StockMarket. *Journal of Finance*. June 2007, Vol. 62 (No. 3), pp 1139-1168
- [47] Thaler, Richard (2000). 'Mental Accounting Matters. In (eds.) Daniel Kahneman and Amos Trevisky., *Choices, Values, and Frames*. Cambridge, New York: Cambridge University Press.,
- [48] Traboulsi, Hayssam., & Khurshid Ahmad, Automatic Construction of a Thesaurus of Proper Names : A Local Grammar-based Approach, *8th Intext Nooj Workshop 2005 - Formaliser les langues avec l'ordinateur: De INTEX à Nooj, Besançon (France), May 30-June 1, 2005*, edited by Svetla Koeva, Denis Maurel & Max Silberstein , Presses universitaires de Franche-Comté, 2007, pp261 – 280
- [49] Wilson, T., Wiebe, J., & Hoffmann, P. (2005). "Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis." *Proc. Conf. on Human Language Technology and Empirical Methods in Natural Language Processing* Vancouver, British Columbia, Canada. pp 347-354

