

Sentic Panalogy: Swapping Affective Common Sense Reasoning Strategies and Foci

Erik Cambria, Daniel Olsher, Kenneth Kwok

{cambria, olsher, kenkwok}@nus.edu.sg

Cognitive Science Programme, Temasek Laboratories

National University of Singapore, 5A Engineering Drive 1, Singapore 117411

<http://sentic.net>

Abstract

An important difference between traditional AI systems and human intelligence is our ability to harness common sense knowledge gleaned from a lifetime of learning and experiences to inform our decision-making and behavior. This allows humans to adapt easily to novel situations where AI fails catastrophically for lack of situation-specific rules and generalization capabilities. In order for machines to exploit common sense knowledge in reasoning as humans do, moreover, we need to endow them with human-like reasoning strategies. In problem-solving situations, in particular, several analogous representations of the same problem should be maintained in parallel while trying to solve it so that, when problem-solving begins to fail while using one representation, the system can switch to one of the others. Sentic panalogy is a technique that aims to emulate such process by exploiting graph-mining and dimensionality-reduction techniques to dynamically interchange both different reasoning strategies and the foci around which such strategies are developed.

Keywords: AI; NLP; Cognitive systems; Sentic computing.

Introduction

Emotions are different Ways to Think (Minsky, 2006) that our mind triggers to deal with different situations we face in our lives. Our decision-making and problem-solving skills, in fact, are strictly dependent on both our common sense knowledge about the world and the appraisal associated with this (Scherer, Shorr, & Johnstone, 2001). The capability to accordingly compress and exploit such information, which we term affective common sense reasoning (Cambria, Olsher, & Kwok, 2012), is a fundamental component in human experience, cognition, perception, learning, and communication.

For this reason, we cannot prescind from emotions in the development of intelligent user interfaces: if we want computers to be really intelligent, not just have the veneer of intelligence, we need to give them the ability to recognize, understand, and express emotions. Furthermore, in order not to get stuck and to be able to tackle different problems from different perspectives, an intelligent machine should not have a unique way to deal with a task, but rather be endowed with different reasoning strategies and with the capability to accordingly switch among these.

This work further develops a recently proposed approach (Cambria, Mazzocco, Hussain, & Durrani, 2011) for the emulation of the human capability to switch between different perspectives and find novel ways to look at things. Such approach is inspired by Minsky's notion of 'panalogy' (parallel analogy), which states that several analogous representations of the same problem should be maintained in parallel while trying to solve it (Minsky, 2006).

To show the effectiveness of the proposed approach, termed sentic panalogy, we employ it for the natural language processing (NLP) task of sentiment analysis, for which a faceted and nuanced analysis is mostly needed.

The structure of the paper is as follows: the first section provides some background information on sentiment analysis; the second section introduces the concept of affective common sense reasoning and explains why and how this can aid sentiment analysis; the third and fourth sections describe the implementation of the switch among different strategies and among the foci around which such strategies are developed, respectively; the fifth section provides an evaluation of the proposed approach; the last section, finally, comprises concluding remarks and future directions.

Sentiment Analysis

Sentiment analysis is a branch of the broad field of text data mining and refers generally to the process of extracting interesting and non-trivial patterns or knowledge from unstructured text documents. It can be viewed as an extension of data mining or knowledge discovery from (structured) databases (Fayyad, Piatetsky, & Smyth, 1996; Simoudis, 1996). As the most natural form of storing information is text, sentiment analysis is believed to have a commercial potential higher than that of data mining. Sentiment analysis, however, is also a much more complex task as it involves dealing with text data that are inherently unstructured and fuzzy. It is a multi-disciplinary research area that involves the adoption of techniques in fields such as text analysis, information retrieval and extraction, auto-categorization, machine learning, clustering, and visualization.

Most of the existing approaches to opinion mining and sentiment analysis rely on the extraction of a vector representing the most salient and important text features, which is later used for classification purposes. Some of the most commonly used features are term frequency (Wu, Luk, Wong, & Kwok, 2008) and presence (Pang, Lee, & Vaithyanathan, 2002). The latter, in particular, is a binary-valued feature vectors in which the entries merely indicate whether a term occurs or not, and formed a more effective basis for polarity classification. This is indicative of an interesting difference between typical topic-based text categorization. While a topic is more likely to be emphasized by frequent occurrences of certain keywords, overall sentiment may not usually be highlighted through repeated use of the same terms.

Differently from topics, in fact, sentiments can often be expressed in a more subtle manner, making it difficult to be identified by specific keywords, especially when considering multiple domains. Humans readers do not face such difficulty as they can infer the cognitive and affective information associated with natural language text through their affective common sense knowledge, that is, obvious or widely accepted things that people normally know about the world, but which are usually left unstated in discourse, e.g., that people smile when they are happy and things fall downwards (and not upwards). An important feature of affective common sense reasoning, in fact, is the sensitivity to nuanced readings of natural language.

A sentence can be read differently depending on nuances in opinionated text and such nuanced reading can lead to markedly different reasoning trajectories. The first step in human cognitive and affective information processing, in fact, is in an appraisal of the current situation (Scherer et al., 2001). In order to accordingly infer semantics and sentics (Cambria, Benson, Eckl, & Hussain, 2012), i.e., the cognitive and affective information associated with natural language text, next-generation sentiment analysis methods need to go beyond a mere word-level analysis and use affective common sense reasoning to better grasp the conceptual rules that govern sentiment and the clues that can convey these concepts from realization to verbalization in the human mind.

Affective Common Sense Reasoning

Current thinking in cognitive psychology suggests that humans process information at a minimum of two distinct levels. There is extensive evidence for the existence of (at least) two processing systems within the human brain, one that involves fast, parallel, unconscious processing, and one that involves slow, serial, more conscious processing (Kirkpatrick & Epstein, 1992; Chaiken & Trope, 1999; Smith & DeCoster, 2000; Epstein, 2003; Kahneman, 2011). Dual-process models of automatic and controlled social cognition have been proposed in nearly every domain of social psychology.

Evidence from neurosciences supports this separation, with identifiably different brain regions involved in each of the two systems (Lieberman, 2007). Such systems, which we term U-level (unconscious) and C-level (conscious), can operate simultaneously or sequentially, and are most effective in different contexts. The former, in particular, works intuitively, effortlessly, globally, and emotionally. The latter, in turn, works logically, systematically, effortfully, and rationally. According to different contexts and purposes, moreover, the systems should be capable to dynamically swap both different reasoning strategies and the foci around which such strategies are developed.

In this work, we emulate such dual-process model through an ensemble application of dimensionality-reduction and graph-mining techniques on AffectNet (Cambria & Hussain, 2012), an affective common sense knowledge base built upon WordNet-Affect (WNA) (Strapparava & Valitutti, 2004), a

linguistic resource for the lexical representation of affect, and ConceptNet (Havasi, Speer, & Alonso, 2007), a semantic network of common sense knowledge. In particular, multi-dimensionality reduction techniques are employed on AffectNet to dynamically configure it and, hence, to model the switch between different reasoning strategies, while graph mining and clustering methods are applied to model the switch between the foci around which those strategies are developed, in order to accordingly exploit the different facets of the affective common sense knowledge base.

Swapping Reasoning Strategies

To some extent, our reasoning capability can be re-conducted to the identification of useful patterns in our acquired knowledge about the world. Our experience and common sense knowledge is likely to be organized in our mind as interconnected concepts and events and most of these links are weighted by affective information, as we tend to forget or hardly recall memories that are not associated with any kind of positive or negative emotion. Therefore, the human capacity to envision possible outcomes of a decision might lie both in the capability of crawling the semantic network of concepts we have acquired through experience (C-level), and in the capability of summarizing the huge amount of inputs and outputs of previous situations to find useful patterns that might work at the present time (U-level).

The latter capability, in cognitive science, is termed ‘compression’ and refers to transforming diffuse and distended conceptual structures that are less congenial to human understanding so that they become better suited to our human-scale ways of thinking. Compression is hereby implemented by representing affective common sense knowledge in a way that it is neither too concrete nor too abstract with respect to the detail granularity needed for performing a particular task.

To this end, we first generate a matrix representation of AffectNet by applying blending (Havasi, Speer, Pustejovsky, & Lieberman, 2009), a technique that performs inference over multiple sources of data simultaneously, taking advantage of the overlap between them. In particular, the alignment of ConceptNet and WNA yields A , a matrix in which common sense and affective knowledge coexist, i.e., a matrix $14,301 \times 117,365$ whose rows are concepts (e.g., ‘dog’ or ‘bake cake’), whose columns are either common sense and affective features (e.g., ‘isA-pet’ or ‘hasEmotion-joy’), and whose values indicate truth values of assertions. Therefore, in A , each concept is represented by a vector in the space of possible features whose values are positive for features that produce an assertion of positive valence (e.g., ‘a penguin is a bird’), negative for features that produce an assertion of negative valence (e.g., ‘a penguin cannot fly’), and zero when nothing is known about the assertion. The degree of similarity between two concepts, then, is the dot product between their rows in A . The value of such a dot product increases whenever two concepts are described with the same feature and decreases when they are described by features that are negations of each other.

In particular, we use truncated singular value decomposition (TSVD) (Wall, Rechtsteiner, & Rocha, 2003) in order to obtain a new matrix containing both hierarchical affective knowledge and common sense. The resulting matrix has the form $\tilde{A} = U_k \Sigma_k V_k^T$ and is a low-rank approximation of A , the original data. This approximation is based on minimizing the Frobenius norm of the difference between A and $\tilde{A}$ under the constraint $\text{rank}(\tilde{A}) = k$. For the Eckart-Young theorem (Eckart & Young, 1936), it represents the best approximation of A in the least-square sense, in fact:

$$\begin{aligned} \min_{\tilde{A} | \text{rank}(\tilde{A})=k} \|A - \tilde{A}\| &= \min_{\tilde{A} | \text{rank}(\tilde{A})=k} \|\Sigma - U^* \tilde{A} V\| \\ &= \min_{\tilde{A} | \text{rank}(\tilde{A})=k} \|\Sigma - S\| \end{aligned}$$

assuming that $\tilde{A}$ has the form $\tilde{A} = USV^*$, where S is diagonal. From the rank constraint, i.e., S has k non-zero diagonal entries, the minimum of the above statement is obtained as follows:

$$\begin{aligned} \min_{\tilde{A} | \text{rank}(\tilde{A})=k} \|\Sigma - S\| &= \min_{s_i} \sqrt{\sum_{i=1}^n (\sigma_i - s_i)^2} = \\ &= \min_{s_i} \sqrt{\sum_{i=1}^k (\sigma_i - s_i)^2 + \sum_{i=k+1}^n \sigma_i^2} = \sqrt{\sum_{i=k+1}^n \sigma_i^2} \end{aligned}$$

Therefore, $\tilde{A}$ of rank k is the best approximation of A in the Frobenius norm sense when $\sigma_i = s_i$ ($i = 1, \dots, k$) and the corresponding singular vectors are the same as those of A . If we choose to discard all but the first k principal components, common sense concepts and emotions are represented by vectors of k coordinates: these coordinates can be seen as describing concepts in terms of ‘eigenmoods’ that form the axes of AffectiveSpace, i.e., the basis $e_0, \dots, e_{k-1}$ of the vector space. For example, the most significant eigenmood, e_0 , represents concepts with positive affective valence. That is, the larger a concept’s component in the e_0 direction is, the more affectively positive it is likely to be. Concepts with negative e_0 components, then, are likely to have negative affective valence.

Thus, by exploiting the information sharing property of TSVD, concepts with the same affective valence are likely to have similar features - that is, concepts conveying the same emotion tend to fall near each other in AffectiveSpace. Concept similarity does not depend on their absolute positions in the vector space, but rather on the angle they make with the origin. For example we can find concepts such as ‘beautiful day’, ‘birthday party’, ‘laugh’, and ‘make person happy’ very close in direction in the vector space, while concepts like ‘sick’, ‘feel guilty’, ‘be laid off’, and ‘shed tear’ are found in a completely different direction (nearly opposite with respect to the centre of the space). By reducing the dimensionality of the matrix representation of A , AffectiveSpace compresses the feature space of affective common sense knowledge into one that allows to better gain global insight and human-scale understanding.

The number k of singular values selected for building AffectiveSpace, in fact, is a measure of the trade-off between precision and efficiency in the representation of the affective common sense knowledge base. Switching between different vector space dimensionalities can be seen as looking at the data from many different points of view. Balancing the number of singular values discarded when synthesizing AffectiveSpace, hence, corresponds to calibrate the affective common sense knowledge representation in a way that it is neither too concrete nor too abstract with respect to the detail granularity needed for performing a particular task. Different k values, for example, work differently according to the affective dimension we consider, e.g., for Pleasantness the best k appears to be closer to 100, while for Sensitivity a space of about 70 dimensions appears to be enough for precisely and efficiently represent affective common sense knowledge.

The capability to look at things from a different perspective, moreover, can be emulated by applying different space transformations to AffectiveSpace. The distribution of the values of each AffectiveSpace dimension is bell-shaped, with different centers and different degree of dispersion around them. In order to more uniformly distribute affective common sense knowledge in the vector space, an alternative strategy to represent AffectiveSpace consists in centering the values of the distribution of each dimension on the origin and in mapping dimensions according to a transformation $x \in \mathbb{R} \mapsto x^* \in [-1, 1]$. This transformation is often pivotal for better clustering AffectiveSpace as the vector space tends to have different grades of dispersion of data points across different dimensions, with some space regions more densely populated than others.

The switch to a different space configuration helps to distribute data more uniformly, possibly leading to an improved (or, at least, different) reasoning process. Switching between different space configurations, in fact, changes how each dimension is influent in the vector space representation of AffectNet and, hence, changes how we are looking at the affective common sense knowledge because similarity in AffectiveSpace does not depend on concepts’ absolute position, but rather on the angle they make with the origin of the vector space. In particular, the transformation $x_{ij} \mapsto x_{ij} - \mu_i$ is first applied, being μ_i the average of all values of the i -th dimension. Then a normalization is applied, combining the previous transformation with a new one $x_{ij} \mapsto \frac{x_{ij}}{a \cdot \sigma_i}$, where σ_i is the standard deviation calculated on the i -th dimension and a is a coefficient that can modify the same proportion of data that is represented within a specified interval.

Finally, in order to ensure that all components of the vectors in the defined space are within $[-1, 1]$ (i.e., that the Chebyshev distance between the origin and each vector is smaller or equal to 1), a final transformation $x_{ij} \mapsto s(x_{ij})$ is needed, where $s(x)$ is a sigmoid function. Different choices for the sigmoid function may be made, influencing how ‘fast’ the function approaches 1 while the independent variable approaches infinity.

Combining the proposed transformations, two possible mapping functions are expressed in the following formulae:

$$x_{ij}^* = \tanh\left(\frac{x_{ij} - \mu_i}{a \cdot \sigma_i}\right)$$

$$x_{ij}^* = \frac{x_{ij} - \mu_i}{a \cdot \sigma_i + |x_{ij} - \mu_i|}$$

This space transformation leads to two main advantages, which could be of notable importance depending on the problem being tackled. First, this different space configuration ensures that each dimension is equally important by avoiding that the information provided by dimensions with higher (i.e., more distant from the origin) averages predominates. Second, normalizing according to the standard deviations of each dimension allows a more uniform distribution of data around the origin, leading to a full use of information potential.

Swapping Reasoning Foci

The capability of switching among different Ways to Think can be thought as changing the foci around which we develop our different reasoning strategies. Such approach can be implemented in AffectiveSpace by changing the centroids around which the vector space is clustered. Such a clustering process is implemented by adopting a k -medoids approach (Kaufman & Rousseeuw, 1990) to partition the given observations into k clusters around as many centroids, trying to minimize a given cost function. Differently from the k -means algorithm, which does not pose constraints on centroids, k -medoids do assume that centroids must coincide with k observed points.

The most commonly used algorithm for finding the k medoids is the partitioning around medoids (PAM) algorithm, which determines a medoid for each cluster selecting the most centrally located centroid within the cluster. After selection of medoids, clusters are rearranged so that each point is grouped with the closest medoid. Since k -medoids clustering is a NP-hard problem, different approaches based on alternative optimization algorithms have been developed, though taking risk of being trapped around local minima. We use a modified version of the algorithm recently proposed by Park and Jun (Park & Jun, 2009), which runs similarly to the k -means clustering algorithm.

This has shown to have similar performance when compared to PAM algorithm while taking a significantly reduced computational time. In particular, we have N concepts ($N = 14,301$) encoded as points $x \in \mathbb{R}^p$ ($p = 100$). We want to group them into k clusters and, in our case, we can fix $k = 24$ as we are looking for one cluster for each sentic level of the Hourglass of Emotions (Cambria, Livingstone, & Hussain, 2012), a novel biologically-inspired and psychologically-motivated emotion categorization model, based on Plutchik's studies on human emotions (Plutchik, 2001), that can potentially describe any human emotion in terms of four independent but concomitant dimensions, whose different levels of activation make up the total emotional state of the mind.

Specifically, we need to cluster AffectiveSpace four times, once for each dimension. According to the Hourglass categorization model, in fact, each concept can convey, at the same time, more than one emotion (which is why we get compound emotions) and this information can be expressed via a sentic vector specifying the concept's affective valence in terms of Pleasantness, Attention, Sensitivity, and Aptitude. Therefore, given that the distance between two points in AffectiveSpace is defined as $D(a, b) = \sqrt{\sum_{i=1}^p (a_i - b_i)^2}$, the used algorithm, applied for each of the four affective dimensions, can be summarized as follows:

1. Each centroid $C_n \in \mathbb{R}^{100}$ ($n = 1, 2, \dots, k$) is set as one of the six concepts corresponding to each s in the current affective dimension
2. Assign each record x to a cluster Ξ so that $x_i \in \Xi_n$ if $D(x_i, C_n) \leq D(x_i, C_m)$ $m = 1, 2, \dots, k$
3. Find a new centroid C for each cluster Ξ so that $C_j = x_i$ if $\sum_{x_m \in \Xi_j} D(x_i, x_m) \leq \sum_{x_m \in \Xi_j} D(x_h, x_m) \quad \forall x_h \in \Xi_j$
4. Repeat step 2 and 3 until no changes on centroids are observed

After such a clustering process (performed at U-level), concepts that are semantically and affectively related to the input data can be intuitively retrieved by analogy and unconsciously crop out to the C-level. According to the initial centroid we choose, the final clusterization of AffectiveSpace can be very different. Hence, the way such initial medoids are chosen can be re-conducted to the human capability to switch between different perspectives to grasp the different facets of a problem.

At C-level, moreover, reasoning is performed by exploiting AffectNet's connectivity to find semantically and affectively related concepts by means of spectral association (Havasi, Speer, & Holmgren, 2010). Spectral association is a technique that involves assigning activation to seed concepts and applying an operation that spreads their values across the graph. This operation, an approximation of many steps of spreading activation, transfers the most activation to concepts that are connected to the seed concepts by short paths or many different paths in affective common sense knowledge.

Seed concepts can also be associated with negative activation values in order to reduce the spreading operation in the parts of the graph we are specifically not interested in. If we want to find concepts semantically related to 'bank' as a financial institution without getting concepts related to 'river bank', for example, we can set as positive seeds concepts like 'money', 'savings', or 'investment', and, as negative seeds, concepts like 'river', 'water', or 'shore'. The outcomes of spectral association can be very different according to which seeds we select as starting points for the spreading activation steps. Since spectral association involves TSVD, results also depend on the number k of singular values selected.

While choosing different k values can be seen as developing different reasoning strategies, choosing different seeds can be associated to changing the foci around which those strategies are developed. Through spectral association, positive and negative values of these concepts are spread across the graph representation of AffectNet, resulting in a set of contextually semantic related instances. Letting a machine switch between such seeds according to its own intuition (e.g., concepts obtained through AffectiveSpace at U-level) can be re-conducted to the human capability to change the foci around which different reasoning strategies are developed and, hence, to iterate on the ways to look at a problem until one that works is found.

Evaluation

In order to efficiently and timely swap different reasoning strategies and foci, we perform all the computations (relative to the most significant configurations) a priori and save the results in a semantic-aware format, using an approach previously adopted for building SenticNet (Cambria, Havasi, & Hussain, 2012). The result is a system for affect recognition that has multiple ways to deal with natural language semantics and sentics. We tested sentic panalogy on a benchmark for affective common sense knowledge (BACK) built by applying CF-IOF (concept frequency - inverse opinion frequency), a technique similar to TF-IDF, on a 5,000 blogpost database extracted from LiveJournal¹, a virtual community of users who keep a blog, journal, or diary. An interesting feature of this website is that bloggers are allowed to label their posts with both a category and a mood tag, by choosing from predefined categories and mood themes.

CF-IOF identifies common domain-dependent semantics in order to evaluate how important a concept is to a set of opinions concerning the same topic. Firstly, the frequency of a concept c for a given domain d is calculated by counting the occurrences of the concept c in the set of available d -tagged opinions and dividing the result by the sum of number of occurrences of all concepts in the set of opinions concerning d . This frequency is then multiplied by the logarithm of the inverse frequency of the concept in the whole collection of opinions, that is:

$$CF-IOF_{c,d} = \frac{n_{c,d}}{\sum_k n_{k,d}} \log \sum_k \frac{n_k}{n_c}$$

where $n_{c,d}$ is the number of occurrences of concept c in the set of opinions tagged as d , n_k is the total number of concept occurrences, and n_c is the number of occurrences of c in the whole set of opinions. A high weight in CF-IOF is reached by a high concept frequency in a given domain and a low frequency of the concept in the whole collection of opinions. Specifically, we exploited CF-IOF weighting to filter out common concepts in the LiveJournal corpus and detect relevant mood-dependent semantics for each of the Hourglass sentic levels. The result was a benchmark of 2000 af-

fective concepts that were screened by 21 English-speaking students who were asked to map each concept to the 24 different emotional categories that form the Hourglass of Emotions. BACK's concepts were compared with the classification results obtained by applying the AffectiveSpace process, spectral association, and sentic panalogy. Results showed that sentic panalogy achieves +9.7% and +6.2% accuracy than the standard (i.e., 100-dimensional) AffectiveSpace process and the default (i.e., fixed on the Hourglass sentic levels) spectral association, respectively.

Brain-Inspired Sentiment Analysis

In order to test sentic panalogy also within a real-world scenario, we developed a brain-inspired software engine for sentiment analysis. This software engine consists of four main components: a pre-processing module, a semantic parser, a target spotting module, and an affect interpreter.

The pre-processing module firstly interprets all the affective valence indicators usually contained in opinionated text such as cross-linguistic onomatopoeias, exclamation words, degree adverbs, and emoticons. Secondly, it lemmatizes text and splits the opinion into single clauses according to grammatical conjunctions and punctuation. Then, the semantic parser deconstructs text into concepts using a lexicon based on sequences of lexemes that represent multiple-word concepts extracted from AffectNet.

The target spotting module aims to individuate one or more sentiment targets, e.g., people, places, events, and ideas, from the input concepts. This is done by projecting the retrieved concepts into both the graph and the vector space representation of AffectNet, in order to assign these to a specific conceptual class. The categorization does not consist in simply labeling each concept, but also in assigning a confidence score to each category label, which is directly proportional to the value of belonging to a specific conceptual cluster (number of steps in the graph and dot product in the vector space). The affect interpreter, similarly, projects the retrieved concepts into the vector space representation of AffectNet, in order to assign these to a specific affective class and, hence, calculates polarity in terms of the Hourglass dimensions.

Approach Comparison

In order to evaluate the different facets of the engine from different perspectives, we used a PatientOpinion² dataset and compared results obtained using standard AffectiveSpace, default spectral association, and sentic panalogy. The resource is a dataset obtained from PatientOpinion, a social enterprise pioneering an on-line feedback service for users of the UK national health service to enable people to share their recent experience of local health services on-line. It is a manually tagged dataset of 2,000 patient opinions that associates to each post a category (namely, clinical service, communication, food, parking, staff, and timeliness) and a positive or negative polarity.

¹<http://livejournal.com>

²<http://patientopinion.org.uk>

Category Label	Affective-Space	Spectral Association	Sentic Panalogy
clinical service	62.4%	71.5%	80.0%
communication	62.3%	59.8%	71.2%
food	70.7%	69.6%	78.5%
parking	56.3%	53.7%	69.4%
staff	58.5%	49.2%	63.9%
timeliness	69.2%	61.9%	75.1%

Table 1: F-measures relative to PatientOpinion evaluation.

We used it to test the detection of opinion targets and the polarity associated with these (F-measure values are reported in Table 1).

Conclusion

Sentic panalogy is novel approach to affective common sense reasoning inspired by Minsky's notion of parallel analogy. It employs different KR strategies and reasoning techniques to maintain several analogous representations of the same problem so that, when a particular strategy begins to fail, the system can switch to one of the others. In the future, we plan to develop heuristics to swap reasoning strategies and foci in real-time, rather than performing all the computations a priori, in order to pave the way for more brain-inspired approaches to affective common sense reasoning.

References

- Cambria, E., Benson, T., Eckl, C., & Hussain, A. (2012). Sentic PROMs: Application of sentic computing to the development of a novel unified framework for measuring health-care quality. *Expert Systems with Applications*, <http://dx.doi.org/10.1016/j.eswa.2012.02.120>.
- Cambria, E., Havasi, C., & Hussain, A. (2012). SenticNet 2: A semantic and affective resource for opinion mining and sentiment analysis. In *FLAIRS*. Marco Island.
- Cambria, E., & Hussain, A. (2012). *Sentic computing: Techniques, tools, and applications*. Berlin Heidelberg: Springer.
- Cambria, E., Livingstone, A., & Hussain, A. (2012). The hourglass of emotions. In *Cognitive behavioral systems*. Berlin Heidelberg: Springer.
- Cambria, E., Mazzocco, T., Hussain, A., & Durrani, T. (2011). Switching between different ways to think: Multiple approaches to affective common sense reasoning. In *LNCIS* (Vol. 6800, p. 56-69). Berlin: Springer-Verlag.
- Cambria, E., Olsher, D., & Kwok, K. (2012). Sentic activation: A two-level affective common sense reasoning framework. In *AAAI*. Toronto.
- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford.
- Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3), 211-218.
- Epstein, S. (2003). Cognitive-experiential self-theory of personality. In *Comprehensive handbook of psychology* (Vol. 5, p. 159-184). Hoboken, NJ: Wiley & Sons.
- Fayyad, U., Piatetsky, G., & Smyth, P. (1996). From data mining to knowledge discovery: An overview. In *Advances in knowledge discovery and data mining* (p. 1-36). Cambridge: MIT Press.
- Havasi, C., Speer, R., & Alonso, J. (2007). ConceptNet 3: A flexible, multilingual semantic network for common sense knowledge. In *RANLP*. Borovets.
- Havasi, C., Speer, R., & Holmgren, J. (2010). Automated color selection using semantic knowledge. In *AAAI CSK*. Arlington.
- Havasi, C., Speer, R., Pustejovsky, J., & Lieberman, H. (2009). Digital intuition: Applying common sense using dimensionality reduction. *IEEE Intelligent Systems*, 24(4), 24-35.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
- Kaufman, L., & Rousseeuw, P. (1990). *Finding groups in data: An introduction to cluster analysis*. Wiley-Interscience.
- Kirkpatrick, L., & Epstein, S. (1992). Cognitive experiential self-theory and subjective probability: Further evidence for two conceptual systems. *Journal of Personality and Social Psychology*, 63, 534-544.
- Lieberman, M. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, 58, 259-89.
- Minsky, M. (2006). *The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind*. New York: Simon & Schuster.
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *EMNLP* (p. 79-86). Philadelphia.
- Park, H., & Jun, C. (2009). A simple and fast algorithm for k-medoids clustering. *Expert Systems with Applications*, 36(2), 3336-3341.
- Plutchik, R. (2001). The nature of emotions. *American Scientist*, 89(4), 344-350.
- Scherer, K., Shorr, A., & Johnstone, T. (2001). *Appraisal processes in emotion: Theory, methods, research*. Canary: Oxford University Press.
- Simoudis, E. (1996). Reality check for data mining. *IEEE Expert*, 11(5).
- Smith, E., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychological Review*, 4(2), 108-131.
- Strapparava, C., & Valitutti, A. (2004). WordNet-Affect: An affective extension of WordNet. In *LREC*. Lisbon.
- Wall, M., Rechtsteiner, A., & Rocha, L. (2003). Singular value decomposition and principal component analysis. In *A practical approach to microarray data analysis* (p. 91-109). Springer.
- Wu, H., Luk, R., Wong, K., & Kwok, K. (2008). Interpreting TF-IDF term weights as making relevance decisions. *ACM Transactions on Information Systems*, 26(3).