# Application of Common Sense Computing for the Development of a Novel Knowledge-Based Opinion Mining Engine

A thesis submitted in accordance with the requirements of
the University of Stirling for the degree of Doctor of Philosophy

by
**Erik Cambria**

Principal Supervisor: Amir Hussain (University of Stirling, UK)
Additional Supervisor: Catherine Havasi (MIT Media Laboratory, USA)
Industrial Supervisor: Chris Eckl (Sitekit Solutions Ltd, UK)

Department of Computing Science & Mathematics
University of Stirling, Scotland, UK

December 2011

# Declaration

I, Erik Cambria, hereby declare that this work has not been submitted for any other degree at this University or any other institution and that, except where reference is made to the work of other authors, the material presented is original.

Erik Cambria

# Abstract

The ways people express their opinions and sentiments have radically changed in the past few years thanks to the advent of social networks, web communities, blogs, wikis and other online collaborative media. The distillation of knowledge from this huge amount of unstructured information can be a key factor for marketers who want to create an image or identity in the minds of their customers for their product, brand, or organisation. These online social data, however, remain hardly accessible to computers, as they are specifically meant for human consumption. The automatic analysis of online opinions, in fact, involves a deep understanding of natural language text by machines, from which we are still very far.

Hitherto, online information retrieval has been mainly 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 their words. But when it comes to interpreting sentences and extracting meaningful information, their capabilities are known to be very limited. Existing approaches to opinion mining and sentiment analysis, in particular, can be grouped into three main categories: keyword spotting, in which text is classified into categories based on the presence of fairly unambiguous affect words; lexical affinity, which assigns arbitrary words a probabilistic affinity for a particular emotion; statistical methods, which calculate the valence of affective keywords and word co-occurrence frequencies on the base of a large training corpus. Early works aimed to classify entire documents as containing overall positive or negative polarity, or rating scores of reviews.

Such systems were mainly based on supervised approaches relying on manually labelled samples, such as movie or product reviews where the opinionist's overall positive or negative attitude was explicitly indicated. However, opinions and sentiments do not occur only at document level, nor they are limited to a single valence or target. Contrary or complementary attitudes toward the same topic or multiple topics can be present across the span of a document. In more recent works, text analysis granularity has been taken down to segment and sentence level, e.g., by using presence of opinion-bearing lexical items (single words or n-grams) to detect subjective sentences, or by exploiting association rule mining for a feature-based analysis of product reviews. These approaches, however, are still far from being able to infer the cognitive and affective information associated with natural language as they mainly rely on knowledge bases that are still too limited to efficiently process text at sentence level.

In this thesis, common sense computing techniques are further developed and applied to bridge the semantic gap between word-level natural language data and the concept-level opinions conveyed by these. In particular, the ensemble application of graph mining and multi-dimensionality reduction techniques on two common sense knowledge bases was exploited to develop a novel intelligent engine for open-domain opinion mining and sentiment analysis. The proposed approach, termed sentic computing, performs a clause-level semantic analysis of text, which allows the inference of both the conceptual and emotional information associated with natural language opinions and, hence, a more efficient passage from (unstructured) textual information to (structured) machine-processable data.

The engine was tested on three different resources, namely a Twitter hashtag repository, a LiveJournal database and a PatientOpinion dataset, and its performance compared both with results obtained using standard sentiment analysis techniques and using different state-of-the-art knowledge bases such as Princeton's WordNet, MIT's ConceptNet and Microsoft's Probase. Differently from most currently available opinion mining services, the developed engine does not base its analysis on a limited set of affect words and their co-occurrence frequencies, but rather on common sense concepts and the cognitive and affective valence conveyed by these. This allows the engine to be domain-independent and, hence, to be embedded in any opinion mining system for the development of intelligent applications in multiple fields such as Social Web, HCI and e-health. Looking ahead, the combined novel use of different knowledge bases and of common sense reasoning techniques for opinion mining proposed in this work, will, eventually, pave the way for development of more bio-inspired approaches to the design of natural language processing systems capable of handling knowledge, retrieving it when necessary, making analogies and learning from experience.

# Dedication

In memory of John McCarthy (September 4th, 1927 - October 24th, 2011), who helped design the foundation of today's Internet-based computing and coined the term for a frontier of research he helped pioneer, AI.

# Acknowledgements

The humble accomplishment of this thesis would not have been possible without the contribution of many individuals, to whom I express my appreciation and gratitude. Firstly, I am deeply indebted to my research supervisors, specifically: Amir Hussain, my principal supervisor and founding Head of COSIPRA Lab at Stirling, who gave me the opportunity to work on this exciting project, guided me every step of the way, funded my research and training visits to the project's prestigious partner institutions, and was an immense source of inspiration throughout; Catherine Havasi, whose guidance, encouragement and support in the past three years has been simply invaluable; and Chris Eckl, who helped me to look at my research from many different points of view.

I am also grateful to all the other mentors and colleagues who supported me during my short term scientific missions and internships, in particular Robert Speer, Kenneth Arnold, Dustin Smith, Jason Alonso and Henry Lieberman, who assisted me in getting started, using and further developing common sense computing tools (both during my research visits at MIT Media Lab and remotely), Campbell Grant, for his critical opinions and sound advice about the commercial aspects of my research work, Anna Esposito, for her unwavering support within the COST 2102 program, Thomas Mazzocco and Marco Grassi, for their invaluable research contributions, Joseph Lyons, for his help in the design and refinement of the Hourglass model, and James Munro, for the support and the data provided for the development of patient centred applications.

Special thanks also go to Praphul Chandra and Sudhir Dixit, who helped me expand the horizons of my research during my internship at Hewlett-Packard Labs India (within the Innovations for the Next Billion Customers Initiative) in Bangalore, Tariq Durrani, Tieniu Tan, Cheng-Lin Liu and Chengqing Zong, for their guidance during my research visit at the National Laboratory of Pattern Recognition (NLPR) in the Institute of Automation of the Chinese Academy of Sciences (within the China-Scotland SIPRA programme) in Beijing, and Haixun Wang and Yangqiu Song, for helping me improve my skills and expertise in the field of knowledge-based systems during my internship at Microsoft Research Asia (within the Probase Project) in Beijing.

A last, but not least, acknowledgement goes to my family and all the old and new friends who, in the past three years, have cheered me up in difficult moments, celebrated with me for my achievements and never blamed me for being often far away from them. This PhD was the best experience of my life as it gave me the possibility to: work shoulder to shoulder with scientists from top research institutes, get in touch with both far West and far East cultures, meet special people who will always be part of my life, and even exchange ideas with beautiful minds such as Tim Berners-Lee, the inventor of the Web, Bebo White, Web pioneer in the United States, and Marvin Minsky, one of the fathers of AI.

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

**3NF** Third Normal Form

**AI** Artificial Intelligence

**AKAPU** Average Knowledge Acquired Per User

**ANN** Artificial Neural Network

**API** Application Programming Interface

**BACK** Benchmark for Affective Common sense Knowledge

**BCNF** Boyce-Codd Normal Form

**CF-IOF** Concept Frequency - Inverse Opinion Frequency

**CMC** Computer Mediated Communication

**CRF** Conditional Random Field

**CRM** Customer Relationship Management

**DAG** Directed Acyclic Graph

**DAU** Daily Active User

**DL** Description Logic

**ECA** Embodied Conversational Agent

**ELM** Extreme Learning Machine

**EQ** Emotional Quotient

**FMRI** Functional Magnetic Resonance Imaging

**FOAF** Friend Of A Friend

**FOL** First Order Logic

**GUI** Graphic User Interface

**GWAP** Game With A Purpose

**HCI** Human Computer Interaction

**HEO** Human Emotion Ontology

**HRQoL** Health Related Quality of Life

**HTML** Hyper Text Markup Language

**HVS** Human Visual System

**ICA** Independent Component Analysis

**IM** Instant Messaging

**IT** Information Technology

**IUI** Intelligent User Interface

**KNN** K-Nearest Neighbor

**KR** Knowledge Representation

**JSON** JavaScript Object Notation

**LSA** Latent Semantic Analysis

**LSR** Label Sequential Rules

**MAU** Monthly Active User

**MDS** Multi-Dimensional Scaling

**MLP** Multi-Layer Perceptron

**MMO** Massively Multiplayer Online

**MVE** Minimum Volume Ellipsoid

**NB** Naïve Bayes

**NELL** Never-Ending Language Learning

**NMF** Non-negative Matrix Factorisation

**NLP** Natural Language Processing

**NP** Nondeterministic Polynomial

**OCR** Optical Character Recognition

**OMCS** Open Mind Common Sense

**OMR** Ontology for Media Resources

**OWL** Ontology Web Language

**PAM** Partitioning Around Medoids

**PCA** Principal Component Analysis

**POS** Part Of Speech

**PROM** Patient Reported Outcome Measure

**RDBMS** Relational Database Management Systems

**RDF** Resource Description Framework

**RDFS** Resource Description Framework Schema

**SBoC** Small Bag of Concepts

**SKOS** Simple Knowledge Organisation System

**SQL** Structured Query Language

**SNN** Sentic Neural Network

**SVD** Singular Value Decomposition

**SVM** Support Vector Machine

**TF-IDF** Term Frequency - Inverse Document Frequency

**TMS** Truth Maintenance System

**TSVD** Truncated Singular Value Decomposition

**UGC** User Generated Content

**UML** Unified Modeling Language

**VSM** Vector Space Model

**XML** Extensible Markup Language

**W3C** World Wide Web Consortium

**WNA** WordNet-Affect

# Chapter 1

# Introduction

*We can understand almost anything,*

*but we can't understand how we understand.*

*Albert Einstein*

In a world in which millions of people express their opinions about commercial products in blogs, wikis, forums, chats, and social networks, the distillation of knowledge from this huge amount of unstructured information can be a key factor for marketers who want to create an image or identity in the minds of their customers for their product, brand, or organisation [1]. The automatic analysis of online opinions, however, involves a deep understanding of natural language text by machines, from which we are still very far [2]. Online information retrieval, in fact, is still mainly based on algorithms relying on the textual representation of web-pages [3].

Such algorithms are very good at retrieving texts, splitting them into parts, checking the spelling, and counting their words. But when it comes to interpreting sentences and extracting useful information for users, their capabilities are still very limited. In this thesis, common sense computing techniques are further developed and applied to bridge the cognitive and affective gap between word-level natural language data and the concept-level opinions conveyed by these. In particular, two common sense knowledge bases were designed, together with a novel emotion categorisation model, and graph mining and multi-dimensionality reduction techniques were applied on them to infer cognitive and affective information from natural language and, hence, develop opinion-mining systems in fields such as Social Web, HCI, and e-health.

The structure of the thesis is as follows: this chapter presents motivations, aims and contributions of the research hereby presented, chapter 2 illustrates the state of the art of opinion mining and sentiment analysis together with past and recent developments in the field of common sense computing, chapter 3 shows the methods employed to build the knowledge bases on which the opinion mining engine is based, chapter 4 explains in detail the strategies adopted to perform reasoning on such knowledge bases, chapter 5 discusses how to exploit the knowledge bases for the development of opinion mining systems in different fields, chapter 6, eventually, comprises concluding remarks and future work.

## 1.1 The Thesis

This thesis is the result of an industrial CASE (Cooperative Awards in Science and Engineering) research project, funded by the UK Engineering and Physical Sciences Research Council (EPSRC grant reference No. EP/G501750/1), which was born from the collaboration between the University of Stirling, the MIT Media Laboratory (Cambridge, USA) and Sitekit Labs (Portree, Scotland), the research branch of Sitekit Solutions Ltd., a software vendor specialised in content management system (CMS) development. The initial aim of the project was to further develop and apply software agent and natural language processing (NLP) technologies in order to blend the Open Mind database with any given ontology and, hence, build a novel intelligent software engine for text auto-categorisation. Soon after the start of the project, it has been consensually chosen to focus the research work on the fields of opinion mining and sentiment analysis, which have recently become more and more popular for their potential implications on areas such as e-commerce, e-tourism, and e-health. Although commonly used interchangeably to denote the same field of study, opinion mining and sentiment analysis actually focus on polarity detection and emotion recognition, respectively. Since the identification of sentiment is often exploited for detecting polarity, however, the two fields are usually combined under the same umbrella or even used as synonyms.

The primary researcher of the project was Erik Cambria, working under the principal academic supervision of Amir Hussain, founding Head of the Cognitive Signal Image Processing and Control Research (COSIPRA) Laboratory in the Department of Computing Science and Mathematics at Stirling University, and with industrial supervision of Chris Eckl, research director at Sitekit Labs. The research has been carried out in collaboration with Catherine Havasi of MIT Media Laboratory, who was part of the research team that pioneered the Common Sense Computing Initiative, which was further developed as part of the project in the context of opinion mining and sentiment analysis. In this section, motivations for the thesis are illustrated (subsection 1.1.1), together with the main aims of the research (subsection 1.1.2) and the original contributions arising from the doctoral work (subsection 1.1.3).

### 1.1.1 Motivations

Opinions play a primary role in decision-making processes. Whenever people need to make a choice, they are interested in hearing others' opinions. When this choice involves consuming valuable resources (e.g., spending time and money to buy products or services), in particular, people strongly rely on their peers' past experiences. Just a few years ago, the main sources for collecting such information were friends, acquaintances and, in some cases, specialised magazine or web sites.

The advent of Web 2.0 has provided people with new tools, e.g., forums, blogs, social network and contents sharing services, that allow them to create and share, in a time and cost efficient way, their own contents, ideas and opinions with virtually the millions of people connected to the World Wide Web. This has made available by click a new and oceanic source of information and opinions and has provided a powerful communication medium to share knowledge and to get advantage from others' experiences [4].

Currently, over 75,000 new blogs are created daily, along with 1.2 million new posts each day, and more and more people in the modern world rely on opinions, reviews and recommendations collected from these and related websites. The Web has made available the opinions of a vast pool of people that are neither our personal acquaintances nor well-known professional critics. People, in fact, are not just naturally keen on listening to others' advice but also naturally inclined to give others advice. Web users are often happy to share both their positive and negative real-world experiences for different reasons, e.g., because they benefited from others' reviews and want to give back to the community, because they seek for a sense of togetherness in adversity, for cathartic complaining, for supporting a product they really like, because it is a way to express themselves, because they think their opinions are important for others.

When people have a strong feeling about a specific product or service they tried, they feel like expressing it. If they loved it, they want others to enjoy it. If they hated it, they want to warn others away. This huge amount of useful information, however, is mainly unstructured, that is in natural language, as it is specifically produced for human consumption and, hence, it is not directly machine-processable. The opportunity to capture the opinions of the general public about social events, political movements, company strategies, marketing campaigns, and product preferences has raised more and more interest both in the scientific community, for the exciting open challenges, and in the business world, for the remarkable fallouts in marketing and financial market prediction.

This has led to the emerging fields of opinion mining and sentiment analysis, which deal with information retrieval and knowledge discovery from text using data mining and NLP techniques to distil knowledge and opinions from the huge amount of information on the World Wide Web. Mining opinions and sentiments from natural language, however, is an extremely difficult task as it involves a deep understanding of most of the explicit and implicit, regular and irregular, syntactical and semantic rules proper of a language.

Opinion mining is a branch of the broad field of text data mining [5] 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 [6, 7]. As the most natural form of storing information is text, opinion mining is believed to have a commercial potential higher than that of data mining. Opinion mining, 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-categorisation, machine learning, clustering and visualisation.

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 [8] and presence [9]. The latter is a binary-valued feature vectors in which the entries merely indicate whether a term occurs (value 1) or not (value 0) formed a more effective basis for review polarity classification. This is indicative of an interesting difference between typical topic-based text categorisation and polarity classification. While a topic is more likely to be emphasised by frequent occurrences of certain keywords, overall sentiment may not usually be highlighted through repeated use of the same terms.

Other term-based features are often added to the features vector. Position is one of these, in consideration of how the position of a token in a text unit can affect the way in which the token affect the sentiment of the text. Also the presence n-grams, typically bi-grams and tri-grams are often taken into account as useful features. Some methods also relies on the distance between terms. Part of speech (POS) information (nouns, adjectives, adverbs, verbs, etc.) is also commonly exploited in general textual analysis as a basic form of word sense disambiguation [10]. Certain adjectives, in particular, have been proved to be good indicators of sentiment and sometimes have been used to guide feature selection for sentiment classification. In other works, eventually, the detection of sentiments was performed through selected phrases, which were chosen via a number of pre-specified POS patterns, most including an adjective or an adverb [11]. All such approaches mainly rely on parts of text in which opinions and sentiments are explicitly expressed, e.g., polarity terms, affect words and their co-occurrence frequencies. Opinions and sentiments, however, are often conveyed implicitly through context and domain dependent concepts, which make purely syntactical approaches ineffective.

To this end, novel approaches that go beyond mere word-level sentiment analysis are needed. Such approaches should employ new techniques capable to better grasp the conceptual rules that govern sentiment and the clues that can convey these concepts from realisation to verbalisation in the human mind. Next-generation opinion mining systems need broader and deeper common sense knowledge bases and more cognitive and affective inspired reasoning methods, in order to better understand natural language opinions and sentiments and, hence, more efficiently bridge the gap between (unstructured) textual information and (structured) machine-processable data.

In this context, a novel approach at the crossroads between affective computing and common sense computing is proposed. Such an approach, termed sentic computing [12], exploits both computer and social sciences to better recognise, interpret and process opinions and sentiments over the Web. What led to the development of sentic computing, primarily, is the need for better accuracy in sentiment analysis when switching between different domains. Currently available keyword-based approaches, in fact, may perform nicely on a specific dataset but they have very low accuracy if the domain changes. Because sentic computing is based on latent/implicit features associated with concepts, it allows open-domain opinion mining. The novelty of the approach, in particular, lies in:

1. its multi-disciplinarity (not only computational, but also biologically-inspired and psychologically-motivated);
2. its semantic-based analysis (not only based on word co-occurrence frequencies, but also on the cognitive and affective information associated with natural language);
3. its fine-grained classification (not only at document, page or paragraph level, but also at sentence and clause level).

To achieve this, sentic computing involves the use of AI and Semantic Web techniques, for knowledge representation and inference; mathematics, for carrying out tasks such as graph mining and multi-dimensionality reduction; linguistics, for discourse analysis and pragmatics; psychology, for cognitive and affective modelling; sociology, for understanding social network dynamics and social influence; finally ethics, for understanding related issues about the nature of mind and the creation of emotional machines.

### 1.1.2 Aims

Today, opinion mining and sentiment analysis find applications in several different scenarios and there is a good number of companies, large and small, that include the analysis of opinions and sentiments as part of their mission. In current product review websites, such as Epinions[1], Yelp[2], and RateItAll[3], feedback and reviews are explicitly solicited within the web interface. Opinion mining techniques can be exploited for the creation and automated upkeep of review and opinion aggregation websites, in which opinions are continuously gathered from the Web and not restricted to just product reviews but also to wider topics such as political issues and brand perception. Opinion mining and sentiment analysis have also a great potential as sub-component technology for other system. They can enhance the capabilities of customer relationship management (CRM) and recommendation systems allowing, for example, to find out which

[1] http://epinions.com
[2] http://yelp.com
[3] http://rateitall.com

features customers are particularly interested in or to exclude from the recommendations items that have received very negative feedbacks [13, 14]. Similarly they can be used in email or other types of communication to detect and exclude 'flames', i.e., overly heated or antagonistic language, and to enhance anti-spam systems. Also, online systems that display advertisements as sidebars can use opinion mining techniques to detect web-pages that contain sensitive content inappropriate for ads placement [15].

Business intelligence is also one of the main factors behind corporate interest in the field of sentiment analysis [16]. Nowadays, companies invest more and more money in marketing strategies and they are constantly interested in both collecting and predicting the opinions and the attitudes of the general public towards their products and brands. The design of automatic tools capable to crawl reviews and opinions over the Web in real-time and to create condensed versions of them represents one of the most active research and development area. Several companies, in fact, already provide tools to track public viewpoints on a large scale by offering graphical summarisations of trends and opinions in the blogosphere (Table 1.1). The development of such systems, moreover, is not only important for commercial purposes but also for government intelligence applications able to monitor increases in hostile or negative communications [17]. All of these tools, however, are still mainly keyword based and, hence, often fail to meet the gold standards of human annotators. The fundamental aim of this thesis is to go beyond such approaches by developing two common sense knowledge bases to bridge the cognitive and affective gap between word-level natural language data and the concept-level opinions conveyed by these. Unlike keyword-based methods, sentic computing uses affective ontologies and common sense reasoning tools for a concept-level analysis of natural language text. Specifically, the ensemble application of graph mining and multi-dimensionality reduction techniques is employed, together with a novel emotion categorisation model, on two common sense knowledge bases, in order to design an open-domain opinion mining engine capable to infer the cognitive and affective information associated with natural language text.

| Company | Founded | Headquarters | Web Link |
|---|---|---|---|
| Vocus | 1992 | USA | http://vocus.com |
| Kantar | 1993 | UK | http://www.kantar.com |
| Cymphony | 1996 | USA | http://www.cymfony.com |
| Alterian | 1997 | USA | http://alterian.com |
| Factiva | 1999 | USA | http://dowjones.com/factiva |
| Brandimensions | 2001 | Canada | http://brandprotect.com |
| Attensity | 2000 | USA | http://attensity.com |
| Converseon | 2001 | USA | http://converseon.com |
| Lithium | 2001 | USA | http://lithium.com |
| Lexalytics | 2003 | USA | http://lexalytics.com |
| MotiveQuest | 2003 | USA | http://www.motivequest.com |
| Visible Technologies | 2003 | USA | http://visibletechnologies.com |
| Evolve24 | 2004 | USA | http://evolve24.com |
| Clarabridge | 2005 | USA | http://clarabridge.com |
| Collective Intellect | 2005 | USA | http://collectiveintellect.com |
| Radian6 | 2006 | Canada | http://radian6.com |
| Rapid-I | 2006 | UK | http://rapid-i.com |
| Luminoso | 2011 | USA | http://lumino.so |

Table 1.1: List of most popular companies that are leveraging sentiment analysis tools to track and dissect how consumers feel about products and services of their own and also of the competition.

Such engine has been exploited for the development of emotion-sensitive systems in fields such as social data mining, multimedia management, personalisation and persuasion, human-computer interaction, intelligent user interfaces, social media marketing, and patient-centred applications. In order to evaluate the different facets of the engine from different perspectives, three different resources, namely a Twitter[4] hashtag repository, a LiveJournal[5] database and a PatientOpinion[6] dataset, were used and results obtained using Princeton's WordNet[7], MIT's ConceptNet[8] and Microsoft's Probase[9] were compared. The first resource is a collection of 3,000 tweets crawled from Bing[10] web repository by exploiting Twitter hashtags as category labels, which is useful to test the engine's target spotting performance. In particular, hashtags about electronics (e.g., IPhone, XBox, Android and Wii), companies (e.g., Apple, Microsoft and Google),

[4] http://twitter.com
[5] http://livejournal.com
[6] http://patientopinion.org.uk
[7] http://wordnet.princeton.edu
[8] http://conceptnet5.media.mit.edu
[9] http://research.microsoft.com/probase
[10] http://bing.com

countries, cities, operative systems and cars were selected. In order to test the resource's consistency and reliability, a manual evaluation of 100 tweets was performed, which showed that hashtags are accurate to 89%.

The second resource is a 5,000 blogpost database extracted from LiveJournal, a virtual community of more than 23 millions 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. Since the indication of mood tags is optional, posts are likely to reflect the true mood of the authors, which is not always true for category tags. After a manual evaluation of 200 posts, in fact, the category tags turned out to be very noisy (53% accuracy). The mood tags, however, showed a good enough reliability (89% accuracy) so they were used to test the engine's affect recognition performance. The third resource, eventually, is a dataset obtained from PatientOpinion, a social enterprise pioneering an online feedback service for users of the UK national health service to enable people to share their recent experience of local health services online. 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. It was used to test the detection of opinion targets and the polarity associated with these.

There are no ethical issues involved in the data used in the experimentation as tweets, blogposts, and patient opinions were all anonymised. In order to guarantee full anonymity, moreover, the text associated with tweets, blogposts, and patient opinions has never been wholly reported in the proposed tables and examples.

### 1.1.3 Original Contributions

Relying solely on traditional methods to develop computer systems with a new set of affect-sensitive functionalities is insufficient [18] because today user emotions are still far from being on the radar of computing methods. This is where insights gleaned from a century and a half of scientific study on human emotions can become useful for the development of affect-sensitive interfaces. Despite the extensive literature in emotion research, however, the affective computing literature has been primarily driven by computer scientists and AI researchers who have remained agnostic to the controversies inherent in the underlying psychological theory. Instead, they have focused their efforts on the technical challenges of developing emotion-sensitive computer interfaces. However, ignoring the important debates has significant limitations because a functional affective computing application can never be completely divorced from underlying emotion theory [19].

Blending scientific theories of emotion with the practical engineering goals of analysing sentiments in natural language text and developing affect-sensitive interfaces is one of the main contributions of this thesis. Recently, many research activities focusing on the extraction of cognitive and affective information from natural language text have gained ground under the umbrella of opinion mining and sentiment analysis. The reason of this trend lies on the ever-growing amount of valuable data available through the Web in the form of news, reviews, blogs, chats, tweets, etc. Sentiment analysis, however, is a multi-faceted and multi-disciplinary problem that requires a deep understanding of natural language.

Existing reported solutions and currently available systems are still far from perfect or fail to meet the satisfaction level of the end users. The main issue may be that there are many conceptual rules that govern sentiment and the possibly unlimited clues that can convey these concepts from realisation to verbalisation in the human mind. Recent efforts in this context have been carried about through research works published in reputed conferences through special tracks and workshops, e.g., *TREC-BLOG* tracks since 2006, *Sentiment and Subjectivity in Text* workshop in COLING-ACL 2006, *SemEval 2007 Task#14: Affective Text*, *TAC 2008 Opinion Summarisation* task, *Emotion, Metaphor, Ontology and Terminology* (EMOT) in LREC 2008, *Social Data on the Web* (SDoW) workshop from 2008, *Workshop on Opinion Mining and Sentiment*

*Analysis* (WOMSA) in 2009, *Topic-Sentiment Analysis for Mass Opinion Measurement* (TSA) in CIKM 2009, *Computational Approaches to Analysis and Generation of Emotion in Text* in NAACL 2010, *Workshop on Computational Approaches to Subjectivity and Sentiment Analysis* (WASSA) in ECAI 2010 and in ACL 2011.

The research work hereby presented has laid the basis for a novel multi-disciplinary approach to opinion mining and sentiment analysis, namely sentic computing, that exploits both computer and social sciences to better recognise, interpret and process opinions and sentiments over the Web. In particular, two common sense knowledge bases for concept-level opinion mining have been developed, together with novel graph mining and multi-dimensionality reduction techniques to perform reasoning on it, to enable the analysis of text not only at document, page or paragraph level but also at sentence and clause level. Evidence of the impact of the approach is found in the presence of sentic computing in high impact factor journals and top AI conferences (see section 1.2), and in its adoption by several leading American, British, and Asian companies, including: Zoral Inc., Luminoso Inc., Abies Ltd., Patient Opinion Ltd., Sitekit Solutions Ltd., HP Labs India, and Microsoft Research Asia. For these reasons, sentic computing has also been recently put forward as impact case study to the UK Research Excellence Framework (REF) by the University of Stirling.

## 1.2 Publications Arising

The following papers have resulted from the research presented in this thesis. In particular, subsection 1.2.1 lists the publications relevant to the thesis work that have been accepted for publication as books, journal papers, book chapters, conference and workshop proceedings, subsection 1.2.2 cites the research work that has been submitted and is currently under review, subsection 1.2.3 finally names those papers that are being currently edited and have not been submitted yet.

### 1.2.1 Accepted for Publication

1. E. Cambria, M. Grassi, A. Hussain and C. Havasi. *Sentic Computing for Social Media Marketing*[11]. In press: Multimedia Tools and Applications, Springer-Verlag (2012)

2. E. Cambria, T. Mazzocco, A. Hussain and C. Eckl. *Sentic Medoids: Organising Affective Common Sense Knowledge in a Multi-Dimensional Vector Space*[12]. LNCS, vol. 6677, pp. 601-610. Springer-Verlag, Berlin Heidelberg (2011)

3. E. Cambria, A. Hussain and C. Eckl. *Bridging the Gap Between Structured and Unstructured Health-Care Data Through Semantics and Sentics*[13]. In: Proceedings of ACM WebSci, Koblenz (2011)

4. E. Cambria, I. Hupont, A. Hussain, E. Cerezo and S. Baldassarri. *Sentic Avatar: Multi-Modal Affective Conversational Agent with Common Sense*[14]. LNCS, vol. 6456, pp. 81-95. Springer-Verlag, Berlin Heidelberg (2011)

5. E. Cambria, A. Hussain and C. Eckl. *Sentic Robots.* In: Proceedings of ICMC, pp. 150, Venice (2011)

6. A. Lascu, E. Cambria, M. Grassi and S. Negulescu. *Human Semiotics Ontology.* In: Proceedings of ICMC, pp. 152, Venice (2011)

7. M. Grassi, E. Cambria, A. Hussain and F. Piazza. *Sentic Web: A New Paradigm for Managing Social Media Affective Information*[15]. Cognitive Computation 3(3), pp. 480-489, Springer-Verlag (2011)

8. E. Cambria, R. Speer, C. Havasi and A. Hussain. *SenticNet: A Publicly Available Semantic Resource for Opinion Mining*[16]. In: Proceedings of AAAI CSK, pp. 14-18, Arlington (2010)

[11] http://springerlink.com/content/q1vq625w2x27x4r7
[12] http://springerlink.com/content/y87t46v473528w4x
[13] http://journal.webscience.org/478/1/94_paper.pdf
[14] http://springerlink.com/content/v408824460u4kl46
[15] http://springerlink.com/content/v730647036834122
[16] http://www.aaai.org/ocs/index.php/FSS/FSS10/paper/download/2216/2617.pdf

9. E. Cambria, P. Chandra, A. Sharma and A. Hussain. *Do Not Feel The Trolls*[17]. In: Proceedings of ISWC, Shanghai (2010)

10. E. Cambria, A. Hussain, T. Durrani, C. Havasi, C. Eckl and J. Munro. *Sentic Computing for Patient Centered Applications*[18]. In: Proceedings of IEEE ICSP, pp. 1279-1282, Beijing (2010)

11. E. Cambria, A. Hussain, C. Havasi and C. Eckl. *SenticSpace: Visualising Opinions and Sentiments in a Multi-Dimensional Vector Space*[19]. LNAI, vol. 6279, pp. 385-393. Springer-Verlag, Berlin Heidelberg (2010)

12. E. Cambria, A. Hussain, C. Havasi, C. Eckl and J. Munro. *Towards Crowd Validation of the UK National Health Service*[20]. In: Proceedings of ACM WebSci, Raleigh (2010)

13. E. Cambria, A. Hussain, C. Havasi and C. Eckl. *Sentic Computing: Exploitation of Common Sense for the Development of Emotion-Sensitive Systems*[21]. LNCS, vol. 5967, pp. 148-156. Springer-Verlag, Berlin Heidelberg (2010)

14. E. Cambria, A. Hussain, C. Havasi and C. Eckl. *AffectiveSpace: Blending Common Sense and Affective Knowledge to Perform Emotive Reasoning*[22]. In: Proceedings of CAEPIA, pp. 32-41, Seville (2009)

15. E. Cambria, A. Hussain, C. Havasi and C. Eckl. *Common Sense Computing: From the Society of Mind to Digital Intuition and Beyond*[23]. LNCS, vol. 5707, pp. 252-259. Springer-Verlag, Berlin Heidelberg (2009)

16. E. Cambria, A. Hussain, C. Havasi and C. Eckl. *Application of Common Sense*

[17] http://sdow.semanticweb.org/2010/pub/sdow2010_paper_1.pdf
[18] http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5657072
[19] http://springerlink.com/content/t6515hw286334534
[20] http://journal.webscience.org/352/2/websci10.pdf
[21] http://springerlink.com/content/9305u22257427j24
[22] http://scholar.tdg-seville.info/Resources/Cambria2009.pdf
[23] http://springerlink.com/content/x24367q25p221p75

*Computing to Enable the Development of Next-Generation Semantic Web Applications*[24]. In: Proceedings of ACM WebSci, Athens (2009)

### 1.2.2 Under Review

1. E. Cambria and A. Hussain. *Sentic Album: Content, Concept and Context Based Online Personal Photo Management System*. Submitted to: Cognitive Computation, Springer-Verlag

2. E. Cambria, A. Hussain, C. Havasi, and C. Eckl. *SenticNet 2: A Semantic and Affective Resource for Opinion Mining and Sentiment Analysis*. Submitted to: IEEE Transactions on Affective Computing

3. E. Cambria, A. Hussain, T. Benson, and C. Eckl. *Sentic PROMs: Application of Sentic Computing to the Development of a Novel Unified Framework for Measuring Health-Care Quality*. Submitted to: Elsevier Expert Systems with Applications

4. E. Cambria, A. Hussain and C. Eckl. *Open Mind Common Sentics: Collecting Affective Common Sense Knowledge*. Submitted to: LNCS, Springer-Verlag, Berlin Heidelberg

5. E. Cambria, A. Hussain and A. Livingstone. *The Hourglass of Emotions*. Submitted to: LNCS, Springer-Verlag, Berlin Heidelberg

6. E. Cambria, A. Hussain, Y. Song and H. Wang. *Blending Common and Common Sense Knowledge for Open-Domain Sentiment Analysis*. Submitted to: KR12, Rome

7. E. Cambria and A. Hussain. *Sentic Agents*. Submitted to: ICACII11, Taipei

8. E. Cambria, A. Hussain, H. Atassi, A. Abel and M. Grassi. *Towards IMACA: Intelligent Multimodal Affective Conversational Agent*. Submitted to: LREC12,

[24] `http://journal.webscience.org/240`

Istanbul

9. I. Hupont, E. Cambria, A. Hussain, E. Cerezo, and S. Baldassarri. *Sentic Maxine: Multi-Modal Affective Fusion and Emotional Paths*. Submitted to: ACM Transactions on Interactive Intelligent Systems

10. E. Cambria, M. Grassi and A. Hussain. *Sentic Computing for Social Media Analysis, Representation and Retrieval*. Submitted to: Computer Communications and Networks, Springer-Verlag

### 1.2.3 In Preparation

1. E. Cambria and A. Hussain. *A Two-Level Architecture for Affective Common Sense Reasoning*. To be submitted to: AAAI12, Toronto

2. E. Cambria, A. Hussain, and P. Chandra. *Clustering Social Networks Using Interaction Semantics and Sentics*. To be submitted to: ISNN12, Shenyang

3. E. Cambria, J. Zhang and A. Hussain. *Application of Machine Translation Techniques for Common Sense Knowledge Transfer*. To be submitted to: CSK12, Boston

4. E. Cambria, T. Mazzocco and A. Hussain. *Sentic Neural Networks: a Novel Cognitive Model for Affective Common Sense Reasoning*. To be submitted to: BICS12, Shenyang

5. E. Cambria, A. Hussain and E. Yang. *Application of Sentic Computing for the Development of Emotion-Sensitive Robots*. To be submitted to: IEEE ICSP12, Beijing

6. Q. Wang, E. Cambria and A. Hussain. *Common-Sense-Aided Pattern Recognition*. To be submitted to: BICS12, Shenyang

7. E. Cambria, A. Hussain and G. Huang. *ELM-Based Affective Common Sense Reasoning*. To be submitted to: ELM12, Singapore

# Chapter 2

# Background

*The good opinion of mankind,*
*like the lever of Archimedes,*
*with the given fulcrum, moves the world.*

*Thomas Jefferson*

The World Wide Web represents one of the most revolutionary applications in the history of computing and human communication, which is keeping on changing how information is disseminated and retrieved, how business is conducted and how people communicate with each other. As the dimension of the Web increases, the technologies used in its development and the services provided to its users are developing constantly. Even if just few years have passed, in fact, Web 1.0's static and read-only HTML pages seem now just an old memory. Today the Web has become a dynamic and interactive reality in which more and more people actively participate by creating, sharing and consuming contents. In this way, the World Wide Web configures itself not only as a 'Web of data' but also as a 'Web of people' where data and users are interconnected in an unbreakable bond.

This chapter shows how and why online opinions are important in the Web 2.0 era (section 2.1) and illustrates existing approaches and depths of analysis in mining and characterising such opinions (section 2.2). Eventually, the chapter comprises a background section on common sense computing, which is hereby exploited to go beyond merely syntactical approaches to sentiment analysis (section 2.3), and some concluding remarks (section 2.4).

## 2.1 Opinion Mining and Sentiment Analysis

The passage from a read-only to a read-write Web made users more enthusiastic about interacting, sharing and collaborating through social networks, online communities, blogs, wikis and other online collaborative media. In the last years this collective intelligence has spread to many different areas of the Web, with particular focus on fields related to our everyday life such as commerce, tourism, education and health.

The online review of commercial services and products, in particular, is an action that users usually perform with pleasure, to share their opinions about services they have received or products they have just bought, and it constitutes immeasurable value for other potential buyers. This trend opened new doors to enterprises that want to reinforce their brand and product presence in the market by investing in online advertising and positioning, that is in social media marketing. The reasons why opinion mining is attracting so much attention from both the academic and the business world, in particular, can be found in the dynamics behind the buzz mechanism (subsection 2.1.1), in the motivating factors that gave birth to the field (subsection 2.1.2), and in the sub-tasks that make it different from standard information retrieval (subsection 2.1.3).

### 2.1.1 The Buzz Mechanism

What mainly makes social media marketing work is the buzz mechanism [20]. A buzz replicates a message through user-to-user contact, rather than purchasing some advertising or promoting a press release. The message does not have to necessarily deal with the product. Many successful viral campaigns, in fact, have spread thanks to a compelling message, with the company logo included incidentally. At the heart of buzz is an understanding that the natural, spontaneous networks that comprise the social universe are the most effective means of reaching people in a meaningful way. The power of marketing lies, therefore, not in pushing information to the masses but in effectively tapping those individuals who wield influence over others. The marketers who are winning are the ones using consumers and culture to their advantage, crafting messages with consumers rather than throwing messages at them.

In confirmation of the growing interest in this novel approach to marketing, several academic and commercial tools, e.g., OASYS[1] [21], ESSE [22], Luminoso[2] [23], Factiva[3], NM Incite[4], Attensity[5], and Converseon[6], have been developed to provide companies (and users) with a way to analyse the blogosphere on a large scale in order to extract information about the trend of the opinions relative to their products. Nevertheless most of the existing tools and the research efforts are limited to a polarity evaluation or a mood classification according to a very limited set of emotions. In addition, such methods mainly rely on parts of text in which emotional states are explicitly expressed and hence they are unable to capture opinions and sentiments that are expressed implicitly.

### 2.1.2 Origins and Peculiarities

The term 'opinion mining' first appears in a paper by Dave et al. [24] dated 2003, which envisioned the ideal opinion mining tool as capable to "process a set of search results for a given item, generating a list of product attributes (quality, features, etc.) and aggregating opinion about each of them (poor, mixed, good)". From this early definition, the term opinion mining has been later extended to refer more generally

[1] `http://oasys.umiacs.umd.edu/oasys`
[2] `http://lumino.so`
[3] `http://dowjones.com/factiva`
[4] `http://nmincite.com`
[5] `http://attensity.com`
[6] `http://converseon.com`

to the computational techniques for extracting, classifying, understanding and assessing the opinions expressed in various online news sources, social media comments and other user-generated contents (UGCs). The introduction of the term 'sentiment' to the automatic analysis of evaluative text and tracking of the predictive judgements was first introduced in 2001 by Das and Chen [25] and Tong [26] in the context of market sentiment analysis. In the context of NLP, the term sentiment can be intended either as the emotions or the polarity conveyed by text. Strictly speaking, sentiment analysis consists in inferring affective information from text while opinion mining mainly concerns polarity detection. However, since the identification of sentiment, affect, subjectivity, and other emotional states is often propaedeutic to polarity detection [27], opinion mining and sentiment analysis are strictly connected and, therefore, commonly used interchangeably to denote the same field of study.

The manifesto of opinion mining and sentiment analysis as a unified field can be seen in the extensive review paper published by Pang and Lee [16] in 2008. This survey covers techniques and approaches that promise to directly enable opinion-oriented information-seeking systems. The authors' focus is on methods that seek to address the new challenges raised by sentiment-aware applications, as compared to those that are already present in more traditional fact-based analysis. They include material on summarisation of evaluative text and on broader issues regarding privacy, manipulation, and economic impact that the development of opinion-oriented information-access services gives rise to. To the inexpert eye, opinion mining and sentiment analysis might look like the same as fields such as traditional text mining or fact-based analysis. Moreover, since sentiment classification deals with a relatively small number of classes, it might look like an easy task compared to text auto-categorisation.

Opinion mining, however, is a very complex task even at its more basic level of sentiment polarity classification, which is a case of binary classification. The extraction of opinion polarity from text can be performed by comparing words extracted from text with a set of keywords with positive valence (e.g., love, wonderful, best, great, superb, still, beautiful) and negative valence (e.g., bad, worst, stupid, waste, boring), as in the case of topic-based binary classification. The identification of a right set of keywords for mining opinions, however, is not a trivial task. Even when machine learning techniques are employed to select keywords from training corpora, the level of accuracy is still very low if compared to the performance of typical topic-based binary classification [9]. The main reason is that, differently from topics, sentiments can often be expressed in a more subtle manner, making it difficult to be identified by any of a sentence or document's terms when considered in isolation.

In addition, sentiment and subjectivity are quite context and domain dependent. This is true not only for changes in vocabulary but also because even the exact same expression can indicate different sentiment in different domains. The concept 'go read the book', for example, most likely indicates positive sentiment for book reviews, but negative sentiment for movie reviews; as well as the adjective unpredictable may have a negative orientation in a car review (e.g., 'unpredictable steering'), but it could have a positive orientation in a movie review (e.g., 'unpredictable plot').

### 2.1.3 Sub-Tasks

One of the most common sub-tasks of opinion mining is polarity classification and the assignment of degrees of positivity, that is, given an opinionated piece of text wherein it is assumed that the overall opinion is about one single issue or item, classify the opinion as falling under one of two opposing sentiment polarities, or locate its position on the continuum between these two polarities. Much work on sentiment polarity classification has been conducted in the context of reviews of evaluative opinions (e.g., 'thumbs up' versus 'thumbs down' or 'like' versus 'dislike').

In addition, polarity classification can be also applied to identifying 'pro and con' expressions that can be used in individual reviews to evaluate the pros and cons that have influenced the judgements of a product and that make such judgements more trustworthy. Another instance of binary sentiment classification is agreement detection, that is, given a pair of text documents, deciding whether they should receive the same or differing sentiment-related labels. The more general problem of rating inference, where one must determine the author's evaluation with respect to a multi-point scale (e.g., one to five stars for a review) can be viewed as a multi-class text categorisation problem. Other common sub-tasks of opinion mining and sentiment analysis are subjectivity detection and opinion identification. The capability of distinguishing if a text, or parts of it, are subjective or objective can be particularly beneficial for a more effective sentiment classification. Mihalcea et al. showed evidence that the complexity of this task is superior than subsequent polarity classification [28]. Wilson et al. remarked how classifying a piece of text as expressing a neutral opinion for rating inference does not equal classifying that piece of text as objective [29]. A piece of text can also have a polarity without necessarily containing an opinion, for example a news article can be classified into good or bad news without being subjective.

The classification of a piece of text as subjective or objective can be useful in several situations. For example, being able to distinguish in opinionated texts where the authors do explicitly express their sentiment through statements (e.g., "this laptop is great") and where they provide objective information (e.g., "the laptop has long battery life") is used to help determine the overall sentiment. Hatzivassiloglou and Wiebe examined the effects of adjective orientation and gradability on sentence subjectivity to detect if a sentence is subjective [30] while other projects address subjectivity detection at sub-sentence level. Wiebe et al. presented a comprehensive survey of subjectivity recognition using different clues and features [31].

Typically, sentiment analysis is performed over an on-topic document, e.g., on the result of a topic-based search engine. However, several studies suggested that managing these two task jointly can be beneficial for the overall performance. According to Riloff et al., topic-based text filtering and subjectivity filtering are complementary, in the context of experiments in information extraction [32]. For example, off-topic passages of a document could contain irrelevant affective information and result misleading for the global sentiment polarity about the main topic. Also, a document can contain material on multiple topics that may be of interest to the user. In this case, it is therefore necessary to identify the topics and separate the opinions associated with each of them. Several other researches in sentiment analysis focus on non-topic based categorisation, for example to classify documents according to their genre [33] and their style [34]. Also authorship and publisher identification are other relevant examples [35, 36]. Another problem that has been considered in intelligence and security settings is the detection of deceptive language. Affect detection, eventually, is also a task that is gaining a growing attention from different perspectives and for different applications.

Sentiment analysis has been traditionally more focused on the extraction of the valence of textual sample (i.e., positive/negative or bad/good) rather than assigning a particular emotion category to text. However, the classification of multimedia resources according to their mood and emotional content is also quite common. The advent of Web 2.0 has pushed the users at the centre of the Web universe, providing them revolutionary tools that have changed the way people communicate and express themselves, their ideas and emotions. People spend more and more time using the Web not only for work but also for expressing their opinions on blogs and forums, chatting and organising events through social networks, and even for living a Second Life[7]. Therefore, the Web contains more and more affective content. The awareness that the capability to manage such affective content can be exploited for the development of next-generation web applications is dragging a growing attention also in sentiment analysis for affect extraction from textual Web content.

## 2.2 Main Approaches to Opinion Mining

Several approaches have been developed for the general task of mapping a given piece of text to a label belonging to a predefined set of categories, or to a real number representative of a polarity degree. Such approaches and their performance, however,

[7] `http://secondlife.com`

are strictly bound to the considered domain of application and to the related topics. Moreover, most of the literature on sentiment analysis has focused on text written in English and consequently most of the resources developed, such as lexicons with sentiment labels, are in English. Adapting such resources to other languages can be considered as a domain adaptation problem. This section discusses the evolution of different approaches from heuristics to discourse structure (subsection 2.2.1), from coarse to fine grained analysis (subsection 2.2.2), from keyword to concept level opinion mining (subsection 2.2.3).

### 2.2.1 From Heuristics to Discourse Structure

Several unsupervised learning approaches rely on the creation of a sentiment lexicon in an unsupervised manner that is later used to determine the degree of positivity (or subjectivity) of a text unit. The crucial component is, therefore, the creation of the lexicon via the unsupervised labelling of words or phrases with their sentiment polarity or subjectivity [16]. This lexicon can be used to identify the *prior polarity* or the *prior subjectivity* of terms or phrases, to use towards further identifying contextual polarity or subjectivity. Early works were mainly based on linguistic heuristics. For example, Hatzivassiloglou and McKeown's technique [37] was built on the fact that, in the case of polarity classification, the two classes of interest represent opposites, and 'opposition constraints' can be used to help labelling decisions.

Other works propagated the valence of seed words, for which the polarity is known, to terms that co-occur with them in general text or in dictionary glosses, or to synonyms and words that co-occur with them in other WordNet-defined relations. A collective labelling approach can also be applied to opinion about product features. Popescu and Etzioni [38] proposed an iterative algorithm that, starting from a global word label computed over a large collection of generic topic text, gradually tried to re-define such label first to one that is specific to a review corpus then to one that is specific to a given product feature, and finally to one that is specific to the particular context in which the word occurs. Also Snyder and Barzilay [39] exploited the idea of utilising discourse information to aid the inference of relationships between product attributes. They designed a linear classifier for predicting whether all aspects of a product are given the same rating, and combined such prediction with that of individual-aspect classifiers, in order to minimise a certain loss function. Regression techniques are often employed for the prediction of the degree of positivity in opinionated documents such as product reviews. Regression, in fact, allows to implicitly model similarity relationships between classes that correspond to points on a scale, such as the number of 'stars' given by a reviewer [16]. Modelling discourse structure, such as twists and turns in documents, contributes to a more effective overall sentiment labelling.

Early works attempted to partially address this problem via incorporating location information in the feature set [40]. More recent studies have underlined that position is particularly relevant in the context of sentiment summarisation. In particular, in contrast to topic-based text summarisation, where the incipits of articles usually serve as a strong baseline, the last $n$ sentences of a review have been shown to serve as a much better summary of the overall sentiment of the document, and to be almost as good as the $n$ (automatically-computed) most subjective sentences [40]. Joshi and Rose, for example, explored how features based on syntactic dependency relations can be utilised to improve performance on opinion mining [41]. Using a transformation of dependency relation triples, they convert them into 'composite back-off features' that generalise better than the regular lexicalised dependency relation features.

### 2.2.2 From Coarse to Fine Grained

The evolution of research works in the field of opinion mining and sentiment analysis can be seen not only in the use of more and more sophisticated techniques but also in the different depths of analysis adopted. Early works, in fact, aimed to classify entire documents as containing overall positive or negative polarity [9] or rating scores (e.g., 1-5 stars) of reviews [42]. These were mainly supervised approaches relying on manually labelled samples, such as movie or product reviews where the opinionist's overall positive or negative attitude was explicitly indicated. However, opinions and sentiments do not occur only at document level, nor are they limited to a single valence or target. Contrary or complementary attitudes toward the same topic or multiple topics can be present across the span of a document. Later works adopted a segment level opinion analysis aiming to distinguish sentimental from non-sentimental sections, e.g., by using graph-based techniques for segmenting sections of a document on the basis of their subjectivity [40], or by performing a classification based on some fixed syntactic phrases that are likely to be used to express opinions [11], or by bootstrapping using a small set of seed opinion words and a knowledge base such as WordNet [43].

In recent works, text analysis granularity has been taken down to sentence level, e.g., by using presence of opinion-bearing lexical items (single words or n-grams) to detect subjective sentences [44, 45], or by using semantic frames defined in FrameNet [46] for identifying the topics (or targets) of sentiment [47], or by exploiting association rule mining [48] for a feature-based analysis of product reviews [49]. Commonly, a certain degree of continuity exists in subjectivity labels of adjacent sentences, as an author usually does not switch too frequently between being subjective and being objective. Hence, some works also propose a collective classification of the document based on assigning preferences for pairs of nearby sentences [42, 50].

All such approaches, however, are still far from being able to infer the cognitive and affective information associated with natural language as they mainly rely on semantic knowledge bases which are still too limited to efficiently process text at sentence level. Moreover, text analysis granularity might still not be enough as a single sentence may express more than one opinion [29].

### 2.2.3 From Keywords to Concepts

Existing approaches can be grouped into four main categories, with few exceptions: keyword spotting, lexical affinity, statistical methods and hand-crafted models. Keyword spotting is the most naïve approach and probably also the most popular because of its accessibility and economy. Text is classified into affect categories based on the presence of fairly unambiguous affect words like 'happy', 'sad', 'afraid' and 'bored'. Elliott's Affective Reasoner [51], for example, watches for 198 affect keywords, e.g., 'distressed', 'enraged', plus affect intensity modifiers, e.g., 'extremely', 'somewhat', 'mildly', plus a handful of cue phrases, e.g., 'did that', 'wanted to'. Other popular sources of affect words are Ortony's Affective Lexicon [52], which groups terms into affective categories, and Wiebe's linguistic annotation scheme [53]. The weaknesses of this approach lie in two areas: poor recognition of affect when negation is involved and reliance on surface features. About its first weakness, while the approach can correctly classify the sentence "today was a happy day" as being happy, it is likely to fail on a sentence like "today wasn't a happy day at all". About its second weakness, the approach relies on the presence of obvious affect words which are only surface features of the prose.

In practice, a lot of sentences convey affect through underlying meaning rather than affect adjectives. For example, the text "My husband just filed for divorce and he wants to take custody of my children away from me" certainly evokes strong emotions, but uses no affect keywords, and therefore, cannot be classified using a keyword spotting approach. Lexical Affinity is slightly more sophisticated than keyword spotting as, rather than simply detecting obvious affect words; it assigns arbitrary words a probabilistic 'affinity' for a particular emotion. For example, 'accident' might be assigned a 75% probability of being indicating a negative affect, as in 'car accident' or 'hurt by accident'. These probabilities are usually trained from linguistic corpora [54, 55, 56, 57]. Though often outperforming pure keyword spotting, there are two main problems with the approach. First, lexical affinity, operating solely on the word-level, can easily be tricked by sentences like "I avoided an accident" (negation) and "I met my girlfriend by accident" (other word senses). Second, lexical affinity probabilities are often biased toward text of a particular genre, dictated by the source of the linguistic corpora. This makes it difficult to develop a reusable, domain-independent model.

Statistical methods, such as latent semantic analysis (LSA) and support vector machine (SVM), have been popular for affect classification of texts and have been used by researchers on projects such as Goertzel's Webmind [58], Pang's movie review classifier [9], and many others [42, 49, 59, 60, 61]. By feeding a machine learning algorithm a large training corpus of affectively annotated texts, it is possible for the systems to not only learn the affective valence of affect keywords as in the keyword spotting approach, but such a system can also take into account the valence of other arbitrary keywords (like lexical affinity), punctuation and word co-occurrence frequencies.

However, statistical methods are generally semantically weak, meaning that, with the exception of obvious affect keywords, other lexical or co-occurrence elements in a statistical model have little predictive value individually. As a result, statistical text classifiers only work with acceptable accuracy when given a sufficiently large text input. So, while these methods may be able to affectively classify user's text on the page or paragraph level, they do not work well on smaller text units such as sentences.

## 2.3 Towards Machines with Common Sense

Communication is one of the most important aspects of human life. Communicating has always a cost in terms of energy and time, since information needs to be encoded, transmitted and decoded, and sometimes these factors can even make the difference between life and death. This is why people, when communicating with each other, provide just the useful information and take the rest for granted. This 'taken for granted' information is what we call common sense – obvious things people normally know and usually leave unstated. Common sense is not the kind of knowledge that we can find in Wikipedia[8] but it consists in all the basic relationships among words, concepts, phrases and thoughts that allow people to communicate with each other and face everyday life problems. It is a kind of knowledge that sounds obvious and natural to us but it is actually daedal and multi-faceted.

[8] `http://wikipedia.org`

The illusion of simplicity comes from the fact that, as each new group of skills matures, we build more layers on top of them and tend to forget about the previous layers. Common sense, in fact, is not a simple thing. Instead, it is an immense society of hard-earned practical ideas, of multitudes of life-learned rules and exceptions, dispositions and tendencies, balances and checks [62]. This section discusses the importance of common sense for the development of intelligent systems (subsection 2.3.1) and illustrates different knowledge representation strategies (subsection 2.3.2). The section also refers to a survey on common sense computing, proposed by Cambria et al. [63], to present the evolution of such research field from logic-based approaches (subsection 2.3.3) to more recent methods based on natural language techniques (subsection 2.3.4), e.g., sentic computing (subsection 2.3.5).

### 2.3.1 The Importance of Common Sense

Concepts are the glue that holds our mental world together [64]. Without concepts, there would be no mental world in the first place [65]. Doubtless to say, the ability to organise knowledge into concepts is one of the defining characteristics of human mind. Of the different sorts of semantic knowledge that are researched, arguably the most general and widely applicable kind is knowledge about the everyday world that is possessed by all people, i.e., common sense knowledge. While to the average person the term common sense is regarded as synonymous with good judgement, to the AI community it is used in a technical sense to refer to the millions of basic facts and understandings possessed by most people, e.g., "a lemon is sour", "to open a door, you must usually first turn the doorknob", "if you forget someone's birthday, they may be unhappy with you".

Common sense knowledge, thus defined, spans a huge portion of human experience, encompassing knowledge about the spatial, physical, social, temporal, and psychological aspects of typical everyday life. Because it is assumed that every person possesses common sense, such knowledge is typically omitted from social communications, such as text. A full understanding of any text then, requires a surprising amount of common sense, which currently only people possess. Common sense knowledge is what we learn and what we are taught about the world we live in during our formative years, in order to better understand and interact with the people and the things around us. Common sense is not universal but cultural and context dependent. The importance of common sense can be particularly appreciated when travelling to far away places, where sometimes it is necessary to almost entirely reset oneself's common sense knowledge in order to get integrated.

Despite the language barrier, in fact, moving to a new place involves facing habits and situations that might go against what we consider basic rules of social interaction or things we were taught by our parents, such as eating with (one, right only) hands, sharing food (rather than ordering your own dish), slurping on noodle-like food (and sometimes also drinks), crossing the road despite the heavy traffic, squatting when tired, removing shoes at home, growing long nails on your last fingers or bargaining on anything you need to buy (sometimes even at the supermarket). This can happen also the other way around, that is when you do something perfectly in line with your common sense that violates the local norms, e.g., cheek kissing as a form of greeting.

Common sense is the knowledge (usually acquired in early stages of our lives) concerning all the social, political, economic and environmental aspects of the society we live in. Machines, as they never got the chance to live a life, have no common sense at all and, hence, they know nothing about us. To help us work, computers must get to know what our jobs are. To entertain us, they need to know what we like. To take care of us, they have to know how we feel. To understand us, they must think as we think. Today, in fact, computers do only what they are programmed to do. They only have one way to deal with a problem and, if something goes wrong, they get stuck. Nowadays we have programs that exceed the capabilities of world experts but are not one able to do what a three years old child can do. It is because machines have no goals, no hopes, no fears; they do not know the meaning of things.

Computers can only do logical things, but meaning is an intuitive process – it cannot be reduced to zeros and ones. We need to transmit to computers our common sense knowledge of the world because soon there will not be enough human workers to perform the necessary tasks for our rapidly ageing population. To face this AI emergency[9], we will have to give them physical knowledge of how objects behave, social knowledge of how people interact, sensory knowledge of how things look and taste, psychological knowledge about the way people think, and so on. But having a database of millions of common sense facts will not be enough: we will also have to teach computers how to handle this knowledge, retrieve it when necessary, learn from experience, in a word we will have to give them the capacity for common sense reasoning.

### 2.3.2 Knowledge Representation

From its very beginning, AI has rested on a foundation of formal representation of knowledge. Knowledge representation (KR) is a research area that directly addresses languages for representation and the inferences that go along with them. One of the central questions of KR research is in what form knowledge is to be expressed. One of the most popular representation strategies is first order logic (FOL), a deductive system that consists of axioms and rules of inferences and can be used to formalise relationally rich predicates and quantification [66].

FOL supports syntax, semantics and, to a certain degree, pragmatics expressions. Syntax specifies the way groups of symbols are to be arranged, so that the group of symbols is considered properly formed. Semantics specify what well-formed expressions are supposed to mean. Pragmatics specifies how contextual information can be leveraged to provide better correlation between different semantics, for tasks such as word sense disambiguation. Logic, however, is known to have the problem of monotonicity. The set of entailed sentences can only increase as information is added to the knowledge base. This violates a common property of human reasoning, i.e., changing one's mind.

Solutions such as default and linear logic serve to address parts of these issues.

[9] `http://tinyurl.com/ai-crisis`

Default logic is proposed by Raymond Reiter to formalise default assumptions, e.g., "all birds fly" [67]. However, issues arise when default logic formalise facts that are true in the majority of cases but not always, e.g., "penguins do not fly". Linear logic, or constructive logic, was developed by Arend Heyting [68]. It is a symbolic logical system that preserves justification, rather than truth, and supports rejecting the weakening and contraction rules. It excels in careful deductive reasoning and is suitable in situations that can be posed precisely. As long as a scenario is static and can be detailedly described, in fact, situation-specific rules can perfectly model it but, when it comes to capture a dynamic and uncertain real-world environment, logical representation usually fails for lack of generalisation capabilities. Accordingly, it is not natural for human to encode knowledge in logical formalisation.

Another standard KR strategy, based on FOL, is the use of relational databases. The idea is to describe a database as a collection of predicates over a finite set of variables and describing constraints on the possible values. Structured query language (SQL) [69] is the database language designed for the retrieval and management of data in relational database management systems (RDBMS) [70]. Commercial (e.g., Oracle[10], Sybase[11], Microsoft SQL Server[12]) and open-source (e.g., mySQL[13]) implementations of RDBMS are available and they are commonly used in the IT industry. Relational database design requires a strict process called normalisation to ensure that the relational database is suitable for general purpose querying and the relational database is free of database operations anomalies. Third normal form (3NF) [71] is stricter than first and second normal forms and less strict as compared to Boyce-Codd normal form (BCNF) [72], fourth and fifth normal forms. Stricter normal forms means that the database design is more structured and hence requires more database tables. The advantage is that the overall design looks more organised. The disadvantage is the performance trade-off when joint table SQL queries are invoked. Relational database

[10] http://oracle.com
[11] http://sybase.com
[12] http://microsoft.com/sqlserver
[13] http://mysql.com

design, moreover, does not directly address representation of parent-child relationship in the object-oriented paradigm, subjective degrees of confidence and temporal dependent knowledge.

A popular KR strategy, especially among Semantic Web researchers, is production rule [73]. A production rule system keeps a working memory of on-going memory assertions. This working memory is volatile and keeps a set of production rules. A production rule comprises an antecedent set of conditions and a consequent set of actions (i.e., IF <conditions> THEN <actions>). The basic operation for a production rule system involves a cycle of three steps ('recognise', 'resolve conflict' and 'act') that repeats until no more rules are applicable to working memory. The step 'recognise' identifies the rules whose antecedent conditions are satisfied by the current working memory. The set of rules identified is also called the conflict set. The step 'resolve conflict' looks into the conflict set and selects a set of suitable rules to execute. The step 'act' simply executes the actions and updates the working memory. Production rules are modular. Each rule is independent from others, allowing rules to be added and deleted easily. Production rule systems have simple control structure and the rules are easy for human to understand. This is because rules are usually derived from observation of expert behaviour or expert knowledge, thus the terminology used in encoding the rules tend to resonate with human understanding. However, there are issues with scalability when production rule systems get larger. Significant maintenance overhead is required to maintain systems with thousands of rules.

Another prominent KR strategy among Semantic Web researchers is the ontology web language (OWL)[14], an XML-based vocabulary that extends resource description framework (RDF)[15] and resource description framework schema (RDFS)[16] to provide a more comprehensive ontology representation, such as the definition of classes, relationships between classes, properties of classes and constraints on relationships between classes and properties of classes. RDF supports subject-predicate-object model that

[14] http://w3.org/TR/owl-overview
[15] http://w3.org/TR/PR-rdf-syntax
[16] http://w3.org/2001/sw/wiki/RDFS

makes assertion about a resource. Reasoning engines have been developed to check for semantic consistency and help to improve ontology classification. OWL is a W3C recommended specification and comprises three dialects: OWL-Lite, OWL-DL and OWL-Full. Each dialect has a different level of expressiveness and reasoning capabilities. OWL-Lite is the least expressive compared to OWL-Full and OWL-DL. It is suitable for building ontologies that only require classification hierarchy and simple constraints and, for this reason, it provides the most computationally efficient reasoning. OWL-DL is more expressive than OWL-Full but more expressive than OWL-Lite. It has restrictions on the use of some of the description tags, hence, computation formed by a reasoning engine on OWL-DL ontologies can be completed in a finite amount of time [74]. OWL-DL is so named due to its correspondence with description logic. It is also the most commonly used dialect for representing domain ontology for Semantic Web applications. OWL-Full is the complete language and is useful for modelling a full representation of a domain. However, the trade-off for OWL-Full is the high complexity of the model that can result in sophisticated computation that may not complete in finite time. In general, OWL requires strict definition of static structures, hence, it is not suitable for representing knowledge that requires subjective degrees of confidence, but rather for representing declarative knowledge. OWL, moreover, does not allow to easily represent temporal dependent knowledge.

Another well-known way to represent knowledge is to use networks. Bayesian networks [75], for example, provide a means of expressing joint probability distributions over many interrelated hypotheses. Bayesian network is also called belief network. All variables are represented using directed acyclic graph (DAG). The nodes of a DAG represent variables. Arcs are causal connections between two variables where the truth of the former directly affects the truth of the latter. A Bayesian network is able to represent subjective degrees of confidence. The representation explicitly explores the role of prior knowledge and combines evidence of the likelihood of events. In order to compute the joint distribution of the belief network, there is a need to know $\Pr(P|\text{parents}(P))$

for each variable $P$. It is difficult to determine the probability of each variable P in the belief network. Hence, it is also difficult to scale and maintain the statistical table for large scale information processing problem. Bayesian network also has limited expressiveness, which is only equivalent to the expressiveness of proposition logic. For this reason, semantic networks are more often used for KR (Fig. 2.1).

A semantic network [76] is a graphical notation for representing knowledge in patterns of interconnected nodes and arcs. There are six types of networks, namely definitional networks, assertional networks, implicational networks, executable networks, learning networks and hybrid networks. A definitional network focuses on is-a relationships between a concept and a newly defined sub-type. The resulting network is called a generalisation, which supports the rule of inheritance for copying properties defined for a super-type to all of its sub-types. Definitions are true by definition, hence the information in definitional networks is often assumed to be true.

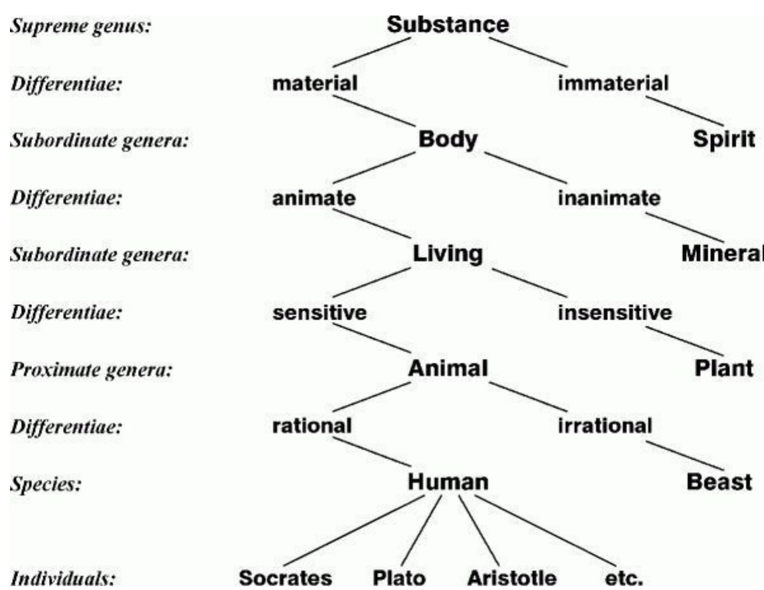


Figure 2.1: Tree of Porphyry. Porphyry presented the basis of Aristotle's thought as a tree-like scheme of dichotomous divisions, which indicates that the process continues until the lowest species is reached.

Assertional networks are meant to assert propositions and the information is assumed to be contingently true. Contingent truth means that the proposition is true in some but not in all the worlds. The proposition also has sufficient reason in which the reason entails the proposition, e.g., "the stone is warm" with the sufficient reasons being "the sun is shining on the stone" and "whatever the sun shines on is warm". Contingent truth is not the same as the truth that is assumed in default logic. Contingent truth is closer to the truth assumed in model logic.

Implicational networks use implication as the primary relation for connecting nodes. They are used to represent patterns of beliefs, causality or inferences. Methods for realising implicational networks include Bayesian networks and logic inferences used in a truth maintenance system (TMS). By combinations of forward and backward reasoning, a TMS propagates truth-values to nodes whose truth-value is unknown. Executable networks contains mechanisms implement in run-time environment such as message passing, attached procedure (e.g., data-flow graph) and graph transformation that can cause change to the network. Learning networks acquire knowledge from examples by adding and deleting nodes and links, or by modifying weights associated with the links. Learning networks can be modified in three ways: rote memory, changing weights and restructuring.

As for the rote memory, the idea is to add information without making changes to the current network. Exemplar methods can be found in relational database. For example, Patrick Winston used a version of relational graphs to describe structures, such as arches and towers [77]. When his program was given positive and negative examples of each type of structure, it would generalise the graphs to derive a definitional network for classifying all the types of structures that were considered. The idea of changing weights, in turn, is to modify the weights of links without changing the network structure for the nodes and links. Exemplar methods can be found in neural networks. As for restructuring, eventually, the idea is to create fundamental changes to the network structure for creative learning. Methods include case-based reasoning. The learning system uses rote memory to store various cases and associated action such as course of action. When a new case is encountered, the system finds those cases that are most similar to the new one and retrieves the outcome. To organise the search and evaluate similarity, the learning system must use restructuring to find common patterns in the individual cases and use those patterns as the keys for indexing the database. Hybrid networks combine two or more of the previous techniques. Hybrid networks can be a single network. They can also be separate but closely interacting networks.

Sowa used unified modelling language (UML) as an example to illustrate a hybrid semantic network. Semantic networks are very expressive. The representation is flexible and can be used to express different paradigm such as relational model and hierarchical relationship. The challenge is at implementation level. For example, it is difficult to implement hybrid semantic network, which requires an integration of different methods.

### 2.3.3 History

What magical trick makes us intelligent? - Marvin Minsky was wondering more than two decades ago - The trick is that there is no trick. The power of intelligence stems from our vast diversity, not from any single, perfect principle [62]. Human brain, in fact, is a very complex system, maybe the most complex in nature. The functions it performs are the product of thousands and thousands of different subsystems working together at the same time. Common sense computing involves trying to emulate such mechanism and, in particular, exploiting common sense knowledge to improve computers' understanding of the world. Before Minsky, many AI researchers started to think about the implementation of a common sense reasoning machine.

The very first person who seriously started thinking about the creation of such a machine was maybe Alan Turing when, in 1950, he first raised the question "can machines think?". But he never managed to answer that question, he just provided a method to gauge artificial intelligence, the famous Turing test. The notion of common sense in AI is actually dated 1958, when John McCarthy, in his seminal paper 'Programs with Common Sense' [78], proposed a program for solving problems by manipulating sentences in formal language. The main aim of the ‘advice taker’, this was the name of the program, was to try to automatically deduce for itself a sufficiently wide class of immediate consequences of anything it was told and what it already knew. In this paper, McCarthy stressed the importance of finding a proper method of representing expressions in the computer since, in order for a program to be capable of learning something; it must first be capable of being told.

He also developed the idea of creating a property list for each object in which are listed the specific things people usually know about it. It was the first attempt to build a common sense knowledge base but, more important, it was the epiphany of the need of common sense to move forward in the technological evolution. In 1959, McCarthy went to MIT and started, together with Minsky, the MIT Artificial Intelligence Project. They both were aware of the need for AI of a common sense reasoning approach but while McCarthy was more concerned with establishing logical and mathematical foundations for it, Minsky was more involved with theories of how we actually reason using pattern recognition and analogy.

These theories were organised some years later with the publication of the Society of Mind [62], a masterpiece of AI literature, which consists in an illuminating vision of how the human brain works. Minsky sees the mind made of many little parts called ‘agents’, each mindless by itself but able to lead to true intelligence when working together. These groups of agents, called ’agencies’, are responsible to perform some type of function, such as remembering, comparing, generalising, exemplifying, analogising, simplifying, predicting, and so on. The most common agents are the so called ‘K-lines’ whose task is simply to activate other agents: this is a very important issue since agents are all highly interconnected and activating a K-line can cause a significant cascade of effects. To Minsky, in fact, mental activity ultimately consists in turning individual agents on and off: at any time only some agents are active and their combined activity constitutes the ‘total state’ of the mind.

K-lines are a very simple but powerful mechanism since they allow entering a particular configuration of agents that formed a useful society in a past situation. This is how we build and retrieve our problem solving strategies in our mind; this is how we should develop our problem solving strategies in our programs. In 1990, McCarthy put together seventeen papers to try to define common sense knowledge by using mathematical logic in such a way that common sense problems could be solved by logical reasoning. Deductive reasoning in mathematical logic has the so-called monotonicity property: if we add new assumptions to the set of initial assumptions, there may be some new conclusions, but every sentence that was a deductive consequence of the original hypotheses is still a consequence of the enlarged set. Much human reasoning is monotonic as well, but some important human common sense reasoning is not. For example, if someone is asked to build a birdcage, this person concludes that it is appropriate to put a top on it, but when he/she learns the further fact that the bird is a penguin he/she no longer draws that conclusion. McCarthy formally described this assumption that things are as expected unless otherwise specified, with the 'circumscription method' of non-monotonic reasoning: a minimisation similar to the closed world assumption that what is not known to be true is false.

In the same years, a similar attempt to give a shape to common sense knowledge was done by Ernest Davis [79]. He tried to develop an ad hoc language for expressing common sense knowledge and inference techniques for carrying out common sense reasoning in specific domains such as space, time, quantities, qualities, flows, goals, plans, needs, beliefs, intentions, actions and interpersonal relations. Thanks to his and McCarthy's knowledge formalisations, the first steps were set towards the expression of common sense facts in a way that would have been suitable for inclusion in a general purpose database and hence towards the development of programs with common sense. Minsky's theory of human cognition, in particular, was welcomed with great enthusiasm by the AI community and gave birth to many attempts to build common sense knowledge bases and develop systems capable of common sense reasoning.

The most representative projects are Cyc [80], Doug Lenat's logic-based repository of common sense knowledge, WordNet [81], Christiane Fellbaum's universal database of word senses, and ThoughtTreasure [82] , Erik Mueller's story understanding system. Cyc is one of the first attempts to assemble a massive knowledge base spanning human common sense knowledge. Initially started by Doug Lenat in 1984, this project utilises knowledge engineers who hand-craft assertions and place them into a logical framework using CycL, Cyc's proprietary language. Cyc's knowledge is represented redundantly at two levels: a frame language distinction (epistemological level), adopted for its efficiency, and a predicate calculus representation (heuristic level), needed for its expressive power to represent constraints. While the first level keeps a copy of the facts in the uniform user language, the second level keeps its own copy in different languages and data structures suitable to be manipulated by specialised inference engines. Knowledge in Cyc is also organised into 'microtheories', resembling Minsky's agencies, each one with its own knowledge representation scheme and sets of assumptions. These microtheories are linked via 'lifting rules' that allow translation and communication of expressions between them.

Begun in 1985 at Princeton University, WordNet is a database of words, primarily nouns, verbs and adjectives. It has been one of the most widely used resources in computational linguistics and text analysis for the ease in interfacing it with any kind of application and system. The smallest unit in WordNet is the word/sense pair, identified by a 'sense key'. Word/sense pairs are linked by a small set of semantic relations such as synonyms, antonyms, is-a superclasses, and words connected by other relations such as part-of. Each synonym set, in particular, is called 'synset': it consists in the representation of a concept, often explained through a brief gloss, and represents the basic building block for hierarchies and other conceptual structures in WordNet.

Erik Mueller's ThoughtTreasure is a story understanding system with a great variety of common sense knowledge about how to read and understand children's stories. It was inspired by Cyc and is similar to Cyc in that it has both natural language and common sense components. But whereas Cyc mostly uses logic, ThoughtTreasure uses multiple representations schemes: grids for stereotypical settings, finite automata for rules of device behaviour and mental processes, logical assertions for encyclopaedic facts and linguistic knowledge. ThoughtTreasure's lexicon is similar to WordNet but, while world knowledge is explicitly excluded from WordNet, ThoughtTreasure contains also concepts that are not lexicalised in English like 'going to the pub' or 'eating at the restaurant', which are very important for common sense reasoning.

### 2.3.4 The Open Mind Common Sense Project

Using logic-based reasoning can solve some problems in computer programming. However, most real-world problems need methods better at matching patterns and constructing analogies, or making decisions based on previous experience with examples, or by generalising from types of explanations that have worked well on similar problems in the past [83]. In building intelligent systems we have to try to reproduce our way of thinking: we turn ideas around in our mind to examine them from different perspectives until we find one that works for us. From this the need of using several representations, each integrated with its set of related pieces of knowledge, to be able to switch from one to another when one of them fails. The key, in fact, is using different representations to describe the same situation.

Minsky blames our standard approach to writing a program for common sense computing failures. Since computers appeared, our approach to solve a problem has always consisted in first looking for the best way to represent the problem, and then looking for the best way to represent the knowledge needed to solve it and finally looking for the best procedure for solving it. This problem-solving approach is good when we have to deal with a specific problem but there is something basically wrong with it: it leads us to write only specialised programs that cope with solving only that kind of problem. This is why, today, we have millions of expert programs but not even one that can be actually defined intelligent.

From here comes the idea of finding a heterogeneous ways to represent common sense knowledge and to link each unit of knowledge to the uses, goals, or functions that each knowledge-unit can serve. This non-monotonic approach reasserted by Minsky was adopted soon after by Push Singh within the Open Mind Common Sense (OMCS) project [84]. Initially born from an idea of David Stork [85], the project differs from previous attempts to build a common sense database for the innovative way to collect knowledge and represent it (Fig. 2.2). OMCS is a second-generation common sense database. Knowledge is represented in natural language, rather than using a formal logical structure, and information is not hand-crafted by expert engineers but spontaneously inserted by online volunteers. The reason why Lenat decided to develop an ad hoc language for Cyc is that vagueness and ambiguity pervade English and computer reasoning systems generally require knowledge to be expressed accurately and precisely. However, as expressed in the Society of Mind, ambiguity is unavoidable when trying to represent the common sense world. No single argument, in fact, is always completely reliable but, if we combine multiple types of arguments, we can improve the robustness of reasoning as well as we can improve table stability by providing it with many small legs in place of just one very big leg. This way information is not only more reliable but also stronger. If a piece of information goes lost, we can still access the whole meaning, exactly as the table keeps on standing up if we cut out one of the small legs.

Diversity is, in fact, the key of OMCS' success: the problem is not choosing a representation in spite of another but it is finding a way for them to work together in one system. The main difference between acquiring knowledge from the general public and acquiring it from expert engineers is that the general public is likely to leave as soon as they encounter something boring or difficult. The key is letting people do what they prefer to do. Different people in fact like to do different things: some like to enter new items, some like to evaluate items, others like to refine items. For this reason, OMCS is based on a distributed workflow model where the different stages of knowledge acquisition could be performed separately by different participants.

The system, in fact, was designed to allow users to insert new knowledge via both

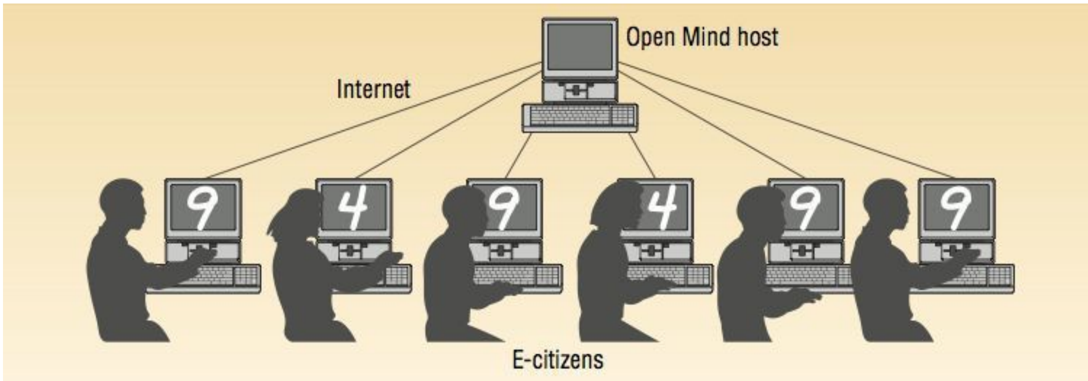


Figure 2.2: An Open Mind project on OCR: handwritten characters are presented to e-citizens whose judgements (here 4 versus 9) are returned to the Open Mind host and used to train the classifier.

template-based input and free-form input, tag concepts, clarify properties and validate assertions. But, since giving so much control to users can be dangerous, a fixed set of pre-validated sentences were meant to be presented to them from time to time, in order to assess their honesty, and the system was designed in a way that allowed users to reciprocally control each other by judging samples of each other's knowledge. OMCS exploits a method termed cumulative analogy [86], a class of analogy-based reasoning algorithms that leverage existing knowledge to pose knowledge acquisition questions to the volunteer contributors.

When acquiring knowledge online, the stickiness of the website is of primary importance. The best way to involve users in this case is making them feel that they are contributing to the construction of a thinking machine and not just a static database. To do this, OMCS first determines what other topics are similar to the topic the user is currently inserting knowledge for, and then it uses cumulative analogy to generate and present new specific questions about this topic. Because each statement consists of an object and a property, the entire knowledge repository can be visualised as a large matrix, with every known object of some statement being a row and every known property being a column.

Cumulative analogy is performed by first selecting a set of nearest neighbours, in terms of similarity, of the treated concept and then by projecting known properties of this set onto not known properties of the concept and presenting them as questions (Fig. 2.3). The replies to the knowledge acquisition questions formulated by analogy are immediately added to the knowledge repository, affecting the similarity calculations. This way users can see the system's behaviour improve or change as a result of the entered knowledge and be more tempted to participate.

A more generalised way to deal with the information contained in the Open Mind corpus is AnalogySpace [87], a process that applies singular value decomposition (SVD) on the matrix representation of the common sense knowledge base, in order to reduce its dimensionality and capture the most important correlations. The entries in the resulting matrix are positive or negative numbers, depending on the reliability of the assertions, and their magnitude increases logarithmically with the confidence score. Applying SVD on this matrix causes it to describe other features that could apply to known concepts by analogy: if a concept in the matrix has no value specified for a feature owned by many similar concepts, then by analogy the concept is likely to have that feature as well.

| Objects | Properties *(with simplified form)* | | | | | |
|---|---|---|---|---|---|---|
| | ... | contains knowledge *contain knowledge* | has pages *have page* | is cold *be cold* | is for reading *be read* | ... |
| ⋮ | | ⋮ | ⋮ | ⋮ | ⋮ | |
| book | ... | x | x | | x | ... |
| ice | ... | | - | x | | ... |
| **newspaper** | ... | **x?** | **x** | | **x** | ... |
| magazine | ... | x | x | | x | ... |
| ⋮ | | ⋮ | ⋮ | ⋮ | ⋮ | |

Figure 2.3: The cumulative analogy process allows to perform comparisons between concepts in a knowledge base, represented as a matrix, and hence to infer new information about similar concepts.

A way to visualise and understand AnalogySpace is provided by Luminoso [23], a tool that allows to interactively explore the dimensionality-reduced semantic space of common sense knowledge by 'grabbing' its data points and, hence, view their associated text and statistics. This operation also allows highlighting the point's neighbourhood of semantically associated data points, providing clues for reasons as to why the points were classified along the dimensions they were. The AnalogySpace process, eventually, is naturally extended by the 'blending' technique [88], a new method to perform inference over multiple sources of data simultaneously, taking advantage of the overlap between them. Blending consists in an alignment phase of two datasets and of a linear combination of them to be able to apply principal component analysis (PCA) on the resulting matrix. This enables common sense to be used as a basis for inference in a wide variety of systems and applications so that they can achieve digital intuition about their own data, making assumptions and conclusions based on the connections between that specific data and the general common sense that people have.

### 2.3.5 Sentic Computing

Sentic computing is a multi-disciplinary approach to sentiment analysis, recently proposed by Cambria et al. [12], at the crossroads between affective computing and common sense computing. In the field of opinion mining, in fact, not only common sense knowledge but also emotional knowledge is important to grasp both the cognitive and affective information (termed semantics and sentics) associated with natural language opinions and sentiments. Although scientific research in the area of emotion stretches back to the 19th century when Charles Darwin and William James proposed theories of emotion that continue to influence thinking today [89, 90], the injection of affect into computer technologies is much more recent. During most of the last century, research on emotions was conducted by philosophers and psychologists, whose work was based on a small set of emotion theories that continue to underpin research in this area. The first researchers to try linking text to emotions were actually social psychologists and anthropologists who tried to find similarities on how people from different cultures communicate [91]. This research was also triggered by a dissatisfaction with the dominant cognitive view centred around humans as 'information processors' [92]. Later on, in the 1980s, researchers such as Turkle [93] began to speculate about how computers might be used to study emotions. Systematic research programs along this front began to emerge in the early 1990s. For example, Scherer [94] implemented a computational model of emotion as an expert system. A few years later, Picard's landmark book affective computing [95] prompted a wave of interest among computer scientists and engineers looking for ways to improve human-computer interfaces by coordinating emotions and cognition with task constraints and demands. Picard described three types of affective computing applications:

1. Systems that detect the emotions of the user;
2. Systems that express what a human would perceive as an emotion;
3. Systems that actually 'feel' an emotion.

Although touching upon HCI [96] and affective modelling [97], sentic computing primarily focuses on affect detection from text. Affect detection is critical because an affect-sensitive interface can never respond to users' affective states if it cannot sense their affective states. Affect detection need not be perfect but must be approximately on target. Affect detection is, however, a very challenging problem because emotions are constructs (i.e., conceptual quantities that cannot be directly measured) with fuzzy boundaries and with substantial individual difference variations in expression and experience. To overcome such a hurdle, sentic computing builds upon a biologically-inspired and psychologically-motivated affective categorisation model, proposed by Cambria et al. [98], that can potentially describe the full range of emotional experiences in terms of four independent but concomitant dimensions, whose different levels of activation make up the total emotional state of the mind. In sentic computing, whose term derives from the Latin *sentire* (root of words such as sentiment and sentience) and *sensus* (intended both as capability of feeling and as common sense), the analysis of natural language is based on affective ontologies and common sense reasoning tools, which enable the analysis of text not only at document, page or paragraph level but also at sentence and clause level. In particular, sentic computing involves the use of AI and Semantic Web techniques, for knowledge representation and inference; mathematics, for carrying out tasks such as graph mining and multi-dimensionality reduction; linguistics, for discourse analysis and pragmatics; psychology, for cognitive and affective modelling; sociology, for understanding social network dynamics and social influence; finally ethics, for understanding related issues about the nature of mind and the creation of emotional machines.

In this thesis, sentic computing tools and techniques are exploited for the design of two common sense knowledge bases for open-domain sentiment analysis and for the development of reasoning strategies for handling such knowledge base. The knowledge representation and inference strategies hereby developed by means of sentic computing, eventually, are employed for designing an opinion mining engine that is exploited to develop emotion-sensitive systems in fields such as social data mining, multimedia management, personalisation and persuasion, human-computer interaction, intelligent user interfaces, social media marketing, and patient-centred applications.

## 2.4 Conclusions

This chapter has shown how and why, today, online opinions are crucial both for companies to succeed in selling their products and services, and for individuals to properly choose among these. The dynamics behind the buzz mechanism were discussed, together with the motivating factors that gave birth to the field of opinion mining, and the tasks that make it different from standard information retrieval (section 2.1). The chapter also illustrated the approaches and depths of analysis in mining and characterising opinions, in order to map a given piece of text to a label belonging to a predefined set of categories or to a real number representative of a polarity degree.

Specifically, the chapter discussed the evolution of different approaches from heuristics to discourse structure, from coarse to fine grained analysis, and from keyword to concept level opinion mining (section 2.2). Eventually, the chapter explained the importance of common sense for the development of intelligent systems, illustrated different knowledge representation strategies, and presented the evolution of common sense computing from logic-based methods to more recent approaches based on natural language techniques (section 2.3). Among such approaches, in particular, sentic computing is hereby developed and exploited to go beyond merely syntactical approaches to sentiment analysis. Specifically, sentic computing tools and techniques are employed to design an affective common sense knowledge base and a common knowledge base, on which different techniques are employed for the extraction of semantics and sentics from natural language text (Fig. 2.4).

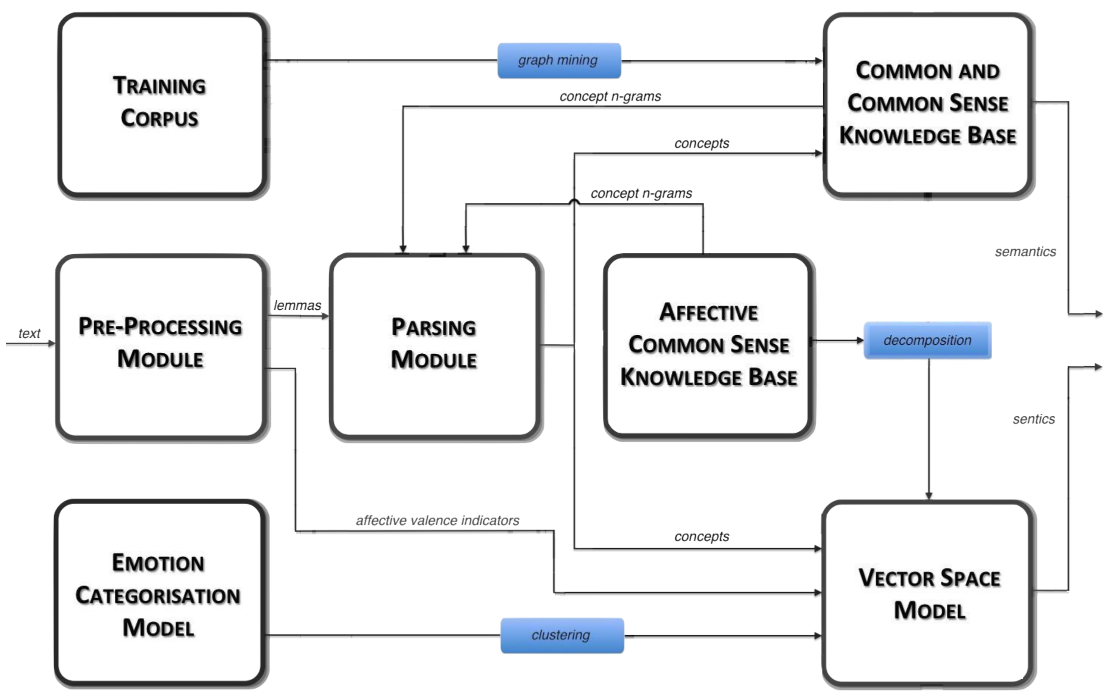


Figure 2.4: Opinion mining engine preview. Different techniques, e.g., graph mining and dimensionality reduction, are employed on two common sense knowledge bases for open-domain sentiment analysis.

# Chapter 3

# Sentic Knowledge Base Design

*You can know the name of a bird in all the languages of the world,*
*but when you're finished, you'll know absolutely nothing whatever about the bird.*
*So let's look at the bird and see what it's doing – that's what counts.*
*I learned very early the difference between*
*knowing the name of something and knowing something.*

*Richard Feynman*

In standard human-to-human communication, people usually refer to existing facts and circumstances and build new useful, funny or interesting information on the top of those. This common knowledge comprehends information usually found in news, articles, debates, lectures, etc. (factual knowledge) but also principles and definitions that can be found in collective intelligence projects such as Wikipedia (vocabulary knowledge). Attempts to build a common knowledge base are countless and comprehend both resources crafted by human experts or community efforts, such as WordNet and Freebase [99], a large collaborative knowledge base consisting of metadata composed mainly by its community members, and automatically-built knowledge bases, such as Wiki-

Taxonomy [100], a taxonomy extracted from Wikipedia's category links, YAGO [101], a semantic knowledge base derived from Wikipedia, WordNet and GeoNames[1], and Never-Ending Language Learning (NELL), CMU's semantic machine learning system [102]. Common knowledge, however, is just a surface layer of human communication and alone is not enough for understanding natural language. In order to join in a conversation and actively participate in it, we need to have some information about the concepts and the context this conversation is based on but, before that, we need to be able to speak. Trying to understand natural language by simply relying on common knowledge would be pretty much like trying to join a conversation about food, as a one-year old child: we would almost perfectly know the debated topics but we would not really know how to relate these concepts to each other in order to fully understand or form a sentence.

These semantic relationships are what we call common sense knowledge and consist in knowing that people are younger than their parents, that a butcher is unlikely to be a vegetarian, that people do not like being repeatedly interrupted, that if you hold a knife by its blade then it may cut you, that if you drop paper into a flame then it will burn, that people pay taxi drivers to drive them to places, that people generally sleep at night, and so forth. Computers do not know such things, as they never had the opportunity to live a life and to experience or be taught the meaning of words, objects and actions, and how these relate to each other. If we want machines to really understand natural language, hence, we need to provide them with such knowledge. Building a common sense knowledge base, however, is not easy as common sense is typically omitted from social communications.

For Grice's theory of pragmatics [103], in fact, when people communicate with each other, they tend not to provide information which is obvious or extraneous. This is true both for face-to-face communication and for asynchronous communication (e.g., on the Web), which makes the collection of common sense an extremely difficult task. This chapter shows how to build a resource of common sense knowledge and how to exploit

[1] `http://geonames.org`

this to develop an affect-sensitive knowledge base for opinion mining (section 3.1). In order to accordingly categorise affect in such knowledge base, moreover, a novel emotion categorisation model (section 3.2) is developed and crowd sourcing techniques are exploited to build an intelligent user interface (IUI) for the collection of new affective common sense knowledge (section 3.3). Eventually, a knowledge base of common and common sense knowledge is also built for improving the topic-spotting capabilities of the final system (section 3.4), and some concluding remarks are provided (section 3.5).

## 3.1 AffectNet: An Affective Common Sense Knowledge Base

Opinions and sentiments are often conveyed implicitly through context and domain dependent concepts, rather than through specific affect words. Hence, in order to semantically and affectively analyse natural language text for opinion mining, we need to rely on a knowledge base containing both the cognitive and affective information (semantics and sentics) associated with concepts. To this end, a semantic network of common sense knowledge (subsection 3.1.1) is merged with a linguistic resource for the lexical representation of affective knowledge, in order to obtain a new knowledge base in which concepts are interrelated by both common sense and affective features (subsection 3.1.2).

### 3.1.1 ConceptNet

ConceptNet [104] is a semantic resource structurally similar to WordNet, but whose scope of contents is general world knowledge, in the same vein as Cyc. Instead of insisting on formalising common sense reasoning using mathematical logic [105], ConceptNet uses a new approach: it represents data in the form of a semantic network and makes it available to be used in natural language processing. The prerogative of ConceptNet, in fact, is contextual common sense reasoning: while WordNet is optimised for lexical categorisation and word-similarity determination, and Cyc is optimised for formalised logical reasoning, ConceptNet is optimised for making practical context-based inferences over real-world texts. In ConceptNet, WordNet's notion of node in the semantic network is extended from purely lexical items (words and simple phrases with atomic meaning) to include higher-order compound concepts, e.g., 'satisfy hunger', 'follow recipe', to represent knowledge around a greater range of concepts found in everyday life (Table 3.1). Moreover WordNet's repertoire of semantic relations is extended from the triplet of synonym, is-a, and part-of, to a repertoire of twenty semantic relations including, for example, EffectOf (causality), SubeventOf (event hierarchy), CapableOf (agent's ability), MotivationOf (affect), PropertyOf, LocationOf. ConceptNet's knowledge is also of a more informal, defeasible and practically valued nature. For example, WordNet has formal taxonomic knowledge that 'dog' is a 'canine', which is a 'carnivore', which is a 'placental mammal'; but it cannot make the practically oriented member-to-set association that 'dog' is a 'pet'.

| **Term** | **WordNet Hypernyms** | **ConceptNet Assertions** |
|---|---|---|
| Cat | Feline; Felid; Adult male; Man; Gossip; Gossiper; Gossipmonger; Rumormonger; Rumourmonger; Newsmonger; Woman; Adult female; Stimulant; Stimulant drug; Excitant; Tracked vehicle; ... | Cats can hunt mice; Cats have whiskers; Cats can eat mice; Cats have fur; cats have claws; Cats can eat meat; cats are cute; ... |
| Dog | Canine; Canid; Unpleasant woman; Disagreeable woman; Chap; Fellow; Feller; Lad; Gent; Fella; Scoundrel; Sausage; Follow, ... | Dogs are mammals; A dog can be a pet; A dog can guard a house; You are likely to find a dog in kennel; An activity a dog can do is run; A dog is a loyal friend; A dog has fur; ... |
| Language | Communication; Auditory communication; Word; Higher cognitive process; Faculty; Mental faculty; Module; Text; Textual matter; | English is a language; French is a language; Language is used for communication; Music is a language; A word is part of language; ... |
| iPhone | N/A; | An iPhone is a kind of telephone; An iPhone is a kind of computer; An IPhone can display your position on a map; An IPhone can send and receive emails; An IPhone can display the time; ... |
| Birthday gift | Present; | Card is birthday gift; Present is birthday gift; Buying something for a loved one is for a birthday gift; ... |

Table 3.1: Comparison between WordNet and ConceptNet. While WordNet synsets contain vocabulary knowledge associated with concepts, ConceptNet assertions convey knowledge about what such concepts are used for.

ConceptNet also contains a lot of knowledge that is defeasible, i.e., it describes something that is often true but not always, e.g., EffectOf('fall off bicycle', 'get hurt'), which is something we cannot leave aside in common sense reasoning. Most of the facts interrelating ConceptNet's semantic network are dedicated to making rather generic connections between concepts. This type of knowledge can be brought back to Minsky's K-lines as it increases the connectivity of the semantic network and makes it more likely that concepts parsed out of a text document can be mapped into ConceptNet.

ConceptNet is produced by an automatic process, which first applies a set of extraction rules to the semi-structured English sentences of the OMCS corpus, and then applies an additional set of 'relaxation' procedures, i.e., filling in and smoothing over network gaps, to optimise the connectivity of the semantic network (Fig. 3.1). In ConceptNet2, a new system for weighting knowledge was implemented, which scores each binary assertion based on how many times it was uttered in the OMCS corpus, and on how well it can be inferred indirectly from other facts in ConceptNet.

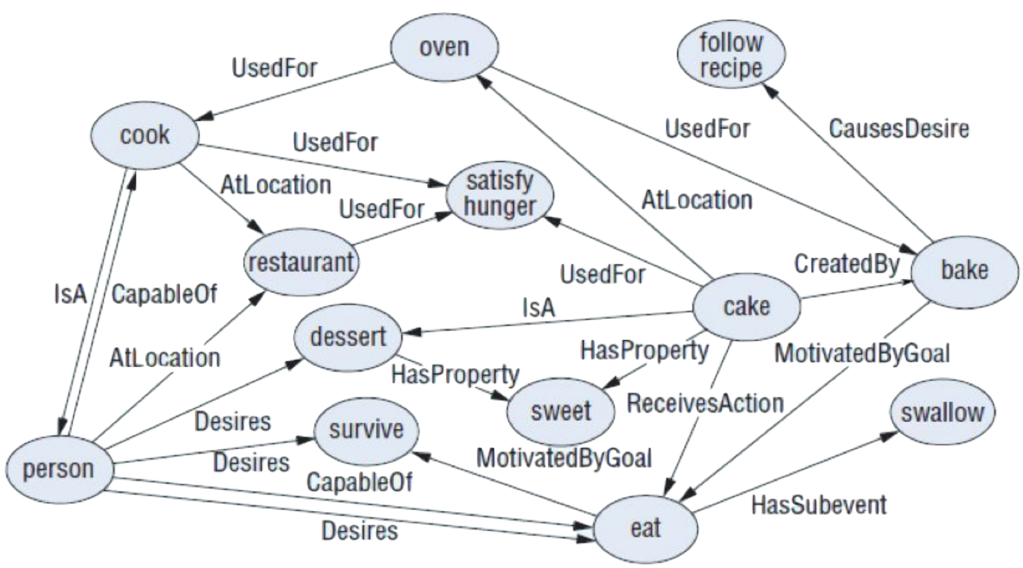


Figure 3.1: ConceptNet represents the information in the Open Mind corpus as a directed graph where nodes are concepts and labelled edges are assertions of common sense that interconnect them.

In ConceptNet3 [106], users can also participate in the process of refining knowledge by evaluating existing statements on Open Mind Commons [107], the new interface for collecting common sense knowledge from users over the Web. By giving the user many forms of feedback and using inferences by analogy to find appropriate questions to ask, Open Mind Commons can learn well-connected structures of common sense knowledge, refine its existing knowledge, and build analogies that lead to even more powerful inferences. ConceptNet4 includes data that was imported from the online game Verbosity. It also includes the initial import of the Chinese ConceptNet.

ConceptNet5, eventually, contains knowledge from English Wikipedia, specifically from DBPedia, which extracts knowledge from the info-boxes that appear on articles, and ReVerb, a machine-reading project extracting relational knowledge from the actual text of each article. It also includes a large amount of content from the English Wiktionary, including synonyms, antonyms, translations of concepts into hundreds of languages, and multiple labelled word senses for many English words. ConceptNet5 contains more dictionary-style knowledge coming from WordNet and some knowledge about people's intuitive word associations coming from games with a purpose (GWAP). Previous versions of ConceptNet have been distributed as idiosyncratic database structures plus some software to interact with them. ConceptNet5 is not a piece of software or a database: it is a graph. To be precise, it is a hypergraph, i.e., it has edges about edges. Each statement in ConceptNet, in fact, has justifications pointing to it, explaining where it comes from and how reliable the information seems to be.

### 3.1.2 Affective Blending

In many cultures (e.g., Chinese), the concepts of 'heart' and 'mind' used to be expressed by the same word (心) as it was believed that consciousness and thoughts came from the cardiac muscle. In human cognition, in fact, thinking and feeling are mutually present: emotions are often the product of our thoughts as well as our reflections are often the product of our affective states.

Emotions are intrinsically part of our mental activity and play a key role in communication and decision-making processes. Emotion is a chain of events made up of feedback loops. Feelings and behaviour can affect cognition, just as cognition can influence feeling. Emotion, cognition and action interact in feedback loops and emotion can be viewed in a structural model tied to adaptation [108]. There is actually no fundamental opposition between emotion and reason. In fact, it may be argued that reason consists of basing choices on the perspectives of emotions at some later time. Reason dictates not giving in to one's impulses because doing so may cause greater suffering later [109]. Reason does not necessarily imply exertion of the voluntary capacities to suppress emotion. It does not necessarily involve depriving certain aspects of reality of their emotive powers. On the contrary, our voluntary capacities allow us to draw more of reality into the sphere of emotion. They allow one's emotions to be elicited not merely by the proximal, or the perceptual, or that which directly interferes with one's actions, but by that which in fact touches on one's concerns, whether proximal or distal, whether occurring now or in the future, whether interfering with one's own life or that of others. Cognitive functions serve emotions and biological needs.

Information from the environment is evaluated in terms of its ability to satisfy or frustrate needs. What is particularly significant is that each new cognitive experience that is biologically important is connected with an emotional reaction such as fear, pleasure, pain, disgust or depression [110]. Emotions, in fact, are special states shaped by natural selection to adjust various aspects of our organism in order to make it better face particular situations, e.g., anger evolved for reaction, fear evolved for protection and affection evolved for reproduction. For these reasons, we cannot prescind from emotions in the development of intelligent systems: if we want computers to be really intelligent, not just have the veneer of intelligence, we need to give them the ability to recognise, understand and express emotions. To this end, it is useful to build a knowledge base that contains not only common sense concepts, but also the affective information associated with these.

ConceptNet is a good source of common sense knowledge but alone is not enough for sentiment analysis tasks as it specifies how concepts are semantically related to each other but often lacks connections between concepts that convey the same kind of emotion or similar polarity. To overcome such a hurdle, WordNet-Affect (WNA) [111], a linguistic resource for the lexical representation of affective knowledge developed starting from WordNet, is used. WNA is built by assigning to a number of WordNet synsets one or more affective labels (a-labels). In particular, the affective concepts representing emotional states are identified by synsets marked with the a-label 'emotion', but there are also other a-labels for concepts representing moods, situations eliciting emotions or emotional responses. WNA was developed in two stages. The first consisted of the identification of a first core of affective synsets. The second step consisted of the extension of the core with the relations defined in WordNet.

ConceptNet and WNA are blended together by combining the matrix representations of the two knowledge bases linearly into a single matrix, in which the information between the two initial sources is shared. The first step to create the affective blend is to transform the input data so that it can all be represented in the same matrix. To do this, the lemma forms of ConceptNet concepts are aligned with the lemma forms of the words in WNA and map the most common relations in the affective knowledge base into ConceptNet's set of relations, e.g., Hypernym into *IsA* and Holonym into *PartOf*. In particular, ConceptNet is first converted into a matrix by dividing each assertion into two parts: a concept and a feature, where a feature is simply the assertion with the first or the second concept left unspecified such as 'a wheel is part of' or 'is a kind of liquid'. The entries in the resulting matrix are positive or negative numbers, depending on the reliability of the assertions, and their magnitude increases logarithmically with the confidence score. WNA, similarly, is represented as a matrix where rows are affective concepts and columns are features related to these. The result of aligning the matrix representations of ConceptNet and WNA is a new affective semantic network, in which common sense concepts are linked to a hierarchy of affective domain labels.

| Mood | LiveJournal Posts | ConceptNet Concepts |
|---|---|---|
| Happy | Finally I got my student cap ! I am officially high school graduate now ! Our dog Tanja, me, Timo (our art teacher) and EmmaMe, Tanja, Emma and Tiia Only two weeks to Japan !! | student; school graduate; Japan |
| Happy | I got a kitten as an early birthday gift on Monday. Abby was smelly, dirty, and knawing on the metal bars of the kitten carrier though somewhat calm when I picked her up. We took her. She threw up on me on the ride home and repeatly keeps sneesing in my face. | kitten; birthday gift; metal bar; face |
| Sad | Hi. Can I ask a favor from you? This will only take a minute. Please pray for Marie, my friends' dog a labrador, for she has canine distemper. Her lower half is paralysed and she's having locked jaw. My friends' family is feeding her through syringe. | friends; dog; labrador; canine distemper; jaw; syringe |
| Sad | my uncle paul passed away on febuary 16, 2008. he lost his battle with cancer. i remember spending time with him and my aunt nina when they babysat me. we would go to taco bell and i would get nachos. | uncle; battle; cancer; aunt; taco bell; nachos |

Table 3.2: Some examples of LiveJournal posts where affective information is not conveyed explicitly through affect words. Such implicit information, however, can be inferred by analysing the semantics and sentics.

In such semantic network, proposed by Cambria et al. [98] and termed AffectNet, common sense and affective knowledge are in fact combined, not just concomitant, i.e., everyday life concepts like 'have breakfast', 'meet people' or 'watch tv' are linked to affective domain labels like 'joy', 'anger' or 'surprise'. Such knowledge base results very useful when performing tasks such as emotion recognition or polarity detection from natural language text as opinions and sentiments are often conveyed implicitly through context and domain dependent concepts, rather than through specific affect words (Table 3.2).

## 3.2 The Hourglass of Emotions: A Novel Emotion Categorisation Model

The study of emotions is one of the most confused (and still open) chapters in the history of psychology. This is mainly due to the ambiguity of natural language, which does not facilitate the description of mixed emotions in an unequivocal way. Love and other emotional words like anger and fear, in fact, are suitcase words (many different meanings packed in), not clearly defined and meaning different things to different people [112]. Hence, more than 90 definitions of emotions have been offered over the past century and there are almost as many theories of emotion, not to mention a complex array of overlapping words in our languages to describe them. Some categorisations include cognitive versus non-cognitive emotions, instinctual (from the amygdala) versus cognitive (from the prefrontal cortex) emotions, and also categorisations based on duration, as some emotions occur over a period of seconds (e.g., surprise), whereas others can last years (e.g., love). The James-Lange theory posits that emotional experience is largely due to the experience of bodily changes [90]. Its main contribution is the emphasis it places on the embodiment of emotions, especially the argument that changes in the bodily concomitants of emotions can alter their experienced intensity.

Most contemporary neuroscientists endorse a modified James-Lange view, in which bodily feedback modulates the experience of emotion [113]. In this view, emotions are related to certain activities in brain areas that direct our attention, motivate our behaviour, and determine the significance of what is going on around us. Pioneering works by Broca [114], Papez, [115] and MacLean [116] suggested that emotion is related to a group of structures in the centre of the brain called limbic system (or paleomammalian brain), which includes the hypothalamus, cingulate cortex, hippocampi, and other structures. More recent research, however, has shown that some of these limbic structures are not as directly related to emotion as others are, while some non-limbic structures have been found to be of greater emotional relevance [117].

For tasks such as emotion recognition and polarity detection, it is key to have a model capable of finely describing the affective information associated with natural language concepts. To this end, a novel emotion categorisation model is proposed that goes beyond mere categorical and dimensional approaches (subsection 3.2.1) by representing affective states both through emotional labels and through four independent but concomitant dimensions that can potentially describe the full range of emotional experiences (subsection 3.2.2).

### 3.2.1 Categorical Versus Dimensional Approaches

Philosophical studies on emotions date back to ancient Greeks and Romans. Following the early Stoics, for example, Cicero enumerated and organised the emotions into four basic categories: *metus* (fear), *aegritudo* (pain), *libido* (lust) and *laetitia* (pleasure). Studies on evolutionary theory of emotions, in turn, were initiated in the late 19th century by Darwin [89]. His thesis was that emotions evolved via natural selection and therefore have cross-culturally universal counterparts. In the early 1970s, Ekman found evidence that humans share six basic emotions: happiness, sadness, fear, anger, disgust and surprise [118]. Few tentative efforts to detect non-basic affective states, such as fatigue, anxiety, satisfaction, confusion, or frustration, have been also made [119, 120, 121, 122, 123, 124] (Table 3.3). In 1980, Averill put forward the idea that emotions cannot be explained strictly on the basis of physiological or cognitive terms. Instead, he claimed that emotions are primarily social constructs; hence, a social level of analysis is necessary to truly understand the nature of emotion [125].

The relationship between emotion and language (and the fact that the language of emotion is considered a vital part of the experience of emotion) has been used by social constructivists and anthropologists to question the universality of Ekman's studies, arguably because the language labels he used to code emotions are somewhat US-centric. In addition, other cultures might have labels that cannot be literally translated to English (e.g., some languages do not have a word for fear [126]).

| **Author** | **#Emotions** | **Basic Emotions** |
|---|---|---|
| Ekman | 6 | anger, disgust, fear, joy, sadness, surprise |
| Parrot | 6 | anger, fear, joy, love, sadness, surprise |
| Frijda | 6 | desire, happiness, interest, surprise, wonder, sorrow |
| Plutchik | 8 | acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise |
| Tomkins | 9 | desire, happiness, interest, surprise, wonder, sorrow |
| Matsumoto | 22 | joy, anticipation, anger, disgust, sadness, surprise, fear, acceptance, shy, pride, appreciate, calmness, admire, contempt, love, happiness, exciting, regret, ease, discomfort, respect, like |

Table 3.3: Some existing definition of basic emotions. The most widely adopted model for affect recognition is Ekman's, although is one of the poorest in terms of number of emotions.

For their deep connection with language and for the limitedness of the emotional labels used, all such categorical approaches usually fail to describe the complex range of emotions that can occur in daily communication. The dimensional approach [127], in turn, represents emotions as coordinates in a multi-dimensional space. For both theoretical and practical reasons, more and more researchers like to define emotions according to two or more dimensions. An early example is Russell's circumplex model [128], which uses the dimensions of arousal and valence to plot 150 affective labels (Fig. 3.2). Similarly, Whissell considers emotions as a continuous 2D space whose dimensions are evaluation and activation [129]. The evaluation dimension measures how a human feels, from positive to negative. The activation dimension measures whether humans are more or less likely to take some action under the emotional state, from active to passive (Fig. 3.3). In her study, Whissell assigns a pair of values <activation, evaluation> to each of the approximately 9,000 words with affective connotations that make up her Dictionary of Affect in Language.

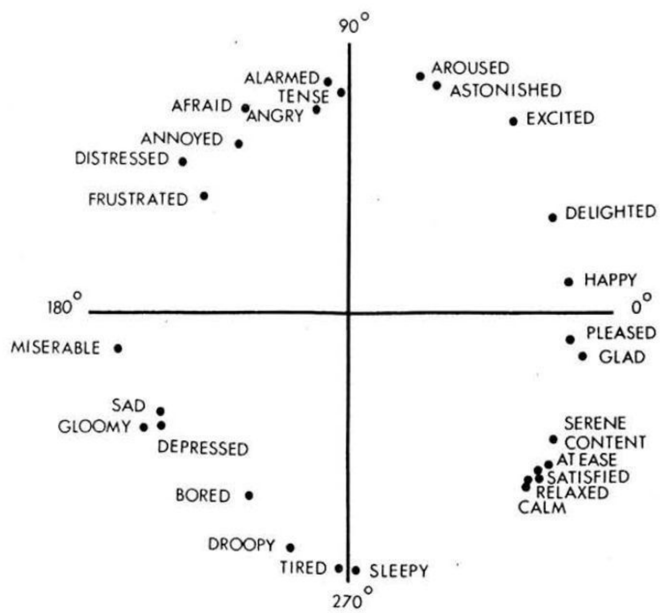


Figure 3.2: Russell's circumplex model is one of the earliest examples of dimensional emotion representations. In the snippet, direct circular scaling coordinates are provided for 28 affect words.

Another bi-dimensional model is Plutchik's wheel of emotions, which offers an integrative theory based on evolutionary principles [108]. Following Darwin's thought, the functionalist approach to emotions holds that emotions have evolved for a particular function, such as to keep the subject safe [109, 130]. Emotions are adaptive as they have a complexity born of a long evolutionary history and, although we conceive emotions as feeling states, Plutchik says the feeling state is part of a process involving both cognition and behaviour and containing several feedback loops. He created a wheel of emotions in 1980, which consisted of 8 basic emotions and 8 advanced emotions each composed of 2 basic ones. In such model, the vertical dimension represents intensity and the radial dimension represents degrees of similarity among the emotions.

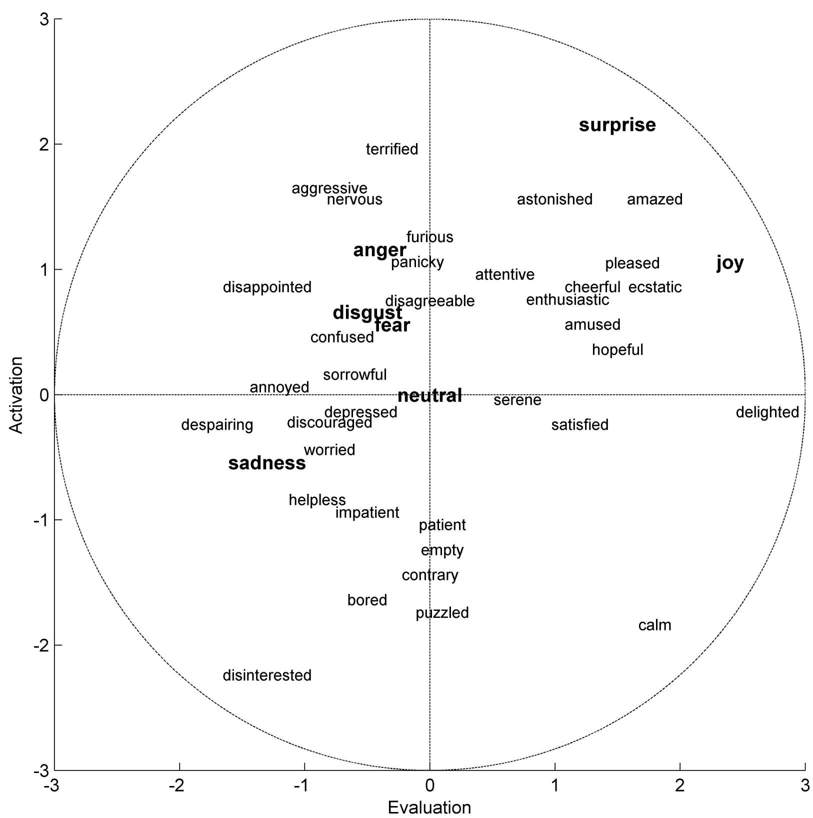


Figure 3.3: Whissell's model is a bi-dimensional representation of emotions, in which affect words are displayed. The diagram shows the position of some of these words in the <activation, evaluation> space.

Besides bi-dimensional approaches, a commonly used set for emotion dimension is the <arousal, valence, dominance> set, which is known in the literature also by different names, including <evaluation, activation, power> and <pleasure, arousal, dominance> [131]. Recent evidence suggests there should be a fourth dimension: Fontaine et al. reported consistent results from various cultures where a set of four dimensions is found in user studies, namely <valence, potency, arousal, unpredictability> [132]. Dimensional representations of affect are attractive mainly because they provide a way of describing emotional states that is more tractable than using words. This is of particular importance when dealing with naturalistic data, where a wide range of emotional states occurs. Similarly, they are much more able to deal with non-discrete emotions and variations in emotional states over time [133], since in such cases changing from one universal emotion label to another would not make much sense in real life scenarios.

Dimensional approaches, however, have a few limitations. Although the dimensional space allows to compare affect words according to their reciprocal distance, it usually does not allow making operations between these, e.g., for studying compound emotions. Most dimensional representations, moreover, do not model the fact that two or more emotions may be experienced at the same time. Eventually, all such approaches work at word level, which makes them unable to grasp the affective valence of multiple-word concepts.

### 3.2.2 A New Cognitive Model for Representing Human Emotions

The Hourglass of Emotions, proposed by Cambria et al. [98], is an affective categorisation model inspired by Plutchik's studies on human emotions [108]. It reinterprets Plutchik's model by organising primary emotions around four independent but concomitant dimensions, whose different levels of activation make up the total emotional state of the mind. Such a reinterpretation is inspired by Minsky's theory of the mind, according to which brain activity consists of different independent resources and that emotional states result from turning some set of these resources on and turning another set of them off [112]. This way, the model can potentially synthesise the full range of emotional experiences in terms of Pleasantness, Attention, Sensitivity, and Aptitude, as the different combined values of the four affective dimensions can also model affective states we do not have a specific name for, due to the ambiguity of natural language and the elusive nature of emotions.

The main motivation for the design of the model is the concept-level inference of the cognitive and affective information associated with text. Such faceted information is needed, within sentic computing, for a feature-based sentiment analysis, where the affective common sense knowledge associated with natural language opinions has to be objectively assessed. Therefore, the Hourglass model systematically excludes what are variously known as self-conscious or moral emotions such as pride, guilt, shame, embarrassment, moral outrage, or humiliation [134, 135, 136, 137]. Such emotions, in fact, present a blind spot for models rooted in basic emotions, because they are by definition contingent on subjective moral standards. The distinction between guilt and shame, for example, is based in the attribution of negativity to the self or to the act. So, guilt arises when believing to have done a bad thing, and shame arises when thinking to be a bad person. This matters because in turn, these emotions have been shown to have different consequences in terms of action tendencies. Likewise, an emotion such as *schadenfreude* is essentially a form of pleasure, but it is crucially different from pride or happiness because of the object of the emotion (the misfortune of another that is not caused by the self), and the resulting action tendency (do not express).

However, since the Hourglass model currently focuses on the objective inference of affective information associated with natural language opinions, appraisal-based emotions are not taken into account within the present version of the model. The Hourglass model, in fact, is a biologically-inspired and psychologically-motivated model based on the idea that emotional states result from the selective activation/disactivation of different resources in the brain. Each such selection changes how we think by changing our brain's activities: the state of anger, for example, appears to select a set of resources that help us react with more speed and strength while also suppressing some other resources that usually make us act prudently. Evidence of this theory is also given by several fMRI experiments showing that there is a distinct pattern of brain activity that occurs when people are experiencing different emotions. Zeki and Romaya, for example, investigated the neural correlates of hate with an fMRI procedure [138]. In their experiment, people had their brains scanned while viewing pictures of people they hated. The results showed increased activity in the medial frontal gyrus, right putamen, bilaterally in the premotor cortex, in the frontal pole, and bilaterally in the medial insula of the human brain. Also the activity of emotionally enhanced memory retention can be linked to human evolution [139]. During early development, in fact, responsive behaviour to environmental events is likely to have progressed as a process of trial-and-error.

Survival depended on behavioural patterns that were repeated or reinforced through life and death situations. Through evolution, this process of learning became genetically embedded in humans and all animal species in what is known as 'fight or flight' instinct [140]. The primary quantity we can measure about an emotion we feel is its strength. But, when we feel a strong emotion, it is because we feel a very specific emotion. And, conversely, we cannot feel a specific emotion like fear or amazement without that emotion being reasonably strong. For such reasons, the transition between different emotional states is modelled, within the same affective dimension, using the function $G(x) = -\frac{1}{\sigma\sqrt{2\pi}}e^{-x^2/2\sigma^2}$, for its symmetric inverted bell curve shape that quickly rises up towards the unit value (Fig. 3.4).

In particular, the function models how the level of activation of each affective dimension varies from the state of 'emotional void' (null value) to the state of 'heightened emotionality' (unit value). Justification for assuming that the Gaussian function (rather than a step or simple linear function) is appropriate for modelling the variation of emotion intensity is based on research into the neural and behavioural correlates of emotion, which are assumed to indicate emotional intensity in some sense.

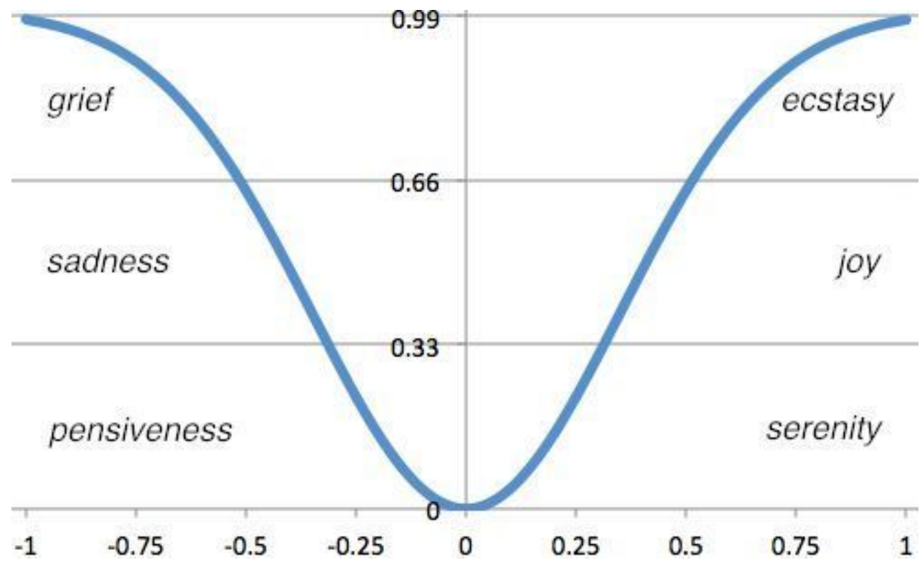


Figure 3.4: The Pleasantness emotional flow. Within each affective dimension, the passage from a sentic level to another is regulated by a Gaussian function that models how stronger emotions induce higher emotional sensitivity.

In fact, nobody genuinely knows what function subjective emotion intensity follows, because it has never been truly or directly measured [141]. For example, the so-called Duchenne smile (a genuine smile indicating pleasure) is characterised by smooth onset, increasing to an apex, and a smooth, relatively lengthy offset [142]. More generally, Klaus Scherer has argued that emotion is a process characterised by non-linear relations among its component elements - especially physiological measures, which typically look Gaussian [143]. Emotions, in fact, are not linear [108]: the stronger the emotion, the easier it is to be aware of it.

Mapping this space of possible emotions leads to a hourglass shape (Fig. 3.5). It is worth to note that, in the model, the state of 'emotional void' is a-dimensional, which contributes to determine the hourglass shape. Total absence of emotion, in fact, can be associated with the total absence of reasoning (or, at least, consciousness) [144], which is not an envisaged mental state as, in human mind, there is always something going on. The Hourglass of Emotions, in particular, can be exploited in the context of HCI to measure how much respectively: the user is amused by interaction modalities

(Pleasantness), the user is interested in interaction contents (Attention), the user is comfortable with interaction dynamics (Sensitivity), the user is confident in interaction benefits (Aptitude). Each affective dimension, in particular, is characterised by six levels of activation (measuring the strength of an emotion), termed 'sentic levels', which represent the intensity thresholds of the expressed/perceived emotion. These levels are also labelled as a set of 24 basic emotions [108], six for each of the affective dimensions, in a way that allows the model to specify the affective information associated with text both in a dimensional and in a discrete form (Table 3.4).

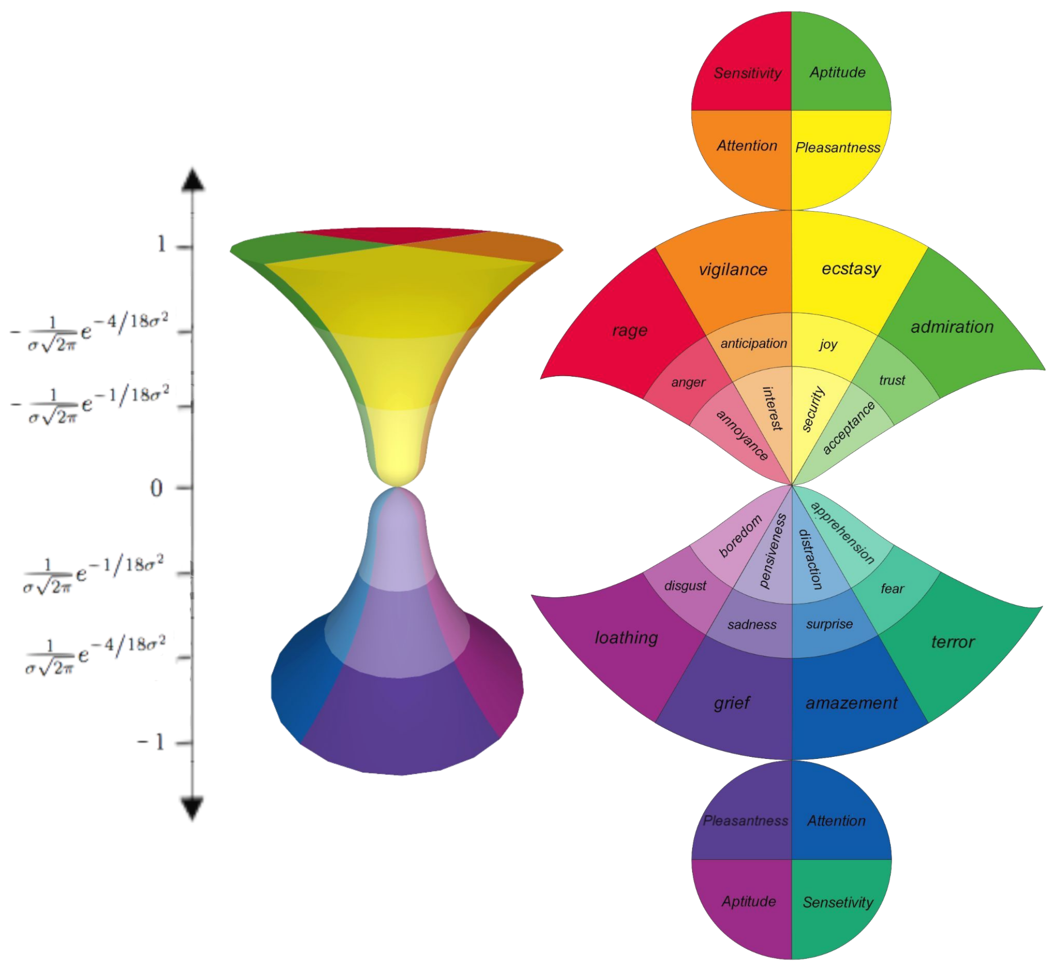


Figure 3.5: The 3D model and the net of the Hourglass of Emotions: since affective states are represented according to their strength (from strongly positive to null to strongly negative), the model assumes a hourglass shape.

| Interval | Pleasantness | Attention | Sensitivity | Aptitude |
|---|---|---|---|---|
| [G(1), G(2/3)) | ecstasy | vigilance | rage | admiration |
| [G(2/3), G(1/3)) | joy | anticipation | anger | trust |
| [G(1/3), G(0)) | serenity | interest | annoyance | acceptance |
| (G(0), –G(1/3)] | pensiveness | distraction | apprehension | boredom |
| (–G(1/3), –G(2/3)] | sadness | surprise | fear | disgust |
| (–G(2/3), –G(1)] | grief | amazement | terror | loathing |

Table 3.4: The sentic levels of the Hourglass model. Labels are organised into four affective dimensions with six different levels each, whose combined activity constitutes the 'total state' of the mind.

The dimensional form, in particular, is termed 'sentic vector' and it is a four-dimensional $float$ vector that can potentially synthesise the full range of emotional experiences in terms of Pleasantness, Attention, Sensitivity, and Aptitude. In the model, the vertical dimension represents the intensity of the different affective dimensions, i.e., their level of activation, while the radial dimension represents k-lines [62] that can activate configurations of the mind, which can either last just a few seconds, e.g., surprise, or years, for long-term emotional states such as love.

The model follows the pattern used in colour theory and research in order to obtain judgements about combinations, i.e., the emotions that result when two or more fundamental emotions are combined, in the same way that red and blue make purple. Hence, some particular sets of sentic vectors have special names as they specify well-known compound emotions (Fig. 3.6).

For example, the set of sentic vectors with a level of Pleasantness $\in$ [G(2/3), G(1/3)), i.e., joy, a level of Aptitude $\in$ [G(2/3), G(1/3)), i.e., trust, and a minor magnitude of Attention and Sensitivity, are termed 'love sentic vectors' since they specify the compound emotion of love (Table 3.5).

| | **Attention>0** | **Attention<0** | **Aptitude>0** | **Aptitude<0** |
|---|---|---|---|---|
| **Pleasantness>0** | optimism | frivolity | love | gloat |
| **Pleasantness<0** | frustration | disapproval | envy | remorse |
| **Sensitivity>0** | aggressiveness | rejection | rivalry | contempt |
| **Sensitivity<0** | anxiety | awe | submission | coercion |

Table 3.5: The second-level emotions generated by pairwise combination of the sentic levels of the Hourglass model. The co-activation of different levels gives birth to different compound emotions, e.g., love and frustration.

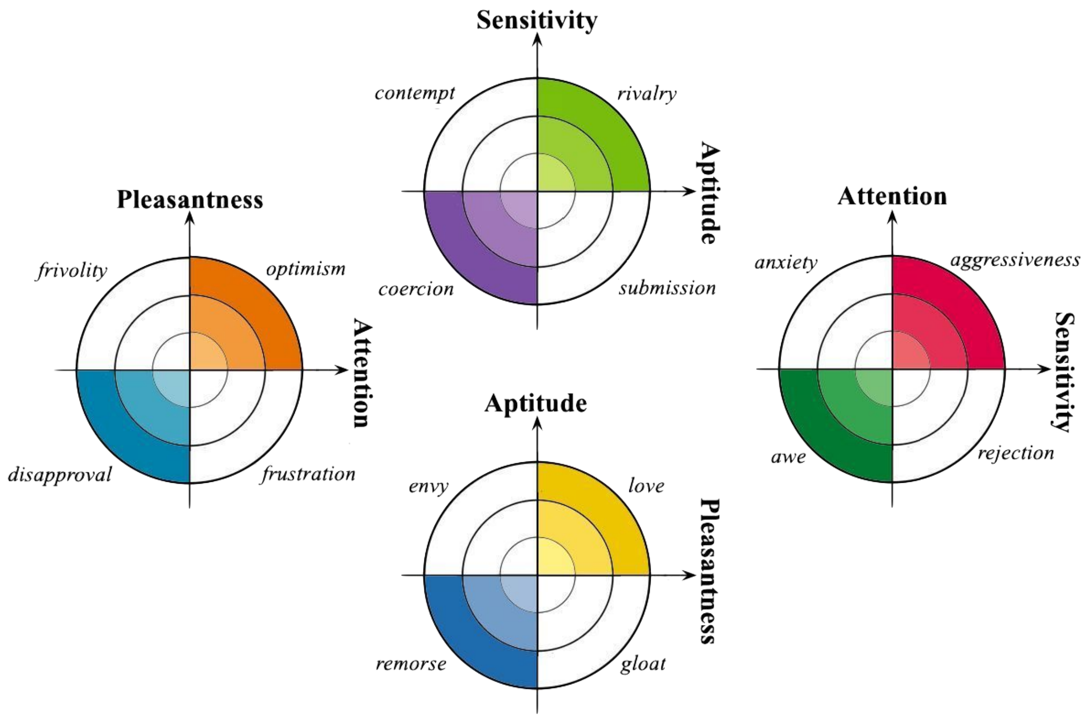


Figure 3.6: Hourglass compound emotions of second level: by combining basic emotions pairwise it is possible to obtain complex emotions resulting from the activation of two of the four affective dimensions.

More complex emotions can be synthesised by using three, or even four, sentic levels, e.g., joy + trust + anger = jealousy. Therefore, analogous to the way primary colours combine to generate different colour gradations (and even colours we do not have a name for), the primary emotions of the Hourglass model can blend to form the full spectrum of human emotional experience. Beyond emotion detection, the Hourglass model is also used for polarity detection tasks. Since polarity is strongly connected to attitudes and feelings, in fact, it is defined in term of the four affective dimensions, according to the formula proposed by Cambria et al. [145]:

$$p = \sum_{i=1}^{N} \frac{Pleasantness(c_i) + |Attention(c_i)| - |Sensitivity(c_i)| + Aptitude(c_i)}{3N} \quad (3.1)$$

where $c_i$ is an input concept and $N$ the total number of concepts, and 3 the normalisation factor (as the Hourglass dimensions are defined as float $\in$ [-1,+1]). In the formula, Attention is taken in absolute value since both its positive and negative intensity values correspond to positive polarity values (e.g., 'surprise' is negative in the sense of lack of Attention but positive from a polarity point of view). Similarly, Sensitivity is taken in negative absolute value since both its positive and negative intensity values correspond to negative polarity values (e.g., 'anger' is positive in the sense of level of activation of Sensitivity but negative in terms of polarity). The formula can be seen as one of the first attempts to show a clear connection between emotion recognition (sentiment analysis) and polarity detection (opinion mining).

## 3.3 Open Mind Common Sentics: An Emotion-Sensitive IUI

In AffectNet, the general common sense knowledge contained in ConceptNet is exploited to spread affective information from selected affect seeds to other concepts. Besides exploiting the emotional content of the Open Mind corpus, AffectNet is also enriched by collecting new affective common sense knowledge through state-of-the-art crowd sourcing and games with a purpose (GWAP) techniques (subsection 3.3.1). In particular, Open Mind Common Sentics is developed. Open Mind Common Sentics is an emotion-sensitive IUI that serves both as a platform for affective common sense acquisition and as a publicly available NLP tool for extracting the cognitive and affective information associated with short texts (subsection 3.3.2).

### 3.3.1 Games for Knowledge Acquisition

The Casual Games Association[2] reports more than 200 million casual gamers worldwide this year. People play games for different reasons, e.g., to relax, to be entertained, for the need of competition and to be thrilled [146]. Additionally, they want to be challenged, both on a mental and on a skill-based level. Such army of gamers could

[2] http://casualgamesassociation.org

be exploited for performing tasks that are relatively easy to complete by humans, but computationally rather infeasible to solve [147]. The idea is to integrate such tasks as goal of games [148] by producing a win-win situation where people have fun playing games while actually doing something useful. The nature of these games, in fact, focuses on exploiting player inputs to both create meaningful data and provide a funnier game experience [149]. Such human-based computational power can be exploited for tasks such as video annotation, e.g., OntoTube[3] [150], PopVideo[4] (Fig. 3.7), Yahoo's Videotaggame [151], and Waisd[5] [152], in which two players have to timely agree on a set of tags about the same streaming YouTube[6] video. Similarly, in ESP game[7] [153] and Google Image Labeler (before being discontinued last September) players have to consensually guess content objects or properties of random images by simultaneously typing what they see.

Other games for image annotation include Matchin[8] [154], which focuses on image perceived quality by asking players to pairwise choose the picture they like better, Phetch [155], a game that collects explanatory descriptions of images in order to improve accessibility of the Web for the visually impaired by letting a player describe an image and others retrieve it using an image search engine, Peekaboom [156], which focuses on locating objects within images by letting a player reveal specific parts of an image in order for the other to guess the correct object name, Squigl[9], in which players have to spot objects in images previously annotated within ESP Game, and Picture This, which asks players to choose, among a set of images, the one that best suits the given query. Among games for image annotation, there are also games for streamlining the robustness evaluation of CAPTCHAs, namely: Magic Bullet[10] [157], a team game in which players need to agree on the meaning of CAPTCHAs, and TagCaptcha[11] [158],

[3] `http://ontogame.sti2.at/games`
[4] `http://gwap.com/gwap/gamesPreview/popvideo`
[5] `http://waisda.nl`
[6] `http://youtube.com`
[7] `http://gwap.com/gwap/gamesPreview/espgame`
[8] `http://gwap.com/gwap/gamesPreview/matchin`
[9] `http://gwap.com/gwap/gamesPreview/squigl`
[10] `http://homepages.cs.ncl.ac.uk/jeff.yan/mb.htm`
[11] `http://dolphin.unige.ch/tagcaptcha`

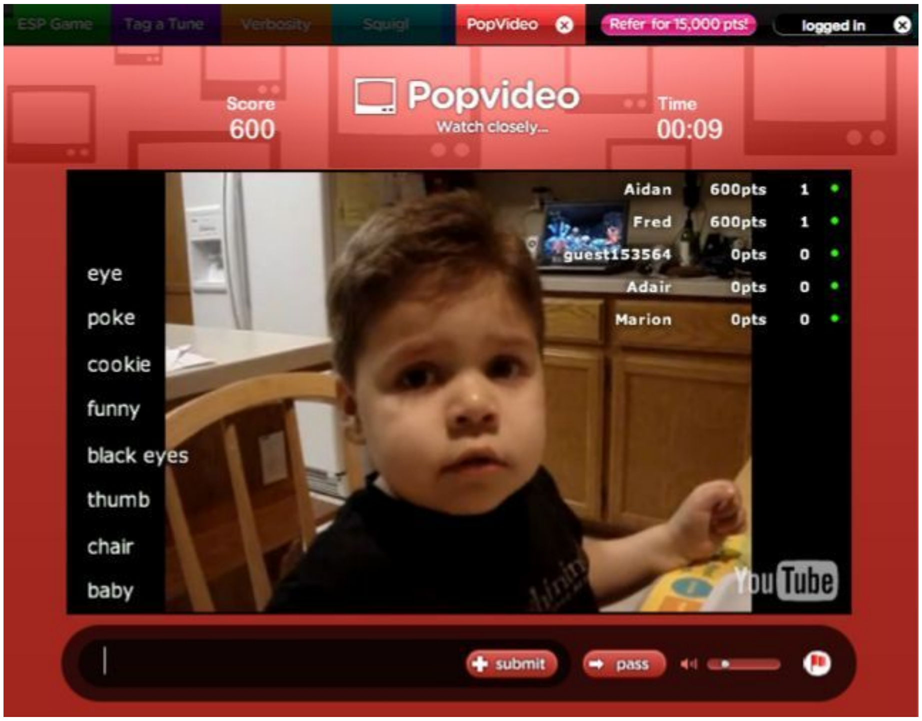


Figure 3.7: A screenshot of PopVideo. Two or more players are shown the same video, which originates from various sources, and have to timely describe the objects that appear in the video in order earn points.

where players are asked to quickly describe CAPTCHA images with one word each. GWAPs are also exploited to automatically tag music with deeper semantic labels. HerdIt[12] [159], for example, asks players have to accomplish different tasks related to the song they are listening to, while in Tagatune[13] [160] two players have to listen to an audio file and describe to the other what they are hearing, in order for him/her to decide whether the game has played the same soundtrack to both or not. Several games have also been designed for text annotation. Verbosity[14] [161], for example, is a real time quiz game for collecting common sense facts. In the game, two players take different roles at different times: a narrator, who has to describe a word using templates, and a

[12] http://herdit.org/music
[13] http://gwap.com/gwap/gamesPreview/tagatune
[14] http://gwap.com/gwap/gamesPreview/verbosity

guesser, who has to guess such word in the shortest time possible. Sentiment Quiz[15], instead, gathers information about the polarity associated to words. It asks its players to evaluate random words on a five grade scale, from very negative over neutral to very positive. Phrase Detectives[16] tries to identify relationships between phrases and other phrases or sentences. Another approach to collecting common sense knowledge is the FACTory Game[17] [80], published by Cycorp. FACTory randomly chooses facts from Cyc and presents them to players, in order for them to guess whether a statement is true, false, or does not make sense. A variant of the FACTory game is the Concept Game on Facebook[18] [162], which collects common sense knowledge by proposing random assertions to users in a slot machine fashion and asking them to decide whether this assertion is meaningful or not (Fig. 3.8).

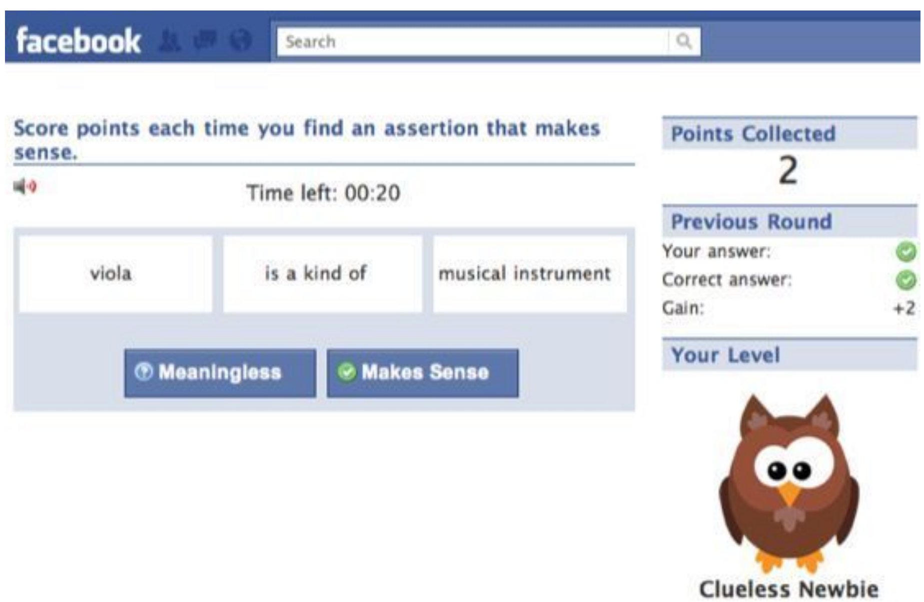


Figure 3.8: A screenshot of Concept Game, a turn-based single player game taking advantage of the Facebook platform for adding competitive flavors and social aspects to the collection of common sense knowledge.

[15]http://www.facebook.com/ecoresearch.sentiment.quiz
[16]http://anawiki.esex.ac.uk/phrasedetective
[17]http://game.cyc.com
[18]http://apps.facebook.com/conceptgame

Page Hunt[19] [163] is a GWAP for the annotation of websites. It allows to index web sites and hence to improve the search index of a search engine (Microsoft Bing). The player gets assigned a random website and is asked to describe it with keywords. The game then shows players the top five page hits for the entered keywords and they are rewarded depending on how high ranked the previously assigned web-page is in the result set. Other GWAPs engage players in building ontologies. OntoPronto [150], for example, is a quiz game for vocabulary building that attempts to build a huge domain ontology from Wikipedia articles. This is achieved by mapping random articles to the most specific class of the Proton ontology[20] using the *subClassOf* relationship.

Virtual Pet Game[21] [164] aims to construct a semantic network that encodes common sense knowledge. The game is built on top of PPT, a popular Chinese bulletin board system that is accessible through a terminal interface. Each player owns a pet, which they should take care of by asking and answering questions. The pet in this game is just a substitute for other players, who receive such questions and answers, and have to respond or validate them.

Rapport Game[22] [164], similarly to Virtual Pet Game, exploits player labour for constructing a semantic network that encodes common sense knowledge. Rapport Game, however, is built on top of Facebook and uses direct interaction between players. An interesting game for the creation of formal domain ontologies from Linked Open Data is Guess What?![23] [165]. Given a seed concept, a player has to find a matching URI in DBpedia, Freebase and OpenCyc.

The resulting labels/URIs are analysed by simple NLP tools in order to identify expressions that can be translated into logical operators and break down complex descriptions into small fragments. The game starts with the most general fragment and, at each round, a more specific fragment is connected to it through a logical operator, with players having to guess the concept described by it. There are GWAPs that try

[19] `http://pagehunt.msrlivelab.com`
[20] `http://proton.semanticweb.org`
[21] `http://agents.csie.ntu.edu.tw/commonsense/cate2_1_en.html`
[22] `http://apps.facebook.com/conceptnet`
[23] `http://nitemaster.de/guesswhat/manual.html`

to align ontologies. Wordhunger[24], for example, is a web-based application that maps WordNet synsets to Freebase. Each game round consists of a WordNet term and up to three suggested possible Freebase articles, among which players have to select the most fitting ( or pass or select 'no match').

SpotTheLink[25] is a two player game focusing on the alignment of random concepts from the DBpedia Ontology[26] to the Proton upper ontology. Each player has to select Proton concepts that are either the same as or more specific than a randomly selected DBpedia concept. The data generated by SpotTheLink is a SKOS mapping between the concepts of the two input ontologies. Based on Wikipedia, there are three Wikiracing game, The Wiki Game[27], Wikispeedia[28] and WikipediaMaze[29], where the objective is to find connections between two Wikipedia articles by clicking links within the text. WikipediaGame and Wikispedia focus on completing the race faster and with fewer clicks than other players. In WikipediaMaze, instead, players are allowed to create races for each other and are incentivised to create and play races by earning badges.

### 3.3.2 Collecting Affective Common Sense Knowledge

Human emotions and their modelling are increasingly understood to be a crucial aspect in the development of intelligent systems [12, 95, 112]. Emotions are a basic part of human communication and have therefore to be taken into account for the development of more effective interfaces for human-machine communication such as chat systems, e-house, e-learning, e-health or emphatic voice boxes. Besides general common sense knowledge, the Open Mind corpus also contains affective information, e.g., "a gift is for celebrating a birthday" or "making a mistake causes embarrassment", as common sense encompasses, among many other aspects, also knowledge about the emotional facets of typical everyday life. However, the amount of affective information contained

[24] http://wordhunger.freebaseapps.com
[25] http://semanticgames.org/2011/11/spotthelink
[26] http://dbpedia.org/Ontology
[27] http://thewikigame.com
[28] http://www.cs.mcgill.ca/~rwest/wikispeedia
[29] http://www.wikimaze.me

in the Open Mind corpus is still very limited, as relationships such as ArisesEmotion, MakesFeel or AffectivelyRelated are missing from the set of properties that are currently used for collecting pieces of common sense knowledge from the public.

Since computers now have the ability to search vast amounts of data in little time, the use of a search engine to collect the affective information needed is pretty tempting. To this end, different lexical patterns, termed sentic patterns, are used for extracting affective information from the Web. Such patterns re built using label sequential rules (LSRs), which are generated from sequential patterns in data mining [166]. A rule is of the form $X \rightarrow Y$, where $Y$ is a sequence and $X$ is a sequence produced from $Y$ by replacing some of its items with wildcards, denoted by a '*', which can match any item. During the learning process, each segment is converted to a sequence. Each sequence element is a word, which is represented by both the word itself and its POS tag in a set.

In the training data, all concepts are manually labelled and replaced by the label *$concept*. A concept can be expressed with a noun (NN), adjective (JJ), verb (VB) or adverb (RB). The labels and their POS tags used in mining LSRs are {*$concept*, NN}, {*$concept*, JJ}, {*$concept*, VB} and {*$concept*, RB}, where *$concept* denotes a concept to be extracted. For example, the sentence segment "chocolate makes me feel happy" is turned into the sequence <{chocolate, NN}{make (me|you) feel, VB}{happy, JJ}>. After labelling, it becomes <{*$concept*, NN}{make (me|you) feel, VB}{happy, JJ}>. All the resulting sequences are then used to mine LSRs. A typical rule, for example, is <{*, NN}{put (me|you) on, VB}{cloud nine, NN}> → <{*$concept*, NN}{put (me|you) on, VB}{cloud nine, NN}> confidence = 80%, where the confidence is the conditional probability, $\Pr(Y|X)$, which measures the accuracy of the rule. Concept extraction is performed by matching the patterns with each sentence segment in a new web-page to extract affective information about concepts contained AffectNet. That is, the words in the sentence segment that both match *$concept* in a pattern and an AffectNet concept are extracted. In the pattern match, only the right-hand side of each rule is used. In rule generation, both the right- and the left-hand sides are needed to compute the conditional probability or confidence. Such patterns yield some useful data, however, they are not good enough for the purposes of the proposed study for three reasons. Firstly, most of the affective common sense knowledge aimed to be collected is so obvious that no one has bothered to record it. Secondly, there exists incorrect knowledge on the Web (for example, the query "plastic can love" returns 33,200 results on Google, while "plastic cannot love" returns just a few links). Thirdly, the text on the Web is unstructured and turning it into a directly useful format is a non-trivial task.

For these reasons, crowd sourcing techniques are mainly used. Distributed online knowledge acquisition projects have become quite popular in the past years. Examples include Freebase[30], with its 1,450 concepts, WikiTaxonomy, counting 127,000 concepts, YAGO[31], with 149,162 instances, NELL[32], containing 959,654 beliefs, ProBase[33], Microsoft's universal probabilistic ontology, and the different projects associated with the Open Mind Initiative, e.g., OMCS, Open Mind Word Expert [167], an active learning system that aims to create large annotated corpora, and Open Mind Indoor Common Sense [168], which aims to develop intelligent mobile robots for use in home and office environments.

In a similar fashion to the Open Mind family of distributed knowledge capture projects, Open Mind Common Sentics, proposed by Cambria et al. [12], aims to collect affective common sense knowledge for sentiment analysis (Fig. 3.9). Whereas previous approaches have mainly relied on paid experts or unpaid volunteers, much stronger emphasis is hereby put on creating a system that is appealing to a large audience of people, regardless of whether or not they are interested in contributing to AI. The fundamental aim of Open Mind Common Sentics, in fact, is to transform as much as possible the activity of entering knowledge into an enjoyable interactive process.

[30] http://freebase.com
[31] http://mpi-inf.mpg.de/yago-naga/yago
[32] http://rtw.ml.cmu.edu/rtw
[33] http://research.microsoft.com/probase

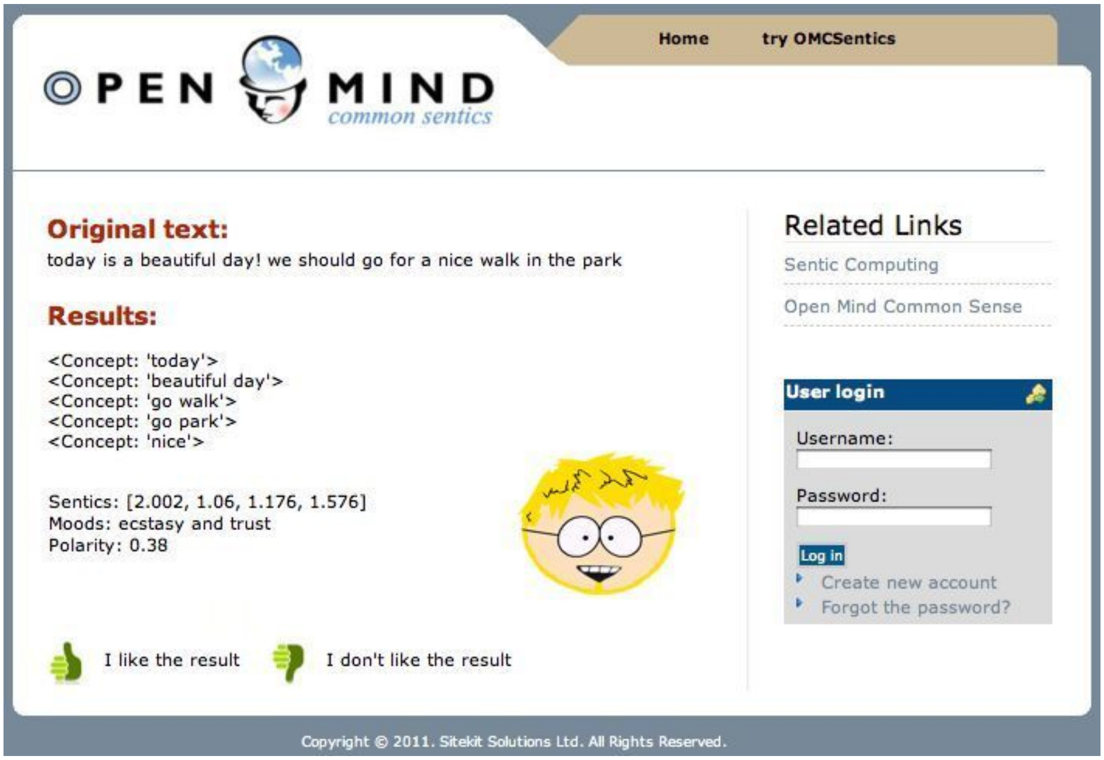


Figure 3.9: A typical output obtained by inserting some text into Open Mind Common Sentics interface. The system provides a list of extracted concepts, their valence and polarity, a list of sentics, and a polarity value.

To this end, the system adopts a two-fold strategy: crowd sourcing, that is challenge volunteers over the Web through mood-spotting and fill-in-the blank questions, in the same wake as Open Mind Commons [107], and GWAPs, that is engage users through online games, in the same wake as Verbosity and ESP game. In particular, the mood-spotting questions consist in asking users to select an emoticon according the overall affect they can infer from a given sentence. The fill-in-the blank questions, in turn, are sentences to be completed such as “opening a Christmas gift makes feel ___”. Sometimes, one or more taboo affect concepts are shown in the game window.

Such concepts are entries that have been validated a sufficient amount of times; hence they are not valid input any more (in order to collect synonyms or alternatives of a given affective common sense concept). As for the GWAPs, the Hourglass Game was developed (Fig. 3.10). The Hourglass Game is a speed game consisting in selecting from the Hourglass model the sentic level that is most likely associated with a given affective concept. Players earn points not only according to accuracy but also quickness in clicking on the right area of the Hourglass. The game is quite engaging, although very simple, and players like to challenge each other to see who has higher emotional quotient (EQ) but users are not too keen on playing more than once.

What is lacking from most of crowd-sourcing and GWAP techniques, in fact, is stickiness. GWAPs can be fun to play for a relatively short period of time but then players are not too keen on returning. In other words, GWAPs generally have a pretty low sticky factor. The sticky factor is defined as the amount of daily active users (DAUs) of an application divided by the number of monthly active users (MAUs). MAU is the most-quoted measure of a game's size, but it is effective only to discuss size or reach, not engagement. DAU, in turn, can be a very valuable number as it relates how much activity a game is seeing on a daily basis, but it falls into the same trap as MAU in that it does not discriminate between retention and acquisition. The single-most important metric for engagement is stickiness, i.e., DAU/MAU, which allows to more accurately calculate repeat visits and average knowledge acquired per user (AKAPU). A key for driving the sticky factor, besides great game play, is the ability of the application to prompt users to reach out to their friends, e.g., via stories and pictures about their gameplay.

To this end, Sentic Pet is being developed (Fig. 3.11). Sentic Pet is a massively multiplayer online (MMO) game in which players have to raise and take care of their own pets. Unlike old-style tamagotchi games, in Sentic Pet, raising and caring pets is not about cleaning, feeding and petting them, but rather training them, both at mental and skill level, by playing mini-GWAPs.

Targeting players of a wide age range (10 to 50 year old), the game should appeal anyone who enjoyed and enjoys PetVille[34] or FarmVille[35]. Players start from level 1

[34] `http://petville.com`
[35] `http://farmville.com`

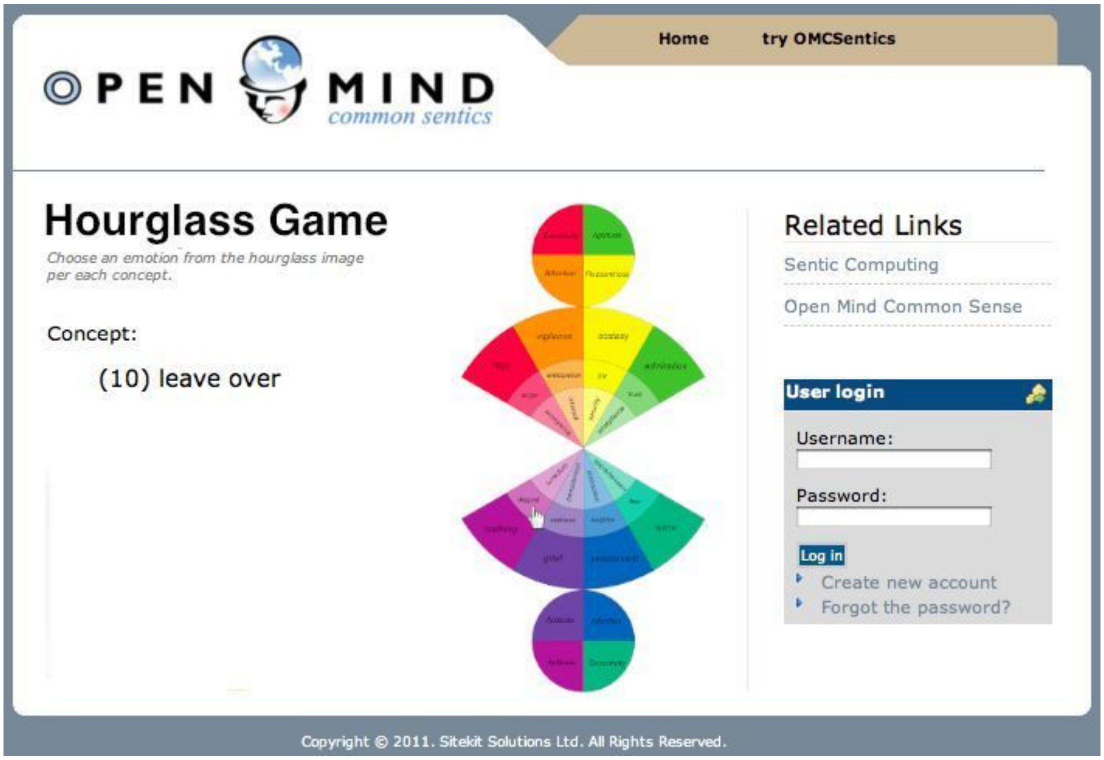


Figure 3.10: A screenshot of the Hourglass Game. The GWAP aims to collect affective common sense knowledge from the general public by engaging users in a game where precision and speed are awarded.

with their pet being a baby born having very little affective common sense knowledge. The game involves balancing two main activities: training the pet and testing its skills by challenging other players. Training does not involve simply teaching the pet new knowledge, but also refining acquired knowledge. Challenges can be taken both at mental and skill level, which involve different kinds of activities. At mental level, for example, pets can be challenged according to different modalities, e.g., affective vocabulary learning (in the same wake of Verbosity) or affective meaning of images (in the same wake of ESP game).

Pets can level-up according to the combination of IQ (light bulb icon) and EQ (heart icon) points earned playing the mini-GWAPs. Data are validated by majority and reputation, that is, the confidence score with which a piece of affective common sense knowledge is saved into the knowledge base depends on how many players validated it and on the expertise level of these. Open Mind Common Sentics is an example of an emerging class of games that can be considered ‘human algorithms’, since humans act as processing nodes for problems that computers cannot yet solve. By providing an incentive for players, a large quantity of computing power can be gained and harnessed for multiple applications. Constructing a complete affective common sense database would be extremely beneficial for many communities, e.g., sentiment analysis and HCI, and Open Mind Common Sentics can be highly effective in doing so.

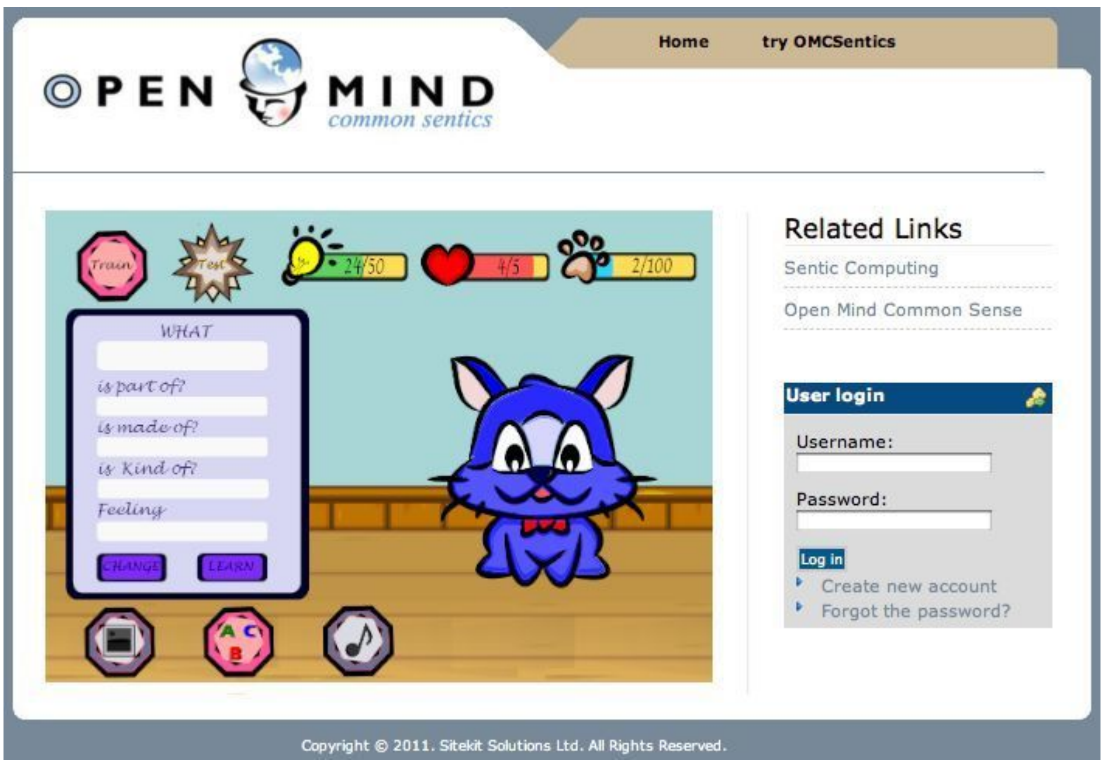


Figure 3.11: A screenshot of Sentic Pet. Icons in the up-right corner specify the abilities of pets, while icons in the up-left corner allow players to train or test these according to different modalities (bottom-left corner).

## 3.4 Isanette: A Common and Common Sense Knowledge Base

Common sense is the key to natural language understanding. However, it is not the only kind of knowledge involved in standard human-to-human communication. When communicating with each other, in fact, people usually refer to existing facts and circumstances and build new useful, funny or interesting information on the top of those. This common knowledge comprehends information usually found in news, articles, debates, lectures, etc. (factual knowledge) but also principles and definitions that can be found in collective intelligence projects such as Wikipedia (vocabulary knowledge).

Attempts to build a common knowledge base are countless and comprehend both resources crafted by human experts or community efforts, such as WordNet and Freebase, and automatically-built knowledge bases, such as WikiTaxonomy, YAGO, and ProBase, which is the most comprehensive resource with its 12 million concepts. In this thesis work, in fact, the richness of Probase (subsection 3.4.1) is exploited to enhance the topic-spotting capabilities of the opinion mining engine. In particular, Probase is blended with ConceptNet in order to build possibly the most comprehensive resource of common and common sense knowledge for sentiment analysis (subsection 3.4.2).

### 3.4.1 Probase

Probase is a research prototype that aims to build a unified taxonomy of worldly facts from web data and search log data [169]. The taxonomy consists of about 8 million concepts (e.g., *painter*), 17 million instances (e.g., “Leonardo da Vinci”), attributes and values (e.g., “Leonardo’s birthday is April 15, 1452”), and relationships (e.g., “Mona Lisa is painted by Leonardo”). Such information is learned iteratively from 1.68 billion web-pages in Bing web repository. The taxonomy is probabilistic, which means that every claim in ProBase is associated with some probabilities that model the claim’s correctness, ambiguity and other characteristics. The probabilities are derived from evidences found in web data, search log data, and other available data.

The core taxonomy consists of the *IsA* relationships extracted by using syntactic patterns such as the Hearst patterns [170]. For example, a part of text like "... artists such as Pablo Picasso ..." is considered as a piece of evidence for the claim that "Pablo Picasso" is an instance of the concept *artist*. Next, given a concept C, syntactic patterns such as "What is the A of B" are used to find its attributes (where B is an instance of C, and A is the attribute of interest). For example, sentences such as "What is the capital of China?" and "What is the GDP of Japan?" suggest that "capital" and "GDP" are candidate attributes of concept *country* (Fig. 3.12). Furthermore, every claim in Probase is associated with a few scores that model the consensus, typicality, ambiguity, and other characteristics of the claim.

In Table 3.6, Probase is compared with WordNet and Freebase. WordNet specialises in the linguistics of English words. For the word *cat*, WordNet has detailed descriptions of its various senses, although many of them are rarely used, or even unknown to many people (e.g., *gossip* and *woman* as concepts for *cat*). Also, it does not contain information for entities such as *IBM*, which is not considered as a word.

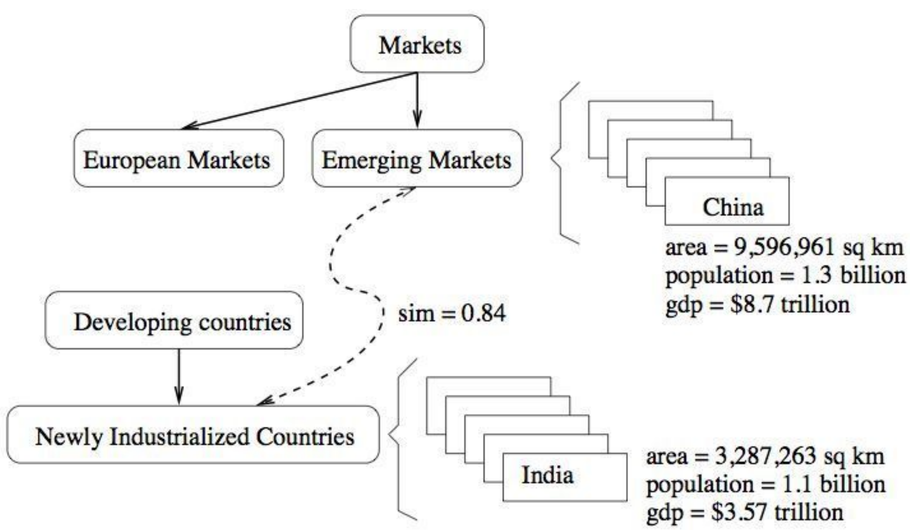


Figure 3.12: A snippet of Probase. The taxonomy consists of concepts (e.g., *emerging markets*), instances (e.g., "China"), attributes and values (e.g., "China's population is 1.3 billion"), and relationships between concepts.

Freebase, on the other hand, contains limited number of concepts for the word *cat*. In fact, the categories there are biased and sometimes inaccurate. For example, Freebase's concept space is biased toward entertainment, media related concepts. More importantly, the categories in WordNet and Freebase are not ranked or scored, and users cannot tell the difference in terms of their importance or typicality. In comparison, the concepts in Probase are more consistent with human's common knowledge. Concepts such as *gossip* and *woman* for *cat* are either not included or ranked very low because people rarely associate them with *cat*. In addition, for a word such as *language*, Probase indicates it can be both an instance on its own or an attribute of some concepts. Thus, Probase not only represents the largest taxonomy currently available but it also provides information that is missing in commonly used resources such as WordNet and Freebase. For this reasons, Probase is selected as the common knowledge base to be exploited by the proposed opinion mining engine for the detection of topics in opinionated text.

### 3.4.2 Building the Instance-Concept Matrix

Common knowledge is factual and vocabulary knowledge that is known by everyone, on which people base their arguments when communicating with each other. Common knowledge, however, is just a surface layer of human communication and alone is not enough for understanding natural language. To this end, common sense knowledge is needed. Common sense knowledge is very important as it spans a huge portion of human experience. It is, however, typically omitted from social communications, hence cannot be retrieved from the Web.

As a subsumption common knowledge base, in fact, Probase lacks information like "a dog is a best friend" (rather than simply an 'animal') or "a rose is a kind of meaningful gift" (rather than simply a kind of 'flower'), i.e., common sense that is not usually stated in web-pages (or, at least, not that often to be extracted by Hearst patterns with a high enough confidence score). To overcome this problem, Probase is enriched with complementary hyponym-hypernym common sense knowledge from ConceptNet.

| Term | WordNet Hypernyms | Freebase Types | Probase Concepts |
|---|---|---|---|
| Cat | Feline; Felid; Gossip; Gossiper; Gossipmonger; Rumormonger; Rumourmonger; Newsmonger; Stimulant; Stimulant drug; Excitant; Tracked vehicle; ... | TV episode; Creative work; Musical recording; Organism classification; Musical release; Book; Musical album; Film character; Publication; Character species; Animal; Domesticated animal; ... | Animal; Pet; Species; Mammal; Small animal; Thing; Mammalian species; Small pet; Animal species; Carnivore; Domesticated animal; Companion animal; Exotic pet; Vertebrate; ... |
| IBM | N/A | Companies listed on the New York Stock Exchange; Cloud computing providers; Companies based in Westchester County, New York; Multinational companies; Software companies of the United States; Top 100 US Federal Contractors; ... | Company; Vendor; Client; Corporation; Organisation; Manufacturer; Industry leader; Firm; Brand; Large company; Fortune 500 company; Technology company; Supplier; Software vendor; Global company; Technology company; ... |
| Language | Communication; Auditory communication; Word; Higher cognitive process; Faculty; Mental faculty; Module; Text; Textual matter; | Written work; Musical recording; Musical artist; Musical album; Literature subject; Query; Periodical; Type profile; Journal; Quotation subject; Broadcast genre; Periodical subject; Video game content descriptor; ... | **Instance of**: Cognitive function; Knowledge; Cultural factor; Cultural barrier; Cognitive process; Cognitive ability; Cultural difference; Ability; **Attribute of**: Film; Area; Book; Publication; Magazine; Work; Program; Media; ... |

Table 3.6: Comparison between Probase and two popular common knowledge bases. Probase overcomes problems such as incompleteness of WordNet hypernyms or inaccuracy and biased knowledge of Freebase.

In particular, all the assertions involving *IsA* relationships with a non-null confidence score, such as “dog is man’s best friend” or “a birthday party is a special occasion”, are extracted from the Open Mind corpus. Such assertions are exploited to generate a directed graph of about 15,000 nodes (interconnected by *IsA* edges), representing subsumption common sense knowledge.

One of the advantages of this new semantic network is that it allows to discriminate nodes between instance and concept, although in ConceptNet there is no such distinction. Instances, in fact, are represented by nodes with high out-degree and null in-degree, while concepts are nodes with high in-degree and null or low (as concepts can also be instances of other concepts) out-degree. Another advantage of this representation is that, since the new network is an unlabelled directed graph, the semantic relatedness of concepts is retained when applying dimensionality reduction techniques. Building a vector space representation of standard ConceptNet, in fact, is good in a meta-heuristic sense but we lose information about the relations governing how concepts self-organise in the multi-dimensional space, that is, we know that concepts close to each other are semantically related but we cannot say which specific relations make these concepts similar. If we apply dimensionality reduction on the *IsA* version of ConceptNet, instead, we know that common sense concepts self-organise according to the *IsA* relationship, e.g., if 'dog' and 'cat' are close to each other in the resulting vector space is because they are hypernyms of the same concepts (more details on how to handle the knowledge bases can be found in the next chapter).

The ensemble of Probase and ConceptNet, proposed by Cambria et al. [171] and termed Isanette (*IsA* net), is obtained by first representing Microsoft's knowledge base as a 2,715,218 $\times$ 1,331,231 matrix. Such hypernym-hyponym matrix is built out of 23,066,575 *IsA* triples with the form <instance, concept, confidence score>, which are reorganised as rows (e.g., 'pablo picasso'), columns (e.g., 'artist') and entries (e.g., '0.91') of the matrix, respectively. Performing reasoning on this matrix as it is would not be very convenient as it is a very large and fat matrix that contains noise and multiple forms, since all of the evidences are automatically extracted from the Web. To this end, the matrix is firstly cleaned by applying different NLP techniques and, secondly, its consistency is enhanced and its sparseness further reduced by adding complementary common sense knowledge. In particular, three main issues need to be solved, namely: multiple word forms, multiple concept forms and low connectivity.

The first issue was addressed by processing both subjects and objects of triples with OMCS lemmatiser, which groups together the different inflected forms of words (different cases, plurals, verb tenses, etc.) so that they can be stored in Isanette as a single item. In case of duplicates, the triple with higher confidence score was considered. To perform multiple concept form reconciliation, both word similarity and multi-dimensionality reduction techniques are exploited. The concept 'barack obama', for example, appears in the triples in many different forms such as 'president obama', 'mr barack obama', 'president barack obama', etc. Trying to disambiguate this kind of instances *a priori*, by simply using word similarity, could be dangerous as concepts like 'buy christmas present' and 'present christmas event' have very different meanings, although they have high word similarity. Hence, an *a posteriori* concept deduplication is performed by exploiting concept semantic relatedness, after Isanette is built.

That is, concepts with high word similarity are merged together just if they are close enough to each other in the vector space generated from Isanette. As for the multiple concept senses, dimensionality reduction techniques are employed for both *a priori* and *a posteriori* disambiguation. In particular, for the former a vector space of concepts is built and dissimilarities between these are exploited to find instances likely to have multiple meanings. For the latter, the vectors corresponding to the instances found in an ambiguous sentence are averaged together to perform a context-dependent coarse sense disambiguation (more details in the next chapter).

As for Isanette's connectivity, in order to apply dimensionality reduction techniques on it for finding similar patterns, the matrix is better to be as less sparse as possible. To this end, we firstly hapax legomena, that is, instances/concepts with singular out-/in-degree, are to be discarded. These nodes can be useful for specific tasks such as finding the meaning of uncommon instances or give an example of a rare concept. For more general reasoning tasks, however, hapax legomena are very bad as they enlarge dimensionality without providing overlapping information that can be useful for finding similar patterns and perform analogies.

In this work, not only hapax legomena are discarded but also the other nodes with low connectivity, in order to heavily reduce Isanette's sparseness. In particular, a trial-and-error approach was used and the best trade-off between size and sparseness was achieved by setting the minimum node connectivity equal to 10. This cut-off operation leaves out almost 40% of nodes and makes Isanette a strongly connected core in which common and common sense knowledge coexist, i.e., a matrix 340,000 $\times$ 200,000 whose rows are instances such as ‘birthday party’ and ‘china’, whose columns are concepts like ‘special occasion’ and ‘country’, and whose values indicate truth values of assertions.

## 3.5 Conclusions

This chapter has shown how ConceptNet was blended with a linguistic resource for the lexical representation of affective knowledge, in order to obtain a new knowledge base in which concepts are interrelated by both common sense and affective features (section 3.1). In order to accordingly categorise affect in such knowledge base, moreover, the chapter presented a novel emotion categorisation model that goes beyond mere categorical and dimensional approaches by representing affective states both through emotional labels and through four independent but concomitant dimensions that can potentially describe the full range of emotional experiences (section 3.2).

Moreover, the chapter discussed how to enrich the developed affective common sense knowledge base through LSR, crowd sourcing, and GWAP techniques (section 3.3), and, finally, how to build possibly the most comprehensive resource of common and common sense knowledge for sentiment analysis, from ConceptNet and Probase (section 3.4). In order to effectively exploit the information contained in the developed knowledge bases for opinion mining tasks, next chapter will illustrate how to employ different dimensionality reduction and graph mining techniques for knowledge inference and retrieval.

# Chapter 4

# Sentic Knowledge Base Handling

*A perfect intelligence would not confine itself to one order of thought,*
*but would simultaneously regard a group of objects*
*as classified in all the ways of which they are capable.*

*Stanley Jevons*

Providing a machine with physical knowledge of how objects behave, social knowledge of how people interact, sensory knowledge of how things look and taste and psychological knowledge about the way people think, is not enough to make it intelligent. Having a database of millions of concepts is not very useful for a computer, unless it is able to conveniently use such knowledge base. Our ability to use common sense knowledge, in fact, highly depends on being able to do common sense reasoning.

Machines need to be taught not just common sense knowledge itself but also strategies for handling it, retrieving it when necessary, and learning from experience. To this end, adequately broad and deep common sense knowledge bases are to be developed, as well as reasoning methods that exhibit the features of human thinking, including the ability to reason with knowledge that is true by default, reason rapidly across a broad range of domains, tolerate uncertainty in the available knowledge, take decisions under incomplete knowledge, and perhaps revise that belief or decision when complete knowledge becomes available. It is also important to develop new kinds of cognitive architectures able to support multiple reasoning methods and representations. If a machine is able to represent knowledge and perform reasoning in many different ways, in fact, it can switch among different points of view and find one that works, rather than getting stuck when something goes wrong.

In this chapter, a novel affective common sense knowledge visualisation and analysis system is presented (section 4.1), together with a new clustering method for organising and categorising such knowledge (section 4.2), and a technique that combines dimensionality reduction and graph mining techniques on two different reasoning levels (section 4.3). In addition, this chapter shows how such methods are employed to design a publicly available semantic resource for opinion mining (section 4.4) and provides a summary and some concluding remarks (section 4.5).

## 4.1 AffectiveSpace: A Novel Representation of Affective Common Sense

The best way to solve a problem is to already know a solution for it. But, if we have to face a problem we have never met before, we need to use our intuition. Intuition can be explained as the process of making analogies between the current problem and the ones solved in the past to find a suitable solution. Marvin Minsky attributes this property to the so called 'difference-engines' [62]. This particular kind of agents operates by recognising differences between the current state and the desired state, and acting to reduce each difference by invoking K-lines that turn on suitable solution methods. This kind of thinking is maybe the essence of our supreme intelligence since in everyday life no two situations are ever the same and have to perform this action continuously. To emulate such process, AffectiveSpace, a novel affective common sense knowledge visualisation and analysis system proposed by Cambria et al. [172], is used. The human mind constructs intelligible meanings by continuously compressing over vital relations [173]. The compression principles aim to transform diffuse and distended conceptual structures to more focused versions so as to become more congenial for human understanding.

To this end, principal component analysis (PCA) has been applied on the matrix representation of AffectNet. In particular, Truncated Singular Value Decomposition (TSVD) has been preferred to other dimensionality reduction techniques for its simplicity, relatively low computational cost, and compactness. TSVD, in fact, is particularly suitable for measuring the cross-correlations between affective common sense concepts as it uses an orthogonal transformation to convert the set of possibly correlated common sense features associated with each concept into a set of values of uncorrelated variables (the principal components of the SVD). By using Lanczos' method [174], moreover, the generalisation process is relatively fast (a few seconds), despite the size and the sparseness of AffectNet. As the dimensions of such a matrix grow, however, PCA might cease to be a good solution. To this end, different techniques, e.g., independent component analysis (ICA), random projections, and non-negative matrix factorisation (NMF) are being investigated.

At the present time, TSVD is applied over the concept-feature matrix in order to conveniently reduce its dimensionality and capture the most important correlations. The objective of such compression is to allow many details in the blend of ConceptNet and WordNet-Affect (WNA) to be removed such that the blend only consists of a few essential features that represent the global picture. Applying TSVD on AffectNet, in fact, causes it to describe other features that could apply to known affective concepts by analogy: if a concept in the matrix has no value specified for a feature owned by many similar concepts, then by analogy the concept is likely to have that feature as well. In other words, concepts and features that point in similar directions, and therefore have high dot products, are good candidates for analogies. This section describes in detail how to build AffectiveSpace (subsection 4.1.1), how to set it up according to the desired trade-off between precision and efficiency (subsection 4.1.2), and how to generate different configurations of it, depending on the problem being tackled (subsection 4.1.3).

### 4.1.1 Building the Vector Space

A pioneering work on understanding and visualising the affective information associated to natural language text was conducted by Osgood et al. [91]. Osgood used multi-dimensional scaling (MDS) to create visualisations of affective words based on similarity ratings of the words provided to subjects from different cultures. Words can be thought of as points in a multi-dimensional space, and the similarity ratings represent the distances between these words. MDS projects these distances to points in a smaller dimensional space (usually two or three dimensions). Similarly, AffectiveSpace aims to grasp the semantic and affective similarity between different concepts by plotting them into a multi-dimensional vector space.

Differently from Osgood's space, however, the building blocks of AffectiveSpace are not simply a limited set of similarity ratings between affect words, but rather millions of confidence scores related to pieces of common sense knowledge linked to a hierarchy of affective domain labels. Rather than merely determined by a few human annotators and represented as a word-word matrix, in fact, AffectiveSpace is built upon an affective common sense knowledge base, namely AffectNet, represented as a concept-feature matrix. After performing TSVD on such matrix, hereby termed $A$ for the sake of conciseness, a low-rank approximation of it is obtained, that is, a new matrix $\tilde{A} = U_k \, \Sigma_k \, V_k^T$. This approximation is based on minimising the Frobenius norm of the difference between $A$ and $\tilde{A}$ under the constraint $rank(\tilde{A}) = k$. For the Eckart–Young theorem [175], it represents the best approximation of $A$ in the least-square sense, in fact:

$$\min_{\tilde{A}|rank(\tilde{A})=k} |A - \tilde{A}| = \min_{\tilde{A}|rank(\tilde{A})=k} |\Sigma - U^* \tilde{A} V| = \min_{\tilde{A}|rank(\tilde{A})=k} |\Sigma - S| \tag{4.1}$$

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:

$$\min_{\tilde{A}|rank(\tilde{A})=k} |\Sigma - S| = \min_{s_i} \sqrt{\sum_{i=1}^{n} (\sigma_i - s_i)^2} \tag{4.2}$$

$$\min_{s_i} \sqrt{\sum_{i=1}^{n} (\sigma_i - s_i)^2} = \min_{s_i} \sqrt{\sum_{i=1}^{k} (\sigma_i - s_i)^2 + \sum_{i=k+1}^{n} \sigma_i^2} = \sqrt{\sum_{i=k+1}^{n} \sigma_i^2} \tag{4.3}$$

Therefore, $\tilde{A}$ of rank $k$ is the best approximation of $A$ in the Frobenius norm sense when $\sigma_i = s_i$ $(i = 1, ..., k)$ and the corresponding singular vectors are the same as those of $A$. If all but the first $k$ principal components are discarded, 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 $e_0$,...,$e_{k-1}$ of the vector space (Fig. 4.1). 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 concepts such as 'beautiful day', 'birthday party' and 'make person happy' are found very close in direction in the vector space, while concepts like '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).

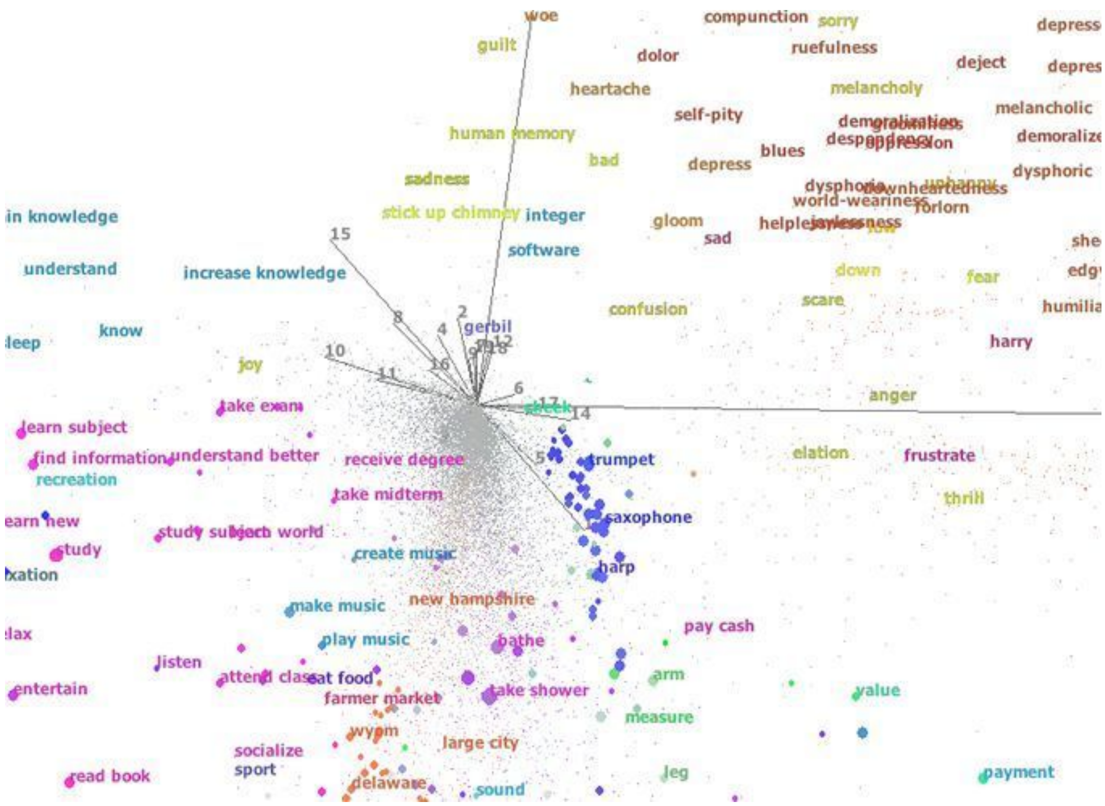


Figure 4.1: A sketch of AffectiveSpace. Affectively positive concepts (in the bottom-left corner) and affectively negative concepts (in the up-right corner) are floating in the multi-dimensional vector space.

### 4.1.2 Adjusting the Vector Space

The key to perform common sense reasoning is to find a good trade-off for representing knowledge. Since in life no two situations are ever the same, no representation should be too concrete, or it will not apply to new situations, but, at the same time, no representation should be too abstract, or it will suppress too many details. ConceptNet already supports different representations in fact it maintains different ways of conveying the same idea with redundant concepts, e.g., ‘car’ and ‘automobile’, that can be reconciled through background linguistic knowledge, if necessary. Within AffectiveSpace, this knowledge representation trade-off can be seen in the choice of the vector space dimensionality.

The number $k$ of singular values selected to build AffectiveSpace, in fact, is a measure of the trade-off between precision and efficiency in the representation of the affective common sense knowledge base. The bigger is $k$, the more precisely AffectiveSpace represents AffectNet's knowledge but generating the vector space is slower, and so is computing dot products between concepts. The smaller is $k$, on the other hand, the more efficiently AffectiveSpace represents affective common sense knowledge both in terms of vector space generation and of dot product computation. However, too few dimensions risk not to correctly represent AffectNet as concepts defined with too few features tend to be too close to each other in the vector space and, hence, not easily distinguishable and clusterable. In order to find a good $k$, AffectiveSpace was tested on a benchmark for affective common sense knowledge (BACK) built by applying CF-IOF (concept frequency - inverse opinion frequency) on the 5,000 posts of the LiveJournal corpus. CF-IOF is a technique, proposed by Cambria et al. [176], that 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:

$$CF\text{-}IOF_{c,d} = \frac{n_{c,d}}{\sum_k n_{k,d}} \log \sum_k \frac{n_k}{n_c} \tag{4.4}$$

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, CF-IOF weighting was exploited to filter out common concepts in the LiveJournal corpus and to detect relevant mood-dependent semantics for each of the Hourglass sentic levels.

| **Level** | **Label** | **Frequency** |
|---|---|---|
| -G(1) | grief | 14.3% |
| -G(2/3) | sadness | 19.8% |
| -G(1/3) | pensiveness | 11.4% |
| 0 | neutral | 10.5% |
| +G(1/3) | serenity | 20.6% |
| +G(2/3) | joy | 18.3% |
| +G(1) | ecstasy | 5.1% |

Table 4.1: Distribution of concepts through the Pleasantness dimension. The affective information associated with most concepts appear to concentrate around the centre of the Hourglass, rather than its extremes.

The result was a benchmark of 2000 affective concepts that were screened by 21 English-speaking students who were asked to evaluate the level $b$ associated to each concept $b \in \Theta = \{\theta \in \mathbb{Z} \mid -1 \leq \theta \leq 1\}$ (each integer corresponding to a level of the Hourglass model) for each of the four affective dimensions (i.e., Pleasantness, Attention, Sensitivity, and Aptitude). Results obtained were averaged (Table. 4.1).

BACK's concepts were compared with the classification results obtained by applying the AffectiveSpace process using different values of $k$, from 1 to 250. As shown in Fig. 4.2, the best trade-off is achieved at 100, as selecting more than 100 singular values does not improve accuracy significantly.

### 4.1.3 Transforming the Vector Space

The distribution of the values of each AffectiveSpace dimension is bell-shaped, with different centres and different degree of dispersion around them. Affective common sense concepts, in fact, tend to be close to the origin of the vector space (Fig. 4.3). In order to more uniformly distribute concept density in AffectiveSpace, an alternative strategy to represent the vector space was investigated. Such strategy consists in centring 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

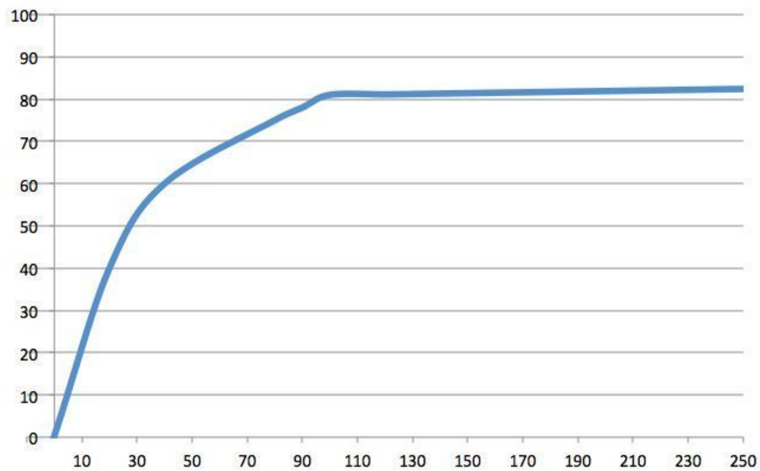


Figure 4.2: Accuracy values achieved by testing AffectiveSpace on BACK, with dimensionality spanning from 1 to 250. The best trade-off between precision and efficiency is obtained by using 100 singular values.

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. In particular, the transformation $x_{ij} \mapsto x_{ij} - \mu_i$ is first applied, being $\mu_i$ the average of all values of the $i$-th dimension. Then a normalisation is applied, combining the previous transformation with a new one $x_{ij} \mapsto \frac{x_{ij}}{a \cdot \sigma_i}$, where $\sigma_i$ 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 4.5 and 4.6:

$$x^*_{ij} = tanh\left(\frac{x_{ij} - \mu_i}{a \cdot \sigma_i}\right) \tag{4.5}$$

$$x^*_{ij} = \frac{x_{ij} - \mu_i}{a \cdot \sigma_i + |x_{ij} - \mu_i|} \tag{4.6}$$

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, normalising 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.

## 4.2 Sentic Medoids: Clustering Affective Common Sense Concepts

Sentic Medoids is a technique proposed by Cambria et al. [177] that adopts a $k$-medoids approach [178] to partition affective common sense concepts in AffectiveSpace into $k$ clusters around as many centroids, trying to minimise a given cost function. Differently from the $k$-means algorithm [179], which does not pose constraints on centroids, $k$-medoids do assume that centroids must coincide with $k$ observed points. This section introduces the standard approach to $k$-medoids clustering (subsection 4.2.1) and describes the algorithm developed for clustering AffectiveSpace (subsection 4.2.2).

### 4.2.1 Partitioning Around Medoids

Clustering is the process of grouping a set of objects into classes or clusters so that objects within a cluster have similarity in comparison to one another, but are dissimilar to objects in other clusters [180]. Well known techniques for performing non-hierarchical

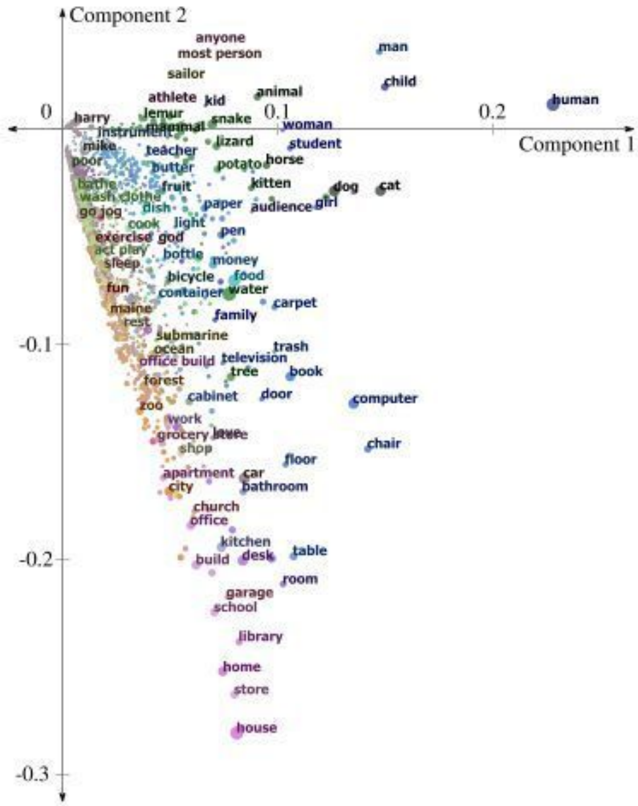


Figure 4.3: A two-dimensional projection (first and second eigenmoods) of AffectiveSpace. From this visualisation, it is evident that concept density is usually higher near the centre of the space.

clustering are $k$-means [181] and partitioning around medoids (PAM) [178]. The $k$-means approach finds the $k$ centroids, where the coordinate of each centroid is the mean of the coordinates of the objects in the cluster and assigns every object to the nearest centroid. Unfortunately, $k$-means clustering is sensitive to the outliers and a set of objects closest to a centroid may be empty, in which case centroids cannot be updated. For this reason, $k$-medoids are sometimes used, where representative objects are considered instead of centroids.

In many clustering problems, in fact, one is interested in the characterisation of the clusters by means of typical objects, which represent the various structural features of objects under investigation. Because it uses the most centrally located object in a cluster, $k$-medoids clustering is less sensitive to outliers compared with $k$-means. Among many algorithms for $k$-medoids clustering, PAM is one of the most widely used. The algorithm, proposed by Kaufman and Rousseeuw [178], first computes $k$ representative objects, termed medoids. A medoid can be defined as that object of a cluster, whose average dissimilarity to all the objects in the cluster is minimal. PAM 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. Compared to $k$-means, PAM operates on the dissimilarity matrix of the given dataset. It is more robust, because it minimises a sum of dissimilarities instead of a sum of squared Euclidean distances. A particularly nice property is that PAM allows clustering with respect to any specified distance metric. In addition, the medoids are robust representations of the cluster centres, which is particularly important in the common context that many elements do not belong well to any cluster.

### 4.2.2 Centroid Selection

Since $k$-medoids clustering is a NP-hard problem [182], different approaches based on alternative optimisation algorithms have been developed, though taking risk of being trapped around local minima. Among many algorithms for $k$-medoids clustering, PAM is known to be most powerful. However, PAM also has a drawback that it works inefficiently for large data sets due to its complexity. To this end, a modified version of the algorithm recently proposed by Park and Jun [183] was used, 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, AffectiveSpace contains $N$ concepts ($N = 14,301$) encoded as points $x \in \mathbb{R}^p (p = 50)$. They need to be grouped into $k$ clusters and, in this specific case, $k$ can be fixed to 24, as one cluster for each sentic level $s$ of the Hourglass model is being searched (Fig. 4.4). Generally, the initialisation of clusters for clustering algorithms is a problematic task as the process often risks to get stuck into local optimum points, depending on the initial choice of centroids [184]. However, in this case, the concepts that are currently used as centroids for clusters are selected as initial centroids, as they specify the emotional categories we want to organise AffectiveSpace into.

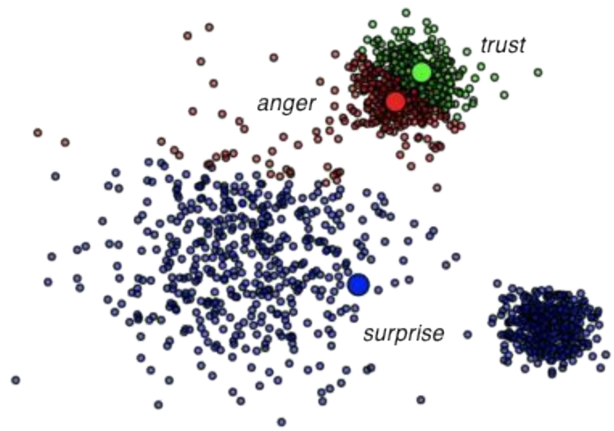


Figure 4.4: Sentic medoids conveniently organise AffectiveSpace by iteratively selecting appropriate centroids, in order to minimise the sum of dissimilarities between these and the other concepts within the same cluster.

For this reason, what is usually seen as a limitation of the algorithm can be seen as advantage for this approach, since what is being searched are not the 24 centroids leading to the best 24 clusters, but indeed for the 24 centroids identifying the required 24 sentic levels (i.e., the centroids should not be 'too far' from the ones currently used). In particular, as the Hourglass affective dimensions are independent but concomitant, AffectiveSpace needs to be clustered four times, once for each dimension. According to the Hourglass categorisation 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 $D(a, b) = \sqrt{\sum_{i=1}^{p} (a_i - b_i)^2}$ (note that the choice of Euclidean distance is arbitrary), the used algorithm, applied for each of the four affective dimensions, can be summarised as follows:

1. Each centroid $C_n \in \mathbb{R}^{50}\,(n = 1, 2, ..., k)$ 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 $\Xi$ so that $x_i \in \Xi_n$ if $D(x_i, C_n) \leq D(x_i, C_m)$ $m = 1, 2, ..., k$

3. Find a new centroid C for each cluster $\Xi$ 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

Note that condition posed on steps 2 and 3 may occasionally lead to more than one solution. Should this happen, the proposed model will randomly choose one of them. This clusterisation of AffectiveSpace allows to calculate, for each common sense concept $x$, a four-dimensional sentic vector that defines its affective valence in terms of a degree of fitness $\mathbf{f}(x)$ where $f_a = D(x, C_j) \quad C_j | D(x, C_j) \leq D(x, C_k) \; a = 1, 2, 3, 4 \quad k = 6a$-$5, 6a$-$4, ..., 6a$.

## 4.3 Sentic Activation: A Two-Level Affective Reasoning Framework

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 two (or more) processing systems within the human brain, one that involves fast, parallel, unconscious processing, and one that involves slow, serial, more conscious processing [185, 186, 187, 188]. 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 [189]. Such systems, termed 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 (subsection 4.3.1). The latter, in turn, works logically, systematically, effortfully, and rationally (subsection 4.3.2).

### 4.3.1 Unconscious Reasoning

In recent years, neuroscience has contributed a lot to the study of emotions through the development of novel methods for studying emotional processes and their neural correlates. In particular, new methods used in affective neuroscience, e.g., fMRI, lesion studies, genetics, electro-physiology, paved the way towards the understanding of the neural circuitry that underlies emotional experience and of the manner in which emotional states influence health and life outcomes. A key contribution in the last two decades has been to provide evidence against the notion that emotions are subcortical and limbic, whereas cognition is cortical.

This notion was reinforcing the flawed Cartesian dichotomy between thoughts and feelings [190]. There is now ample evidence that the neural substrates of cognition and emotion overlap substantially [191]. Cognitive processes, such as memory encoding and retrieval, causal reasoning, deliberation, goal appraisal, and planning, operate continually throughout the experience of emotion. This evidence points to the importance of considering the affective components of any human-computer interaction [19]. Affective neuroscience, in particular, has provided evidence that elements of emotional learning can occur without awareness [192] and elements of emotional behaviour do not require explicit processing [193]. Affective information processing, in fact, mainly takes place at unconscious level (U-level) [188].

Reasoning, at this level, relies on experience and intuition, which allow considering issues intuitively and effortlessly. Hence, rather than reflecting upon various considerations in sequence, the U-level forms a global impression of the different issues. In addition, rather than applying logical rules or symbolic codes (e.g., words or numbers), the U-level considers vivid representations of objects or events. Such representations are laden with the emotions, details, features, and sensations that correspond to the objects or events. Such human capability of summarising the huge amount of inputs and outputs of previous situations to find useful patterns that might work at the present time is hereby implemented by means of AffectiveSpace. By reducing the dimensionality of the matrix representation of AffectNet, in fact, AffectiveSpace compresses the feature space of affective common sense knowledge into one that allows to better gain global insight and human-scale understanding.

In cognitive science, the term 'compression' 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 achieved by balancing the number of singular values discarded when synthesising AffectiveSpace, in a way that the affective common sense knowledge representation is neither too concrete nor too abstract with respect to the detail granularity needed for performing a particular task. The reasoning-by-analogy capabilities of AffectiveSpace, hence, are exploited at U-level to achieve digital intuition about the input data. In particular, the vector space representation of affective common sense knowledge is clustered according the Hourglass model using the sentic medoids technique, in a way that 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.

### 4.3.2 Conscious Reasoning

U-level and C-level are two conceptual systems that operate by different rules of inference. While the former operates emotionally and intuitively, the latter relies on logic and rationality. In particular, the C-level analyses issues with effort, logic, and deliberation rather than relying on intuition. Hence, while at U-level the vector space representation of AffecNet is exploited to intuitively guess semantic and affective relations between concepts, at C-level associations between concepts are made according to the actual connections between different nodes in the graph representation of affective common sense knowledge. Memory is not a 'thing' that is stored somewhere in a mental warehouse and can be pulled out and brought to the fore. Rather, it is a potential for reactivation of a set of concepts that together constitute a particular meaning. Associative memory involves the unconscious activation of networks of association–thoughts, feelings, wishes, fears, and perceptions that are connected, so that activation of one node in the network leads to activation of the others [194].

Sentic activation aims to implement such process through the ensemble application of multi-dimensionality reduction and graph mining techniques. Specifically, the semantically and affectively related concepts retrieved by means of AffectiveSpace at U-level are fed into AffectNet in order to crawl it according to how such seed concepts are interconnected to each other and to other concepts in the semantic network. To this end, spectral association [195] is employed. Spectral association is a technique that assigns values, or activations, to seed concepts and spreads their values across the AffectNet 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. These related concepts are likely to have similar affective values. This can be seen as an alternate way of assigning affective values to all concepts, which simplifies the process by not relying on an outside resource such as WNA. In particular, a matrix $A$ that relates concepts to other concepts, instead of their features, is built and the scores are added up over all relations that relate one concept to another, disregarding direction. Applying $A$ to a vector containing a single concept spreads that concept's value to its connected concepts. Applying $A^2$ spreads that value to concepts connected by two links (including back to the concept itself). But the desired operation is to spread the activation through any number of links, with diminishing returns, so the operator wanted is:

$$1 + A + \frac{A^2}{2!} + \frac{A^3}{3!} + ... = e^A \qquad (4.7)$$

This odd operator, $e^A$, can be calculated because $A$ can be factored. $A$ is already symmetric, so instead of applying Lanczos' method [174] to $AA^T$ and getting the SVD, it can be applied directly to $A$ to obtain the spectral decomposition $A = V\Lambda V^T$. As before, this expression can be raised to any power and everything but the power of $\Lambda$ cancelled. Therefore, $e^A = Ve^{\Lambda}V^T$. This simple twist on the SVD allows to calculate spreading activation over the whole matrix instantly. As with the SVD, these matrices can be truncated to $k$ axes and therefore space can be saved while generalising from similar concepts. The matrix can also be rescaled so that activation values have a maximum of 1 and do not tend to collect in highly-connected concepts such as 'person', by normalising the truncated rows of $Ve^{\Lambda/2}$ to unit vectors, and multiplying that matrix by its transpose to get a rescaled version of $Ve^{\Lambda}V^T$.

Spectral association can spread not only positive, but also negative activation values. Hence, unconscious reasoning at U-level is exploited not only to retrieve concepts that are most semantically and affectively related, but also concepts that are most likely to be unrelated with the input data (lowest dot product). While the former are exploited to spread semantics and sentics across the AffectNet graph, the latter are used to contain such activation in a way that potentially unrelated concepts (and their twins) do not get triggered. Such brain-inspired ensemble application of dimensionality reduction and graph mining techniques (hereby referred as unconscious and conscious reasoning, respectively) allows sentic activation to more efficiently infer semantics and sentics from natural language text. In fact, sentic activation was tested on the benchmark for affective common sense knowledge (BACK) built by means of CF-IOF. In particular, BACK's concepts were compared with the classification results obtained by applying the AffectiveSpace process (U-level), spectral association (C-level) and sentic activation (ensemble of U-level and C-level). Results showed that sentic activation achieves +13.9% and +8.2% accuracy than the AffectiveSpace process and spectral association, respectively.

## 4.4 SenticNet: A Publicly Available Semantic Resource for Opinion Mining

SenticNet is a publicly available resource for opinion mining, proposed by Cambria et al. [196], which exploits both AI and Semantic Web techniques to infer the polarity associated with common sense concepts and represent it in a semantic-aware format. In particular, SenticNet uses dimensionality reduction to calculate the affective valence of a set of Open Mind concepts and represent it in a machine-accessible and machine-processable format. The result is a publicly available resource for mining opinions from natural language text at a semantic, rather than just syntactic, level. This section shows motivations and techniques for building such resource (subsection 4.4.1) and explains how it is encoded in a Semantic Web aware format (subsection 4.4.2) and how it can be exploited for NLP tasks (subsection 4.4.3).

### 4.4.1 Building SenticNet

The development of SenticNet was inspired by SentiWordNet [197], a lexical resource in which each WordNet synset is associated to three numerical scores describing how objective, positive and negative the terms contained in the synset are. Each of the three scores ranges from 0.0 to 1.0, and their sum is 1.0 for each synset. This means that a synset may have non-zero scores for all the three categories, which would indicate that the corresponding terms have, in the sense indicated by the synset, each of the three opinion-related properties only to a certain degree.

The method used to develop SentiWordNet is based on the quantitative analysis of the 'glosses' associated to synsets, and on the use of the resulting vector representations for semi-supervised synset classification. The three scores are derived by combining the results produced by a committee of eight ternary classifiers, all characterised by similar accuracy levels but different classification behaviour. SentiWordNet currently represents a good resource for opinion mining, however, it contains a lot of noise and it mainly provides opinion polarity at syntactical level, leaving out polarity information for common sense knowledge concepts such as ‘accomplish goal’, ‘bad feeling’, ‘celebrate special occasion’, ‘lose temper’ or ‘be on cloud nine’, which are usually found in natural language text to express positive and negative viewpoints.

To this end, SenticNet was developed. SenticNet is a publicly available resource for opinion mining that aims to create a collection of commonly used ‘polarity concepts’, that is common sense concepts with relatively strong positive or negative polarity. Differently from SentiWordNet (which also includes null polarity terms), in fact, SenticNet does not contain concepts with neutral or almost neutral polarity, i.e., concepts with polarity magnitude close to zero. Moreover, while SentiWordNet stores three values for each synset, in SenticNet each concept $c$ is associated to just one value $p_c$, i.e., a float $\in$ [-1,1] representing its polarity, in order to avoid redundancy and more easily represent SenticNet as a semantic network. Therefore, in SenticNet, concepts like ‘make good impression’, ‘look attractive’, ‘show appreciation’ or ‘good deal’ are likely to have $p_c$ very close to 1 while concepts such as ‘being fired’, ‘leave behind’ or ‘lose control’ are likely to have $p_c \approx$ -1. Polarity values are assigned to Open Mind concepts by means of sentic activation. In particular, the sentic levels of the Hourglass model are used as input concepts for the U-level, which uses its digital intuition to infer semantically and affectively related concepts. Such seed concepts then crop out to the C-level that exploits spectral association to accordingly spread their activation through the graph structure of AffectNet.

### 4.4.2 Encoding SenticNet

After retrieving polarity concepts through sentic activation, they need to be reorganised in a way that they can be represented in a unique and consistent resource. Possible conflicts are handled by discarding duplicate concepts with smaller polarity magnitude since bigger concept polarity values usually correspond to more reliability (higher dot products in the vector space). Since concepts are usually strongly related to just one or two affective dimensions (most of compound emotions are in fact given by summing just two elementary emotions), the average magnitude is pretty low.

Therefore, in order to obtain more homogeneous and intelligible polarity values, a normalisation process over SenticNet is conducted before storing its contents in a Semantic Web aware format. In order to represent SenticNet in a machine-accessible and machine-processable way, results are encoded in RDF triples using a XML syntax. In particular, concepts are identified using the ConceptNet Web API and statements, which have the form concept-hasPolarity-polarityValue, are encoded in RDF/XML format on the base of the human emotion ontology (HEO) [198], a high level ontology for human emotions that supplies the most significant concepts and properties which constitute the centrepiece for the description of every human emotion. The main purpose of HEO is to create a description framework that could grant at the same time enough flexibility, by allowing the use of a wide and extensible set of descriptors to represent all the main features of an emotion, and interoperability, by allowing to map concepts and properties belonging to different emotion representation models. HEO was developed in OWL description logic in order to allow a taxonomical organisation of emotion categories and properties restriction to link emotion description made both by category and dimension.

In HEO, for example, Ekman's archetypal emotion of 'joy' represents a superclass for Plutchik's emotions of 'ecstasy', 'joy' and 'serenity'. Using property restriction, Plutchik's emotion of 'joy' can also be defined as an emotion that 'has Pleasantness some float $\in$ [+G(1/3),+G(2/3)]', 'interest' as an emotion that 'has Attention $\in$ [0,+G(1/3)]' and 'love' as an emotion that 'has Pleasantness some float $\in$ [0,+G(1)] and Aptitude some float $\in$ [0,+G(1)]'. In this way, querying a database that support OWL description logic inference for basic emotions of type 'joy' will return not only the emotions expressly encoded as Ekman's archetypal emotions of type 'joy', but also the emotions encoded as Plutchik's basic emotion of type 'joy' and emotions that 'have Pleasantness some float $\in [+G(1/3), +G(2/3)]$'.

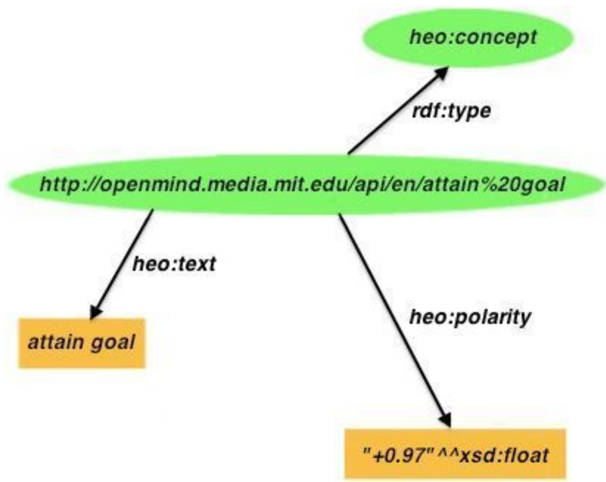


Figure 4.5: SenticNet is a Semantic Web aware AI resource in which Open Mind concepts are identified using the ConceptNet Web API, and statements are encoded in RDF/XML format on the base of HEO.

### 4.4.3 Working with SenticNet

SenticNet, freely available at `http://cs.stir.ac.uk/~eca/sentics`, currently contains more than 5,700 polarity concepts (nearly 40% of Open Mind corpus). It is very easy to interface SenticNet with any kind of opinion mining application and, especially if used within Open Mind software (for a full correspondence of concepts), it is a precise polarity detection tool. In particular, after deconstructing text into concepts (more details in chapter 5), SenticNet can be used to associate polarity values to these and, hence, infer the overall polarity of a clause, sentence, paragraph or document by averaging such values. SenticNet's capacity of detecting opinion polarity was compared with SentiWordNet's over a collection of 2,000 patient opinions, of which 57% are labelled as negative, 32% as positive and the rest as neutral. After extracting concepts from each opinion, relative polarity values were searched in SentiWordNet and SenticNet and compared with the dataset labels, in order to compute recall and precision rates as evaluation metrics. Results showed SenticNet to be much more accurate than SentiWordNet. The former, in particular, can identify positive opinions with much higher precision (79.1% against 53.8%) and significantly better recall rate (58.4%

against 46.5%), for a total F-measure value of 67.1% versus 49.8%. In SenticNet 2.0, currently under-development, the whole Open Mind corpus is being labelled with polarity values and a list of mood and sentic values is being associated with each common sense concept, in order to provide the public with a comprehensive semantic resource for easily extracting affective information from natural language text.

## 4.5 Conclusions

This chapter has shown how the application of dimensionality reduction techniques on the matrix representation of AffectNet yields a vector space of affective common sense knowledge, which can be accordingly configured, depending on the desired trade-off between precision and efficiency, and on the problem being tackled (section 4.1). So far, TSVD appears to be a good method for generalising the information contained in AffectNet but it is very expensive in both computing time and storage, as it requires costly arithmetic operations such as division and square root in the computation of rotation parameters. This is a big issue because AffectNet is keeping on growing, in parallel with the continuously extended versions of ConceptNet. To this end, alternative multi-dimensionality reduction techniques, e.g., independent component analysis (ICA) and random projections, are currently being explored.

The rest of the chapter illustrated a new PAM-based clustering method for organising and categorising such vector space (section 4.2) and how the ensemble application of dimensionality reduction and graph mining techniques can be exploited to emulate conscious and unconscious reasoning processes (section 4.3). In addition, this chapter showed how the developed methods are employed to design a publicly available semantic resource for opinion mining (section 4.4). In order to assess how the developed reasoning techniques can be effectively exploited for tackling real-world problems, next chapter will explore multiple ways to combine such techniques for the design of intelligent applications in fields such as Social Web, HCI, and e-health.

# Chapter 5

# Sentic Knowledge Base Exploitation

*Knowing is not enough; we must apply.*

*Willing is not enough; we must do.*

*Johann Wolfgang von Goethe*

The amount of data available on the Web is growing exponentially. These data, however, are mainly in an unstructured format and, hence, not machine-processable and machine-interpretable. What is called collective intelligence today is actually just collected intelligence as the value of user contributions is simply in their being collected together and aggregated into community or domain specific sites. True collective intelligence can emerge if the data collected from all those people is aggregated and recombined to create new knowledge and new ways of learning that individual humans cannot do by themselves [199]. So far, online information retrieval has mainly relied on keyword-based algorithms, which have proved to have important limitations, e.g., the inability to recognise topical authority that humans recognise effortlessly without the explicit words being in the content. In order to let machines better understand natural language and, hence, conveniently analyse and aggregate opinions and sentiments over the Web, we need to provide them with both adequately broad common sense knowledge bases and reasoning methods to efficiently handle these. This chapter describes how the knowledge bases, and the reasoning tools built on the top of them, are exploited for the design of an intelligent opinion mining engine (section 5.1) and, hence, for the development of applications in fields such as Social Web (section 5.2), HCI (section 5.3), and e-health (section 5.4). The chapter, eventually, ends with some concluding remarks (section 5.5).

## 5.1 Opinion Mining Engine: A Semantics and Sentics Extraction Tool

In order to effectively mine and analyse opinions and sentiments, it is necessary to bridge the gap between unstructured natural language data and structured machine-processable data. To this end, an intelligent software engine has been proposed by Cambria et al. [145] that aims to extract the semantics and sentics, that is the cognitive and affective information, associated with natural language text, in a way that the opinions and sentiments in it contained can be more easily aggregated and interpreted. The engine exploits graph mining and multi-dimensionality reduction techniques on Isanette and AffectNet respectively, and it is based on the Hourglass model (Fig. 5.1).

Several other affect recognition and sentiment analysis systems [200, 201, 202, 203, 204, 205, 206] are based on different emotion categorisation models, which generally comprise a relatively small set of categories (Table 5.1). The Hourglass of Emotions, in turn, allows the opinion mining engine to classify affective information both in a categorical way (according to a wider number of emotion categories) and in a dimensional format (which facilitates comparison and aggregation). Such engine, in particular, consists of four main components: a pre-processing module, which performs a first skim of the opinion (subsection 5.1.1), a semantic parser, whose aim is to extract concepts from the opinionated text (subsection 5.1.2), the Isanette module, for inferring the semantics associated with the given concepts (subsection 5.1.3), and the AffectiveSpace module, for the extraction of sentics (subsection 5.1.4). Eventually, this section illustrates an output example of the engine, given a short natural language sentence as input (subsection 5.1.5), and provides a thorough evaluation of the system (subsection 5.1.6).

### 5.1.1 Pre-Processing Module

The pre-processing module firstly exploits linguistic dictionaries to interpret all the affective valence indicators usually contained in opinionated text, e.g., special punctuation, complete upper-case words, cross-linguistic onomatopoeias, exclamation words, degree adverbs, and emoticons. Secondly, the module detects negation and spreads it in a way that it can be accordingly associated to concepts during the parsing phase. Handling negation is an important concern in opinion- and sentiment-related analysis, as it can reverse the meaning of a statement.

Such task, however, is not trivial as not all appearances of explicit negation terms reverse the polarity of the enclosing sentence and that negation can often be expressed in rather subtle ways, e.g., sarcasm and irony, which are quite difficult to detect. Lastly, the module converts text to lower-case and, after lemmatising it, splits the opinion into single clauses according to grammatical conjunctions and punctuation.

### 5.1.2 Semantic Parser

The semantic parser deconstructs text into concepts using a lexicon based on sequences of lexemes that represent multiple-word concepts extracted from AffectNet and Isanette. These n-grams are not used blindly as fixed word patterns but exploited as reference for the module, in order to extract multiple-word concepts from information-rich sentences. So, differently from other shallow parsers, the module can recognise complex concepts also when irregular verbs are used or when these are interspersed with adjective and adverbs, e.g., the concept ‘buy christmas present’ in the sentence “I bought a

| Study | Techniques | Model | Corpora | Knowledge Base |
|---|---|---|---|---|
| [202] | NB, SVM | 2 categories | Political articles | None |
| [203] | LSA, MLP, NB, KNN | 3 categories | Dialogue turns | ITS interaction |
| [206] | Cohesion indices | 4 categories | Dialogue logs | ITS interaction |
| [204] | VSM, NB, SVM | 5 categories | ISEAR | ConceptNet |
| [205] | WN presence, LSA | 6 categories | News stories | WNA |
| [200] | WN presence | 6 categories | Chat logs | WNA |
| [201] | Winnow linear, C4.5 | 7 categories | Children stories | None |
| [172] | VSM, KNN | 24 categories | LiveJournal | ConceptNet, WNA |
| [145] | VSM, $k$-means | 24 categories | YouTube, LiveJournal | ConceptNet, WNA, HEO |
| [207] | VSM, $k$-means | 24 categories | LiveJournal, PatientOpinion | ConceptNet, WNA |
| [171] | VSM, $k$-medoids | 24 categories | Twitter, LiveJournal, PatientOpinion | ConceptNet, Probase |

Table 5.1: An overview of recent model-based affect recognition and sentiment analysis systems. Studies are divided by techniques applied, number of categories of the model adopted, corpora and knowledge base used.

lot of very nice Christmas presents". The semantic parser, additionally, provides, for each retrieved concept, the relative frequency, valence and status, that is the concept's occurrence in the text, its positive or negative connotation and the degree of intensity with which the concept is expressed.

For each clause, the module outputs a small bag of concepts (SBoC), which is later on analysed separately by the Isanette and AffectiveSpace modules to infer the cognitive and affective information associated with the input text, respectively. In case any of the detected concepts is found more than once in the vector space (that is, any of the concepts has multiple senses), all the SBoC concepts are exploited for a context-dependent coarse sense disambiguation. In particular, to represent the expected semantic value of the clause as a whole, the vectors corresponding to all concepts in the clause (in their ambiguous form) can be averaged together. The resulting vector does not represent a single meaning but the 'ad hoc category' of meanings that are similar to the various possible meanings of concepts in the clause [208]. Then, to assign the correct sense to the ambiguous concept, the concept sense with the highest dot product (and thus the strongest similarity) with the clause vector is searched.

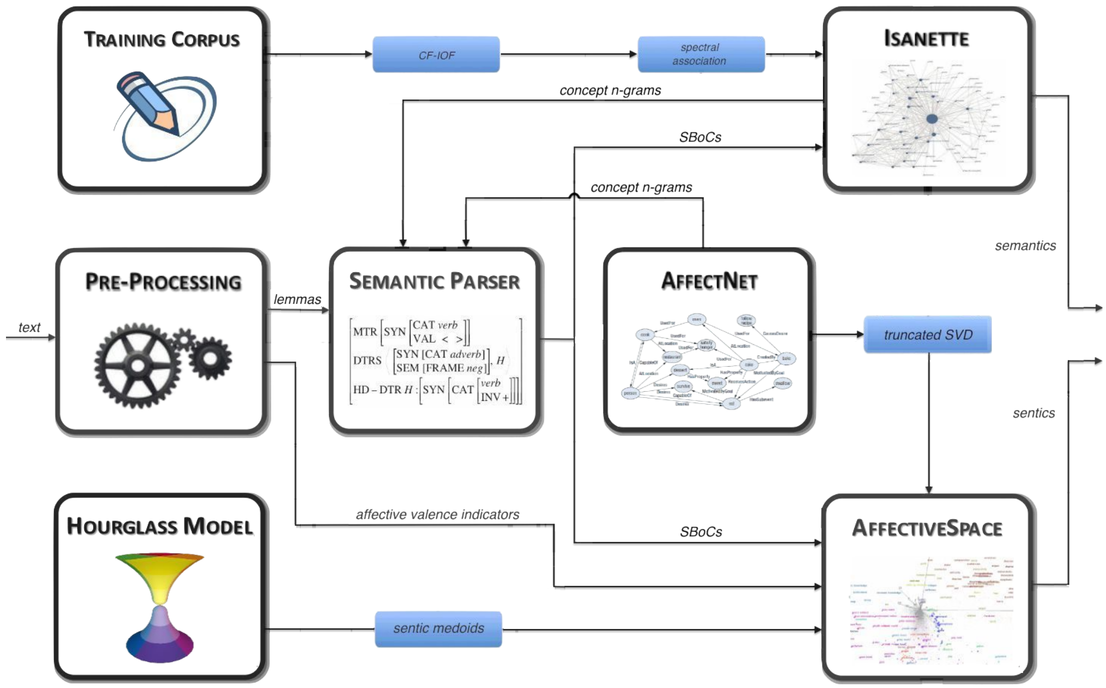


Figure 5.1: Opinion mining engine block diagram. After performing a first skim of the input text, the engine extracts concepts from it and, hence, infers related semantics and sentics.

### 5.1.3 Isanette Module

Once natural language text is deconstructed into concepts, these are given as input to both the Isanette and the AffectiveSpace modules. While the former exploits the graph representation of the affective common sense knowledge base to detect semantics, the latter exploits the vector space representation of Isanette to infer sentics. In particular, the Isanette module applies spectral association for assigning activation to key concepts, that is nodes of the semantic network, which are used as seeds or centroids for classification. Such seeds can simply be the concepts corresponding to the class labels of interest plus their available synonyms and antonyms, if any.

As shown in section **??**, seeds can also be found by applying CF-IOF on a training corpus (when available), in order to perform a classification that is more relevant to the data under analysis. After seeds concepts are identified, the module spreads their values across the Isanette 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. Therefore, the concepts of each SBoC provided by the semantic parser are projected on the matrix resulting from spectral association in order to calculate their semantic relatedness to each seed concept and, hence, their degree of belonging to each different class. Such classification measure is directly proportional to the degree of connectivity between the nodes representing the retrieved concepts and the seed concepts in the affective common sense knowledge graph.

### 5.1.4 AffectiveSpace Module

In the Isanette module, graph-mining techniques are exploited to extract semantics from the concepts retrieved by the semantic parser. Such concepts are also given as input to the AffectiveSpace module, which, in turn, exploits dimensionality reduction techniques to infer the affective information associated with them. To this end, the concepts of each SBoC are projected into AffectiveSpace and, according to their position in the vector space representation of affective common sense knowledge, they are assigned to an affective class defined through the sentic medoids technique.

As well as in the Isanette module, the categorisation does not consist in simply labelling each concept but also in assigning a confidence score to each emotional label, which is directly proportional to the degree of belonging to a specific affective cluster (dot product between the given concept and the relative sentic medoid). As shown in section 3.2.2, such affective information can also be exploited to calculate a polarity value associated with each SBoC provided by the semantic parser, as well as to detect the overall polarity associated with the opinionated text.

### 5.1.5 Output Example

On average, for each 100 words of input text, the pre-processing module extracts 3 affective valence indicators and the semantic parser detects 11 SBoCs. In order for the engine to perform fast real-time opinion mining, the sentic vector associated with each AffectNet concept is calculated a priori and saved in an SQL database. At run-time, then, the sentic vectors relative to each of the concepts composing the SBoC are retrieved from such a database and aggregated to compute the overall affective valence and opinion polarity associated with the specific SBoC. This allows the sentics extraction process to be faster than directly applying the AffectiveSpace process.

Similarly, spectral association is computed a priori on Isanette and the semantic classification of each concept (that is a set of different topic labels and the confidence associated with these) is stored in an SQL database, to speed up the run-time processing of natural language opinions. On average, in fact, while the processing of a 100-word opinion is on the order of tens of seconds when directly applying the AffectiveSpace and Isanette processes, the extraction of semantics and sentics is on the order of seconds when using the corresponding SQL databases. Both resources are on the order of hundreds of megabytes and, hence, easily exportable and embeddable into bigger systems for the development of applications in fields such as Social Web, HCI, and e-health, as shown in the next sections.

As an example of how the software engine works, intermediate and final outputs obtained when a natural language opinion is given as input to the system are hereby examined. In particular, the following tweet was selected: "I think iPhone4 is the top of the heap! OK, the speaker is not the best i hv ever seen bt touchscreen really puts me on cloud 9... camera looks pretty good too!". After the pre-processing and semantic parsing operations, the following SBoCs are obtained:

*SBoC#1:*

*<Concept: 'think'>*

*<Concept: 'iphone4'>*

<*Concept: ‘top heap’*>

*SBoC#2:*

<*Concept: ‘ok’*>

<*Concept: ‘speaker’*>

<*Concept:* !*‘good’*++>

<*Concept: ‘see’*>

*SBoC#3:*

<*Concept: ‘touchscreen’*>

<*Concept: ‘put cloud nine’*++>

*SBoC#4:*

<*Concept: ‘camera’*>

<*Concept: ‘look good’*−−>

These are then concurrently processed by the Isanette and the AffectiveSpace modules, which output the cognitive and affective information associated with each SBoC, both in a discrete way, with one or more labels, and in a dimensional way, with a polarity value $\in$ [-1,+1] (Table 5.2).

### 5.1.6 Evaluation

The developed engine overcomes the main problems faced by state-of-the art keyword-based approaches. Thanks to the exploitation of common sense knowledge bases that allow a concept-level analysis of text, in fact, the opinion mining engine is able to more effectively infer the cognitive and affective information associated with natural language. The validity of the proposed approach, hence, highly depends on two main

| **Opinion Target** | **Category** | **Moods** | **Polarity** |
|---|---|---|---|
| ‘iphone4’ | ‘phones’, ‘electronics’ | ‘ecstasy’, ‘interest’ | +0.71 |
| ‘speaker’ | ‘electronics’, ‘music’ | ‘annoyance’ | -0.34 |
| ‘touchscreen’ | ‘electronics’ | ‘ecstasy’, ‘anticipation’ | +0.82 |
| ‘camera’ | ‘photography’, ‘electronics’ | ‘acceptance’ | +0.56 |

Table 5.2: Structured output example of opinion mining engine. For each clause, the engine detects the opinion target, the category it belongs to, and the affective information associated with it.

factors: the richness of the knowledge bases and the accuracy of the semantic parser. If a concept for which there is no match in the knowledge bases is encountered in the opinionated text, the engine will not be able to infer the semantics and sentics associated with such concept and, hence, it might not be able to correctly perform the overall feature-based sentiment analysis of the input text. For the same reason, the accuracy of the engine process also depends on the concept detection precision of the semantic parser. As explained in subsection 5.1.2, the semantic parser deconstructs text into concepts using a lexicon based on sequences of lexemes that represent multiple-word concepts extracted from AffectNet and Isanette. These n-grams are not used blindly as fixed word patterns but exploited as reference for the module, in order to extract multiple-word concepts from information-rich sentences.

So, differently from other shallow parsers, the module can recognise complex concepts also when irregular verbs are used or when these are interspersed with adjective and adverbs, e.g., the concept 'buy christmas present' in the sentence "I bought a lot of very nice Christmas presents". However, the parser would not be able to extract the same concept from a sentence like "I bought some nice Christmas bells for my mom". Since the current version of the parser is based on n-grams, in fact, it is not able to guess that a bell, in this case, is a Christmas present and, hence, does not detect the concept 'buy christmas present'. Since, moreover, the concept 'buy christmas bell' is not contained in AffectNet, the engine will not be able to infer the semantics and sentics associated with that concept, despite the affective common sense knowledge base actually contains information about it. In this case, however, the semantic parser would be able to extract concepts such as 'buy', 'christmas', and 'bell' and, hence, make a good-enough guess about the semantics and sentics associated to the input text.

On average, in fact, SBoCs mostly consist of single-word or 2-word concepts, hence, the validity of the proposed approach mainly depends on the richness of AffectNet and Isanette. To this end, the evaluation process was performed at knowledge base level. An extensive corroborative evaluation of the opinion mining engine, in fact, was not possible because currently available test datasets in the field of sentiment analysis are still very few and there are no universally recognised benchmarks that can serve as a basis for comparing different techniques or different configurations of the same method.

A comparative evaluation of the engine with state-of-the-art techniques, moreover, was not feasible as most of current approaches are keyword-based, rather than concept-based, and aim for a document- or paragraph-level, rather than sentence- or clause-level, sentiment analysis. Hence, in order to evaluate the different facets of the opinion mining engine from various perspectives, three different resources were used, namely a Twitter hashtag repository, a LiveJournal database, and a PatientOpinion dataset, and results were compared with those obtained by using WordNet, ConceptNet, and Probase, in place of the two developed knowledge bases.

Each of the evaluation resources has been selected to test a specific capability of the engine, according to the different properties of each test dataset. Specifically, the Twitter hashtag repository has been chosen for testing the precision of the engine in inferring semantics, the LiveJournal database has been selected to assess the engine's accuracy in extracting sentics, while the PatientOpinion dataset has been used for evaluating the inference of both semantics and sentics concurrently. The Twitter hashtag repository, in particular, is a subset of the dataset used in Song et al. [209] for short text conceptualisation and consists of 3,000 tweets crawled from Bing web repository by exploiting Twitter hashtags as category labels.

Such hashtags pertain to electronics (e.g., IPhone, XBox, Android, and Wii), companies (e.g., Apple, Microsoft, and Google), countries, cities, operative systems, and cars. In order to test the resource's consistency and reliability, a manual evaluation of 100 tweets was performed, which showed that hashtags are accurate to 91%. The LiveJournal database, in turn, is a variation of the dataset used by Strapparava and Mihalcea [205] for emotion annotation and consists in a collection of 5,000 blogpost database extracted from LiveJournal, a virtual community of more than 23 millions 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 or by creating new ones. Since the indication of mood tags is optional, posts are likely to reflect the true mood of the authors, which is not always true for category tags. After a manual evaluation of 200 posts, in fact, the category tags turned out to be very noisy (53% accuracy). The mood tags, however, showed a good enough reliability (89% accuracy) so they were used to test the engine's affect recognition performance. In order to have full correspondence between LiveJournal mood labels and the activation levels of the Hourglass model, moreover, a pool of 21 native English-speaking students was asked to map each of the 130 mood labels into the 24 emotional labels of the Hourglass model. In order to avoid any bias, students were randomly selected among different faculties in the university of Stirling and they were asked to perform the mapping by playing an ad hoc version of the Hourglass game in a secluded environment. Eventually, a reduced set of 80 moods (those with higher confidence level) was selected for inclusion in the blogpost database.

The PatientOpinion dataset, finally, is a manually tagged evaluation resource kindly provided by James Munro, CEO of PatientOpinion, a social enterprise pioneering an online feedback service for users of the UK National Health Service to enable people to share their recent experience of local health services online. It is a collection 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. From a first manual evaluation of 50 opinions, the dataset appeared to be 100% accurate. It was used to test the detection of opinion targets and the polarity associated with these. Each of the used datasets has different strengths and weaknesses. The Twitter hashtag repository, for example, is written in very formal English and offers a great span of common knowledge and instances of concepts but it is rather domain dependent and almost totally lacks the presence of common sense concepts.

The LiveJournal database, in turn, is more open-domain and affective common sense oriented but it often contains slang terms and grammatical mistakes, which lower the accuracy of clause chunking and concept parsing. The PatientOpinion database is very well-formed and the most accurate of the three datasets in terms of both affect and category labels but it is very domain specific and rather limited in number of entries.

As the key components of the opinion mining engine are the common sense knowledge bases, the evaluation process was performed at knowledge base level. What led to the development of sentic computing, in fact, is the need for better accuracy of sentiment analysis systems when switching between different domains. Currently available keyword-based approaches may perform nicely on a specific dataset but they have very low accuracy if the domain changes. AffectNet and Isanette, in turn, allow sentic computing to perform an open-domain opinion mining, in which text analysis is not simply based on word co-occurrence frequencies but rather on the latent/implicit features associated with concepts.

In order to assess the accuracy of such knowledge bases, hence, a comparison study was carried out by replacing AffectNet and Isanette with state-of-the-art knowledge bases in the opinion mining engine. In particular, WordNet, ConceptNet, and Probase were firstly swapped with Isanette to compare topic spotting performance of the engine on the Twitter hashtag repository. Secondly, the same knowledge bases were swapped with AffectNet to assess emotion recognition capabilities of the system on the LiveJournal database. Thirdly, WordNet, ConceptNet, and Probase were compared with the ensemble of Isanette and AffectNet to concurrently evaluate the engine's topic spotting and polarity detection capabilities on the PatientOpinion dataset. As for the Twitter evaluation, results show that Probase and Isanette perform significantly better than WordNet and ConceptNet, as these lack factual knowledge concepts such as Wii or Ford Focus (Table 5.3). Probase and Isanette topic spotting precision, on the other hand, are comparable as Probase hyponym-hypernym common knowledge is enough for this kind of task. It actually even outperforms Isanette sometimes as this contains just a subset of Probase instances (hub instances) and common sense knowledge does not play a key role in this type of classification.

| Category | WordNet | ConceptNet | Probase | Isanette |
|---|---|---|---|---|
| electronics | 34.2% | 45.4% | 76.8% | 79.3% |
| companies | 26.5% | 51.2% | 82.3% | 82.1% |
| countries | 38.4% | 65.1% | 89.1% | 86.4% |
| cities | 25.9% | 59.5% | 81.5% | 81.8% |
| operative systems | 37.1% | 51.6% | 79.7% | 78.3% |
| cars | 13.3% | 22.9% | 74.8% | 77.1% |

Table 5.3: Precision values relative to Twitter evaluation. Probase and Isanette perform significantly better than WordNet and ConceptNet, as these usually lack factual knowledge concepts.

| Category | WordNet | ConceptNet | Probase | AffectNet |
|---|---|---|---|---|
| joy-sadness | 47.1% | 55.7% | 33.9% | 79.4% |
| anticipation-surprise | 30.4% | 41.1% | 19.7% | 68.2% |
| anger-fear | 43.8% | 49.4% | 25.7% | 74.6% |
| trust-disgust | 27.2% | 39.3% | 12.8% | 69.9% |

Table 5.4: F-measure values relative to LiveJournal evaluation. WordNet, ConceptNet and AffectNet perform consistently better than Probase as this is based on semantic rather than affective relatedness of concepts.

As for the LiveJournal evaluation, the capability of the software engine to properly categorise antithetical affective pairs from the Hourglass model (namely joy-sadness, anticipation-surprise, anger-fear and trust-disgust) was tested. Results show that, in this case, Probase is consistently outperformed by WordNet, ConceptNet and AffectNet as it is based on semantic rather than affective relatedness of concepts (F-measure values are reported in Table 5.4). In Probase graph representation, in fact, instances like 'joy', 'surprise' and 'anger' are all close to each other, although they convey different affective valence, for being associated with the same hyponym-hypernym relationships.

As for the PatientOpinion evaluation, eventually, the ensemble application of Isanette and AffectNet turns out to be the best choice as it represents the best trade-off between common knowledge and affective common sense knowledge, which is particularly needed when aiming to infer both the cognitive and affective information associated with text (F-measure values are reported in Table 5.5). As also shown by previous experiments, in fact, common knowledge is particularly functional for tasks such as open-domain text auto-categorisation while affective common sense knowledge is notably useful for inferring the implicit emotional meaning underpinning words.

| Category | WordNet | ConceptNet | Probase | Isanette+AffectNet |
|---|---|---|---|---|
| clinical service | 35.7% | 49.3% | 56.6% | 82.4% |
| communication | 41.3% | 50.6% | 43.7% | 79.1% |
| food | 39.4% | 45.7% | 40.8% | 81.7% |
| parking | 47.8% | 51.4% | 49.1% | 77.2% |
| staff | 32.5% | 37.5% | 51.2% | 73.8% |
| timeliness | 44.1% | 50.2% | 41.9% | 80.6% |

Table 5.5: F-measure values relative to PatientOpinion evaluation. The ensemble application of Isanette and AffectNet represents the best trade-off between topic spotting and affect recognition.

More evidence of this is given in the next sections, in which the opinion mining engine (or a sub-part of it) is employed for the realisation of different tasks and applications in fields such as social web, HCI, and e-health.

## 5.2 Development of Social Web Systems

With the rise of the Social Web, there are now millions of humans offering their knowledge online, which means that the information is stored, searchable, and easily shared. This trend has created and maintained an ecosystem of participation, where value is created by the aggregation of many individual user contributions. Such contributions, however, are meant for human consumption and, hence, hardly accessible and processable by computers.

Making sense of the huge amount of social data available on the Web requires the adoption of novel approaches to natural language understanding that can give a structure to such data, in a way that they can be more easily aggregated and analysed. The above-described opinion mining engine can be employed in such area for NLP tasks requiring the inference of semantic and/or affective information associated with text, e.g., for the analysis of social network interaction dynamics or for managing online community data and metadata. This section, in particular, shows how the engine can be exploited for the development of a troll filtering system (subsection 5.2.1), a social media marketing tool (subsection 5.2.2), and an online personal photo management system (subsection 5.2.3)

### 5.2.1 Troll Filtering

The democracy of the Web is what made it so popular in the past decades, but such a high degree of freedom of expression also gave birth to negative side effects – the so called 'dark side' of the Web. Be it real or virtual world, in fact, existence of malicious faction among inhabitants and users is inevitable. An example of this, in the Social Web context, is the exploitation of anonymity to post inflammatory, extraneous, or off-topic messages in an online community, with the primary intent of provoking other users into a desired emotional response or of otherwise disrupting normal on-topic discussion.

Such a practice is usually referred as as 'trolling' and the generator of such messages is called 'a troll'. The term was first used in early 1990 and since then a lot of concern has been raised to contain or curb trolls. The trend of trolling appears to have spread a lot recently and it is alarming most of the biggest social networking sites since, in extreme cases such as abuse, has led some teenagers to commit suicide. These attacks usually address not only individuals but also entire communities. For example, reports have claimed that a growing number of Facebook tribute pages had been targeted, including those in memory of the Cumbria shootings victims and soldiers who died in Afghanistan. At present users cannot do much rather than manually delete abusive messages. Current anti-trolling methods, in fact, mainly consist in identifying additional accounts that use the same IP address and blocking fake accounts based on name and anomalous site activity, e.g., users who send lots of messages to non-friends or whose friend requests are rejected at a high rate.

In July 2010, Facebook launched an application that gives users a direct link to advice, help and the ability to report cyber problems to the child exploitation and online protection centre (CEOP). Reporting trouble through a link or a button, however, is a too slow process since social networking websites usually cannot react instantly to these alarms. A button, moreover, does not stop users from being emotionally hurt by trolls and it is more likely to be pushed by people who actually do not need help rather than, for instance, children who are being sexually groomed and do not realise it.

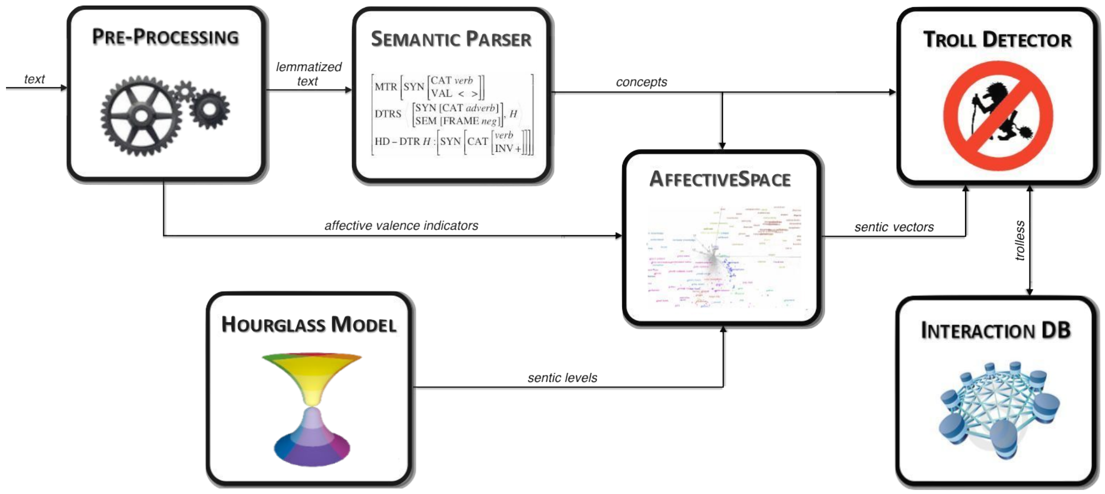


Figure 5.2: Troll filtering process. Once extracted, semantics and sentics are used to calculate blog posts' level of trollness, which is then stored in the interaction database for the detection of malicious behaviours.

A prior analysis of the trustworthiness of statements published on the Web has been presented by Rowe and Butters [210]. Their approach adopts a contextual trust value determined for the person who asserted a statement as the trustworthiness of the statement itself. Their study, however, does not focus on the problem of trolling, but rather on defining a contextual accountability for the detection of web, email, and opinion spam. The main aim of the troll filter proposed by Cambria et al. [211] (Fig. 5.2) is to identify malicious contents in natural language text with a certain confidence level and, hence, automatically block trolls.

To train the system, the concepts most commonly used by trolls are first identified by using the CF-IOF technique and, then, this set is expanded through spectral association. In particular, after analysing a set of 1000 offensive phrases extracted from Wordnik[1], it was found that, statistically, a post is likely to be edited by a troll when its average sentic vector has a high absolute value of Sensitivity and a very low polarity.

[1]`http://wordnik.com`

Hence, the *trollness* $t_i$ associated to a concept $c_i$ is defined as a float $\in [0,1]$ such that:

$$t_i(c_i) = \frac{s_i(c_i) + |Snsit(c_i)| - p_i(c_i)}{3} \tag{5.1}$$

where $s_i$ (float $\in [0,1]$) is the semantic similarity of $c_i$ wrt any of the CF-IOF seed concepts, $p_i$ (float $\in [-1,1]$) is the polarity associated to the concept $c_i$ and 3 is the normalisation factor. Hence, the total *trollness* of a post containing $N$ concepts is defined as:

$$t = \sum_{i=1}^{N} \frac{3\ s_i(c_i) + 4\ |Snsit(c_i)| - Plsnt(c_i) - |Attnt(c_i)| - Aptit(c_i)}{9N} \tag{5.2}$$

This information is stored, together with post type and content plus sender and receiver ID, in an interaction database that keeps trace of all the messages and comments interchanged between users within the same social network. Posts with a high level of *trollness* (current threshold has been set, using a trial-and-error approach, to 60%) are labelled as troll posts and, whenever a specific user addresses more than two troll posts to the same person or community, his/her sender ID is labelled as troll for that particular receiver ID. All the past troll posts sent to that particular receiver ID by that specific sender ID are then automatically deleted from the website (but kept in the database with the possibility for the receiver to either visualise them in an apposite *troll folder* and, in case, restore them). Moreover, any new post with a high level of *trollness* edited by a user labelled as troll for that specific receiver is automatically blocked, i.e., saved in the interaction database but never displayed in the social networking website.

This information, encoded as a sentic vector, is given as input to a troll detector which exploits it, together with the semantic information coming directly from the semantic parser, to calculate the post's *trollness* and, eventually, to detect and block the troll (according to the information stored in the interaction database). As an example of troll filtering process output, a troll post recently addressed to the Indian author, Chetan Bhagat, can be considered: "You can't write, you illiterate douchebag, so quit trying, I say!!!". In this case, there are a very high level of Sensitivity (corresponding sentic level 'rage') and a negative polarity, which give a high percentage of *trollness*, as shown below:

<*Concept:* !*'write'*>

<*Concept: 'illiterate'*>

<*Concept: 'douchebag'*>

<*Concept: 'quit try'*>

<*Concept: 'say'*>

*Semantics: 0.69*

*Sentics: [0.0, 0.17, 0,85, -0.43]*

*Polarity: -0.38*

*Trollness: 0.75*

Because the approach adopted by Rowe and Butters [210] is not directly comparable with the developed troll filtering system, a first evaluation was performed by considering a set of 500 tweets manually annotated as troll and non-troll posts, most of which were fetched from Wordnik. In particular, true positives were identified as posts with both a positive troll-flag and a *trollness* $\in$ [0.6, 1], or posts with both a negative troll-flag and a *trollness* $\in$ [0, 0.6). The threshold has been set to 60% based on trial-and-error over a separate dataset of 50 tweets. Results show that, by using the troll filtering process, inflammatory and outrageous messages can be identified with good precision (82.5%) and decorous recall rate (75.1%). In particular, the F-measure value (78.9%) is significantly high compared to the corresponding F-measure rates obtained by using Isanette and AnalogySpace in place of the AffectiveSpace process (Table 5.6).

| **Metric** | **Isanette** | **AnalogySpace** | **AffectiveSpace** |
|---|---|---|---|
| precision | 57.1% | 69.1% | 82.5% |
| recall | 40.0% | 56.6% | 75.1% |
| F-measure | 47.0% | 62.2% | 78.6% |

Table 5.6: Precision, recall, and F-measure values relative to the troll filter evaluation. The AffectiveSpace process performs consistently better than Isanette and AnalogySpace in detecting troll posts.

However, much better results are expected for the process evaluation at interaction-level, rather than just at post-level. In the next future, in fact, the troll filtering process will be evaluated by monitoring not just single posts but also users' holistic behaviour, i.e., contents and recipients of their interaction, within the same social network.

### 5.2.2 Social Media Marketing

The advent of Web 2.0 made users more enthusiastic about interacting, sharing and collaborating through social networks, online communities, blogs, wikis and other online collaborative media. In the last years, this collective intelligence has spread to many different areas in the Web, with particular focus on fields related to our everyday life such as commerce, tourism, education and health. The online review of commercial services and products, in particular, is an action that users usually perform with pleasure, to share their opinions about services they have received or products they have just bought, and it constitutes immeasurable value for other potential buyers.

This trend opened new doors to enterprises that want to reinforce their brand and product presence in the market by investing in online advertising and positioning, i.e., in social media marketing. In confirmation of the growing interest in social media marketing, several commercial tools have been recently developed to provide companies with a way to analyse the blogosphere on a large scale in order to extract information about the trend of the opinions relative to their products. Nevertheless most of the existing tools and the research efforts are limited to a polarity evaluation or a mood classification according to a very limited set of emotions. In addition, such methods mainly rely on parts of text in which emotional states are explicitly expressed and hence they are unable to capture opinions and sentiments that are expressed implicitly.

To this end, a novel social media marketing tool has been proposed by Cambria et al. [145] to provide marketers with a IUI for the management of social media information at semantic level, able to capture both opinion polarity and affective information associated with UGCs. A polarity value associated to an opinion, in fact, sometimes can be restrictive. Enriching automatic analysis of social media with affective labels such as 'joy' or 'disgust' can help marketers to have a clearer idea of what their customers think about their products. In particular, YouTube was selected as a social media source since, with its over 2 billions views per day, 24 hours of video uploaded every minute and 15 minutes a day spent by the average user, it represents more than 40% of online video market[2]. Specifically, the focus was on video reviews of mobile phones because of the quantity and the quality of the comments usually associated with them.

The social media analysis is performed through three main steps: firstly comments are analysed using the opinion mining engine, secondly the extracted information is encoded on the base of different web ontologies, and eventually the resulting knowledge base is made available for browsing through a multi-faceted classification website. Web resources, and social media resources in particular, represent a peculiar kind of data that is characterised for a deeply interconnected nature. Web itself is in fact based on links that bound together different data and information, and community contributed multimedia resources characterise themselves for the collaborative way in which they are created and maintained. An effective description of such resources needs therefore to capture and manage such interconnected nature, allowing to encode information not only about the resource itself but also about the linked resources into an interconnected knowledge base. Encoding information relative to a market product to analyse its market trends represents a situation in which this approach is particularly suitable and useful. In this case, in fact, it is necessary not only to encode the information relative to product features but also the information about the producer, the consumers and their opinions. To achieve this purpose, Semantic Web techniques are exploited. The Semantic Web initiative by W3C[3] tackles this problem through an appropriate representation of information in the web-page, able to univocally identify resources and encode the meaning of their description. In particular, the Semantic Web uses

[2] `http://viralblog.com/research/youtube-statistics`

[3] `http://w3.org`

uniform resource identifiers (URIs) to univocally identify entities available on the Web as documents or images but not as concepts or properties, and RDF data model to describe such resources in an univocally interpretable format, whose basic building block is an object-attribute-value triple, i.e., a statement.

Resources may be authors, books, publishers, places, people, hotels, rooms, search queries, etc., while properties describe relations between resources such as ‘writtenBy’, ‘age’, ‘title’. Statements assert the properties of resources and their values can be either resources or literals (strings). To provide machine-accessible and machine-processable representations, it is usual to encode RDF triples using XML syntax. Each triple can also be seen as a directed graph with labelled nodes and arcs, where the arcs are directed from the resource (the subject of the statement) to the value (the object of the statement). Each statement describes the graph node or connects it to other nodes, linking together multiple data from different sources without pre-existing schema. It is according to this representation that indeed the Semantic Web in its whole can be envisioned as a Giant Global Graph of Linked Data. RDF, however, does not make assumptions about any particular application domain, nor does it defines the semantics of any domain. For this purpose it is necessary to introduce ontologies.

Ontologies basically deal with knowledge representation and can be defined as formal explicit descriptions of concepts in a domain of discourse (named classes or concepts), properties of each concept describing various features and attributes of the concept (roles or properties), and restrictions on property (role restrictions). Ontologies make possible the sharing of common understanding about the structure of information among people or software agents. In addition, ontologies make possible reasoning, i.e., it is possible, starting from the data and the additional information expressed in the form of ontology, to infer new relationships between data. Different languages have been developed for the design of ontologies, among the most popular there are RDFS (RDF Schema) and OWL (Ontology Web Language). RDFS can be seen as a RDF vocabulary and a primitive ontology language. It offers certain modelling primitives with fixed meaning. Key concepts of RDF are class, subclass relations, property, sub-property relations, and domain and range restrictions. OWL is a language more specifically conceived for ontologies creation. It builds upon RDF and RDFS and a XML-based RDF syntax is used. Instances are defined using RDF descriptions and most RDFS modelling primitives are used. Moreover OWL introduces a number of features that are missing in RDFS such as local scope of property, disjointness of classes, Boolean combination of classes (like union, intersection and complement), cardinality restriction and special characteristics of properties (like transitive, unique or inverse).

The proposed framework for opinions and affective information description aims to be applicable to most of online resources (videos, images, text) coming from different sources, e.g., online video sharing services, blogs and social networks. To such purpose it is necessary to standardise as much as possible the descriptors used in encoding the information about multimedia resources and people to which the opinions refer (considering that every website uses its own vocabulary) in order to make it univocally interpretable and suitable to feed other applications. For this reason, the information relative to multimedia resources and people is encoded using respectively the descriptors provided by OMR[4] (Ontology for Media Resources) and FOAF[5] (Friend of a Friend Ontology). OMR represents an important effort to help circumventing the current proliferation of audio/video meta-data formats, currently carried on by the W3C Media Annotations Working Group. It offers a core vocabulary to describe media resources on the Web, introducing descriptors such as 'title', 'creator', 'publisher', 'createDate' and 'rating'. It defines semantic-preserving mappings between elements from existing formats. This ontology is supposed to foster the interoperability among various kinds of meta-data formats currently used to describe media resources on the Web. FOAF represents a recognised standard in describing people, providing information such as their names, birthdays, pictures, blogs, and especially other people they know, which makes it particularly suitable for representing data that appears on social networks

[4] `http://w3.org/TR/mediaont-10`
[5] `http://www.foaf-project.org`

and communities. OMR and FOAF together supply most of the vocabulary needed for describing media and people and other descriptors are added only when necessary. For example OMR, at least in the current realisation, does not supply vocabulary for describing comments, which are analysed to extract the affective information relative to media.

This ontology is extended by introducing the 'Comment' class, and by defining for it the 'author', 'text' and 'publicationDate' properties. In HEO, properties to link emotions to multimedia resources and people were introduced. In particular, 'hasManifestationInMedia' and 'isGeneratedByMedia' were defined to describe emotions that respectively occur and are generated in media, and the property 'affectPerson' was defined to connect emotions to people. Moreover, to improve the hierarchical organisation of emotions in HEO, WordNet-Affect (WNA) was exploited as an ontology.

Thus, the combination of HEO with WNA, OMR and FOAF provides a complete framework to describe not only multimedia contents and the users that have created, uploaded or interacted with them, but also the opinions and the affective content carried by the media and the way they are perceived by people (Fig. 5.3). As remarked above, due to the way they are created and maintained, community-contributed multimedia resources are very different from standard web-data. One fundamental aspect is constituted by the collaborative way in which such data is created, uploaded and annotated. A deep interconnection emerges in the nature of these data and meta-data, allowing for example to associate videos of completely different genre, but uploaded by the same user, or different users, even living in opposite sides of the world, who have appreciated the same pictures. In the context of social media marketing, this interdependence can be exploited to find similar patterns in customer reviews of commercial products and hence to gather useful information for marketing, sales, public relations, and customer service. Online reviews of electronics products, in particular, usually offer substantial and reliable information about the perceived quality of the products because of the size of electronics online market and the type of customers related to it.

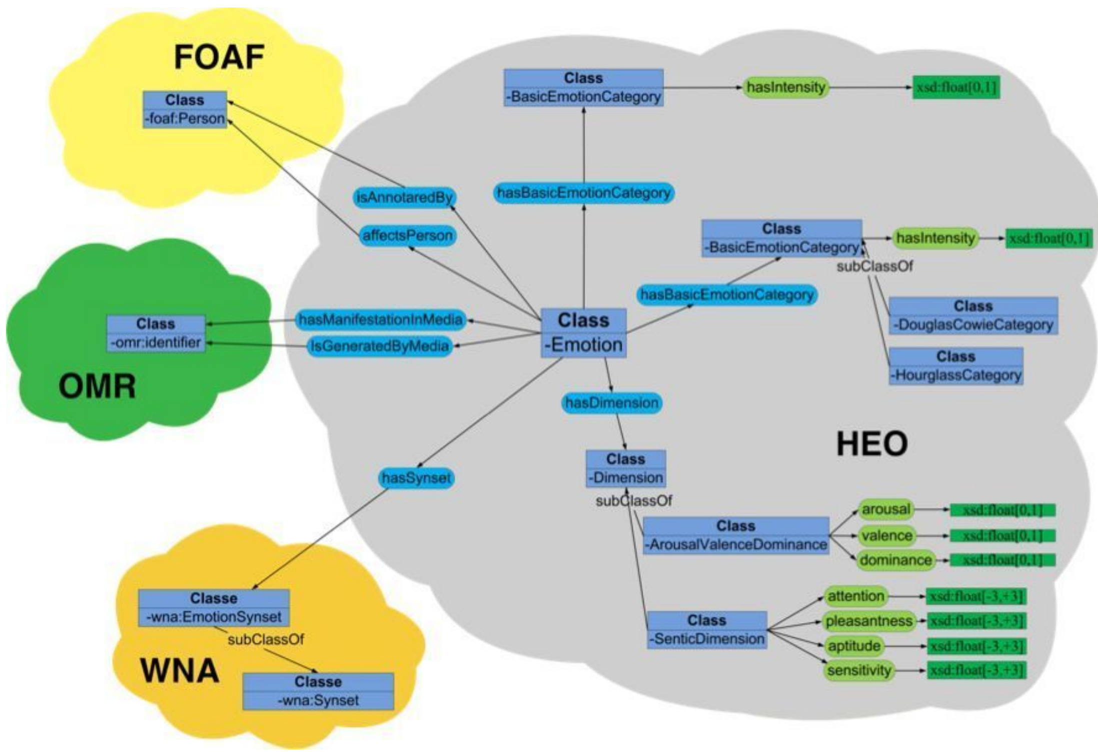


Figure 5.3: Merging different ontologies. The combination of HEO, WNA, OMR and FOAF provides a comprehensive framework for the representation of social media affective information.

To visualise this information, the multi-faceted categorisation paradigm is exploited. Faceted classification allows the assignment of multiple categories to an object, enabling the classifications to be ordered in multiple ways, rather than in a single, predetermined, and taxonomic order. This makes possible to perform searches combining the textual approach with the navigational one. Faceted search, in fact, enables users to navigate a multi-dimensional information space by concurrently writing queries in a text box and progressively narrowing choices in each dimension. For this application, SIMILE Exhibit API[6] is used. Exhibit is a set of Javascript files that allows to easily create rich interactive web-pages including maps, timelines and galleries, with very detailed client-side filtering. Exhibit pages use the multi-faceted classification paradigm to display semantically structured data stored in a Semantic Web aware format, e.g.,

[6] `http://simile-widgets.org/exhibit`

RDF or JavaScript object notation (JSON). One of the most relevant aspects of Exhibit is that, once the page is loaded, the web-browser also loads the entire data set in a lightweight database and performs all the computations (sorting, filtering, etc.) locally on the client-side, providing high performance. Because they are one of the most prolific types of electronic products in terms of data reviews available on the Web, mobile phones were selected as a review target. In particular, a set of 220 models was considered. Such models were ranked as the most popular according to Kelkoo[7], a shopping site featuring online shopping guides and user reviews, from which all the available information about each handset, such as model, brand, input type, screen resolution, camera type, standby time, and weight, was parsed.

This information was encoded in RDF and stored in a Sesame[8] triple-store, a purpose-built database for the storage and retrieval of RDF meta-data. YouTube Data API was then exploited to retrieve from YouTube database the most relevant video reviews for each mobile phone and their relative meta-data such as duration, rating, upload date and name, gender, and country of the uploaders. The comments associated with each video were also extracted and processed by means of sentic computing for emotion recognition and polarity detection. We then encoded the extracted opinions in RDF/XML, using the descriptors defined by HEO, WNA, OMR and FOAF, and inserted them into the triple-store.

Sesame can be embedded in applications and used to conduct a wide range of inferences on the information stored, based on RDFS and OWL type relations between data. In addition, it can also be used in a standalone server mode, much like a traditional database with multiple applications connecting to it. In this way all the knowledge stored inside Sesame can be queried and the results can also be retrieved in a semantic aware format and used for other applications. For the developed demo[9], the information contained in the triple-store was exported into a JSON file to feed the Exhibit application, in order to make it available for being browsed as a unique knowledge base.

[7] `http://kelkoo.co.uk`
[8] `http://openrdf.org`
[9] `http://cs.stir.ac.uk/~eca/sentics/smm`


Figure 5.4: A screenshot of the social media marketing tool. The faceted classification interface allows the user to quickly and efficiently navigate through both the explicit and implicit features of the different products.

Mobile phones are displayed in a dynamic gallery, that can be ordered according to different parameters like model, price and rating, showing technical information jointly with their video reviews and the opinions extracted from the relative comments (Fig. 5.4). Using faceted menus it is possible to explore such information both using the search box, to perform keyword-based queries, and filtering the results using the faceted menus, i.e., by adding or removing constraints on the facet properties.

In this way, it becomes very easy and intuitive to search for mobile phones of interest: users can specify the technical features required using the faceted menus and compare different phones that match such requirements consulting the video reviews and the opinions extracted from the relative comments. In addition it is possible to explore in detail the comments of each video review through a specific Exhibit page in which comments are organised in a timeline and highlighted in different colours, according to the value of their polarity. Moreover, faceted menus allow filtering the comments according to the reviewers' information, e.g., age, gender, and nationality.

Using such a tool a marketer can easily get an insight about the trend of a product, e.g., at the end of an advertising campaign, observing how the number of reviews and the relative satisfaction evolve in time and also monitoring this trend for different campaign targets. In order to evaluate the proposed system both on the level of opinion mining and sentiment analysis, we separately tested its polarity detection accuracy with a set of like/dislike-rated video reviews from YouTube and evaluated its affect recognition capabilities with a corpus of mood-tagged blogs from LiveJournal. In order to evaluate the system in terms of polarity detection accuracy, we exploited YouTube Data API to retrieve from YouTube database the ratings relative to the 220 video reviews previously selected for displaying in the faceted classification interface. On YouTube, in fact, users can express their opinions about videos either by adding comments or by simply rating them using a like/dislike button.

YouTube Data API makes this kind of information available by providing, for each video, number of raters and average rating, i.e., sum of likes and dislikes divided by number of raters. This information is expressed as a float $\in$ [1, 5] and indicates if a video is generally considered as bad (float $\in$ [1, 3]) or good (float $\in$ [3, 5]). This information was compared with the polarity values previously extracted by employing sentic computing on the comments relative to each of the 220 videos. True positives were identified as videos with both an average rating $\in$ [3, 5] and a polarity $\in$ [0, 1] (for positively rated videos), or videos with both an average rating $\in$ [1, 3] and a polarity $\in$ [-1, 0] (for negatively rated videos).

The evaluation showed that, by using the system to perform polarity detection, negatively and positively rated videos (37.7% and 62.3% of the total respectively) can be identified with precision of 97.1% and recall of 86.3% (91.3% F-measure). Since no mood-labelled dataset about commercial products is currently available, the LiveJournal database was used to test the system's affect recognition capabilities. For this test, a reduced set of 10 moods has been considered, i.e., 'ecstatic', 'happy', 'pensive', 'surprised', 'enraged', 'sad', 'angry', 'annoyed', 'scared' and 'bored'.

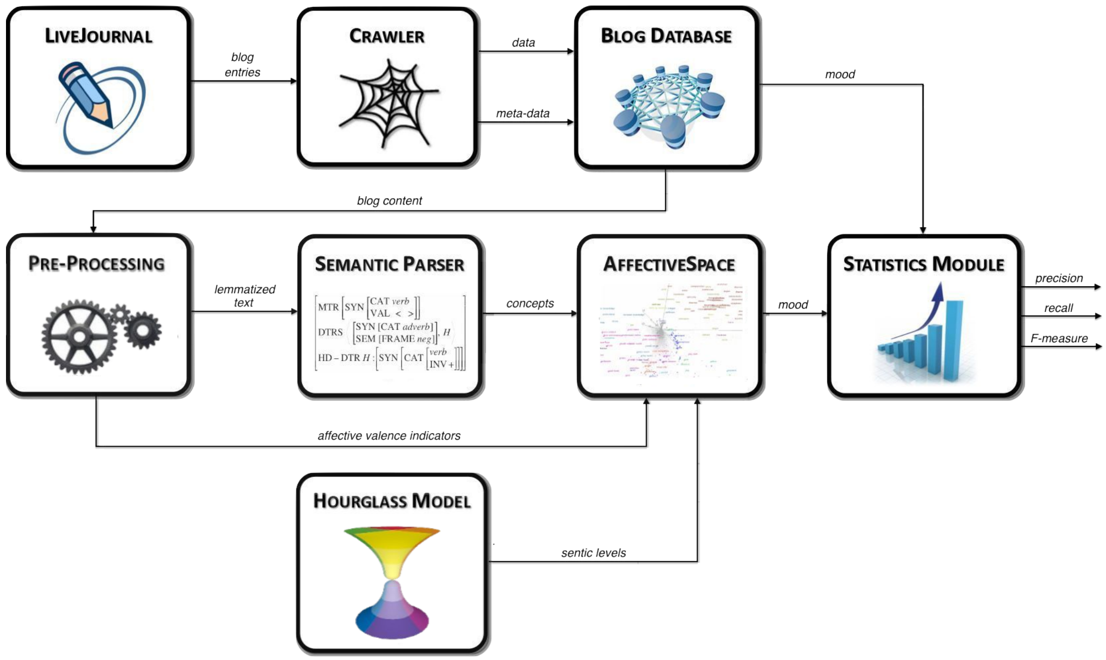


Figure 5.5: Sentics extraction evaluation. The process extracts sentics from natural language posts in the LiveJournal database, and then compare inferred emotional labels with the relative mood-tags in the database.

All LiveJournal accounts have Atom, RSS, and other data feeds which show recent public entries, friend relationships and interests. Unfortunately, there is no possibility to get mood-tagged blog-posts via data feeds, so an ad hoc crawler had to be designed. After retrieving and storing relevant data and meta-data for a total of 5,000 posts, the sentics extraction process was conducted on each of these and its outputs were compared with the relative mood-tags, in order to compute recall and precision rates as evaluation metrics (Fig. 5.5). On average, each post contained around 140 words and, from it, about 4 affective valence indicators and 60 sentic vectors were extracted. According to this information, mood-labels were assigned to each post and compared with the corresponding LiveJournal mood-tags, obtaining very good accuracy for each of the 10 selected moods (Table 5.7). Among these, 'happy' and 'sad' posts were identified with particularly high precision (89.2% and 81.8% respectively) and decorous recall rates (76.5% and 68.4%).

| Mood | Precision | Recall | F-measure |
|---|---|---|---|
| ecstatic | 73.1% | 61.3% | 66.6% |
| happy | **89.2%** | **76.5%** | **82.3%** |
| pensive | 69.6% | 52.9% | 60.1% |
| surprised | 81.2% | 65.8% | 72.6% |
| enraged | 68.9% | 51.6% | 59.0% |
| sad | **81.8%** | **68.4%** | **74.5%** |
| angry | 81.4% | 53.3% | 64.4% |
| annoyed | 77.3% | 58.7% | 66.7% |
| scared | 82.6% | 63.5% | 71.8% |
| bored | 70.3% | 55.1% | 61.7% |

Table 5.7: Evaluation results of the sentics extraction process. Precision, recall and F-measure rates are calculated for ten different moods by comparing the engine output with the relative LiveJournal mood-tag.

The F-measure values obtained, hence, were significantly good (82.3% and 74.5% respectively), especially if compared to the corresponding F-measure rates of the baseline methods (53.2% and 51.3% for keyword spotting, 63.5% and 58.4% for lexical affinity, 69.8% and 62.6% for statistical methods).

### 5.2.3 Sentic Album

Efficient access to online personal pictures requires the ability to properly annotate, organise and retrieve the information associated with them. While the technology to search personal documents has been available for some time, the technology to manage personal images is much more challenging. This is mainly due to the fact that, even if images can be roughly interpreted automatically, many salient features exist only in the user's mind. The only way for a system to accordingly index personal images, hence, is to try to capture and process such features. Existing CBIR systems such as QBIC [212], Virage [213], MARS [214], ImageGrouper [215], MediAssist [216], CIVR [217], EGO [218], ACQUINE [219] and K-DIME [220] have attempted to build IUIs capable of retrieving pictures according to their intrinsic content through statistics, pattern recognition, signal processing, computer vision, support vector machines and neural networks, but these techniques are still too weak to bridge the gap between the data representation and the images' conceptual models in the user's mind.

Image meta search engines such as Webseek [221], Webseer [222], PicASHOW [223], IGroup [224] or Google[10], Yahoo[11] and Bing[12] Images, on the other hand, rely on tags associated with online pictures but, in the case of personal photo management, users are unlikely to expend substantial effort to manually classify and categorise images in the hopes of facilitating future retrieval. Moreover these techniques, since they mainly depend on keyword-based rather than concept-based algorithms, often miss potential connections between keywords expressed through different vocabularies or concepts that exhibit implicit semantic connectedness. In order to properly deal with photo metadata, and hence effectively annotate images it is, in fact, necessary to work at a semantic, rather than syntactic, level.

A good effort in this sense has been made within the development of ARIA [225], a software agent which aims to facilitate the storytelling task by opportunistically suggesting photos which may be relevant to what the user is typing. ARIA goes beyond the naïve approach of suggesting photos by simply matching keywords in a photo annotation with keywords in the story. ARIA applies natural language techniques to the annotation process to extract concepts rather than keywords from the text. A similar approach has been followed by Raconteur [226], a system for conversational storytelling that encourages people to make coherent points, by instantiating large-scale story patterns and suggesting illustrative media. It exploits a large common sense knowledge base to perform NLP in real-time on a text chat between a storyteller and a viewer and recommends appropriate media items from a library. Both these approaches present a lot of advantages since concepts, unlike keywords, are not sensitive to morphological variation, abbreviations or near synonyms. However, simply relying on a semantic knowledge base is not enough to infer the salient features that make different pictures more or less relevant in each user's mind.

To this end, Sentic Album, a multi-tier architecture proposed by Cambria et al. [227], exploits AI and Semantic Web techniques to perform reasoning on different knowledge

[10]http://google.com/images

[11]http://images.search.yahoo.com

[12]http://bing.com/images

bases and, hence, infer both the cognitive and the affective information associated with photo metadata. The system, moreover, supports this concept-level analysis with content and context based techniques, in order to capture all the different aspects of online pictures and, hence, provide users with an IUI that is navigable in real-time through a multi-faceted classification website, since much of what is called problem-solving intelligence is really the ability to identify what is relevant and important in a context and to make that knowledge available just in time [228]. Cognitive and affective processes are tightly intertwined in everyday life [229]. The affective aspect of cognition and communication is recognised to be a crucial part of human intelligence and has been argued to be more fundamental in human behaviour and success in social life than intellect [230, 231].

Emotions, in fact, influence our ability to perform common cognitive tasks, such as forming memories and communicating with other people. A psychological study, for example, showed that people asked to concealed emotional facial expressions in response to unpleasant and pleasant slides remember the slides less well than control participants [232]. Similarly, a study of conversations revealed that romantic partners who were instructed to conceal both facial and vocal cues of emotion while talking about important relationship conflicts with each other, remembered less of what was said than did partners who received no suppression instructions [233]. Many studies have indicated that emotions both seem to improve memory for the gist of an event and to undermine memory for more peripheral aspects of the event [234, 235, 236, 237].

The idea, broadly, is that arousal causes a decrease in the range of cues an organism can take in. This narrowing of attention leads directly to the exclusion of peripheral cues, and this is why emotionality undermines memory for information at the event's edge. At the same time, this narrowing allows a concentration of mental resources on more central materials, and this leads to the beneficial effects of emotion on memory for the event's centre [238]. Hence, rather than assigning particular cognitive and affective valence to a specific visual stimulus, we more often balance the importance of personal pictures is according to how much the information in them contained is pertinent to our lives, goals, and values (or perhaps, the lives and values of people we care about). For this reason, a bad quality picture can be ranked high in the mind of a particular user, if it reminds him/her of a notably important moment or person of his/her life. Events and situations, in fact, are likely to be organised in human mind as interconnected concepts and most of the links relating such concepts are probably weighted by affect, as we tend to better recall memories associated with either very positive or very negative emotions, as well as we usually tend to more easily forget about concepts associated with very little or null affective valence. The problem, when trying to emulate such cognitive and affective processes, is that, while cognitive information is usually objective and unbiased, affective information is rather subjective and argumentative.

For example, while in the cognitive domain 'car' is always a car and there is usually not much discussion about the correctness of retrieving an image showing a tree in an African savannah under the label 'landscape', there might be some discussion about whether the retrieved car is "cool" or just "nice" or whether the found landscape is "peaceful" or "dull" [239]. To this end, Sentic Album applies sentic computing techniques on pictures data and metadata to infer what really matters to each user in different online photos. In particular, the Annotation Module mainly exploits metadata such as descriptions, tags and comments, termed 'conceptual metadata', associated with each image to extract its relative semantics and sentics and, hence, enhance the picture specification with its intrinsic cognitive and affective information. This concept-level annotation procedure is performed through an ensemble of sentic computing tools and techniques, and it is supported with a parallel content and context level analysis. User's personal photo data and metadata are currently pulled from Picasa (through Google Data API[13]) but, in the next future, the breadth of the system is planned to be expanded by interfacing it with more sources, e.g., other online photo sharing services, blogs, and social networks. The annotation module works at three different levels: content, context, and concept. The content based annotation, in par-

[13] http://code.google.com/apis/gdata

ticular, is performed through Python Imaging Library[14] (PIL), an external library for the Python[15] programming language that adds support for opening, manipulating and saving many different image file formats. For every online personal picture, in particular, PIL is exploited to extract luminance and chrominance information and other image statistics, e.g., the total, mean, standard deviation, and variance of the pixel values. The context-based annotation, in turn, exploits information such as timestamp, geolocation and user interaction metadata. Such metadata, termed 'contextual metadata', are processed by the Context Deviser, a sub-module that extracts small bits of information suitable for storing in a relational database for re-use at a later time, i.e., time, date, city and country of caption plus all the relevant user interaction metadata such as number and IDs of friends who viewed, commented or liked the picture. The concept-based annotation represents the core of the module and it is designed by means of sentic computing, which allows the system to go beyond a mere syntactic analysis of the metadata associated with pictures. A big problem of manual image annotation, in fact, is the different vocabulary that different users (or even the same user) can use to describe the content of a picture.

The different expertise and purposes of tagging users, in fact, may result in tags that use various levels of abstraction to describe a resource: a photo can be tagged at the 'basic level' of abstraction [240] as 'cat' or at a superordinate level as 'animal' or at various subordinate levels below the basic level as 'Persian cat' or 'Felis silvestris catus longhair Persian'. To overcome this problem, Sentic Album extends the set of available tags (if any) with related semantics and sentics and, to further expand the cognitive and affective metadata associated with each picture, it extracts additional common sense and affective concepts from its description and comments (if any). In particular, the conceptual metadata is processed by the opinion mining engine (Fig. 5.6). The AffectNet sub-module, specifically, finds matches between the retrieved concepts and those previously calculated using CF-IOF and spectral association.

[14] `http://pythonware.com/products/pil`
[15] `http://python.org`

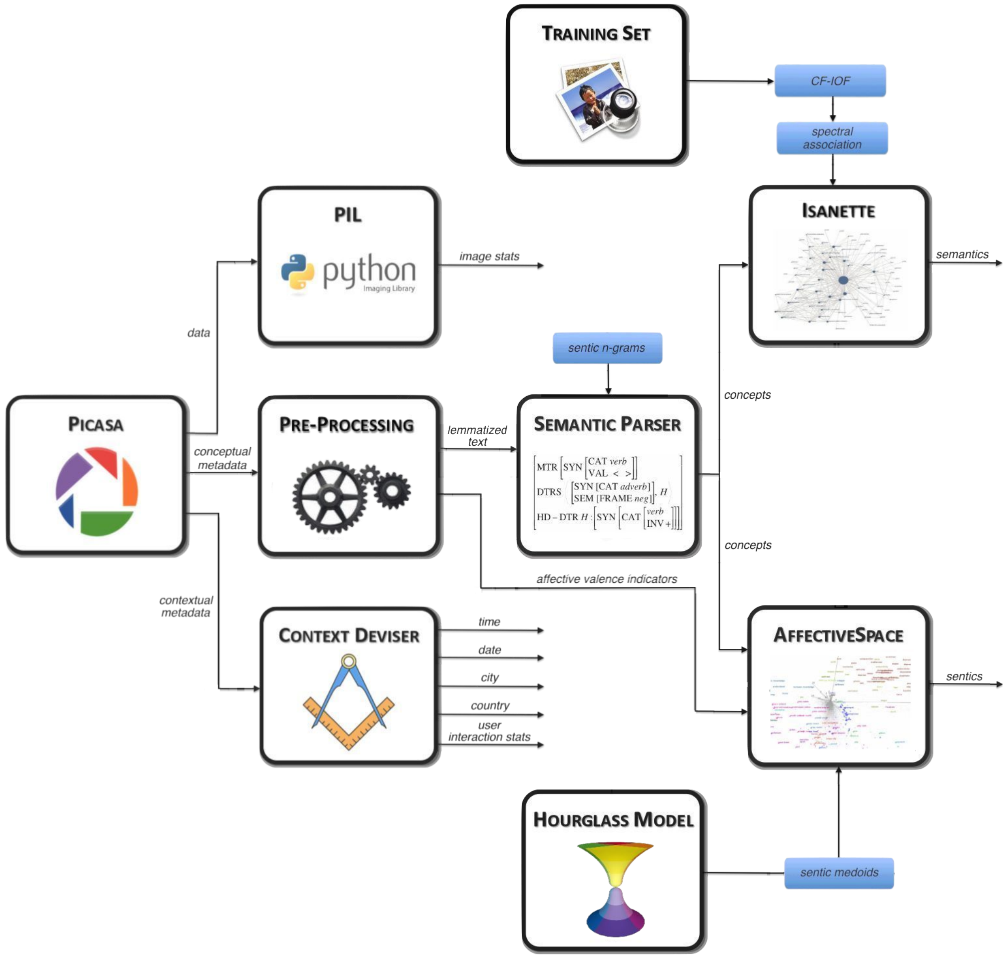


Figure 5.6: Sentic Album's annotation module. Online personal pictures are annotated at three different levels: content level (PIL), concept level (opinion mining engine) and context level (Context Deviser).

CF-IOF weighting is exploited to find seed concepts for a set of a-priori categories, extracted from Picasa's popular tags, meant to cover common topics in personal pictures, e.g., art, nature, friends, travel, wedding, or holiday. Spectral association is then used to expand this set with semantically related common sense concepts. The AffectiveSpace sub-module projects the retrieved concepts into the vector space representation of AffectNet. The multi-dimensional space, clustered with respect to the Hourglass model using sentic medoids, is then exploited to infer the affective valence of the retrieved concepts, in terms of Pleasantness, Attention, Sensitivity, and Aptitude, according to the relative position they occupy in the space.

This information, finally, is also exploited to calculate the overall polarity associated with pictures, which is calculated according to the sentics relative to each retrieved concept. Providing a satisfactory visual experience is one of the main goals for present-day electronic multimedia devices. All the enabling technologies for storage, transmission, compression, rendering should preserve, and possibly enhance, image quality; to do so, quality control mechanisms are required. Systems in charge to automatically assess visual quality are generally known as objective quality metrics. The design of objective quality metrics is a complex task because predictions must be consistent with human visual quality preferences. Human preferences are inherently quite variable and, by definition, subjective; moreover, in the field of visual quality, they stem from perceptual mechanisms that are not fully understood yet. A common choice is to design metrics that replicate the functioning of the human visual system (HVS) to a certain extent, or at least that take into account its perceptual response to visual distortions by means of numerical features [241]. Although successful, these approaches come with a considerable computational cost, which makes them impractical for most real-time applications. Computational intelligence paradigms allow to tackle the quality assessment task from a different perspective, since they aim at mimicking quality perception instead of designing an explicit model of the HVS [242, 243, 244].

In the special case of personal pictures, perceived quality metrics can be computed not only at content level, but also at concept and context level. One of the primary reasons why people take pictures is to remember the emotions they felt on special occasions of their lives. Extracting and storing such affective information can be a key factor in improving future searches, as users seldom want to find photos matching general requirements. Users' criteria in browsing personal pictures, in fact, are more often related to the presence of a particular person in the picture and/or its perceived quality (e.g., to find a good photo of your mother). Satisfying this type of requirement is a tedious task as chronological ordering or classification by event does not help much.

The process usually involves repeatedly trying to think of a matching picture, and then looking for it. An exhaustive search (looking through the whole collection for all of the photos matching a requirement) would normally only be carried out in exceptional circumstances, such as following a death in the family. In order to accordingly rank personal photos, Sentic Album exploits data and metadata associated with them to extract useful information at content, concept and context level and, hence, calculate the perceived quality of online pictures (PQOP), defined as:

$$PQOP(p,u) = 3\frac{Content(p) * Concept(p,u) * Context(p,u)}{Content(p) + Concept(p,u) + Context(p,u)} \tag{5.3}$$

where $Content(p)$, $Concept(p,u)$, and $Context(p,u)$ are float $\in$ [0,1] representing image quality assessment values associated with picture $p$ and user $u$, in terms of visual, conceptual and contextual information, respectively. $Content(p)$, in particular, is computed from numerical features extracted through a reduced-reference framework for objective quality assessment exploiting a circular extreme learning machine (C-ELM)[245] and the colour correlogram [246] of $p$. $Concept(p,u)$, in turn, specifies how much the picture $p$ is relevant to the user $u$ in terms of cognitive and affective information. $Context(p,u)$, eventually, defines the degree of relevance of picture $p$ for user $u$ in terms of time, location and user interaction.

The 3C (Content, Concept and Context) are all equally relevant for measuring how good a personal picture is to the eye of a user. According to the formula, in fact, if any of the 3C is null the PQOP is null as well, even though the remaining elements of the 3C have both maximum value, e.g., a perfect quality picture ($Content(p) = 1$) taken in the hometown of the user on the date of his birthday ($Context(p,u) = 1$) but depicting people he/she does not know and objects/places that are totally irrelevant for him/her ($Concept(p,u) = 0$). The Storage Module is the middle-tier in which the outputs of the Annotation Module are stored, in a way that these can be easily accessible by the Search and Retrieval Module at a latter time. The module stores information relative to photo data and metadata redundantly at three levels:

1. in a relational database fashion
2. in a Semantic Web format
3. in a matrix format

Sentic Album stores information in three main SQL databases (Fig. 5.7), that is a Content DB, for the information relative to data (image statistics), a Concept DB, for the information relative to conceptual metadata (semantics and sentics), and a Context DB, for the information relative to contextual metadata (timestamp, geolocation and user interaction metadata).

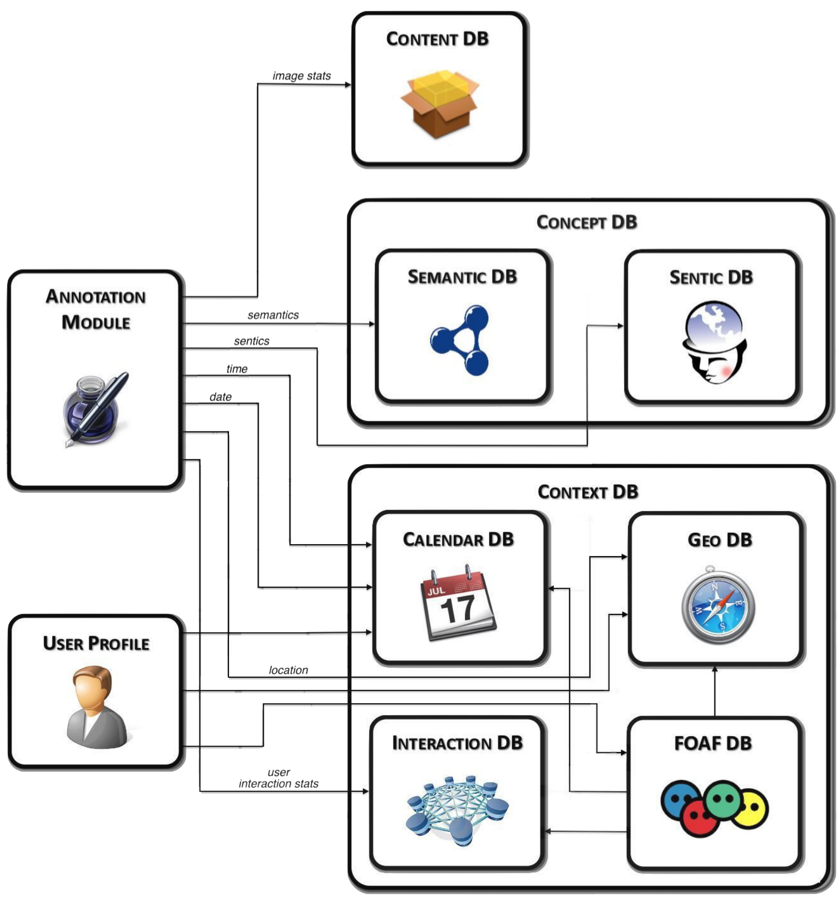


Figure 5.7: Sentic Album's SQL-level storage module. Image statistics are saved into the Content DB, semantics and sentics are stored into the Concept DB, while timestamp and geolocation are saved into the Context DB.

The Concept DB, in particular, consists of two databases, the Semantic DB and the Sentic DB, in which the cognitive and the affective information associated with photo metadata, respectively, are stored. The Context DB, in turn, is divided into four databases, the Calendar, Geo, FOAF (Friend Of A Friend) and Interaction DBs, which contain the information relative to timestamp, geolocation, social links and social interaction respectively. These databases are also integrated with information coming from the web profile of the user such as user's DOB (for the Calendar DB), user's current location (for the Geo DB) or user's list of friends (for the FOAF DB). The FOAF DB, in particular, plays an important role within the Context DB since it provides the other peer databases with information relative to user's social connections, e.g., relatives' birthdays or friends' hometowns. Moreover, the Context DB receives extra contextual information from the inferred semantics. Personal names in the conceptual metadata are recognised by building a dictionary of first names from the Web and combining that with regular expressions to recognise full names. These are added to the database (in the FOAF DB) together with geographical places (in the Geo DB), which are also mined from databases on the Web and added to the parser's semantic lexicon.

As for the Semantic Web format [247], all the information related to pictures' metadata is stored in RDF/XML according to a set of predefined web ontologies. This operation aims to make the description of the semantics and sentics associated with pictures applicable to most online images coming from different sources, e.g., online photo sharing services, blogs, social networks. To further this aim, it is necessary to standardise as much as possible the descriptors used in encoding the information about multimedia resources and people to which the images refer, in order to make it univocally interpretable and suitable to feed other applications. Hence, the ensemble of HEO, OMR, FOAF and WNA is used again.

As for the storage of photo data and metadata in a matrix format, a dataset, termed '3CNet', is built, which integrates the information from the 3C in a unique knowledge base. The aim of this representation is to exploit principal component analysis (PCA) to later organise online personal images in a multi-dimensional vector space (as for AffectiveSpace) and, hence, reason on their similarity. 3CNet, in fact, is an $n \times m$ matrix whose rows are user's personal pictures IDs, whose columns are either content, concept and context features (e.g., 'contains cold colours', 'conveys joy' or 'located in Italy'), and whose values indicate truth values of assertions. Therefore, in 3CNet, each image is represented by a vector in the space of possible features whose values are +1, for features that produce an assertion of positive valence, -1, for features that produce an assertion of negative valence, and 0 when nothing is known about the assertion.

The degree of similarity between two images, then, is the dot product between their rows in 3CNet. The value of such a dot product increases whenever two images are described with the same feature and decreases when they are described by features that are negations of each other. The main aim of the Search and Retrieval Module is to provide users with an IUI that allows them to easily manage, search and retrieve their personal pictures online (Fig. 5.8).

Most of the existing photo management systems let users search for pictures through a keyword-based query, but results are hardly ever good enough since it is very difficult to come up with an ideal query from the user's initial request. The initial idea of an image the user has in mind before starting a search session, in fact, often deviates from the final results he/she will choose [248]. In order to let users start from a sketchy idea and then dynamically refine their search, the multi-faceted classification paradigm is adopted.

Personal images are displayed in a dynamic gallery that can be ordered according to different parameters, either textual or numeric, that is visual features (e.g., colour balance, hue, saturation, brightness and contrast), semantics (i.e., common sense concepts such as 'go jogging' and 'birthday party' but also people and objects contained in the picture), sentics (i.e., emotions conveyed by the picture and its polarity) and contextual information (e.g., time of caption, location and social information such as users who viewed/commented the picture).


Figure 5.8: Sentic Album's search and retrieval module. The IUI allows to browse personal images both by performing keyword-based queries in the search box and by adding or removing constraints on the facet properties.

In particular, NLP techniques similar to those used to process the image conceptual metadata are employed to analyse the text typed in the search box and, hence, perform queries on the SQL databases of the Storage Module. The order of visualisation of the retrieved images is given by the PQOP, so that images containing more relevant information at content, concept and context level are first displayed. If, for example, the user is looking for pictures of his/her partner, Sentic Album firstly proposes photos representing important events such as first date, first childbirth or honeymoon, that is, pictures with high PQOP. Storage Module's 3CNet is also exploited in the IUI, in order to find similar pictures. Towards the end of a search, in fact, the user sometimes is interested in finding pictures similar to one of those so far obtained, even if this does not fulfill the constraints currently set via the facets. To such purpose, every picture is provided with a 'like me' button that opens a new Exhibit window displaying content, concept and context related images, independently from any constraint.

Picture similarity is calculated by means of PCA and, in particular, through TSVD, as for AffectiveSpace. The number of singular values to be discarded (in order to reduce the dimensionality of 3CNet and hence reason on picture similarity) is chosen accordingly to the total number of user's online personal pictures and the amount of available metadata associated with them, i.e., according to size and density of 3CNet. Thus, by exploiting the information sharing property of TSVD, images specified by similar content, concept and context are likely to have similar features and, hence, tend to fall near each other in the built-in vector space. The IUI, eventually, also offers to display images according to date of caption on a timeline. Chronology, in fact, is a key categorisation concept for the management of personal pictures. Having the collection in chronological order is helpful for locating particular photos or events, because it is usually easier to remember when an event occurred relative to other events, than to remember its absolute date and time [249].

Many works dealing with object detection, scene categorisation or content analysis on the cognitive level have been published, trying to bridge the semantic gap between represented objects and high-level concepts associated with them [250], however, where affective retrieval and classification of digital media is concerned, publications, and especially benchmarks, are very few [251]. To overcome the lack of relevant datasets, the performance and the user-friendliness of Sentic Album were tested on a topic and mood tagged evaluation dataset and through a usability test on a pool of 18 Picasa regular users, respectively.

As for the system performance testing, in particular, 1,000 LiveJournal posts with labels matching Picasa tags such as 'friends', 'travel', and 'holiday', were selected in order to collect natural language text that is likely to have the same semantics as the conceptual metadata typical of personal photos. The classification test, hence, concurrently estimated the capacity of the system to infer both the cognitive and affective information (topic and mood tags, respectively) usually associated with online personal pictures (Table 5.8).

| LiveJournal Tag | Precision | Recall | F-measure |
|---|---|---|---|
| art | 62.9% | 55.6% | 59.0% |
| friends | **77.2%** | **65.4%** | **70.8%** |
| wedding | 71.3% | 60.4% | 65.4% |
| holiday | 68.9% | 59.2% | 63.7% |
| travel | **81.6%** | **71.1%** | **75.9%** |
| nature | 67.5% | 61.8% | 64.5% |

Table 5.8: System performance test. Assessment of Sentic Album's accuracy in inferring the cognitive (topic tags) and affective (mood tags) information associated with the conceptual metadata typical of personal photos.

The classification of 'travel' and 'friends' posts, in particular, was performed with a precision of 81.6% and 77.2% and recall rates of 65.4% and 71.1%, respectively. The total F-measure rates, hence, were considerably good (75.9% for 'travel' posts and 70.8% for 'friends' posts) in comparison with the corresponding F-measure rates obtained by using Probase (48.6%) and ConceptNet (58.2%) in place of AffectNet within the annotation module.

As for the usability test, users were asked to freely browse their online personal collections using Sentic Album IUI and to retrieve particular sets of pictures, in order to judge both usability and accuracy of the interface. Common queries included "find a funny picture of your best friend", "search for the shots of your last summer holiday", "retrieve pictures of you with animals", "find an image taken on Christmas 2009", "search for pictures of you laughing" and "find a good picture of your mom". From the test, it emerged that users really appreciate being able to dynamically and quickly set/remove constraints in order to display specific sets of pictures (which they cannot do in Picasa). After the test session, participants were asked to fill-in an online questionnaire in which they had to rate, on a five-level scale, each single functionality of the interface according to their perceived utility. Concept facets and timeline, in particular, resulted to be the most used by participants for search and retrieval tasks (Table 5.9). Users also really appreciated the 'like me' functionality, which most of the time proposed very relevant semantically and affectively related pictures (again not available in Picasa).

| Feature | Not at all | Just a little | Somewhat | Quite a lot | Very much |
|---|---|---|---|---|---|
| Concept facets | 0% | 0% | 5.6% | 5.6% | 88.8% |
| Content facets | 77.8% | 16.6% | 5.6% | 0% | 0% |
| Context facets | 16.6% | 11.2% | 5.6% | 33.3% | 33.3% |
| Search box | 0% | 11.2% | 16.6% | 33.3% | 38.9% |
| Like me | 0% | 5.6% | 5.6% | 16.6% | 72.2% |
| Timeline | 0% | 0% | 0% | 16.6% | 83.4% |
| Sorting | 11.2% | 33.3% | 33.3% | 16.6% | 5.6% |

Table 5.9: Perceived utility of the different interface features by 18 Picasa regular users. Participants particularly appreciated the usefulness of concept facets and timeline, for search and retrieval tasks.

When freely browsing their collections, users were particularly amused by the ability to navigate their personal pictures according to the emotion these conveyed, even though they did not always agree with the results. Additionally, participants were not very happy with the accuracy of the search box, especially if they searched for one particular photo out of the entire collection. However, they always very much appreciated the order in which the pictures were proposed, which allowed them to quickly have all the most relevant pictures available as first results. 83.3% of test users declared that, despite not being as nifty as Picasa, Sentic Album is a very good photo management tool (especially for its semantic faceted search and PQOP functionalities) and they hope they could still be using it because, in the end, what really counts when browsing personal pictures is to find best matches in the shortest amount of time.

## 5.3 Development of HCI Systems

Human computer intelligent interaction is an emerging field aimed at providing natural ways for humans to use computers as aids. It is argued that for a computer to be able to interact with humans it needs to have the communication skills of humans. One of these skills is the affective aspect of communication, which is recognised to be a crucial part of human intelligence and has been argued to be more fundamental in human behaviour and success in social life than intellect [230, 231]. Emotions influence cognition, and therefore intelligence, especially when it involves social decision-making and interaction.

The latest scientific findings indicate that emotions play an essential role in decision-making, perception, learning and more. Most of the past research on affect sensing has considered each sense such as vision, hearing and touch in isolation. However, natural human-human interaction is multi-modal: we communicate through speech and use body language (posture, facial expressions, gaze) to express emotion, mood, attitude, and attention. To this end, a novel fusion methodology is proposed, which is able to fuse any number of unimodal categorical modules, with very different time-scales, output labels and recognition success rates, in a simple and scalable way. In particular, such a methodology is exploited to fuse the outputs of the opinion mining engine with the ones of a facial expression analyser for designing an embodied conversational agent with affective common sense (subsection 5.3.1). This section, moreover, illustrates how the engine can be exploited for the development of HCI applications in fields such as instant messaging (IM) (subsection 5.3.2) and multimedia management (subsection 5.3.3).

### 5.3.1 Sentic Avatar

The capability of perceiving and expressing emotions through different modalities is a key issue for the enhancement of HCI. Natural human-human affective interaction is inherently multi-modal: people communicate emotions through multiple channels such as facial expressions, gestures, dialogues, etc. Although several studies prove that the multi-sensory fusion (e.g., audio, visual and physiological responses) improves robustness and accuracy of human emotion analysis [123, 252, 253], most emotional recognition works still focus on increasing the success rates in sensing emotions from a single channel rather than merging complementary information across channels.

The multi-modal fusion of different affective channels is far from being solved and represents an active and open research issue [254]. Affect recognition from multiple modalities has a short historical background and is still in its first stage. It was not till 1998 that computer scientists attempted to use multiple modalities for recognition of emotions/affective states [255].

The results were promising: using multiple modalities improved the overall recognition accuracy helping the systems function in a more efficient and reliable way. Following the findings in psychology, which suggested that the most significant channel for judging emotional cues of humans is the visual channel of face and body [256], a number of works combine facial expressions and body gestures for affect sensing [257, 258, 259]. Other approaches combine different biological information such as brain signals or skin conductivity for affect sensing [260, 261]. However this research makes use of a single information channel, i.e., a single type of computer input device, and, therefore, must assume the reliability on this channel. For that reason, the trend in recent works is to consider and combine affective information coming from different channels. That way, eventual changes on the reliability of the different information channels are considered.

Recent literature on multi-modal affect sensing is focused on the fusion of data coming from the visual and audio channels. Most of those works make use of the visual channel for body gesture recognition [262] or facial expression classification [263] and the audio channel to analyse non-linguistic audio cues such as laughters [264], coughs [265] or cries [266]. However, very few works fuse information coming from the visual channel with linguistic-based (speech contents) audio affect sensing. With all these new areas of research in affect sensing, a number of challenges have arisen. In particular, the synchronisation and fusion of the information coming from different channels is a big problem to solve. Previous studies fused emotional information either at a decision-level, in which the outputs of the unimodal systems are integrated by the use of suitable expert criteria [267], or at a feature-level, in which the data from both modalities are combined before classification [268].

In any case, the choice of fusion strategy depends on the targeted application. Accordingly, all available multi-modal recognisers have designed and/or used ad hoc solutions for fusing information coming from multiple modalities but cannot accept new modalities without re-defining the whole system. In summary, there is not a general consensus when fusing multiple modalities and systems' scalability is not possible.

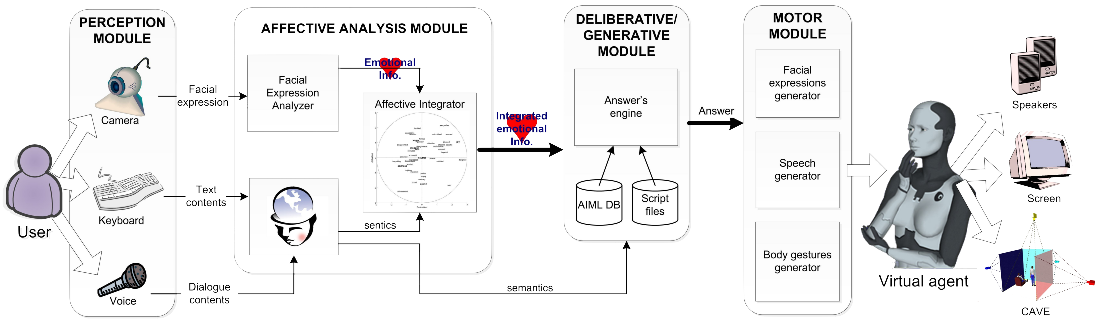


Figure 5.9: Sentic Avatar's architecture. The system mainly consists of two modules for managing the avatar's inputs and outputs, and two modules for performing affective common sense reasoning.

Sentic Avatar is an embodied conversational agent (ECA), proposed by Cambria et al. [96], based on the multi-modal animation engine Maxine [269], which consists of four main modules: Perception, Affective Analysis, Deliberative/Generative, and Motor module (Fig. 5.9). The Perception module simply consists of the hardware necessary to gather the multi-modal information from the user, i.e., keyboard, microphone, and webcam. The Affective Analysis module aims to infer the user's affective state from the different inputs and integrate it. The Deliberative/Generative module is in charge of processing the extracted emotional information to manage the virtual agent's decisions and reactions, which are finally generated by the Motor module.

The Affective Analysis module is in charge of extracting emotions from the textual, vocal, and video inputs and integrating them. It consists of three main parts: the opinion mining engine, for inferring semantics and sentics associated with typed-in text and speech-to-text converted contents, the Facial Expression Analyser, for extracting affective information from video, and the Affective Integrator, for integrating the outputs coming from the two previous modules. The Facial Expression Analyser achieves an automatic classification of the shown facial expressions into discrete emotional categories. It is able to classify the user's emotion in terms of Ekman's six universal emotions (*fear*, *sadness*, *joy*, *disgust*, *surprise* and *anger*) [118] plus *neutral*, giving a membership confidence value to each emotional category.

The face modelling selected as input for the Facial Expression Analyser follows a feature-based approach: the inputs are a set of facial distances and angles calculated from feature points of the mouth, eyebrows and eyes. In fact, the inputs are the variations of these angles and distances with respect to the neutral face. The points are obtained thanks to a real-time facial feature tracking program [270]. Fig. 5.10 shows, on the left side, the correspondence of these points with those defined by the MPEG4 standard. On the right side, in turn, the set of parameters obtained from these points is shown. In order to make the distance values consistent (independently of the scale of the image, the distance to the camera, etc.) and independent of the expression, all the distances are normalised with respect to the distance between the eyes, i.e., the MPEG4 Facial Animation Parameter Unit (FAPU), also called ESo. The choice of angles provides a size invariant classification and saves the effort of normalisation. As regards the classification process itself, the system intelligently combines the outputs of 5 different classifiers simultaneously.

In this way, the overall risk of making a poor selection with a given classifier for a given input is reduced. The classifier combination chosen follows a weighted majority voting strategy, where the voted weights are assigned depending on the performance of each classifier for each emotion. In order to select the best classifiers to combine, the Waikado Environment for Knowledge Analysis (Weka) tool was used [271]. This provides a collection of machine learning algorithms for data mining tasks. From this collection, five classifiers were selected after tuning: RIPPER, MLP, SVM, NB, and C4.5. The selection was based on their widespread use as well as on the individual performance of their Weka implementation. To train the classifiers and evaluate the performance of the system, two different facial emotion databases were used: the FGNET database [272] that provides video sequences of 19 different Caucasian people, and the MMI Facial Expression Database [273] that holds 1280 videos of 43 different subjects from different races (Caucasian, Asian and Arabic). Both databases show Ekman's six universal emotions plus *neutral*.

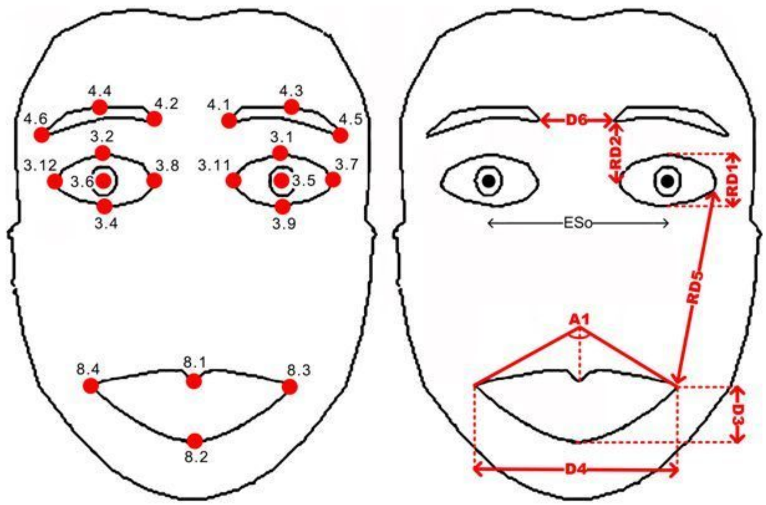


Figure 5.10: Tracked facial feature points according to MPEG4 standard (on the left) and corresponding facial parameters (on the right). The variations of such distances and angles are exploited for affect recognition.

A new database has been built for testing this work with a total of 1500 static frames carefully selected from the apex of the video sequences from the FG-NET and MMI databases. The results obtained when applying the strategy explained previously to combine the scores of the five classifiers are shown in the form of confusion matrix in Table 5.10 (results have been obtained with a 10-fold cross-validation test over the 1500 database images). As can be observed, the success rates for *neutral*, *joy* and *surprise* are very high (84.44%–95.23%).

However, the system tends to confuse *disgust* with *fear*, *anger* with *disgust* and *fear* with *surprise*; therefore, the performance for those emotions is slightly worse. The lowest result of the classification is for *sadness*: it is confused with *neutral* on 67.80% of occasions, due to the similarity of the facial expressions. Confusion between these pairs of emotions occurs frequently in the literature and for this reason many classification works do not consider some of them. Nevertheless, the results can be considered positive as two incompatible emotions (such as *sadness* and *joy* or *fear* and *anger*) are confused on less than 0.2% of occasions.

| *classified as* | disgust | joy | anger | fear | sadness | neutral | surprise |
|---|---|---|---|---|---|---|---|
| disgust | **79.41%** | 0% | 2.39% | 18.20% | 0% | 0% | 0% |
| joy | 4.77% | **95.23%** | 0% | 0% | 0% | 0% | 0% |
| anger | 19.20% | 0% | **74.07%** | 0% | 3.75% | 2.98% | 0% |
| fear | 9.05% | 0% | 0% | **62.96%** | 8.53% | 0% | 19.46% |
| sadness | 0.32% | 0.20% | 1.68% | 0% | **30.00%** | 67.80% | 0% |
| neutral | 0% | 0% | 1.00% | 2.90% | 4.10% | **92.00%** | 0% |
| surprise | 0% | 0% | 0% | 11.23% | 0% | 4.33% | **84.44%** |

Table 5.10: Confusion matrix obtained combining the five classifiers. Success rates for *neutral*, *joy* and *surprise* are very high but the system tends to confuse *disgust* with *fear*, *anger* with *disgust* and *fear* with *surprise*.

Another relevant aspect to be taken into account when evaluating the results is human opinion. The labels provided in the database for training classifiers correspond to the real emotions felt by users although they do not necessarily have to coincide with the perceptions other human beings may have about the facial expressions shown. Undertaking this kind of study is very important when dealing with human-computer interaction, since the system is proved to work in a similar way to the human brain. In order to take into account the human factor in the evaluation of the results, 60 persons were told to classify the 1500 images of the database in terms of emotions. As a result, each one of the frames was classified by 10 different people in 5 sessions of 50 images. The Kappa statistic obtained from raters' annotations is equal to 0.74 (calculated following the formula proposed in [274]), which indicates an adequate inter-rater agreement in the emotional images annotation. With this information, the evaluation of the results was repeated: the recognition was marked as good if the decision was consistent with that reached by the majority of the human assessors. The results (confusion matrix) of considering users' assessment are shown in Table 5.11. As can be seen, the success ratios have considerably increased. Therefore, it can be concluded that the confusions of the algorithms go in the same direction as those of the users: the adopted classification strategy is consistent with human classification.

The opinion mining engine outputs a list of sentic vectors that encompass the affective information associated with text and dialogue contents in terms of Pleasantness, Attention, Sensitivity, and Aptitude, while the Facial Expression Analyser provides

| *classified as* | disgust | joy | anger | fear | sadness | neutral | surprise |
|---|---|---|---|---|---|---|---|
| disgust | **84.24%** | 0% | 2.34% | 13.42% | 0% | 0% | 0% |
| joy | 4.77% | **95.23%** | 0% | 0% | 0% | 0% | 0% |
| anger | 15.49% | 0% | **77.78%** | 0% | 3.75% | 2.98% | 0% |
| fear | 1.12% | 0% | 0% | **92.59%** | 2.06% | 0% | 4.23% |
| sadness | 0.32% | 0.20% | 1.68% | 0% | **66.67%** | 31.13% | 0% |
| neutral | 0% | 0% | 0% | 0.88% | 1.12% | **98.00%** | 0% |
| surprise | 0% | 0% | 0% | 6.86% | 0% | 2.03% | **91.11%** |

Table 5.11: Confusion matrix obtained after considering human assessment. Success ratios considerably increase, meaning that the adopted classification strategy is consistent with human classification.

an affective evaluation of video contents in terms of Ekman's six universal emotions. The evaluation dimension measures how a human feels, from positive to negative. The activation dimension measures whether humans are more or less likely to take some action under the emotional state, from active to passive. To overcome the problem of the integration of the affective information coming from the opinion mining engine and the Facial Expression Analyser, a continuous 2D description of affect is considered.

Specifically, the Whissell space [129] is used: in it, the emotion-related words corresponding to each one of Ekman's six emotions plus *neutral* and to the levels of the Hourglass of Emotions have a specific location. Thanks to this, the sentics of the opinion mining engine and the labels provided by the Facial Expression Analyser can be mapped in the Whissell space: a pair of values <activation, evaluation> can be calculated from the obtained labels, and hence concurrently visualised and compared in the 2D space (Fig. 5.11). In particular, the process of affective integration is achieved through the following three steps.

Firstly, each one of the emotional labels inferred by the opinion mining engine from the video spoken sentence is mapped as a 2D point on to the Whissell space. Secondly, the Facial Expression Analyser outputs the user's emotion in terms of Ekman's six universal emotions (plus *neutral*), giving a membership confidence value to each emotional category. The mapping of its output in the Whissell space is carried out considering each of Ekman's six basic emotions plus *neutral* as 2D weighted points in the <activation, evaluation> space, where the weights are assigned depending on the confidence value obtained for each emotion in the classification process. The final detected emotion is calculated as the centre of mass of the seven weighted points in the Whissell space. That way, the Facial Expression analyser outputs one emotional location in the Whissell space per frame of the studied video sequence.

Finally, the whole set of 2D <activation, evaluation> points obtained from both the opinion mining engine and the Facial Expression Analyser is fitted to the Minimum Volume Ellipsoid (MVE) that better covers the shape of the set of extracted points. The MVE is calculated following the algorithm described by Kumar and Yildrim [275]. The final emotional information outputted by affective analysis module for the whole video sequence is given by the x-y coordinates of the centre of that MVE.

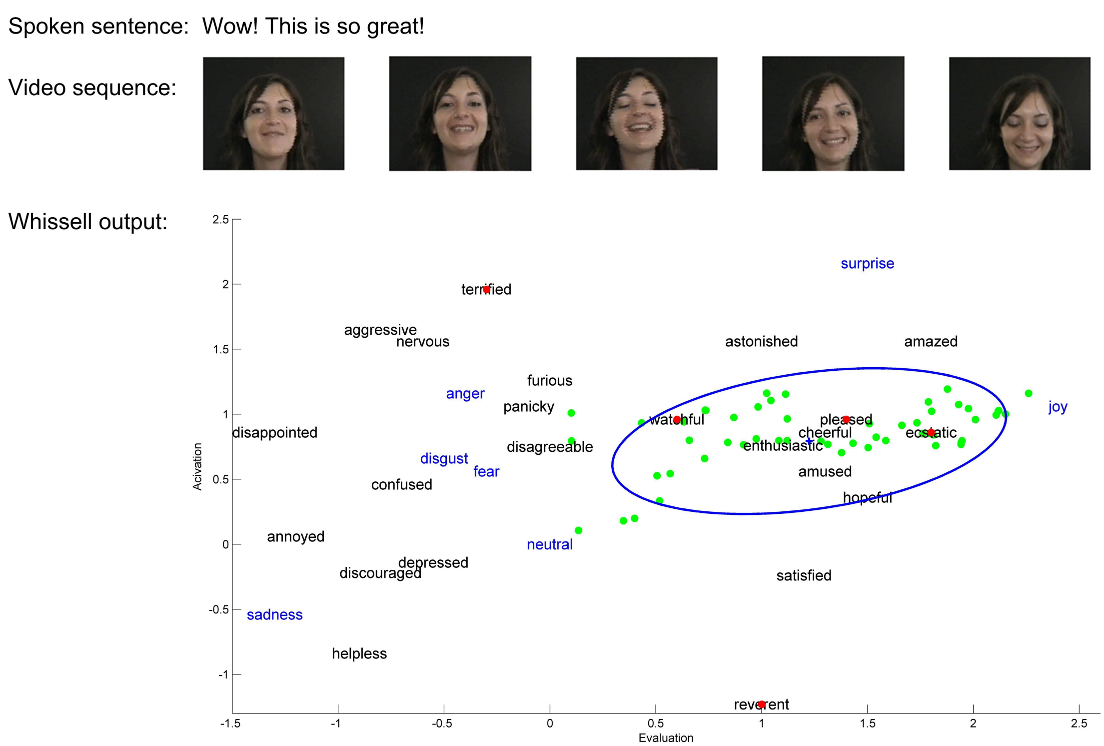


Figure 5.11: A screenshot of the Affective Analysis Module output. The integration of the extracted multi-modal emotional information takes place into the Whissell space, in terms of activation and evaluation.

| metric | disgust | joy | anger | fear | sadness | neutral | surprise |
|---|---|---|---|---|---|---|---|
| precision | 86.3% | 95.4% | 80.0% | 94.0% | 79.8% | 95.6% | 92.0% |
| recall | 81.7% | 99.1% | 95.1% | 85.4% | 91.1% | 72.8% | 96.2% |
| F-measure | 83.9% | 97.2% | 86.9% | 89.5% | 85.1% | 82.7% | 94.0% |

Table 5.12: Precision, recall, and F-measure rates associated with Ekman's basic emotions (plus neutral) obtained by concurrently employing the opinion mining engine and the Facial Expression Analyser.

In order to evaluate the performance of the Affective Analysis module, that is the F-measure rates of the opinion mining engine and the Facial Expression Analyser as an ensemble, an emotion recognition test on 50 selected videos from the HUMAINE database[16] was performed. As shown in Table 5.12, the concurrent exploitation of different modalities generally leads to an improvement in affect detection, as the two subsystem components make uncorrelated errors. When multiple components make uncorrelated errors, in fact, the probability that they all make the same error is the product of their individual error probabilities, resulting in much lower error rates.

### 5.3.2 Sentic Chat

Online communication is an extremely popular form of social interaction. Unlike face-to-face communication, online IM tools are extremely limited in conveying emotions or the context associated with a communication. Users have adapted to this environment by inventing their own vocabulary, e.g., by putting actions within asterisks ("I just came from a shower *shivering*"), by using emoticons, by addressing a particular user in a group communication ("@Ravi"). Such evolving workarounds clearly indicate a latent need for a richer, more immersive user experience in social communication. This problem is addressed by exploiting the semantics and sentics associated with the on-going communication to develop an adaptive user interface (UI) capable to change according to content and context of the online chat. Popular approaches to enhance and personalise computer-mediated communication (CMC) include emoticons, skins, avatars, customisable status messages, etc.

[16] http://humaine-db.sspnet.eu

However, all these approaches require explicit user configuration or action: the user needs to select the emoticon, status-message or avatar, which best represents her. Furthermore, most of these enhancements are static - once selected by the user, they do not adapt themselves automatically. There is some related work on automatically updating the status of the user by analysing various sensor data available on mobile devices [276]. However, most of these personalisation approaches are static and do not automatically adapt.

Sentic chat's approach is unique in that it is: intelligent, as it analyses content and does not require explicit user configuration; adaptive, as the UI changes according to communication content and context; inclusive, as the emotions of one or more participants in the chat session are analysed to let the UI adapt dynamically. The module architecture can be deployed either on the cloud (if the client has low processing capabilities) or on the client (if privacy is a concern). Most IM clients offer a very basic UI for text communication. In sentic chat, the focus is on extracting the semantics and sentics embedded in the text of the chat session to provide a UI that adapts itself to the mood of the communication.

For this prototype application, the weather metaphor was selected, as it is scalable and has previously been used effectively to reflect the subject's mood [277] or content's 'flavor' [278]. In the proposed UI, if the detected mood of the conversation is 'happy', the UI will reflect a clear sunny day. Similarly a gloomy weather reflects a melancholy tone in the conversation (Fig. 5.12). Of course, this is a subjective metaphor - one that supposedly scales well with conversation analysis. In the future, other relevant scalable metaphors could be explored, e.g., colours [195].

The adaptive UI primarily consists of three features: the stage, the actors, and the story. For any mapping these elements pay a crucial role in conveying the feel and richness of the conversation mood, e.g., in the 'happy' conversation the weather 'clear sunny day' will be the stage, the actors will be lush green valley, the rainbow, and the cloud, which may appear or disappear as per the current conversation tone of the story.

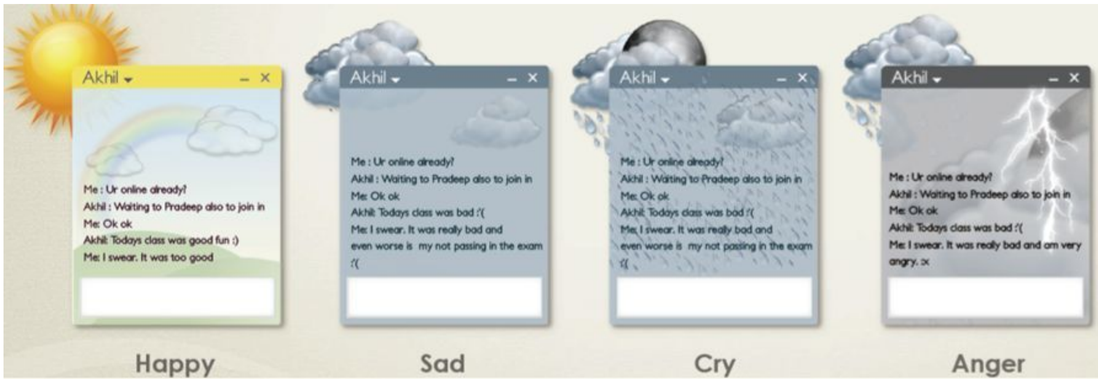


Figure 5.12: A few screenshots of sentic chat IUI. Stage and actors gradually change, according to the semantics and sentics associated with the on-going conversation, in order to provide a more immersive chat experience.

The idea is similar to a visual narrative of the mood the conversation is in; as the conversation goes on the actors may come in or go off as per the tone of the thread. By analysing the semantics and sentics associated with communication content (data) and context (metadata), the UI may adapt to include images of landmarks from remote-user's location (e.g., Times Square), images about concepts in the conversation (pets, education, etc.) or time of day of remote user (e.g., sunrise or dusk).

The effectiveness of Sentic Chat was assessed through a usability test on a group of 6 regular chat users, who were asked to chat to each other pairwise for approximately 10 minutes (for a total of 130 minutes of chat data) and to rate the consistency with the story of both stage and actor alternation during the CMC (Table 5.13).

| **Feature** | **Not consistent** | **Consistent** | **Very consistent** |
|---|---|---|---|
| stage change | 0% | 83.3% | 16.7% |
| actor alternation | 16.8% | 66.6% | 16.7% |

Table 5.13: Values associated with the perceived consistency with chat text of stage change and actor alternation. The evaluation was performed on a 130-minute chat session operated by a pool of 6 regular chat users.

### 5.3.3 Sentic Corner

In a world in which web users are continuously blasted by ads and often compelled to deal with user-unfriendly interfaces, we sometimes feel like we want to evade from the sensory overload of standard web-pages and take refuge in a safe web corner, in which contents and design are in harmony with our current frame of mind. Sentic Corner, proposed by Cambria et al. [279], is an intelligent user interface that dynamically collects audio, video, images and text related to the user's current feelings and activities as an interconnected knowledge base, which is browsable through a multi-faceted classification website. In the new realm of Web 2.0 applications, the analysis of emotions has undergone a large number of interpretations and visualisations, e.g., We Feel Fine[17] [280], MoodView[18], MoodStats[19], and MoodStream[20], which have often led to the development of emotion-sensitive systems and applications.

Nonetheless, today web users still have to almost continuously deal with sensory-overloaded web-pages, pop-up windows, annoying ads, user-unfriendly interfaces, etc. Moreover, even for websites uncontaminated by web spam, the affective content of the page is often totally unsynchronised with the user's emotional state. web-pages containing multimedia information inevitably carry more than just informative content. Behind every multimedia content, in fact, there is always an emotion.

Sentic Corner exploits this concept to build a sort of parallel cognitive/affective digital world in which the most relevant multimedia contents associated to the users' current moods and activities are collected, in order to enable them, whenever they want to evade from sensory-rich, overwrought and earnest web-pages, to take refuge in their own safe web corner. To our knowledge, there is still no published study on the task of automatically retrieving and displaying multimedia contents according to user's moods and activities, although the affective and semantic analysis of video, audio and textual contents have been separately investigated extensively [239, 281, 282].

[17] `http://wefeelfine.org`
[18] `http://moodviews.com`
[19] `http://moodstats.com`
[20] `http://moodstream.gettyimages.com`

The most relevant commercial tool within this area is Moodstream, a mashup of several forms of media, designed to bring users music, images, and video according to the mood they manually select on the web interface. Moodstream aims to create a sort of audio-visual ambient mix that can be dynamically modified by users by selecting from the presets of 'inspire', 'excite', 'refresh', 'intensify', 'stabilise', and 'simplify', e.g., mixtures of mood spectra on the Moodstream mixer such as happy/sad, calm/lively or warm/cool. Users can start with a preset and then mix things up including the type of image transition, whether they want more or less vocals in their music selection and how long images and video will stay, among other settings.

In Moodstream, however, songs are not played entirely but blended into one another every 30 seconds and, even if the user has control on the multimedia flow through the mood presets, he/she cannot actually set a specific mood and/or activity as a core theme for the audio-visual ambient mix. Sentic Corner, on the contrary, uses sentic computing to automatically extract semantics and sentics associated with user's status updates on micro-blogging websites and, hence, to retrieve relevant multimedia contents in harmony with his/her current emotions and motions.

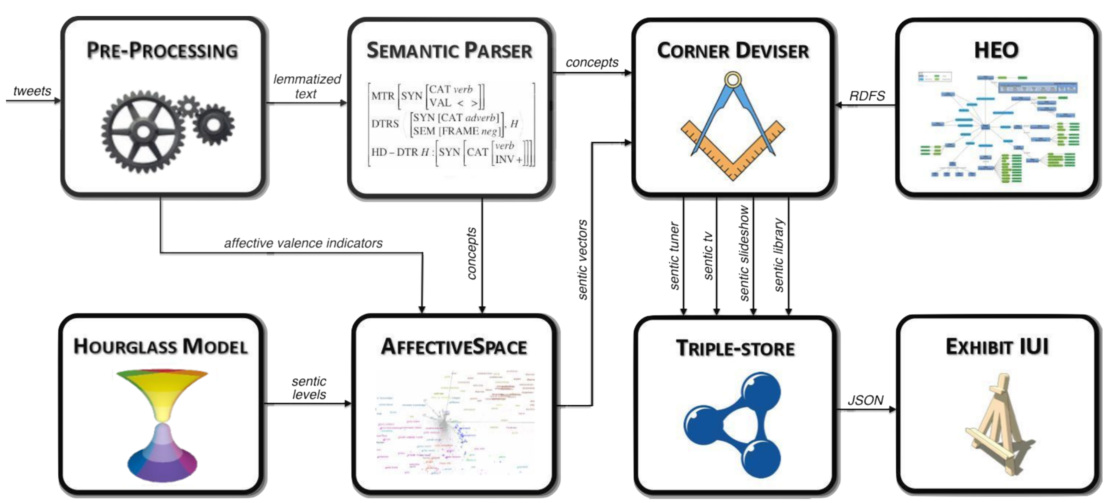


Figure 5.13: Sentic Corner generation process. The semantics and sentics extracted from the user's micro-blogging activity are exploited to retrieve relevant audio, video, visual, and textual information.

The module for the retrieval of semantically and affectively related music is termed Sentic Tuner. The relevant audio information is pulled from Stereomood[21], an emotional online radio that provides music that best suits users' mood and activities. In the web interface, music is played randomly through an online music player with the possibility for the user to play/stop/skip tracks. In Stereomood, music tracks are classified according to some tags that users are supposed to manually choose in order to access a list of semantically or affectively related songs. These tags are either mood-tags (e.g., 'happy', 'calm', 'romantic', 'lonely' and 'reflective') or activity-tags (such as 'reading', 'just woke up', 'dressing up', 'cleaning' and 'jogging'), the majority of which represent cognitive and affective knowledge contained in AffectiveSpace as common sense concepts and emotional labels.

The Sentic Tuner uses the mood-tags as centroids for blending and the activity-tags as seeds for spectral association, in order to build a set of affectively and semantically related concepts respectively, which will be used at run-time to match the concepts extracted from user's micro-blogging activity. The Sentic Tuner also contains a few hundreds *rāgas* (Sanskrit for moods), which are melodic modes used in Indian classical music meant to be played in particular situations (mood, time of the year, time of the day, weather conditions, etc.). It is considered inappropriate to play *rāgas* at the wrong time (it would be like playing Christmas music in July, lullabies at breakfast or sad songs at a wedding) so these are played just when semantics and sentics exactly match time and mood specifications in the *rāgas* database.

Hence, once semantics and sentics are extracted from natural language text through sentic computing, Stereomood API and the *rāgas* database are exploited to select the most relevant tracks to user's current feelings and activities. Sentic TV is the module for the retrieval of semantically and affectively related videos. In particular, the module pulls information from Jinni[22], a new site that allows users to search for video entertainment in many specific ways.

[21] `http://stereomood.com`
[22] `http://jinni.com`

The idea behind Jinni is to reflect how people really think and talk about what they watch. It is based on an ontology developed by film professionals and new titles are indexed with an innovative NLP technology for analysing metadata and reviews. In Jinni, users can choose from movies, TV shows, short films and online videos to find specific genres or what they are in the mood to watch. In particular, users can browse videos by topic, mood, plot, genre, time/period, place, audience and praise. Similarly to the Sentic Tuner, Sentic TV uses Jinni's mood-tags as centroids for blending and the topic-tags as seeds for spectral association in order to retrieve affectively and semantically related concepts respectively. Time-tags and location-tags are also exploited in case relevant time-stamp and/or geo-location information is available within user's micro-blogging activity.

Sentic Corner also offers semantically and affectively related images through the Sentic Slideshow module. Pictures related to the user's current mood and activity are pulled from Fotosearch[23], a provider of royalty free and rights managed stock photography which claims to be the biggest repository of images on the Web. Since Fotosearch does not offer a priori mood-tags and activity-tags, the CF-IOF technique is used on a set of 1000 manually tagged (according to mood and topic) tweets, in order to find seeds for spectral association (topic-tagged tweets) and centroids for blending (mood-tagged tweets). Each of the resulting concepts is used to retrieve mood and activity related images through the Fotosearch search engine. The royalty free pictures, eventually, are saved in an internal database according to their mood and/or activity tag, in a way that they can be quickly retrieved at run-time, depending on user's current feelings and thoughts.

The aim of Sentic Library is to provide book excerpts depending on user's current mood. The module proposes random book passages users should read according to the mood they should be in while reading it and/or what mood they will be in when they have finished. The excerpt database is built according to '1001 Books for Every Mood: A Bibliophile's Guide to Unwinding, Misbehaving, Forgiving, Celebrating, Commiser-

[23] `http://fotosearch.com`

ating' [283], a guide in which the novelist Hallie Ephron serves up a literary feast for every emotional appetite. In the guide, books are labelled with mood-tags such as 'for a good laugh', 'for a good cry' and 'for romance', but also some activity-tags such as 'for a walk on the wild side' or 'to run away from home'.

As for Sentic TV and Sentic Tuner, Sentic Library uses these mood-tags as centroids for blending and the topic-tags as seeds for spectral association. The Corner Deviser exploits the semantic and sentic knowledge bases previously built by means of blending, CF-IOF and spectral association to find matches for the concepts extracted by the semantic parser and their relative affective information inferred by AffectiveSpace. Such audio, video, visual, and textual information (namely Sentic Tuner, Sentic TV, Sentic Slideshow, and Sentic Library) is then encoded in RDF/XML according to HEO and stored in the triple-store (Fig. 5.13).

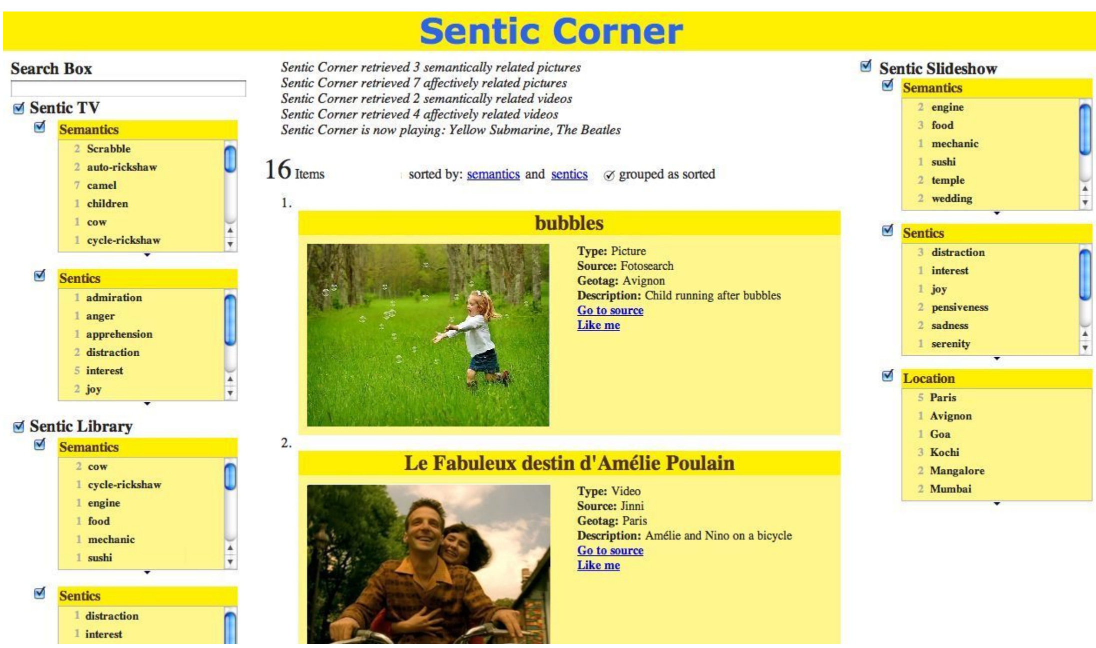


Figure 5.14: Sentic Corner web interface. The multi-modal information obtained by means of Sentic Tuner, Sentic TV, Sentic Slideshow, and Sentic Library is encoded in RDF/XML for multi-faceted browsing.

| Content | Not at all | Just a little | Somewhat | Quite a lot | Very much |
|---|---|---|---|---|---|
| audio | 0% | 11.1% | 11.1% | 44.5% | 33.3% |
| video | 11.1% | 11.1% | 44.5% | 33.3% | 0% |
| visual | 0% | 0% | 22.2% | 33.3% | 44.5% |
| textual | 22.2% | 11.1% | 55.6% | 11.1% | 0% |

Table 5.14: Relevance of audio, video, visual, and textual information assembled by Sentic Corner over 80 tweets. Possibly because of their larger datasets, Sentic Tuner and Sentic Slideshow are the best-performing modules.

In case the sentics detected belong to the lower part of the Hourglass, the multimedia contents searched will have an affective valence opposite to the emotional charge detected, as Sentic Corner aims to restore the positive emotional equilibrium of the user, e.g., if the user is angry he/she might want to calm down. The Exhibit IUI module, eventually, visualises the contents of the Sesame database exploiting the multi-faceted categorisation paradigm (Fig. 5.14).

In order to test the relevance of multimedia content retrieval, an evaluation based on the judgements of 8 regular Twitter users was performed. Specifically, users had to link Sentic Corner to their Twitter accounts and evaluate, over 10 different tweets, how the IUI would react to their status change in terms of relevance of audio, video, visual, and textual information assembled by Sentic Corner. The multimedia contents retrieved turned out to be pretty relevant in most cases, especially for tweets concerning concrete entities and actions (Table 5.14).

## 5.4 Development of E-Health Systems

In health-care, it has long been recognised that, although the health professional is the expert in diagnosing, offering help and giving support in managing a clinical condition, the patient is the expert in living with that condition. Health-care providers need to be validated by somebody outside the medical departments but, at the same time, inside the health-care system. The best candidate for this is not the doctor, the nurse or the therapist but the real end-user of health-care – none other than the patient him/herself. Patient 2.0 is central to understanding the effectiveness and efficiency of services and how they can be improved. The patient is not just a consumer of the health-care system but a quality control manager – his/her opinions are not just reviews of a product/service but more like small donations of experience, digital gifts which, once given, can be shared, copied, moved around the world and directed to just the right people who can use them to improve health-care locally, regionally or nationally. Web 2.0 dropped the cost of voice, of finding others 'like me', of forming groups, of obtaining and republishing information, to zero. As a result, it becomes easy and rewarding for patients and carers to share their personal experiences with the health-care system and to research conditions and treatments.

To bridge the gap between this social information and the structured information supplied by health-care providers, the above-described engine is exploited to extract the semantics and sentics associated with patient opinions over the Web, and hence provide the real end-users of the health system with a common framework to compare, validate and select their health-care providers (section 5.4.1). This section, moreover, shows how the engine can be used as an embedded tool for improving patient reported outcome measures (PROMs) for health related quality of life (HRQoL), that is to record the level of each patient's physical and mental symptoms, limitations and dependency (section 5.4.2).

### 5.4.1 Crowd Validation

As Web 2.0 dramatically reduced the cost of reaching others, forming groups, obtaining and republishing information, today it is easy and rewarding for patients and carers to share their personal experiences with the health-care system. This social information, however, is often stored in natural language text and hence intrinsically unstructured, which makes comparison with the structured information supplied by health-care providers very difficult. To bridge the gap between these data, which though different at structure-level are similar at concept-level, a patient opinion mining tool has been proposed by Cambria et al. [207] to provide the end-users of the health system with a common framework to compare, validate and select their health-care providers. In order to give structure to online patient opinions, both the semantics and sentics associated with these are extracted in a way that it is possible to map them to the fixed structure of health-care data. This kind of data, in fact, usually consists of ratings that associate a polarity value to specific features of health-care providers such as communication, food, parking, service, staff and timeliness.

The polarity can be either a number in a fixed range or simply a flag (positive/negative). In the proposed approach, structure is added to unstructured data by building semantics and sentics on top of it (Fig. 5.15). In particular, given a textual resource containing a set of opinions $O$ about a set of topics $T$ with different polarity $p \in [-1, 1]$, the subset of opinions $o \subseteq O$ is extracted, for each $t \in T$, and $p$ is determined for each $o$. In other words, since each opinion can regard more than one topic and the polarity values associated with each topic are often independent from each other, in order to perform the mapping a set of topics needs to be extracted from each opinion and then, for each topic detected, the polarity associated with it is inferred. Once natural language data are converted to a structured format, each topic expressed in each patient opinion and its related polarity can be aggregated and compared. These can then be easily assimilated with structured health-care information contained in a database or available through an API.

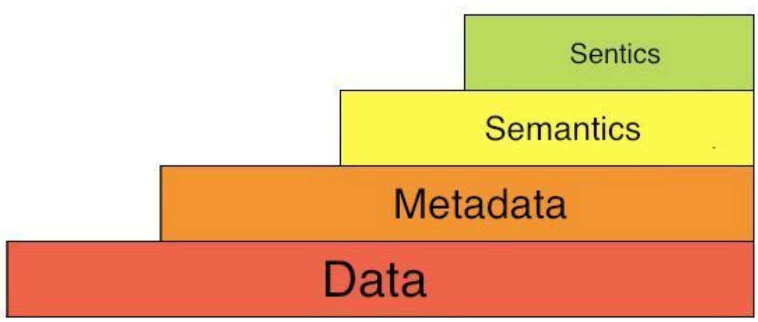


Figure 5.15: The semantics and sentics stack. Semantics are built on the top of data and metadata. Sentics, in turn, are built on the top of semantics, as they the represent the affective information associated with these.

This process, proposed by Cambria et al. [284], is termed 'crowd validation' (Fig. 5.16), because of the feedback coming from the masses, and it fosters next-generation health-care systems, in which patient opinions are crucial in understanding the effectiveness and efficiency of health services and how they can be improved.

Within this work, in particular, the opinion analysis process is used to marshal PatientOpinion's social information in a machine-accessible and machine-processable format and, hence, compare it with the official hospital ratings provided by NHS Choices[24] and each NHS trust. The inferred ratings are used to validate the information declared by the relevant health-care providers, crawled separately from each NHS trust website, and the official NHS ranks, extracted using NHS Choices API.

At the present time, crowd validation cannot be directly tested because of the impossibility to objectively assess the truthfulness of both patient opinions and official NHS ratings.

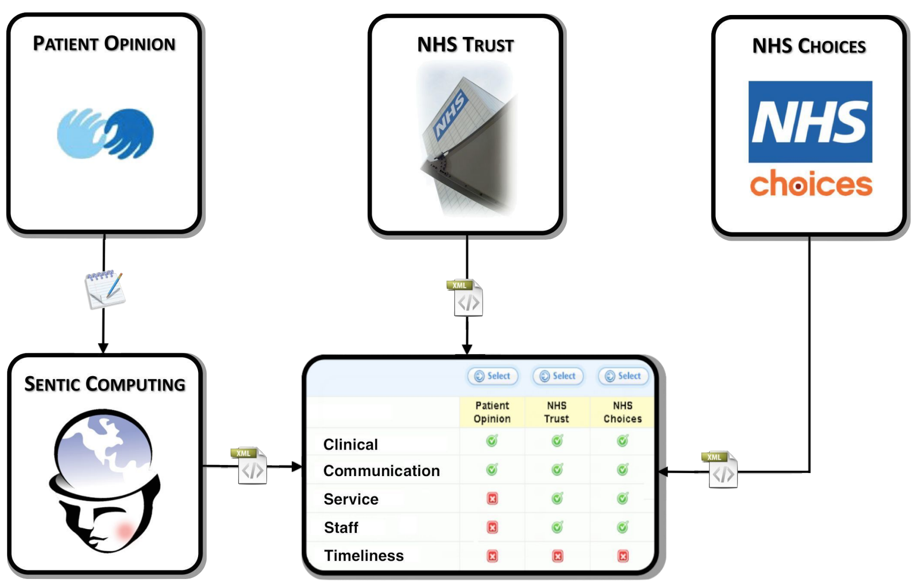


Figure 5.16: The crowd validation schema. PatientOpinion's stories are encoded in a machine-accessible format, in a way that they can be compared with the official ratings provided by NHS Choices and each NHS trust.

[24]http://www.nhs.uk

Given that the patient opinion mining performance of the system has already been tested (Table 5.5), however, an experimental investigation has been performed over a set of 200 patient opinions about three different NHS trusts, for which self-assessed ratings were crawled from each hospital website and official NHS ranks were obtained through NHS Choices API. Results showed an average discrepancy of 39% between official and unofficial ratings, which sounds plausible as, according to Panorama[25], 60% of hospitals inspected in 2010 gave inaccurate information to the government in assessing their own performance.

### 5.4.2 Sentic PROMs

Public health measures such as better nutrition, greater access to medical care, improved sanitation and more widespread immunisation, have produced a rapid decline in death rates in all age groups. Since there is no corresponding decline in birth rates, however, the average age of population is increasing exponentially. If we want health services to keep up with such monotonic growth, we need to automatise as much as possible the way patients access the health-care system, in order to improve both its service quality and timeliness.

Everything we do that does not provide benefit to patients or their families, in fact, is waste. To this end, a new generation of short and easy-to-use tools to monitor patient outcomes and experience on a regular basis has been recently proposed by Benson et al. [285]. Such tools are quick, effective and easy to understand, as they are very structured. However, they leave no space to those patients who would like to say something more.

Patients, in fact, are usually keen on expressing their opinions and feelings in free text, especially if driven by particularly positive or negative emotions. They are often happy to share their health-care experiences for different reasons, e.g., because they seek for a sense of togetherness in adversity, because they benefited from others' opinions and want to give back to the community, for cathartic complaining, for supporting a

[25] `www.bbc.co.uk/programmes/b00rfqfm`

service they really like, because it is a way to express themselves, because they think their opinions are important for others. When people have a strong feeling about a specific service they tried, they feel like expressing it. If they loved it, they want others to enjoy it. If they hated it, they want to warn others away.

Standard PROMs allow patients to easily and efficiently measure their HRQoL but, at the same time, they limit patients' capability and will to express their opinions about particular aspects of the health-care service that could be improved or important facets of their current health status. *Sentic PROMs*, in turn, exploit the ensemble application of standard PROMs and sentic computing to allow patients to evaluate their health status and experience in a semi-structured way, i.e., both through a fixed questionnaire and through free text. Next-generation patients are central to understanding the effectiveness and efficiency of services and how they can be improved. PROMs provide a means of gaining an insight into the way patients perceive their health and the impact that treatments or adjustments to lifestyle have on their quality of life.

Pioneered by Donabedian [286], health status research began during the late 1960s with works focusing on health-care evaluation and resource allocation. In particular, early works mainly aimed to valuate health states for policy and economic evaluation of health-care programmes, but devoted little attention to the practicalities of data collection [287, 288, 289]. Later works, in turn, aimed to develop lengthy health profiles to be completed by patients, leading to the term patient reported outcome [290, 291]. PROMs can provide a new category of real-time health information, which enables every level of the health service to focus on continuously improving those things that really matter to patients.

The benefits of routine measurement of HRQoL include helping to screen for problems, promoting patient-centric care, aiding patients and doctors to take decisions, improving communication amongst multi-disciplinary teams and monitoring progress of individual or groups of patients and the quality of care in a population. However, in spite of demonstrated benefits, routine HRQoL assessment in day-to-day practice remains rare as few patients are willing to spend the time needed to daily fill-in questionnaires, such as *SF-36* [292], *SF-12* [293], *Euroqol EQ-5D* [294] or the *Health Utilities Index* [295]. To overcome this problem, *howRu*, a new generic PROM was recently proposed by [285] for recording the level of each patient's physical and mental symptoms, limitations and dependency on four simple levels. The questionnaire was designed to take no more than a few seconds using electronic data collection and integration with electronic patient records as part of other routine tasks that patients have to do, such as booking appointments, checking in on arrival at clinic, or ordering or collecting repeat medication. The main aim of *howRu* is to use simple terms and descriptions, in order to reduce the risk of ambiguity and to ensure that as many people as possible could use the measure reliably and consistently without training or support.

The same approach has been employed to monitor also patient experience (*howRwe*) and staff satisfaction (*howRus*) on a regular basis. These questionnaires have been proved to be quick, effective and easy to understand, as they are short, rigid and structured. However, such structuredness can be very limiting, as it leaves no space to those patients who would like to say something more about their health or the service they are receiving. Patients, especially when driven by particularly positive or negative emotions, do want to express their opinions and feelings. *Sentic PROMs* allow patients to assess their health status and health-care experience in a semi-structured way by enriching the functionalities of the new PROM tools with the possibility of adding free text (Fig. 5.17). This way, when patients are happy with simply filling-in the questionnaire, they can just leave the input text box blank but, when they feel like speaking out their opinions and feelings, e.g., in the occasion of a particularly positive or negative situation or event, they can now do it in their own words. Hence, *Sentic PROMs* input data, although very similar at concept level, are on two completely different structure levels – structured (questionnaire selection) and unstructured (natural language). As we would like to extract meaningful information from such data, the final aim of *Sentic PROMs* is to format the unstructured input and accordingly aggregate it with the structured data, in order to perform statistical analysis and pattern recognition. In particular, the gap between unstructured and structured data is bridged by means of sentic computing. Among the benefits of questionnaires' structuredness, there are the quickness, effectiveness and ease to use and understand. However, such structuredness involves some drawbacks. A questionnaire, in fact, can limit the possibility to discover new important patterns in the input data and can constrain users to omit important opinions that might be valuable for measuring service quality. In the medical sphere, in particular, patients driven by very positive or very negative emotions are usually willing to detailedly express their point of view, which can be particularly valuable for assessing uncovered points, raising latent problems or redesigning the questionnaire.

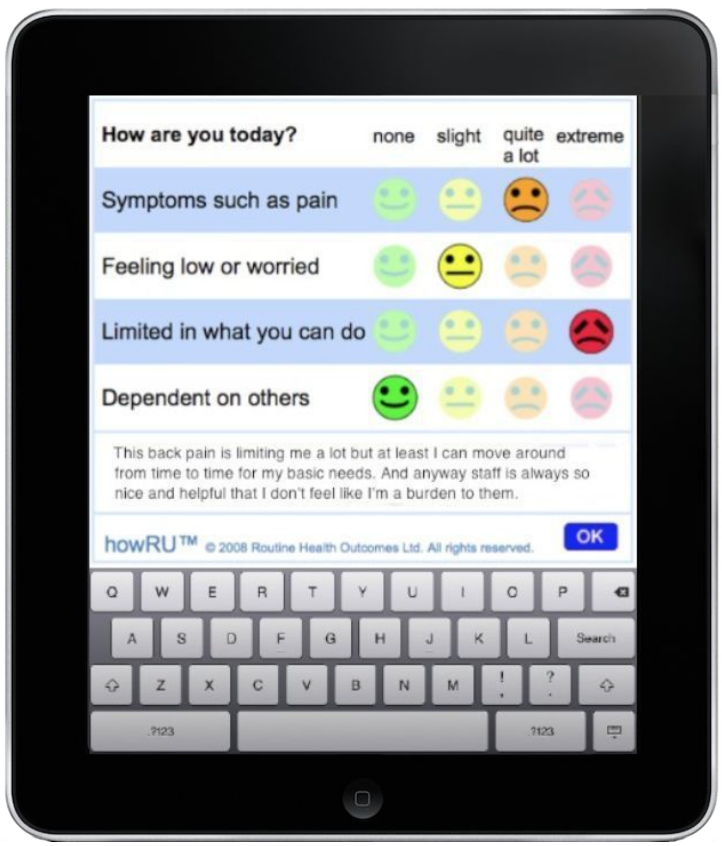


Figure 5.17: *Sentic PROMs* prototype on iPad. The new interface allows patients to assess their health status and health-care experience both in a structured (questionnaire) and unstructured (free text) way.

To this end, *Sentic PROMs* adopt a semi-structured approach that allows patients to assess their health status and health-care experience both by filling in a four-level questionnaire and by adding free text. The two different input methods are not mutually exclusive but complementary. When patients are happy with simply filling-in the questionnaire, they can just leave the input text box blank but, when they feel like speaking out their opinions and feelings, e.g., in the occasion of a particularly positive or negative situation or event, they can do it in their own words.

As a result, the stored input data, although very similar at concept level, are on two completely different structure levels – structured (questionnaire selection) and unstructured (natural language). *Sentic PROMs* aim to format the unstructured input and accordingly aggregate it with the structured data in order to perform statistical analysis and pattern recognition on these and, hence, extract meaningful information. In order to bridge such gap between unstructured and structured patient data, the semantics and sentics associated with natural language text are extracted by means of sentic computing. In particular, semantics are built on the top of patient data and metadata, while sentics are built on the top of semantics, as they the represent the emotions or the polarity conveyed by the detected concepts.

The importance of physio-emotional sensitivity in humans has been proven by recent health research, which has shown that individuals who feel loved and supported by friends and family, or even by a loving pet, tend to have higher survival rates following heart attacks than other cardiac patients who experience a sense of social isolation. Such concept is also reflected in natural language as we use terms such as 'heartsick', 'broken-hearted' and 'heartache' to describe extreme sadness and grief, idioms like 'full of gall' and 'venting your spleen' to describe anger, and expressions such as 'gutless', 'yellow belly' and 'feeling kicked in the gut' to describe shame. Human body contracts involuntarily when it feels emotional pain such as grief, fear, disapproval, shock, helplessness, shame, terror, in the same reflex it does if physically injured. Such gripping reflex normally releases slowly, but if a painful experience is intense, or happens repeatedly, the physio-emotional grip does not release, and constriction is retained in the body. Any repeated similar experience then layers on top of the original unreleased contraction, until we are living with layers of chronic tension, which constricts our bodies.

The mind, in fact, may forget the origin of pain and tension, but the body does not. Besides HRQoL measurement, *Sentic PROMs* aim to monitor also users' physio-emotional sensitivity on a regular basis, as a means of patient affective modelling. In particular, the dimensional affective information coming from both questionnaire data (*howRu* aggregated score) and natural language data (sentic vectors) is stored separately by the system every time patients conclude a *Sentic PROMs* session and plotted on four different bi-dimensional diagrams. Such diagrams represent the pairwise fusion of the four dimensions of the Hourglass model and allow to detect more complex (compound) emotions that can be particularly relevant for monitoring patients' physio-emotional sensitivity, e.g., frustration, anxiety, optimism, disapproval, and rejection.

A preliminary validation study was undertaken to examine the psychometric properties and construct validity of *Sentic PROMs* and to compare these with *SF-12*. In particular, 2,751 subjects with long-term conditions (average age 62, female 62.8%), were classified by *howRu* score, primary condition, number of conditions suffered, age group, duration of illness and area of residence. Across all six classifications, the correlation of the mean *howRu* scores with the mean values of the Physical Components Summary (*PCS-12*), the Mental Components Summary (*MCS-12*) and the sum of *PCS-12 + MCS-12* were generally very high (0.91, 0.45 and 0.97 respectively).

## 5.5 Conclusions

This chapter has shown how the developed common sense knowledge bases, and the reasoning tools built on the top of them, can be exploited for the design of a novel opinion mining engine able to infer the cognitive and affective information associated with natural language text (section 5.1).

In order to assess the capability of such engine to tackle real-world NLP tasks, the process for the extraction of semantics and sentics has been embedded in multiple systems for the development of intelligent applications in different domains. Specifically, the tools and techniques developed within this thesis work have been exploited for the design of Social Web applications, i.e., a troll filtering system, a social media marketing tool, and an online personal photo management system (section 5.2); HCI applications, that is, an embodied conversational agent with affective common sense, an adaptive IM tool, and an emotion-sensitive IUI (section 5.3); and e-health applications, i.e., a framework for assessing the quality of health-care providers and a tool for enhancing PROMs for HRQoL (section 5.4).

All these developed applications demonstrate how the opinion mining engine can be employed in nearly any domain, as it does not rely on domain-dependent keywords, but rather on common sense knowledge bases that allow it to extrapolate the cognitive and affective information associated with natural language text. The tools and techniques employed for the development of such engine are summarised in the following chapter, where a discussion about limitations and future developments of these is also offered.

# Chapter 6

# Concluding Remarks

*It is difficult to be rigorous about whether a machine really knows and thinks,*
*because we are hard put to define these things.*
*We understand human mental processes only slightly better than*
*a fish understands swimming.*

*John McCarthy*

This thesis was the result of an industrial research project born from the collaboration between the University of Stirling, the MIT Media Laboratory and Sitekit Labs. The main aim of the thesis work was to go beyond keyword-based approaches by further developing and applying common sense computing techniques to bridge the cognitive and affective gap between word-level natural language data and the concept-level opinions conveyed by these. This has been pursued through a variety of novel tools and techniques that have been tied together to develop an opinion mining engine for the semantic analysis of natural language opinions and sentiments. Such engine has then been used for the development of intelligent web applications in fields such as Social Web, HCI, and e-health.

This chapter contains a summary of the contributions the thesis work has introduced (section 6.1), a discussion about limitations and future developments of these (section 6.2), and conclusions (section 6.3).

## 6.1 Summary of Contributions

Despite significant progress, opinion mining and sentiment analysis are still finding their own voice as new inter-disciplinary fields. Engineers and computer scientists use machine learning techniques for automatic affect classification from video, voice, text, and physiology. Psychologists use their long tradition of emotion research with their own discourse, models, and methods. This thesis work has assumed that opinion mining and sentiment analysis are research fields inextricably bound to the affective sciences that attempt to understand human emotions. Simply put, the development of affect-sensitive systems cannot be divorced from the century-long psychological research on emotion. The emphasis on the multi-disciplinary landscape that is typical for emotion-sensitive applications and the need for common sense sets this work apart from previous research on opinion mining and sentiment analysis.

In this thesis, a novel approach to opinion mining and sentiment analysis has been developed by exploiting both AI and Semantic Web techniques. In particular, two common sense knowledge bases have been developed and an ensemble application of graph mining and multi-dimensionality reduction techniques has been employed to perform reasoning on them. Such data and techniques have then been exploited for designing a new opinion mining engine able to infer cognitive and affective information from natural language and, hence, for the development of intelligent web applications in fields such as Social Web, HCI, and e-health. This section presents a summary of the techniques (subsection 6.1.1), tools (subsection 6.1.2) and applications (subsection 6.1.3) developed within this thesis work.

### 6.1.1 Techniques

In this thesis, a variety of techniques have been developed for the extraction of semantics and sentics from natural language text, namely:

1. **Affective Blending**: the process for building the matrix representation of AffectNet from ConceptNet and WordNet-Affect;

2. **Sentic Medoids**: a novel clustering technique for organising affective common sense concepts in AffectiveSpace;

3. **Sentic Activation**: a bio-inspired two-level framework that exploits an ensemble application of dimensionality reduction and graph mining techniques;

4. **CF-IOF Weighting**: a technique similar to TF-IDF weighting that evaluates how important a concept is to a set of opinions concerning the same topic;

### 6.1.2 Tools

The above-mentioned techniques have been employed within this research work for the design of a set of tools for the automatic analysis of opinions and sentiments, videlicet:

1. **AffectiveSpace**: a vector space representation of AffectNet for reasoning by analogy on affective common sense knowledge;

2. **The Hourglass of Emotions**: a biologically-inspired and psychologically-motivated model for the representation and the analysis of human emotions;

3. **Open Mind Common Sentics**: an emotion-sensitive IUI for collecting affective common sense knowledge from general public;

4. **SenticNet**: a publicly available semantic resource for opinion mining built using both AI and Semantic Web techniques;

5. **Isanette**: a semantic network of common and common sense knowledge for auto-categorisation built upon ConceptNet and Probase;

6. **Opinion Mining Engine**: an intelligent engine for concept-level open-domain opinion mining and sentiment analysis.

### 6.1.3 Applications

Eventually, the above-mentioned opinion mining tool has been exploited for the development of emotion-sensitive applications in fields such as Social Web, HCI, and e-health, namely:

1. **Troll Filter**: a system for automatically filtering inflammatory and outrageous posts within online communities;

2. **Social Media Marketing Tool**: an intelligent web application for managing social media information about products and services through a faceted interface;

3. **Sentic Album**: a content, concept and context based online personal photo management system;

4. **Sentic Chat**: an IM platform that enriches social communication through semantics and sentics;

5. **Sentic Corner**: an IUI that dynamically collects audio, video, images and text related to the user's emotions and motions;

6. **Sentic Avatar**: an emotion-sensitive avatar built through the integration of sentic computing with a facial emotional classifier;

7. **Crowd Validation**: a process for mining patient opinions and bridging the gap between unstructured and structured health-care data;

8. **Sentic PROMs**: a new framework for measuring health care quality that exploits the ensemble application of standard PROMs and sentic computing.

## 6.2 Limitations and Future Work

The research work carried out in the past three years has put solid bases for the development of a variety of emotion-sensitive systems and applications in the fields of opinion mining and sentiment analysis. One of the main contributions of this thesis has also been the introduction of a novel approach to the analysis of opinions and sentiments, which goes beyond merely keyword-based methods by using common sense reasoning tools and affective ontologies.

The developed techniques, however, are still far from perfect as the common and common sense knowledge base need to be further expanded and the reasoning tools built on the top of them adjusted accordingly. This last section discusses the limitations of such techniques (subsection 6.2.1) and includes their further developments both in the short-term (subsection 6.2.2) and in the long-term (subsection 6.2.3).

### 6.2.1 Limitations

As discussed in subsection 5.1.6, the validity of the proposed approach mainly depends on the richness of the developed knowledge bases. Without a comprehensive resource that encompasses human knowledge, in fact, it is not easy for an opinion mining system to get a hold of the ability to grasp the cognitive and affective information associated with natural language text and, hence, accordingly aggregate opinions in order to make statistics on them. Attempts to encode human common knowledge are countless and comprehend both resources generated by human experts (or community efforts) and automatically-built knowledge bases. The former kinds of resources are generally too limited, as they need to be hand-crafted, the latter too noisy, as they mainly rely on information available on the Web.

The span and the accuracy of knowledge available, however, is not the only limitation of opinion mining systems. Even though a machine "knows 50 million such things"[1], it needs to be able to accordingly exploit such knowledge through different

[1] http://mitworld.mit.edu/video/484

types of associations, e.g., inferential, causal, analogical, deductive, or inductive. For the purposes of this thesis work, singular value decomposition (SVD) appeared to be a good method for generalising the information contained in the common sense knowledge bases but it is very expensive in both computing time and storage, as it requires costly arithmetic operations such as division and square root in the computation of rotation parameters. This is a big issue because both AffectNet and Isanette are keeping on growing, in parallel with the continuously extended versions of ConceptNet and Probase. Moreover, the eigenmoods of AffectiveSpace cannot be easily understood because they are linear combinations of all of the original concept features. Different strategies that clearly show various steps of reasoning might be preferable in the future.

Another limitation of the sentic computing approach is in its typicality. The clearly defined knowledge representation of AffectNet and Isanette, in fact, does not allow to grasp different concept nuances as the inference of semantic and affective features associated with concepts is bounded. New features associated to a concept can be indeed inferred through the AffectiveSpace process but the number of new features that can be discovered after reconstructing the concept-feature matrix is limited to the set of features associated with semantically related concepts (that is, concepts that share similar features). However, depending on the context, concepts might need to be associated with features that are not strictly pertinent to germane concepts. The concept of book, for example, is typically associated to concepts such as newspaper or magazine, as it contains knowledge, has pages, etc. In a different context, however, a book could be used as paperweight, doorstop, or even as a weapon.

Biased (context-dependent) association of concepts is possible through spectral association, in which spreading activation is concurrently determined by different nodes in the graph representation of Isanette. Because concepts are hereby considered atomic and mono-faceted, however, it is not easy for the system to grasp the many different ways a concept can be meaningful in a particular context, as the features associated with each concept identify just its typical qualities, traits, or characteristics.

Finally, another limitation of the proposed approach is in the lack of time representation. Such an issue is not addressed by any of the currently available knowledge bases, including ConceptNet and Probase, upon which AffectNet and Isanette are built. In the context of sentic computing, however, time representation is not specifically needed as the main aim of the opinion mining engine is the passage from unstructured natural language data to structured machine-processable information, rather than genuine natural language understanding.

Every SBoC, in fact, is treated as independent from other SBoCs in the text data, as the goal is to simply infer a topic and a polarity associated with it, rather than understanding the whole meaning of the sentence in correlation with adjacent ones. In some cases, however, time representation might be needed for tasks such as comparative opinion analysis and co-reference resolution.

### 6.2.2 On-Going Work

In order to overcome some of the above-mentioned limitations, the current research work is focusing on expanding AffectNet and Isanette with different kinds of knowledge (e.g., common sense, affective knowledge, common knowledge) coming from external resources, e.g., Cyc, Freebase and Yago. Such operation is not simply convenient for improving the accuracy of the opinion mining engine, but also for reducing the sparseness of the matrix representations of such knowledge bases and, hence, aid dimensionality reduction procedures. In order to overcome the current weaknesses of the semantic parser, moreover, a new parsing tool is being developed.

The parser is based on a construction grammar approach, that is, a constraint-based, generative, and mono-stratal grammatical model, committed to incorporating the cognitive and interactional foundations of language. Construction grammar is also inherently tied to a particular model of the 'semantics of understanding', known as frame semantics, which offers a way of structuring and representing meaning while taking into account the relationship between lexical meaning and grammatical patterning.

New graph-mining and multi-dimensionality reduction techniques are also being explored to perform reasoning on the common sense knowledge bases. In particular, it is being investigated how AffectiveSpace can be built by means of independent component analysis (ICA) and random projections. Moreover, new classification techniques, such as support and relevance vector machines, are being experimented, together with the ensemble application of dimensionality reduction and extreme learning machine (ELM) techniques for emulating fast affective learning and reasoning.

### 6.2.3 Future Work

The developed opinion mining engine is currently being used within Stirling Department of Computing Science and Mathematics for many different projects, e.g., a research project, jointly investigated by Amir Hussain and Saliha Minhas, in collaboration with Khaled Hussainey, from the Department of Accounting and Finance, for automatically processing the information content of operating and financial review (OFR) statements; and a project funded by the UK Engineering and Physical Sciences Research Council, jointly investigated by Amir Hussain and Kamran Farooq, in collaboration with Calum MacRae, from Harvard Medical School, and John Moore, from the MIT Media Laboratory, for the development of next-generation health-care decision support system in the cardiovascular domain.

Other projects include the on-going collaborations with Praphul Chandra, from HP Labs India, in the field of multimedia management, Yangqiu Song and Haixun Wang, from Microsoft Research Asia, in the field of knowledge representation and reasoning, and especially with Cheng-Lin Liu, Chengqing Zong and Qiu-Feng Wang, from the National Laboratory of Pattern Recognition (NLPR) in the Institute of Automation of the Chinese Academy of Sciences, in the field of document analysis and pattern recognition, within the on-going China-Scotland SIPRA programme funded by the Royal Society of Edinburgh (RSE) and the National Science Foundation of China (NSFC).

## 6.3 Conclusions

The textual information available on the Web can be broadly grouped into two main categories: facts and opinions. Facts are objective expressions about entities or events. Opinions are usually subjective expressions that describe people's sentiments, appraisals or feelings towards such entities and events. Much of the existing research on textual information processing has been focused on mining and retrieval of factual information, e.g., information retrieval, Web search, text classification, text clustering and many other text mining and NLP tasks. Little work had been done on the processing of opinions until only recently [2].

One of the main reasons for the lack of study on opinions is the fact that there was little opinionated text available before the recent passage from a read-only to a read-write Web. Before that, in fact, when people needed to make a decision, they typically asked for opinions from friends and family. Similarly, when organisations wanted to find the opinions or sentiments of the general public about their products and services, they had to specifically ask people by conducting opinion polls and surveys. However, with the advent of the Social Web, the way people express their views and opinions has dramatically changed. They can now post reviews of products at merchant sites and express their views on almost anything in Internet forums, discussion groups, and blogs. Such online word-of-mouth behaviour represents new and measurable sources of information with many practical applications.

However, finding opinion sources and monitoring them can be a formidable task because there are a large number of diverse sources, and each source may also have a huge volume of opinionated text. In many cases, in fact, opinions are hidden in long forum posts and blogs. It is extremely time consuming for a human reader to find relevant sources, extract related sentences with opinions, read them, summarise them, and organise them into usable forms. Thus, automated opinion discovery and summarisation systems are needed. Sentiment analysis, also known as opinion mining, grows out of this need. It is a challenging NLP or text mining problem.

Due to its tremendous value for practical applications, there has been an explosive growth of both research in academia and applications in the industry. Due to many challenging research problems and a wide variety of practical applications, opinion mining has been a very active research area in recent years. All the sentiment analysis tasks, however, are very challenging. Our understanding and knowledge of the problem and its solution are still limited. The main reason is that it is a NLP task, and NLP has no easy problems. Another reason may be due to our popular ways of doing research. So far, in fact, researchers have probably relied too much on machine learning algorithms. Some of the most effective machine learning algorithms, e.g., SVM and CRF, in fact, produce no human understandable results such that, although they may achieve improved accuracy, little about how and why is known, apart from some superficial knowledge gained in the manual feature engineering process. All such approaches, moreover, rely on syntactical structure of text, which is far from the way human mind processes natural language.

In this thesis, common sense computing techniques were further developed and applied to bridge the semantic gap between word-level natural language data and the concept-level opinions conveyed by these. In particular, the ensemble application of graph mining and multi-dimensionality reduction techniques was exploited on two common sense knowledge bases to develop a novel intelligent engine for open-domain opinion mining and sentiment analysis. The proposed approach, termed sentic computing, performs a clause-level semantic analysis of text, which allows the inference of both the conceptual and emotional information associated with natural language opinions and, hence, a more efficient passage from (unstructured) textual information to (structured) machine-processable data.

Blending scientific theories of emotion with the practical engineering goals of analysing sentiments in natural language text and developing affect-sensitive interfaces is one of the main contributions of this thesis. Differently from most currently available opinion mining services, in fact, the developed engine does not base its analysis on a limited set of affect words and their co-occurrence frequencies, but rather on common sense concepts and the cognitive and affective valence conveyed by these. This allows the engine to be domain-independent and, hence, to be embedded in any opinion mining system for the development of intelligent applications in multiple fields such as Social Web, HCI and e-health.

Looking ahead, the combined novel use of different knowledge bases and of affective common sense reasoning techniques for opinion mining proposed in this work, will, eventually, pave the way for development of more bio-inspired approaches to the design of intelligent systems capable of handling knowledge, making analogies, learning from experience, perceiving and expressing affect. The question, in fact, is not whether intelligent machines can have emotions, but whether machines can be intelligent without any emotions [112].

# Bibliography

[1] J. Trout and S. Rivkin, *The New Positioning: The Latest on the World's #1 Business Strategy.* New York: McGraw Hill, 1996.

[2] B. Liu, "Sentiment analysis and subjectivity," in *Handbook of Natural Language Processing.* Chapman and Hall/CRC, 2010.

[3] C. Manning, P. Raghavan, and H. Schütze, *An Introduction to Information Retrieval.* Cambridge: Cambridge University Press, 2009.

[4] K. Khan, B. Baharudin, A. Khan, and F. Malik, "Mining opinion from text documents: A survey," in *IEEE DEST*, Istanbul, 2009, pp. 217–222.

[5] M. Hearst, "Text data mining: Issues, techniques, and the relationship to information access," in *UW/MS Workshop on Data Mining*, Berkeley, 1997.

[6] U. Fayyad, G. Piatetsky, and P. Smyth, "From data mining to knowledge discovery: An overview," in *Advances in Knowledge Discovery and Data Mining.* Cambridge: MIT Press, 1996, pp. 1–36.

[7] E. Simoudis, "Reality check for data mining," *IEEE Expert*, vol. 11, no. 5, 1996.

[8] H. Wu, R. Luk, K. Wong, and K. Kwok, "Interpreting TF-IDF term weights as making relevance decisions," *ACM Transactions on Information Systems*, vol. 26, no. 3, 2008.

[9] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," in *EMNLP*, Philadelphia, 2002, pp. 79–86.

[10] Y. Wilks and M. Stevenson, "The grammar of sense: Using part-of-speech tags as a first step in semantic disambiguation," *Journal of Natural Language Engineering*, vol. 4, no. 2, 1998.

[11] P. Turney, "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews," in *ACL*, Philadelphia, 2002, pp. 417–424.

[12] E. Cambria and A. Hussain, *Sentic Computing: Techniques, Tools, and Applications.* Berlin Heidelberg: Springer, 2012.

[13] J. Tatemura, "Virtual reviewers for collaborative exploration of movie reviews," in *IUI*, New Orleans, 2000.

[14] L. Terveen, W. Hill, B. Amento, D. McDonald, and J. Creter, "PHOAKS: A system for sharing recommendations," *Communications of the Association for Computing Machinery*, vol. 40, no. 3, 1997.

[15] X. Jin, Y. Li, T. Mah, and J. Tong, "Sensitive webpage classification for content advertising," in *ACM SIGKDD*, San Jose, 2007, pp. 28–33.

[16] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, pp. 1–135, 2008.

[17] A. Abbasi, "Affect intensity analysis of dark web forums," in *IEEE ISI*, New Brunswick, 2007, pp. 282–288.

[18] K. Boehner, R. Depaula, P. Dourish, and P. Sengers, "Affect: From information to interaction," in *Critical Computing*, Aaehus, 2005, pp. 59–68.

[19] R. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 18–37, 2010.

[20] M. Salzman, I. Matathia, and A. O'Reilly, *Buzz: Harness the Power of Influence and Create Demand.* New York: John Wiley and Sons, 2003.

[21] C. Cesarano, B. Dorr, A. Picariello, D. Reforgiato, A. Sagoff, and V. Subrahmanian, "OASYS: An opinion analysis system," in *AAAI CAAW*, Stanford, 2006.

[22] S. Sood and L. Vasserman, "ESSE: Exploring mood on the web," in *ICWSM*, San Jose, 2009.

[23] R. Speer, C. Havasi, N. Treadway, and H. Lieberman, "Finding your way in a multi-dimensional semantic space with Luminoso," in *IUI*, Hong Kong, 2010.

[24] K. Dave, S. Lawrence, and D. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," in *WWW*, Budapest, 2003.

[25] S. Das and M. Chen, "Yahoo! for Amazon: Extracting market sentiment from stock message boards," in *APFA*, Bangkok, 2001.

[26] R. Tong, "An operational system for detecting and tracking opinions in on-line discussion," in *SIGIR*, New Orleans, 2001.

[27] C. Hsinchun and D. Zimbra, "AI and opinion mining," *IEEE Intelligent Systems*, vol. 25, no. 3, pp. 74–80, 2010.

[28] R. Mihalcea, C. Banea, and J. Wiebe, "Learning multilingual subjective language via cross-lingual projections," in *ACL*, Prague, 2007.

[29] T. Wilson, J. Wiebe, and R. Hwa, "Just how mad are you? Finding strong and weak opinion clauses," in *AAAI*, San Jose, 2004, pp. 761–769.

[30] V. Hatzivassiloglou and J. Wiebe, "Effects of adjective orientation and gradability on sentence subjectivity," in *COLING*, Saarbrücken, 2000.

[31] J. Wiebe, T. Wilson, R. Bruce, M. Bell, and M. Martin, "Learning subjective language," *Computational Linguistics*, vol. 30, no. 3, 2004.

[32] E. Riloff, J. Wiebe, and W. Phillips, "Exploiting subjectivity classification to improve information extraction," in *AAAI*, Pittsburgh, 2005.

[33] A. Finn and N. Kushmerick, "Learning to classify documents according to genre," *Journal of the American Society for Information Science and Technology*, vol. 7, no. 5, 2006.

[34] D. Biber, *Variation across Speech and Writing*. Cambridge: Cambridge University Press, 1988.

[35] F. Mosteller and D. Wallace, *Applied Bayesian and Classical Inference: The Case of the Federalist Papers*. Berlin Heidelberg: Springer-Verlag, 1984.

[36] S. Argamon, M. Koppel, and G. Avneri, "Style-based text categorization: What newspaper am i reading?" in *AAAI Workshop on Text Categorization*, Madison, 1998.

[37] V. Hatzivassiloglou and K. McKeown, "Predicting the semantic orientation of adjectives," in *ACL/EACL*, Madrid, 1997.

[38] A. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in *HLT/EMNLP*, Vancouver, 2005.

[39] B. Snyder and R. Barzilay, "Multiple aspect ranking using the good grief algorithm," in *HLT/NAACL*, Rochester, 2007.

[40] B. Pang and L. Lee, "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts," in *ACL*, Barcelona, 2004, pp. 271–278.

[41] M. Joshi and C. Rose, "Generalizing dependency features for opinion mining," in *ACL/IJCNLP*, Singapore, 2009.

[42] B. Pang and L. Lee, "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales," in *ACL*, Ann Arbor, 2005, pp. 115–124.

[43] J. Kamps, M. Marx, R. Mokken, and M. de Rijke, "Using WordNet to measure semantic orientation of adjectives," in *LREC*, Lisbon, 2004, pp. 1115–1118.

[44] S. Kim and E. Hovy, "Automatic detection of opinion bearing words and sentences," in *IJCNLP*, Jeju Island, 2005, pp. 61–66.

[45] E. Riloff and J. Wiebe, "Learning extraction patterns for subjective expressions," in *EMNLP*, Sapporo, 2003, pp. 105–112.

[46] C. Baker, C. Fillmore, and J. Lowe, "The Berkeley FrameNet project," in *COLING/ACL*, Montreal, 1998, pp. 86–90.

[47] S. Kim and E. Hovy, "Extracting opinions, opinion holders, and topics expressed in online news media text," in *Workshop on Sentiment and Subjectivity in Text*, Sydney, 2006.

[48] R. Agrawal and R. Srikant, "Fast algorithm for mining association rules," in *VLDB*, Santiago de Chile, 1994.

[49] M. Hu and B. Liu, "Mining opinion features in customer reviews," in *AAAI*, San Jose, 2004.

[50] C. Zirn, M. Niepert, H. Stuckenschmidt, and M. Strube, "Fine-grained sentiment analysis with structural features," in *IJCNLP*, Chiang Mai, 2011.

[51] C. D. Elliott, "The affective reasoner: A process model of emotions in a multi-agent system," Ph.D. dissertation, Northwestern University, Evanston, 1992.

[52] A. Ortony, G. Clore, and A. Collins, *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press, 1988.

[53] J. Wiebe, T. Wilson, and C. Cardie, "Annotating expressions of opinions and emotions in language," *Language Resources and Evaluation*, vol. 39, no. 2, pp. 165–210, 2005.

[54] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in *HLT/EMNLP*, Vancouver, 2005.

[55] R. Stevenson, J. Mikels, and T. James, "Characterization of the affective norms for english words by discrete emotional categories," *Behavior Research Methods*, vol. 39, pp. 1020–1024, 2007.

[56] S. Somasundaran, J. Wiebe, and J. Ruppenhofer, "Discourse level opinion interpretation," in *COLING*, Manchester, 2008.

[57] D. Rao and D. Ravichandran, "Semi-supervised polarity lexicon induction," in *EACL*, Athens, 2009, pp. 675–682.

[58] B. Goertzel, K. Silverman, C. Hartley, S. Bugaj, and M. Ross, "The Baby Webmind project," in *AISB*, Birmingham, 2000.

[59] P. Turney and M. Littman, "Measuring praise and criticism: Inference of semantic orientation from association," *ACM Transactions on Information Systems*, vol. 21, no. 4, pp. 315–346, 2003.

[60] A. Abbasi, H. Chen, and A. Salem, "Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums," *ACM Transactions on Information Systems*, vol. 26, no. 3, pp. 1–34, 2008.

[61] L. Velikovich, S. Goldensohn, K. Hannan, and R. McDonald, "The viability of web-derived polarity lexicons," in *NAACL*, Los Angeles, 2010, pp. 777–785.

[62] M. Minsky, *The Society of Mind*. New York: Simon and Schuster, 1986.

[63] E. Cambria, A. Hussain, C. Havasi, and C. Eckl, "Common sense computing: From the society of mind to digital intuition and beyond," in *Biometric ID Management and Multimodal Communication*, ser. Lecture Notes in Computer Science, J. Fierrez, J. Ortega, A. Esposito, A. Drygajlo, and M. Faundez-Zanuy, Eds. Berlin Heidelberg: Springer, 2009, vol. 5707, pp. 252–259.

[64] G. Murphy, *The Big Book of Concepts.* Cambridge: The MIT Press, 2004.

[65] P. Bloom, "Glue for the mental world," *Nature*, vol. 421, pp. 212–213, 2003.

[66] J. Barwise, "An introduction to first-order logic," in *Handbook of Mathematical Logic. Studies in Logic and the Foundations of Mathematics.* North-Holland, 1977.

[67] R. Reiter, "A logic for default reasoning," *Artificial Intelligence*, vol. 13, pp. 81–132, 1980.

[68] A. Heyting, *Intuitionism. An introduction.* North-Holland, 1956.

[69] C. Date and H. Darwen, *A Guide to the SQL Standard.* Reading, Mass: Addison-Wesley, 1993.

[70] E. Codd, "A relational model of data for large shared data banks," *Communications of the ACM*, vol. 13, no. 6, pp. 377–387, 1970.

[71] ——, "Further normalization of the data base relational model," IBM Research Report, New York, Tech. Rep., 1971.

[72] ——, "Recent investigations into relational data base systems," IBM Research Report, New York, Tech. Rep. RJ1385, 1974.

[73] N. Chomsky, "Three models for the description of language," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 113–124, 1956.

[74] L. Lacy, *OWL: Representing Information Using the Web Ontology Language.* Trafford Publishing, 2005.

[75] J. Pearl, "Bayesian networks: A model of self-activated memory for evidential reasoning," UCLA Technical Report, Irvine, Tech. Rep. CSD-850017, 1985.

[76] J. Sowa, "Semantic networks," in *Encyclopedia of Artificial Intelligence.* Stuart Shapiro, 1987.

[77] P. Winston, "Learning structural descriptions from examples," *The Psychology of Computer Vision*, pp. 157–209, 1975.

[78] J. McCarthy, "Programs with common sense," in *Teddington Conference on the Mechanization of Thought Processes*, 1959.

[79] D. Ernest, *Representations of Commonsense Knowledge.* Morgan Kaufmann, 1990.

[80] D. Lenat and R. Guha, *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project.* Boston: Addison-Wesley, 1989.

[81] C. Fellbaum, *WordNet: An Electronic Lexical Database (Language, Speech, and Communication).* The MIT Press, 1998.

[82] E. Mueller, *Natural Language Processing with ThoughtTreasure.* Signifonn, 1998.

[83] M. Minsky, "Commonsense-based interfaces," *Communications of the ACM*, vol. 43, no. 8, pp. 67–73, 2000.

[84] P. Singh, "The open mind common sense project," *KurzweilAI.net*, 2002.

[85] D. Stork, "The open mind initiative." *IEEE Intelligent Systems*, vol. 14, no. 3, pp. 16–20, 1999.

[86] T. Chklovski, "Learner: a system for acquiring commonsense knowledge by analogy," in *K-CAP*, 2003.

[87] R. Speer, C. Havasi, and H. Lieberman, "Analogyspace: Reducing the dimensionality of common sense knowledge," in *AAAI*, 2008.

[88] C. Havasi, R. Speer, J. Pustejovsky, and H. Lieberman, "Digital intuition: Applying common sense using dimensionality reduction," *IEEE Intelligent Systems*, vol. 24, no. 4, pp. 24–35, 2009.

[89] D. Charles, *The Expression of the Emotions in Man and Animals.* John Murray, 1872.

[90] W. James, "What is an emotion?" *Mind*, vol. 34, pp. 188–205, 1884.

[91] C. Osgood, W. May, and M. Miron, *Cross-Cultural Universals of Affective Meaning.* Univ. of Illinois Press, 1975.

[92] C. Lutz and G. White, "The anthropology of emotions," *Ann. Rev. Anthropology*, vol. 15, pp. 405–436, 1986.

[93] S. Turkle, *The Second Self: Computers and the Human Spirit.* Simon & Schuster, 1984.

[94] K. Scherer, "Studying the emotion-antecedent appraisal process: An expert system approach," *Cognition and Emotion*, vol. 7, pp. 325–355, 1993.

[95] R. Picard, *Affective Computing.* Boston: The MIT Press, 1997.

[96] E. Cambria, I. Hupont, A. Hussain, E. Cerezo, and S. Baldassarri, "Sentic avatar: Multimodal affective conversational agent with common sense," in *Toward Autonomous, Adaptive, and Context-Aware Multimodal Interfaces: Theoretical and Practical Issues*, ser. Lecture Notes in Computer Science, A. Esposito, A. Hussain, M. Faundez-Zanuy, R. Martone, and N. Melone, Eds. Berlin Heidelberg: Springer-Verlag, 2011, vol. 6456, pp. 82–96.

[97] E. Cambria, T. Mazzocco, A. Hussain, and T. Durrani, "Switching between different ways to think: Multiple approaches to affective common sense reasoning," in *Analysis of Verbal and Nonverbal Communication and Enactment: The Processing Issues*, ser. Lecture Notes in Computer Science, A. Esposito, A. Vinciarelli, K. Vicsi, C. Pelachaud, and A. Nijholt, Eds. Berlin Heidelberg: Springer, 2011, vol. 6800, pp. 56–69.

[98] E. Cambria, A. Hussain, C. Havasi, and C. Eckl, "SenticSpace: Visualizing opinions and sentiments in a multi-dimensional vector space," in *Knowledge-Based*

*and Intelligent Information and Engineering Systems*, ser. Lecture Notes in Artificial Intelligence, R. Setchi, I. Jordanov, R. Howlett, and L. Jain, Eds. Berlin Heidelberg: Springer, 2010, vol. 6279, pp. 385–393.

[99] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor, "Freebase: A collaboratively created graph database for structuring human knowledge," in *SIGMOD*, Vancouver, 2008, pp. 1247–1250.

[100] S. Ponzetto and M. Strube, "Deriving a large-scale taxonomy from Wikipedia," in *AAAI*, Vancouver, 2007, pp. 1440–1445.

[101] F. Suchanek, G. Kasneci, and G. Weikum, "Yago: a core of semantic knowledge," in *WWW*, Banff, 2007, pp. 697–706.

[102] A. Carlson, J. Betteridge, B. Kisiel, B. Settles, E. Hruschka, and T. Mitchell, "Toward an architecture for never-ending language learning," in *AAAI*, 2010.

[103] P. Grice, *Logic and Conversation*. Academic Press, 1975, vol. Speech Acts, pp. 41–58.

[104] H. Liu and P. Singh, "ConceptNet: A practical commonsense reasoning toolkit," *BT Technology Journal*, 2004.

[105] E. Mueller, *Commonsense Reasoning*. Morgan Kaufmann, 2006.

[106] C. Havasi, R. Speer, and J. Alonso, "ConceptNet 3: A flexible, multilingual semantic network for common sense knowledge," in *RANLP*, Borovets, 2007.

[107] R. Speer, "Open Mind Commons: An inquisitive approach to learning common sense," in *Workshop on Common Sense and Interactive Applications*, Honolulu, 2007.

[108] R. Plutchik, "The nature of emotions," *American Scientist*, vol. 89, no. 4, pp. 344–350, 2001.

[109] N. Frijda, "The laws of emotions," *American Psychologist*, vol. 43, no. 5, 1988.

[110] U. Neisser, *Cognitive Psychology.* Appleton Century Crofts, 1967.

[111] C. Strapparava and A. Valitutti, "WordNet-Affect: An affective extension of WordNet," in *LREC*, Lisbon, 2004.

[112] M. Minsky, *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind.* New York: Simon & Schuster, 2006.

[113] T. Dalgleish, "The emotional brain," *Nature: Perspectives*, vol. 5, pp. 582–589, 2004.

[114] P. Broca, "Anatomie comparée des circonvolutions cérébrales: Le grand lobe limbique," *Rev. Anthropol*, vol. 1, pp. 385–498, 1878.

[115] J. Papez, "A proposed mechanism of emotion," *Neuropsychiatry Clin Neurosci*, vol. 7, pp. 103–112, 1937.

[116] P. Maclean, "Psychiatric implications of physiological studies on frontotemporal portion of limbic system (visceral brain)," *Electroencephalogr Clin Neurophysiol Suppl*, vol. 4, pp. 407–18, 1952.

[117] J. Ledoux, *Synaptic Self.* Penguin Books, 2003.

[118] P. Ekman, T. Dalgleish, and M. Power, *Handbook of Cognition and Emotion.* Chichester: Wiley, 1999.

[119] K. Scherer, "Psychological models of emotion," *The Neuropsychology of Emotion Pages*, pp. 137–162, 2000.

[120] W. Parrott, *Emotions in Social Psychology.* Psychology Press, 2001.

[121] J. Prinz, *Gut Reactions: A Perceptual Theory of Emotion.* Oxford University Press, 2004.

[122] E. Douglas-Cowie, "Humaine deliverable d5g: Mid term report on database exemplar progress," Information Society Technologies, Tech. Rep., 2006.

[123] A. Kapoor, W. Burleson, and R. Picard, "Automatic prediction of frustration," *International Journal of Human-Computer Studies*, vol. 65, pp. 724–736, 2007.

[124] G. Castellano, L. Kessous, and G. Caridakis, "Multimodal emotion recognition from expressive faces, body gestures and speech," in *Doctoral Consortium of ACII*, Lisbon, 2007.

[125] J. Averill, "A constructivist view of emotion," *Emotion: Theory, Research and Experience*, pp. 305–339, 1980.

[126] J. Russell, "Core affect and the psychological construction of emotion," *Psychological Rev.*, vol. 110, pp. 145–172, 2003.

[127] C. Osgood, G. Suci, and P. Tannenbaum, *The Measurement of Meaning*. University of Illinois Press, 1957.

[128] J. Russell, "Affective space is bipolar," *Journal of Personality and Social Psychology*, vol. 37, pp. 345–356, 1979.

[129] C. Whissell, "The dictionary of affect in language," *Emotion: Theory, Research, and Experience*, vol. 4, pp. 113–131, 1989.

[130] A. Freitas and E. Castro, "Facial expression: The effect of the smile in the treatment of depression. empirical study with portuguese subjects," in *Emotional Expression: The Brain and The Face*. University Fernando Pessoa Press, 2009, pp. 127–140.

[131] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychology*, vol. 14, no. 4, pp. 261–292, 1996.

[132] J. Fontaine, K. Scherer, E. Roesch, and P. Ellsworth, "The world of emotions is not two-dimensional," *Psychological Science*, vol. 18, no. 12, pp. 1050–1057, 2007.

[133] T. Cochrane, "Eight dimensions for the emotions," *Social Science Information*, vol. 48, no. 3, pp. 379–420, 2009.

[134] R. Lazarus, *Emotion and Adaptation*. New York: Oxford University Press, 1991.

[135] M. Lewis, "Self-conscious emotions: Embarrassment, pride, shame, and guilt," in *Handbook of Cognition and Emotion*. Guilford Press, 2000, vol. 2, pp. 623–636.

[136] K. Scherer, A. Shorr, and T. Johnstone, *Appraisal Processes in Emotion: Theory, Methods, Research*. Canary: Oxford University Press, 2001.

[137] J. Tracy, R. Robins, and J. Tangney, *The Self-Conscious Emotions: Theory and Research*. The Guilford Press, 2007.

[138] S. Zeki and J. Romaya, "Neural correlates of hate," *PloS one*, vol. 3, no. 10, pp. 35–56, 2008.

[139] L. Cahill and J. McGaugh, "A novel demonstration of enhanced memory associated with emotional arousal," *Consciousness and Cognition*, vol. 4, no. 4, pp. 410–421, 1995.

[140] W. Bradford Cannon, *Bodily Changes in Pain, Hunger, Fear and Rage: An Account of Recent Researches into the Function of Emotional Excitement*. Appleton Century Crofts, 1915.

[141] L. Barrett, "Solving the emotion paradox: Categorization and the experience of emotion," *Personality and Social Psychology Review*, vol. 10, no. 1, pp. 20–46, 2006.

[142] E. Krumhuber and A. Kappas, "Moving smiles: The role of dynamic components for the perception of the genuineness of smiles," *Journal of Nonverbal Behavior*, vol. 29, no. 1, pp. 3–24, 2005.

[143] M. Lewis and I. Granic, *Emotion, Development, and Self-Organization: Dynamic Systems Approaches to Emotional Development*. Cambridge University Press, 2002.

[144] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. Harper Perennial, 1991.

[145] E. Cambria, M. Grassi, A. Hussain, and C. Havasi, "Sentic computing for social media marketing," *Multimedia Tools and Applications*, vol. http://dx.doi.org/10.1007/s11042-011-0815-0, 2011.

[146] PopCap, "2010 social gaming research," Pop Cap Games, Tech. Rep., 2010.

[147] L. von Ahn, M. Blum, N. Hopper, and J. Langford, "Captcha: Using hard ai problems for security," in *EUROCRYPT*, ser. Lecture Notes in Computer Science, E. Biham, Ed. Berlin Heidelberg: Springer, 2003, pp. 294–311.

[148] L. von Ahn, "Games with a purpose," *IEEE Computer Magazine*, vol. 6, pp. 92–94, 2006.

[149] S. Thaler, K. Siorpaes, E. Simperl, and C. Hofer, "A survey on games for knowledge acquisition," Semantic Technology Institute, Tech. Rep., 2011.

[150] K. Siorpaes and M. Hepp, "Ontogame: Weaving the semantic web by online games," in *ESWC*, Tenerife, 2008, pp. 751–766.

[151] R. van Zwol, L. Garcia, G. Ramirez, B. Sigurbjornsson, and M. Labad, "Video tag game," in *WWW*, Beijing, 2008.

[152] M. Addis, L. Boch, W. Allasia, F. Gallo, W. Bailer, and R. Wright, "100 million hours of audiovisual content: Digital preservation and access in the prestoprime project," in *Digital Preservation Interoperability Framework Symposium*, Dresden, 2010.

[153] L. von Ahn and L. Dabbish, "Labeling images with a computer game," in *CHI*, Vienna, 2004, pp. 319–326.

[154] S. Hacker and L. von Ahn, "Matchin: Eliciting user preferences with an online game," in *CHI*, Boston, 2009.

[155] L. von Ahn, S. Ginosar, M. Kedia, R. Liu, and M. Blum, "Improving accessibility of the web with a computer game," in *CHI*, Quebec, 2006, pp. 79–82.

[156] L. von Ahn, R. Liu, and M. Blum, "Peekaboom: A game for locating objects in images," in *CHI*, 2006, pp. 55–64.

[157] J. Yan and S.-Y. Yu, "Magic bullet: a dual-purpose computer game," in *ACM SIGKDD*, Paris, 2009, pp. 32–33.

[158] D. Morrison, S. Maillet, and E. Bruno, "Tagcaptcha: Annotating images with captchas," in *ACM SIGKDD*, Paris, 2009, pp. 44–45.

[159] L. Barrington, D. O'Malley, D. Turnbull, and G. Lanckriet, "User-centered design of a social game to tag music," in *ACM SIGKDD*, Paris, 2009, pp. 7–10.

[160] E. Law, L. von Ahn, R. Dannenberg, and M. Crawford, "Tagatune: A game for music and sound annotation," in *International Conference on Music Information Retrieval*, Vienna, 2007.

[161] L. von Ahn, M. Kedia, and M. Blum, "Verbosity: A game for collecting common sense facts," in *CHI*, Quebec, 2006, pp. 75–78.

[162] A. Herdagdelen and M. Baroni, "The concept game: Better commonsene knowledge extraction by combining text mining and game with a purpose," in *AAAI CSK*, Arlington, 2010.

[163] H. Ma, R. Chandrasekar, C. Quirk, and A. Gupta, "Page hunt: Improving search engines using human computation games," in *SIGIR*, Boston, 2009, pp. 746–747.

[164] Y. Kuo, J. Lee, K. Chiang, R. Wang, E. Shen, C. Chan, and J. Y. Hu, "Community-based game design: Experiments on social games for commonsense data collection," in *ACM SIGKDD*, Paris, 2009, pp. 15–22.

[165] T. Markotschi and J. Volker, "GuessWhat?! - Human intelligence for mining linked data," in *EKAW*, Lisbon, 2010.

[166] B. Liu, M. Hu, and J. Cheng, "Opinion observer: Analyzing and comparing opinions on the web," in *WWW*, Chiba, 2005.

[167] R. Mihalcea, "Open mind word expert: Creating large annotated data collections with web users' help," in *Workshop on Linguistically Interpreted Corpora*, Budapest, 2003.

[168] R. Gupta, M. Kochenderfer, D. Mcguinness, and G. Ferguson, "Common sense data acquisition for indoor mobile robots," in *AAAI*, San Jose, 2004, pp. 605–610.

[169] W. Wu, H. Li, H. Wang, and K. Zhu, "Towards a probabilistic taxonomy of many concepts," MSRA, Tech. Rep. MSR-TR-2011-25, 2011.

[170] M. Hearst, "Automatic acquisition of hyponyms from large text corpora," in *COLING*, Nantes, 1992, pp. 539–545.

[171] E. Cambria, Y. Song, H. Wang, and A. Hussain, "Isanette: A common and common sense knowledge base for opinion mining," in *ICDM*, Vancouver, 2011.

[172] E. Cambria, A. Hussain, C. Havasi, and C. Eckl, "AffectiveSpace: Blending common sense and affective knowledge to perform emotive reasoning," in *CAEPIA*, Seville, 2009, pp. 32–41.

[173] G. Fauconnier and M. Turner, *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities.* Basic Books, 2003.

[174] C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *Journal of Research of The National Bureau of Standards*, vol. 45, no. 4, pp. 255–282, 1950.

[175] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.

[176] E. Cambria, A. Hussain, T. Durrani, C. Havasi, C. Eckl, and J. Munro, "Sentic computing for patient centered application," in *IEEE ICSP*, Beijing, 2010, pp. 1279–1282.

[177] E. Cambria, T. Mazzocco, A. Hussain, and C. Eckl, "Sentic medoids: Organizing affective common sense knowledge in a multi-dimensional vector space," in *Advances in Neural Networks*, ser. Lecture Notes in Computer Science, vol. 6677. Berlin Heidelberg: Springer-Verlag, 2011, pp. 601–610.

[178] L. Kaufman and P. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis.* Wiley-Interscience, 1990.

[179] J. Hartigan and M. Wong, "Algorithm AS 136: A k-means clustering algorithm," *Journal of the Royal Statistical Society*, vol. 28, no. 1, pp. 100–108, 1979.

[180] J. Han, M. Kamber, and A. Tung, "Spatial clustering methods in data mining: A survey," in *Geographic Data Mining and Knowledge Discovery*, T. . Francis, Ed., 2001.

[181] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Fifth Berkeley Symposium on Mathematical Statistics and Probability*, U. of California Press, Ed., 1967, pp. 281–297.

[182] M. Garey and D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness.* Freeman, 1979.

[183] H. Park and C. Jun, "A simple and fast algorithm for k-medoids clustering," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3336–3341, 2009.

[184] R. Duda and P. Hart, *Pattern Classification and Scene Analysis.* New York: John Wiley & Sons Inc, 1973.

[185] L. Kirkpatrick and S. Epstein, "Cognitive experiential self-theory and subjective probability: Further evidence for two conceptual systems," *Journal of Personality and Social Psychology*, vol. 63, pp. 534–544, 1992.

[186] S. Chaiken and Y. Trope, *Dual-Process Theories in Social Psychology.* New York: Guilford, 1999.

[187] E. Smith and J. DeCoster, "Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems," *Personality and Social Psychological Review*, vol. 4, no. 2, pp. 108–131, 2000.

[188] S. Epstein, "Cognitive-experiential self-theory of personality," in *Comprehensive Handbook of Psychology*, T. Millon and M. Lerner, Eds. Hoboken, NJ: Wiley & Sons, 2003, vol. 5, pp. 159–184.

[189] M. Lieberman, "Social cognitive neuroscience: A review of core processes," *Annual Review of Psychology*, vol. 58, pp. 259–89, 2007.

[190] A. Damasio, *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain.* Harcourt, Inc., 2003.

[191] T. Dalgleish, B. Dunn, and D. Mobbs, "Affective neuroscience: Past, present, and future," *Emotion Review*, pp. 355–368, 2009.

[192] A. Ohman and J. Soares, "Emotional conditioning to masked stimuli: Expectancies for aversive outcomes following nonre- cognized fear-relevant stimuli," *Journal of Experimental Psychology: General*, vol. 127, no. 1, pp. 69–82, 1998.

[193] M. Calvo and L. Nummenmaa, "Processing of unattended emotional visual scenes," *Journal of Experimental Psychology: General*, vol. 136, pp. 347–369, 2007.

[194] D. Westen, "Implications of developments in cognitive neuroscience for psychoanalytic psychotherapy," *Harv Rev Psychiatry*, vol. 10, no. 6, pp. 369–73, 2002.

[195] C. Havasi, R. Speer, and J. Holmgren, "Automated color selection using semantic knowledge," in *AAAI CSK*, Arlington, 2010.

[196] E. Cambria, R. Speer, C. Havasi, and A. Hussain, "SenticNet: A publicly available semantic resource for opinion mining," in *AAAI CSK*, Arlington, 2010, pp. 14–18.

[197] A. Esuli and F. Sebastiani, "SentiWordNet: A publicly available lexical resource for opinion mining," in *LREC*, 2006.

[198] M. Grassi, "Developing HEO human emotions ontology," in *Biometric ID Management and Multimodal Communication*, ser. Lecture Notes in Computer Science, J. Fierrez, J. Ortega-Garcia, A. Esposito, A. Drygajlo, and M. Faundez-Zanuy, Eds. Berlin Heidelberg: Springer, 2009, vol. 5707, pp. 244–251.

[199] T. Gruber, "Collective knowledge systems: Where the social web meets the semantic web," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 6, no. 1, pp. 4–13, 12 2007.

[200] C. Ma, A. Osherenko, H. Prendinger, and M. Ishizuka, "A chat system based on emotion estimation from text and embodied conversational messengers," in *Int'l Conf. Active Media Technology*, 2005, pp. 546–548.

[201] C. Alm, D. Roth, and R. Sproat, "Emotions from text: Machine learning for text-based emotion prediction," in *HLT/EMNLP*, 2005, pp. 347–354.

[202] W. Lin, T. Wilson, J. Wiebe, and A. Hauptmann, "Which side are you on? identifying perspectives at the document and sentence levels," in *Conference on Natural Language Learning*, 2006, pp. 109–116.

[203] S. D'Mello, S. Craig, J. Sullins, and A. Graesser, "Predicting affective states expressed through an emote-aloud procedure from autotutor's mixed-initiative dialogue," *Int'l J. Artificial Intelligence in Education*, vol. 16, pp. 3–28, 2006.

[204] T. Danisman and A. Alpkocak, "Feeler: Emotion classification of text using vector space model," in *AISB*, 2008.

[205] C. Strapparava and R. Mihalcea, "Learning to identify emotions in text," in *ACM Symp. Applied Computing*, 2008, pp. 1556–1560.

[206] S. D'Mello, N. Dowell, and A. Graesser, "Cohesion relationships in tutorial dialogue as predictors of affective states," in *Proceeedings of Conf. Artificial Intelligence in Education*, 2009, pp. 9–16.

[207] E. Cambria, A. Hussain, and C. Eckl, "Bridging the gap between structured and unstructured health-care data through semantics and sentics," in *WebSci*, Koblenz, 2011.

[208] C. Havasi, R. Speer, and J. Pustejovsky, "Coarse word-sense disambiguation using common sense," in *In AAAI CSK*, Arlington, 2010.

[209] Y. Song, H. Wang, Z. Wang, H. Li, and W. Chen, "Short text conceptualization using a probabilistic knowledgebase," in *IJCAI*, Barcelona, 2011.

[210] M. Rowe and J. Butters, "Assessing trust: Contextual accountability," in *ESWC*, Heraklion, 2009.

[211] E. Cambria, P. Chandra, A. Sharma, and A. Hussain, "Do not feel the trolls," in *ISWC*, Shanghai, 2010.

[212] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content: the QBIC system," *Computer*, vol. 28, no. 9, pp. 23–32, 1995.

[213] J. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, and C. Shu, "Virage image search engine: An open framework for image management," in *Storage and Retrieval for Still Image and Video Databases*, I. Sethi and R. Jain, Eds., vol. 2670. SPIE, 1996, pp. 76–87.

[214] K. Porkaew and K. Chakrabarti, "Query refinement for multimedia similarity retrieval in MARS," in *ACM International Conference on Multimedia*. New York: ACM, 1999, pp. 235–238.

[215] M. Nakazato, L. Manola, and T. Huang, "ImageGrouper: Search, annotate and organize images by groups," in *Recent Advances in Visual Information Systems*, ser. Lecture Notes in Computer Science, S. Chang, Z. Chen, and S. Lee, Eds. Berlin Heidelberg: Springer, 2002, vol. 2314, pp. 93–105.

[216] N. O'Hare, H. Lee, S. Cooray, C. Gurrin, G. Jones, J. Malobabic, N. O'Connor, A. Smeaton, and B. Uscilowski, "MediAssist: Using content-based analysis and context to manage personal photo collections," in *CIVR*, Tempe, 2006, pp. 529–532.

[217] N. Sebe, Q. Tian, E. Loupias, M. S. Lew, and T. S. Huang, "Evaluation of salient point techniques," in *International Conference on Image and Video Retrieval*. London: Springer-Verlag, 2002, pp. 367–377.

[218] J. Urban and J. M. Jose, "EGO: A personalized multimedia management and retrieval tool," *International Journal of Intelligent Systems*, vol. 21, no. 7, pp. 725–745, 2006.

[219] R. Datta and J. Wang, "ACQUINE: Aesthetic quality inference engine - real-time automatic rating of photo aesthetics," in *International Conference on Multimedia Information Retrieval*, Philadelphia, 2010.

[220] N. Bianchi-Berthouze, "K-DIME: An affective image filtering system," *IEEE Multimedia*, vol. 10, no. 3, pp. 103–106, 2003.

[221] J. Smith and S. Chang, "An image and video search engine for the world-wide web," in *Symposium on Electronic Imaging: Science and Technology*, 1997.

[222] C. Frankel, M. J. Swain, and V. Athitsos, "WebSeer: An image search engine for the world wide web," University of Chicago, Tech. Rep., 1996.

[223] R. Lempel and A. Soffer, "PicASHOW: Pictorial authority search by hyperlinks on the web," in *WWW*, Hong Kong, 2001.

[224] F. Jing, C. Wang, Y. Yao, K. Deng, L. Zhang, and W.-Y. Ma, "IGroup: Web image search results clustering," in *ACM Multimedia*, Santa Barbara, 2006.

[225] H. Lieberman, E. Rosenzweig, and P. Singh, "ARIA: An agent for annotating and retrieving images," *IEEE Computer*, vol. 34, no. 7, pp. 57–62, 2001.

[226] P. Chi and H. Lieberman, "Intelligent assistance for conversational storytelling using story patterns," in *IUI*, Palo Alto, 2011.

[227] E. Cambria and A. Hussain, "Sentic album: Content, concept and context based online personal photo management system," *Cognitive Computation*, 2012.

[228] H. Lieberman and T. Selker, "Out of context: Computer systems that adapt to, and learn from, context," *IBM Systems Journal*, vol. 39, no. 3, pp. 617–632, 2000.

[229] A. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain.* New York: Grossett/Putnam, 1994.

[230] E. Vesterinen, "Affective computing," in *Digital Media Research Seminar*, Helsinki, 2001.

[231] M. Pantic, "Affective computing," in *Encyclopedia of Multimedia Technology and Networking*, vol. 1. Idea Group Reference, 2005, pp. 8–14.

[232] G. Bonanno, A. Papa, K. O'Neill, M. Westphal, and K. Coifman, "The importance of being flexible: The ability to enhance and suppress emotional expressions predicts long-term adjustment," *Psychological Science*, vol. 15, pp. 482–487, 2004.

[233] J. Richards, E. Butler, and J. Gross, "Emotion regulation in romantic relationships: The cognitive consequences of concealing feelings," *Journal of Social and Personal Relationships*, vol. 20, pp. 599–620, 2003.

[234] A. Burke, F. Heuer, and D. Reisberg, "Remembering emotional events," *Memory and Cognition*, vol. 20, pp. 277–290, 1992.

[235] S. Christianson and E. Loftus, "Remembering emotional events: The fate of detailed information," *Cognition and Emotion*, vol. 5, pp. 81–108, 1991.

[236] I. Wessel and H. Merckelbach, "The impact of anxiety on memory for details in spider phobics," *Applied Cognitive Psychology*, vol. 11, pp. 223–231, 1997.

[237] D. Reisberg and F. Heuer, "Memory for emotional events," *Memory and Emotion*, pp. 3–41, 2004.

[238] C. Laney, H. Campbell, F. Heuer, and D. Reisberg, "Memory for thematically arousing events," *Memory and Cognition*, vol. 32, no. 7, pp. 1149–1159, 2004.

[239] A. Hanjalic, "Extracting moods from pictures and sounds: Towards truly personalized TV," *IEEE Signal Processing Magazine*, vol. 23, no. 2, pp. 90–100, 2006.

[240] G. Lakoff, *Women, Fire, and Dangerous Things.* University Of Chicago Press, 1990.

[241] B. Keelan, *Handbook of Image Quality.* New York: Marcel Dekker, 2002.

[242] M. Narwaria and W. Lin, "Objective image quality assessment based on support vector regression," *IEEE Trans. on Neural Networks*, vol. 12, no. 3, pp. 515–519, 2010.

[243] W. Lu, K. Zeng, D. Tao, Y. Yuan, and X. Gao, "No-reference image quality assessment in contourlet domain," *Neurocomputing*, vol. 73, no. 4-6, pp. 784–794, 2012.

[244] J. Redi, P. Gastaldo, I. Heynderickx, and R. Zunino, "Color distribution information for the reduced-reference assessment of perceived image quality," *IEEE*

*Trans. on Circuits and Systems for Video Technology*, vol. 20, no. 12, pp. 1757–1769, 2012.

[245] S. Decherchi, P. Gastaldo, J. Redi, R. Zunino, and E. Cambria, "Circular-ELM for the reduced-reference assessment of perceived image quality," *Neurocomputing*, 2012.

[246] J. Huang, S. Ravi, M. Mitra, W. Zhu, and R. Zabih, "Image indexing using color correlograms," in *IEEE CVPR*, 1997, pp. 762–768.

[247] B. Lee, J. Hendler, and O. Lassila, "The semantic web," *Scientific American*, 2001.

[248] J. Urban, J. Jose, and C. Van Rijsbergen, "An adaptive approach towards content-based image retrieval," *Multimedia Tools and Applications*, vol. 31, pp. 1–28, 2006.

[249] M. Lansdale and E. Edmonds, "Using memory for events in the design of personal filing systems," *International Journal of Man-Machine Studies*, vol. 36, no. 1, pp. 97–126, 1992.

[250] M. Lew, N. Sebe, C. Djeraba, and R. Jain, "Content-based multimedia information retrieval: State of the art and challenges," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 2, no. 1, pp. 1–19, 2006.

[251] J. Machajdik and A. Hanbury, "Affective image classification using features inspired by psychology and art theory," in *International Conference on Multimedia*, Florence, 2010.

[252] S. Gilroy, M. Cavazza, M. Niiranen, E. Andre, T. Vogt, J. Urbain, M. Benayoun, H. Seichter, and M. Billinghurst, "Pad-based multimodal affective fusion," in *ACII09*, Amsterdam, 2009.

[253] Z. Zeng, M. Pantic, G. Roisman, and T. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 1, pp. 39–58, 2009.

[254] H. Gunes, M. Piccardi, and M. Pantic, "From the lab to the real world: Affect recognition using multiple cues and modalities," *Affective Computing: Focus on Emotion Expression, Synthesis, and Recognition*, pp. 185–218, 2008.

[255] J. Riseberg, J. Klein, R. Fernandez, and R. Picard, "Frustrating the user on purpose: Using biosignals in a pilot study to detect the user's emotional state," in *CHI*, Los Angeles, 1998.

[256] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: a meta-analysis," *Psychological Bulletin*, vol. 11, no. 2, pp. 256–274, 1992.

[257] A. Camurri, B. Mazzarino, and G. Volpe, "Analysis of expressive gesture: The eyesweb expressive gesture processing library," in *Gesture Workshop*, Genova, 2003.

[258] H. Gunes and M. Piccardi, "Bi-modal emotion recognition from expressive face and body gestures," *Network and Computer Applications*, vol. 30, no. 4, pp. 1334–1345, 2007.

[259] K. Karpouzis, G. Caridakis, L. Kessous, N. Amir, A. Raouzaiou, L. Malatesta, and S. Kollias, "Modeling naturalistic affective states via facial, vocal and bodily expressions recognition," in *Lecture Notes in Artificial Intelligence*. Springer, 2007, vol. 4451, pp. 92–116.

[260] T. Pun, T. Alecu, G. Chanel, J. Kronegg, and S. Voloshynovskiy, "Brain-computer interaction research at the computer vision and multimedia laboratory," *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 210–213, 2006.

[261] W. Burleson, R. Picard, K. Perlin, and J. Lippincott, "A platform for affective agent research," in *International Conference on Autonomous Agents and Multi-agent Systems*, New York, 2004.

[262] S. Petridis and M. Pantic, "Audiovisual discrimination between laughter and speech," in *ICASSP*, Las Vegas, 2008.

[263] M. Valstar, H. Gunes, and M. Pantic, "How to distinguish posed from spontaneous smiles using geometric features," in *ICMI*, Nagoya, 2007.

[264] K. Truong and D. Van Leeuwen, "Automatic discrimination between laughter and speech," *Speech Communication*, vol. 49, pp. 144–158, 2007.

[265] S. Matos, S. Birring, I. Pavord, and D. Evans, "Detection of cough signals in continuous audio recordings using HMM," *IEEE Trans. on Biomedical Engineering*, vol. 53, no. 6, pp. 1078–1083, 2006.

[266] P. Pal, A. Iyer, and R. Yantorno, "Emotion detection from infant facial expressions and cries," in *International Conference on Acoustics, Speech and Signal Processing*, Dallas, 2006.

[267] J. JongTae, S. SangWook, K. KwangEun, and S. KweeBo, "Emotion recognition method based on multimodal sensor fusion algorithm," in *ISIS*, Sokcho-City, 2007.

[268] C. Shan, S. Gong, and P. McOwan, "Beyond facial expressions: Learning human emotion from body gestures," in *BMVC*, Warwick, 2007.

[269] S. Baldassarri, E. Cerezo, and F. Seron, "Maxine: a platform for embodied animated agents," *Computers and Graphics*, vol. 32, no. 4, pp. 430–437, 2008.

[270] E. Cerezo, I. Hupont, C. Manresa, J. Varona, S. Baldassarri, F. Perales, and F. Seron, "Real-time facial expression recognition for natural interaction," in *Pattern Recognition and Image Analysis*, ser. Lecture Notes in Computer Science. Berlin Heidelberg: Springer-Verlag, 2007, vol. 4478, pp. 40–47.

[271] I. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*. San Francisco: Morgan Kaufmann, 2005.

[272] F. Wallhoff, “Facial expressions and emotion database,” Technische Universitat Munchen, Tech. Rep., 2006.

[273] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, “Web-based database for facial expression analysis,” in *ICME*, Singapore, 2005.

[274] S. Siegel and N. Castellan, *Nonparametric Statistics for the Social Siences*. New York: McGraw-Hill, 1988.

[275] P. Kumar and E. Yildirim, “Minimum-volume enclosing ellipsoids and core sets,” *Journal of Optimization Theory and Applications*, vol. 126, pp. 1–21, 2005.

[276] A. Milewski and T. Smith, “Providing presence cues to telephone users,” in *ACM Conference on Computer Supported Cooperative Work*, 2000.

[277] H. Chang, “Emotion barometer of reading: User interface design of a social cataloging website,” in *International Conference on Human Factors in Computing Systems*, 2009.

[278] E. Pampalk, A. Rauber, and D. Merkl, “Content-based organization and visualization of music archives,” in *ACM International Conference on Multimedia*, 2002.

[279] E. Cambria, A. Hussain, and C. Eckl, “Taking refuge in your personal sentic corner,” in *IJCNLP*. Chiang Mai, 2011.

[280] S. Kamvar and J. Harris, “We feel fine and searching the emotional web,” in *WSDM*, 2011, pp. 117–126.

[281] U. Srinivasan, S. Pfeiffer, S. Nepal, M. Lee, L. Gu, and S. Barrass, “A survey of mpeg-1 audio, video and semantic analysis techniques,” *Multimedia Tools and Applications*, vol. 27, no. 1, pp. 105–141, 2005.

[282] R. Schleicher, S. Sundaram, and J. Seebode, "Assessing audio clips on affective and semantic level to improve general applicability," in *Fortschritte der Akustik - DAGA*, Berlin, 2010.

[283] H. Ephron, *1001 Books for Every Mood: A Bibliophile's Guide to Unwinding, Misbehaving, Forgiving, Celebrating, Commiserating.* Adams Media, Avon, 2008.

[284] E. Cambria, A. Hussain, C. Havasi, C. Eckl, and J. Munro, "Towards crowd validation of the uk national health service," in *WebSci*, Raleigh, 2010.

[285] T. Benson, S. Sizmur, J. Whatling, S. Arikan, D. McDonald, and D. Ingram, "Evaluation of a new short generic measure of health status," *Informatics in Primary Care*, vol. 18, no. 2, pp. 89–101, 2010.

[286] A. Donabedian, "Evaluating the quality of medical care," *The Millbank Memorial Fund Quarterly*, vol. 44, pp. 166–203, 1966.

[287] S. Fanshel and J. Bush, "A health status index and its application to health-services outcomes," *Operations Research*, vol. 18, pp. 1021–1066, 1970.

[288] G. Torrance, W. Thomas, and D. Sackett, "A utility maximisation model for evaluation of health care programs," *Health Services Research*, vol. 7, pp. 118–133, 1972.

[289] A. Culyer, R. Lavers, and A. Williams, "Social indicators: Health," *Social Trends*, vol. 2, pp. 31–42, 1971.

[290] J. Ware, "Scales for measuring general health perceptions," *Health Services Research*, vol. 11, pp. 396–415, 1976.

[291] M. Bergner, R. Bobbitt, S. Kressel, W. Pollard, B. Gilson, and J. Morris, "The sickness impact profile: Conceptual formulation and methodology for the development of a health status measure," *International Journal of Health Services*, vol. 6, pp. 393–415, 1976.

[292] J. Ware and C. Sherbourne, "The MOS 36-item short-form health survey (SF-36). Conceptual framework and item selection," *Medical Care*, vol. 30, pp. 473–83, 1992.

[293] J. Ware, M. Kosinski, and S. Keller, "A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity," *Medical Care*, vol. 34, no. 3, pp. 220–233, 1996.

[294] R. Brooks, "EuroQoL - the current state of play," *Health Policy*, vol. 37, pp. 53–72, 1996.

[295] J. Horsman, W. Furlong, D. Feeny, and G. Torrance, "The health utility index (HUI): Concepts, measurement, properties and applications," *Health and Quality of Life Outcomes*, vol. 1, no. 54, 2003.