A MODEL TO EXTRACT SENTIMENTAL KNOWLEDGE IN A SEMANTIC WEB CONTEXT

Domenico Consoli[1]

Abstract: Nowadays, in a social and semantic context, the web contains millions of data, generic information and opinions. The semantic techonology, using specific ontologies, is very important to represent the web knowledge and to understand the meaning of online textual information. To extract useful knowledge and sentiment (positivity or negativity) of opinions it is necessary to have a model that interpreters this information from both semantic and sentiment point of view. In this paper, after the analysis of the web semantic architecture, we describe a framework of sentiment analysis to interpret and extract the sentiment of opinions expressed by people about a product/service, public administrator or law.

Introduction

Over the years, in Internet we have moved from the first generation of the web (web 1.0) to social (web 2.0) and semantic web (web 3.0); in the future we will arrive to intelligent web (web 4.0). On the web, especially with the advent of social networks, there is a lot of information, generic and opinions on a particular product/service, politician, law or regional rule. It is interesting to gather, process and extract the sentiment of these opinions by semantic technologies of web 3.0. Processing manually this data became very stressful. In addition to the semantic layer, we must also consider the sentimental layer, an algorithm that allow us to assign an emotional weight to opinions of individual sentences thus to make easy their classification.

The structure of the paper is the followings: in the next section we describe the evolution of the web: from the web 1.0 to web 4.0. The third section focuses on the semantic web. In the fourth section a layered architecture is discussed. In the fifth section we describe the concept of ontologies. The sixth section analyses the social and sentimental processing. In the seventh section we describe a sentimental model to process on line opinions. Finally some conclusions are drawn.

1. From The Web 1.0 To Web 4.0

In the first decade, the web was a read-only medium: a large library where receive and read information. The second decades have transformed the web into a writing medium. By blogs and social networks, users have been converted from passive spectators in active actors, publishing information, opinions and suggestions. This is the period of the Web 2.0 or Web of participation, where the users become User Generated Content (UGC) (Strobbe et al., 2010) or Consumer Generated Media (CGM) (Sumi, 2008) that publish contents in Internet. Now it is starting the third decade, the new era of the Web 3.0, era of machine understanding where the information available on-line is understandable also from the computers. In this way, by automatic tools, it is possible to link and aggregate information (mash-up) from different data sources and reuse it in other forms to create new contents and therefore new knowledge.

[1] Dr Domenico Consoli, Department of Economics, Society, Politics, University of Urbino, Italy
e-mail: domenico.consoli@uniurb.it

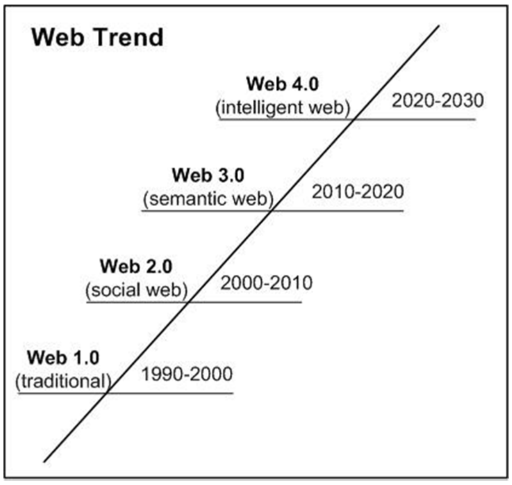
The evolution of the web, from the web 1.0 (human readable) to web 4.0 (intelligent and reasoner web), is represented in Table 1.

Table 1: Web evolution

Web stage	Description
web 1.0	Reading (traditional web)
web 2.0	Writing and generation of contents (social web)
Web 3.0	Machine understandable (semantic web)
Web 4.0	Intelligent and reasoner (intelligent web)

In Figure 1, it is visible the trend of the Web with the period and dominant features: social, semantic and intelligent.

Figure 1: Different ages of the web



The social and collaborative dimension of the web has been emphasized as an information space and a communication tool. By Web 2.0, more social than technical tool, the network has become a more complex interaction among sites and users with a quick explosion of the social context. In social networks, people share opinions, conversations, files, photographs, etc…. The motto of this era is "sharing".

During the social era, the most important changes have been:

- the personal sites have been transformed in blogs that are simple to manage, also for beginners;
- the information management systems became global and shared (i.e Wikipedia, Linux)
- the published information have been shared, re-published by aggregators and available in other platforms (e.g. Mobile phones).

Nowadays, in the web 3.0 era, we think to web as a tool of collaboration between computers. The machines become capable to analyse data, contents, links and transactions between people and computers. Machines will talk with other machines; people will think only to provide inspiration and intuition. In the paper on the semantic web (T. Berners-Lee et al. 2001), affirm that: "The Web was designed as an information space, with the aim of being useful not only for human-human communication, but also machines could have the opportunity to participate and make their contribution". The Semantic Web approach develops languages for expressing information in an accessible way processable by a machine. The idea of the Semantic Web arises from the need to extend, not replace, the current web in order to encourage the exchange of information among human beings, computers and programs.

Semantic technologies are able to extract meaning from information. They are very useful in companies and institutions that need to catalog and manage a large amount of information.We must think to computer, not longer as a simple machine that responds to standard input, but as an intelligent tool, a thinking mind that supports the work and the research of people.

Published documents (HTML pages, files, images, etc..) are associated to information and data (metadata) that specify the semantic context in a format suitable to questions, interpretation and, more generally, to automatic processing.

The passage from the web 3.0 (semantic web) to web 4.0 (intelligent web) is represented from the intelligence module. Web 3.0 works in a semantic context while in the web 4.0 there are intelligence, reasoning and inference. Thanks to web 4.0 in the future, for example, will be possible an automatic booking of hotels and flights to a simple request like this: “I would like arrive in London about at 13.00”. The semantic and intelligent search engine looks for timetable of the 4 London airport: Stansted, Heathrow, Gatwick e Luton and after information processing it decides the flight and then the hotel to book.

2. The Semantic Web

The semantics is that part of linguistics which studies the meaning of single words, sentences and texts. In general, the semantics creates relationship among linguistic expressions with the meanings of the terms.

The Semantic Web is an extension of the current web where computers and people can work in cooperation; it is necessary to consider new models of knowledge representation, which should be readable by humans and also, in large part, by machines.

The term “Semantic Web” was firstly proposed in 2001 by Tim Berners Lee, inventor of the World Wide Web. Since then the term has been associated with the idea of a web where intelligent agents are able to understand the meaning of texts and to directly guide users in the information searching or to replace them in some operations. The semantic web is composed of three main levels: Data, Metadata and Schema or Ontology.

Data is the lowest level and represents the content. Metadata is useful to build a schema that expresses the relations among concepts. A schema is useful to describe a specific domain of information. The metadata should be written in a readable format by the machines to allow syntactic and semantic interoperability. All documents, associated with data and metadata that describe the semantic context, can be automatically interpretated and processed.

Ontology is an exhaustive and rigorous conceptual framework that represents all relevant entities and their relationships contained within a specific domain.

Adding semantics to web contents requires a specific language and more complex rules for expressing data, relationship and reasoning (Eiter et al., 2008). We should be able to write links that describe also the semantic meaning of a link and the destination area (semantic ability).

With the Semantic Web will be possible a more advanced research and other specialized operations such as the building of networks of relationships and connections among documents according to logical hyperlinks. We can search a document basing on the meaning of a concept or several concepts related to each others. For example, if we search, at the moment, with Google the word "book" it returns many occurrences of terms; in a semantic context, the same request will be able to return documents that corresponds only to book.

The Semantic Web should be able, for example, to disambiguate the italian term “espresso”, in a specific context, from a particular train or coffee to drink.

If the semantic web is able to understand the meaning of the contents of a page, it is possible to formulate queries like "What is the cheapest family-friendly hotel in Rome” and provide a single answer, the right one. By semantic web we can create a new generation of applications that can not only estimate the price for a room in a hotel, but also automatically book the cheapest and the closest hotel to a specific place.

In the past, we had repository of documents contained in silos, documents interpretable only from human agents with difficulties in the interoperability and reuse of data and services, with different terminologies (different names for the same objects or same name for different objects) and based on proprietary protocols for the access like to Application Programming Interface (API).

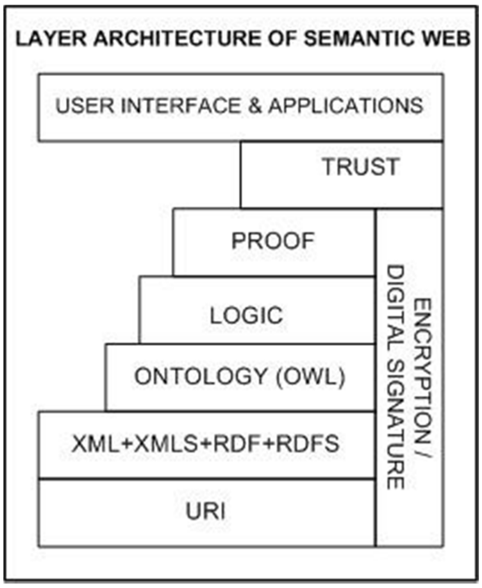
By semantic technology it is possible to make different actions: assigning an unique name to each resource, gathering data from multiple sources without the use of the API protocol, making a mash-up, using specific agents/programs/scripts for data understanding, interpretating the semantics of different links and mapping data into a representation of abstract/logic level.

3. Layered Architecture Of Semantic Web

The Semantic Web presents a layered architecture (Figure 2) described by the W3C (World Wide Web Consortium) (www.w3.org/2001/sw), in the form of a stack of independent layers. Each layer uses or extends the previous. The highest level exchanges data with the lowest.

Figure 2: Architecture of semantic web by W3C

(Adapted by author)



We can resume the layers of the architecture of web semantic (Gerber et al., 2007) (Gerber et al., 2008) in this way:

- URI: to identify the single resources of the knowledge
- Language (XML, OWL,…): to represent the knowledge
- Logic: to acquire new knowledge using the functions of inference
- Proof: to verify that the knowledge obtained, from the logic level, is demonstrable
- Trust: to establish, by the use of digital signatures, that the whole process is based on reliable connections and can be safely used in applications and web services.

At the base of the stack there is the Uniform Resource Identifier (URI) or International Resource Identifier (IRI) to identify an unique document inside the internal network.

A resource presents different elements: *Entity* (object) identified by an URI, *Property* (attribute), relationships among resources and/or atomic values, *Statement* (proposition) that specifies the values of a specific resource's property.

The languages to represent the knowledge (Gomez-Perez and Corcho, 2002) are different: eXtended Markup Language (XML), XML Schema (XMLS), Resource Description Framework (RDF), RDF Schema (RDFS). By XML, RDF languages (Brickley and Guha 2003) it is possible to define a shared vocabulary to identify the model of machine and use extended terms to infer new properties of the concepts or to establish rules for the definition of resources. RDF is a basic tool to encode, exchange and reuse structured metadata and enable the interoperability among applications that exchange information on the web.

Ontology is a shared conceptualization of a specific domain that can be created by Ontology Web Language (OWL) (McGuinness and van Harmelen, 2003) (Dean and Schreiber, 2003). The OWL allows to encode the knowledge in an ontology derived from the standard model of a machine readable. By Rule Interchange Format (RIF) we can share rules among different semantic systems and by Simple Protocol and RDF Query Language (SPARQL) it is possible to make queries on databases. To this level of ontology there is not inference but only knowledge representation.

Logic layer answers to questions and extracts useful information using rules of inference and other logical techniques. The Rule Markup Language (RuleML) (Boley et al., 2001) is a language for expressing inference's rules in the semantic web. The logic layer presents functions of inference that allow to applications to connect with different schemas of ontologies.

The layer of the demonstrations (web of reasoning) establishes an universal language ("Language Proof") to represent the demonstrations. A proof is a sequence of formula with rules of inference, axioms and definitions. With the help of a reasoner, we can extract very sophisticated inferences about data. For example, we can establish that the Dolphin is a subclass of the Mammal class. If Phinker is a Dolphin, then the Reasoner add it in the ontology that represents Phinker as a mammal.

The goal of layer of trust (web of trust) is to ensure, by digital signature or encryption, the authenticity of various statements and to discover their origins. The system may authenticate the digital signature and export the knowledge in other systems that could be incorporated in the semantic web.

4. The Ontologies

In computer science, an ontology is a formal representation of knowledge as a set of concepts, in relationship among themselves, within a specific domain (Flouris et al., 2008). It can be considered as an explicit specialization of a shared conceptualization of a domain.

An ontology is a sort of "skeleton" of knowledge that includes rules that can be used to prove theorems and answer questions about specific instances.

Ontologies organize a domain of knowledge by a schema that divides the domain of ontology in objects' classes with specific roles.

Ontologies will become the most powerful tool available for the Knowledge Management. The same information can be useful in different situations; more often it is useful to recover part of an information. For example, in a mechanics course that contain a section on fluid's mechanics, this part can be used in a course of hydraulics. If different materials are organized in units (learning objects), each unit can be connected to others and reassembled in a new course.

An ontology can be understandable from machines; software agents and humans can exchange knowledge from different ontologies (Khong Cua et al., 2010).

As we have seen in the previous section, an ontology can be created by OWL that is compatibile with web standards XML, XMLS, RDF and RDFS.

Each resource of an ontology must be represented by a triplet: *subject – predicate – object* or *resource – property – value*.

For example: These companies (*subject*) supply (*predicate*) shoes (*object*) / Fiat Panda (*subject*) is a type of (*predicate*) utilitary (*object*) / www.iloveyou.en (*resource*) has as author (*property*) MisterX (*value*).

The resources can be linked between themselves by statements:

<http://mydomain/resource1> <http://mydomain/property1> <http://mydomain/resource2>

In the representation of knowledge we can consider, from the bottom, three levels: physical (database), abstract (logical structure), application (software).

In next figures, physycal (Figure 3) and abstract/logic (Figure 4) layers of a representantion of a part of Book Ontology (Herman, 2010), with original and traducted version of books, are shown.

In Figure 3 we have considered a database with three tables (Table 1, Table 2, Table 3) linked among themselves with relationships (relational database) on the management of books. In Figure 4 we have represented the abstract (logic) layer of a part of a Book Ontology with resources and their properties. For example the f:Aut_book_original represents the property Author of the original book (resource identified by its ISBN code).

Figure 3: The physical layer of a part of Books Ontology

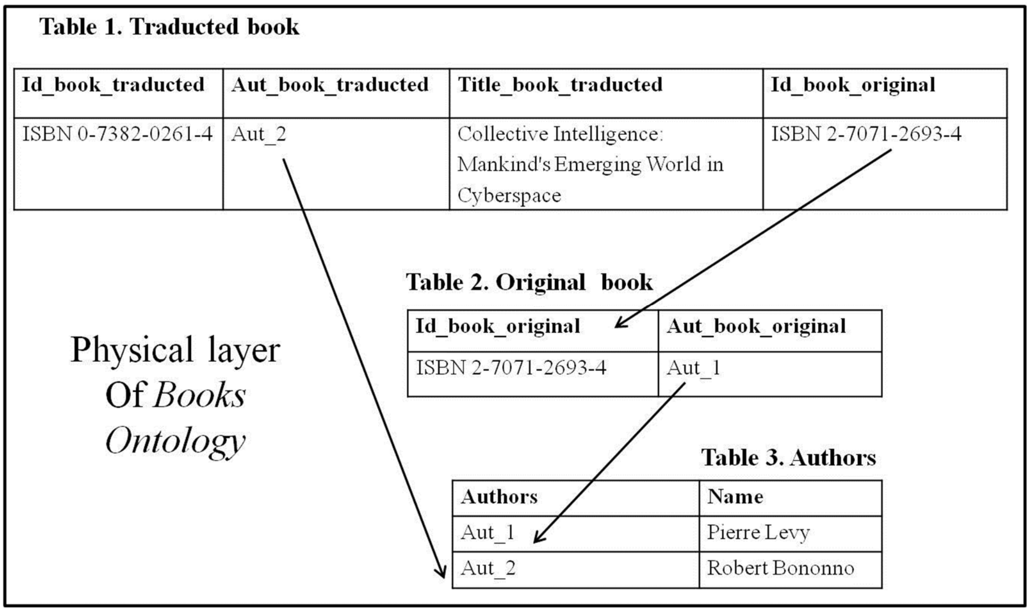
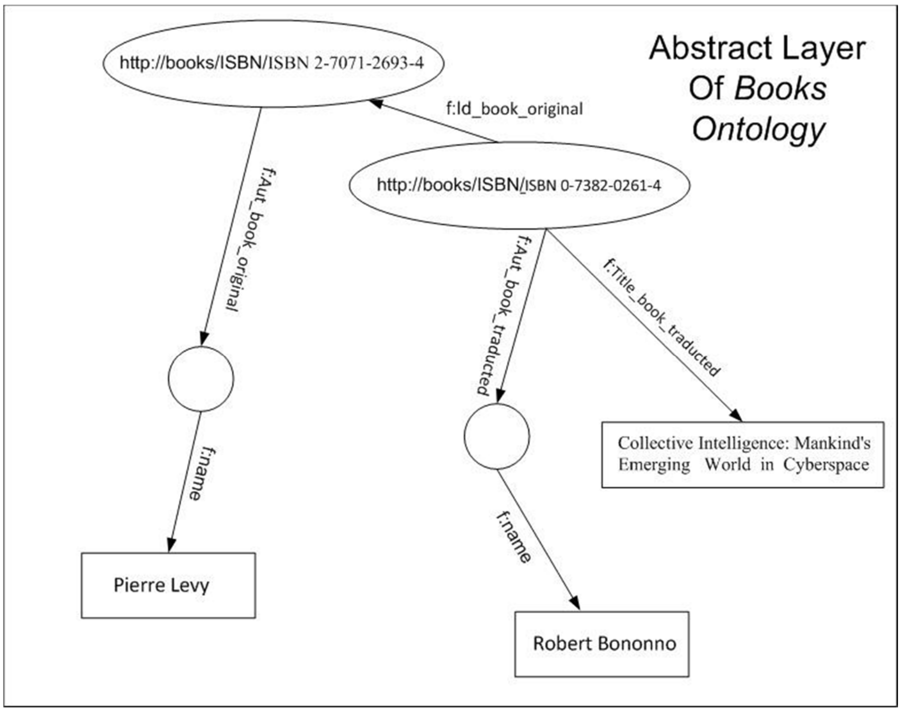


Figure 4: The abstract (logic) layer of a part of Books Ontology.



5. Social and Sentiment Analysis

Surely the semantic technology is a powerful tool to accurately monitor online reputations and opinions (sentiments of users on the network). Social media are the virtual places where interpersonal communications and opinions are spread in public emotions. A virtuous synergy between cognitive and emotional dimension are realized by a relational model that rewards the communities and the sharing of knowledge. Social sciences study and investigate these new online dynamic relationships. Commercial software analyses the sentiment of people on the web to understand their perceptions on a specific brand or product.

By web 2.0 tools, anyone can publish contents, news and express opinions in a context of "participatory culture". Users have an active role as user generated content that contributes to create a great reservoir of ideas, opinions and knowledge.

This mass of data can be interpreted by artificial machines that use "social listening technologies". These technologies enable companies to manage the “voice of customers”, to know how consumers are moving in the purchaising process and to evaluate the effects of a communication campaign on customers’ perceptions. These feedbacks are used, from the companies, to improve their communication strategies and their product/services in a context of customer satisfaction. The opinions in the blogosphere are spread in a viral modality and can overwhelm the brand.

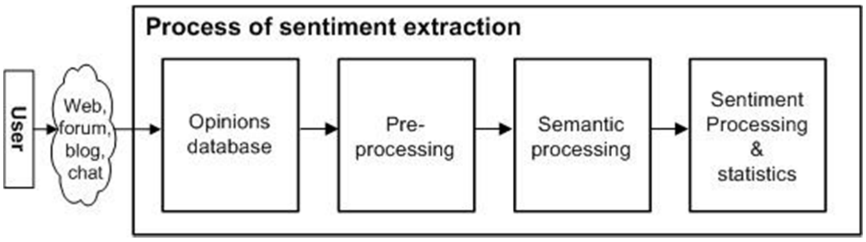
In the study and analysis of social networks, the experiment conducted by Stanley Milgram in the 1960 is very important. He arrived in the city of Omaha in Nebraska and delivered the letters with only the names, without addressees, and gave them to people. Through acquaintances of acquaintances several letters arrived at their destination (30%) by 6 steps (6 degrees of separation). On Facebook, the popularity of a person depends on the number of friends, the popularity of a page from the pages linked to it and a query-log from the time to execute it.

Today, more than ever, the marketing intelligence needs to capture the sentiment of the market and brand reputation. Semantic technology provides an immediate perception of signals that coming from the market and can exploit the strategic opportunities.

6. Sentimental Model

In the Figure 5 a model to extract sentiments, from online opinions, is shown.

Figure 5: Sentimental model



The model is divided in four modules: opinions database, pre-processing, semantic and sentiment processing. These modules are explained in the following sub-sections.

7.1 Opinion databases

The Opinion Database contains online opinions that crawler agents gather from the web (Consoli et al., 2008). Crawler agents are specialized in different protocols (http, https, pop3, imap4, nntp) and are responsible to inspect and retrieve opinions, respectively, from websites, blogs, chats, e-mails, newsgroups and so on. Each agent can be configured with policies to extract texts with advanced techniques of Natural Language Processing (NLP). Alternatively, users may, apriori, select web sites of interest and define wrappers (Kuhlins and Korthaus, 2003) in order to analyse the structure of each site and extract opinions.

7.2 Pre-processing

Since opinions are written in Natural Language, to process them, we need a specific pre-processing. The goal of this phase is to obtain significant words and statements (Consoli et al., 2009a) for each opinion expressed, in a web post. The pre-processing consists of the following steps:

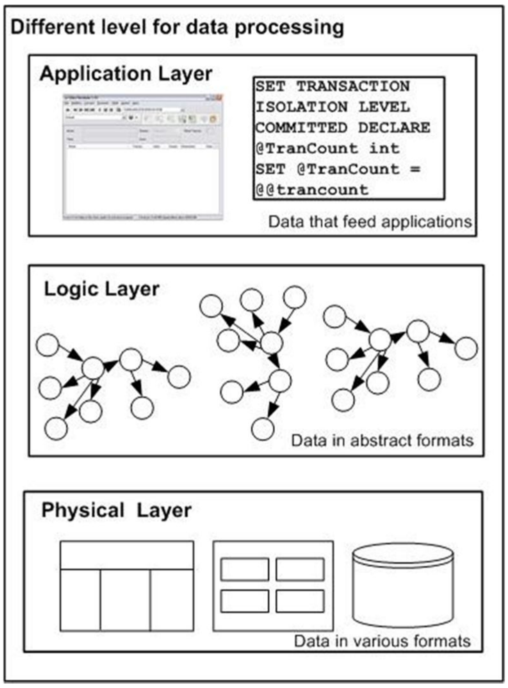
- Sentence extraction. From every post we extract the sentences. In this step we eliminate all interrogative clauses because these clauses don't carry affective information.
- Statement extraction. The goal of this phase is to divide each sentence in statements. A statement is an elementary sub-sentence that expresses a single positive, neutral or negative polarity.

A single sentence can express more than one opinion. For example "The Institute of Secondary School Alfa is good but some teachers are terrible" can be split in two statements: " The Institute of Secondary School Alfa is good" (positive statement) and "some teachers are terrible" (negative statement). To divide these sentences in statements it is necessary to analyse the conjunctions that are present in the sentences. It is need to separate the words in the proximity of those conjunctions that link two propositions with opposite polarity for example "but" (coordinative conjunction) or "although, even, thus, whereas, while" (subordinate conjunction). After the pre-processing phase, opinions and in general textual information are transformed in a matrix terms-documents and then in a numerical array and so it is possible to process this array by specific software.

7.3 Semantic processing

The Semantic processing module uses semantic technology to understand online texts. The framework is shown in Figure 6 where we can see the three levels of information processing.

Figure 6: Online information processing



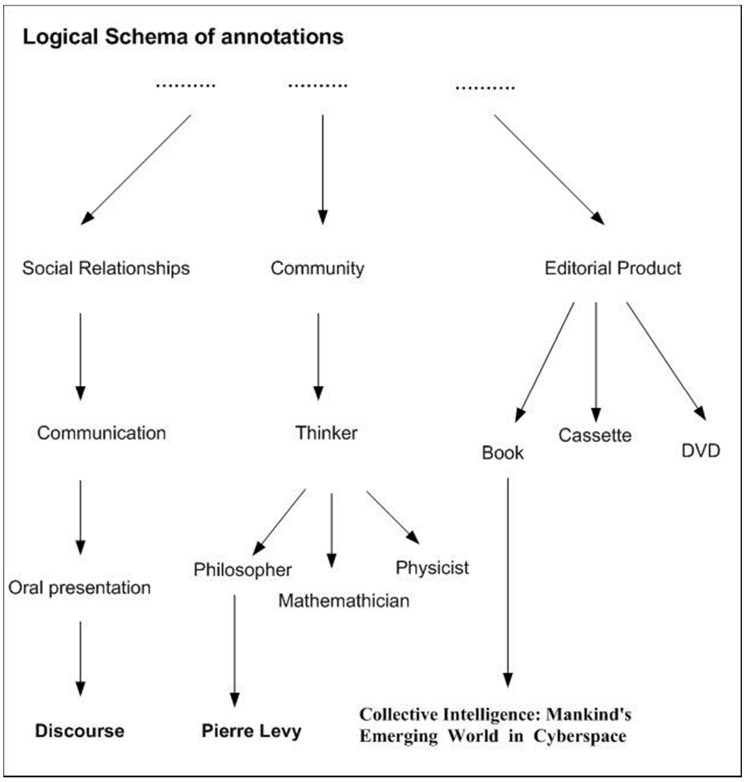
To explain the model, for example, we consider the following sentence: “The book -Collective Intelligence: Mankind's Emerging World in Cyberspace- by Pierre Levy was presented”.

To understand the text from the machine, the first stage is the annotation. After the annotation it is important to map the annotate text in the abstract/logic layer and then, at the end, it is possible to insert it in a physical database.

The procedure of annotation is the following: The book “Collective Intelligence: Mankind's Emerging World in Cyberspace”/**Boook/** by Pierre Levy /**Philosopher**/ was presented /**Oral presentation**/.

The logical schema of annotations, with concepts and relationships to build the ontology, is represented in Figure 7.

Figure 7: Logical Schema of annotations



After the logical mapping we can insert data in a physical database that use linked table (relational database) like in the previous Figure 3.

7.4 Literature and Sentimental Processing

To sentimentally process data we must consider algorithms that match words, nouns, adjectives (part of speech) (Tanawongsuwan, 2010) that are contained inside different “affective” dictionaries and ontologies. In the literature there are several contributions for sentiments and affectivity. In the Keyword Spotting (Boucouvalas and Zhe, 2002) approach the text is classified into affective categories based on the presence of fairly affective words like distressed, happiness and anger. All terms that describe emotional states represent the most direct way to communicate emotions by text. The simplest and most used analysis is based on the search for keywords (e.g. happy, sad, angry, etc).

This is at the basis of the Elliot's Affective Reasoner (Elliott, 1992) that watches for almost 200 affect keywords plus affect intensity modifiers (e.g. extremely, somewhat, mildly). Ortony's Affective Lexicon (Ortony and Collins, 1990) provides a source of affect words grouped into three main emotional classes (affective categories): satisfied/unsatisfied, approved/ disapproved, pleasant/unpleasant.

The weaknesses of this approach are in the negation of sentences and in semantic concepts "Yesterday wasn't a happy day" or "My wife would like divorce and take the custody of my children". In the last sentence there aren't affective words but the phrase certainly evokes strong emotions.

The Lexical Affinity (Valitutti et al., 2004) method, trained from linguistic dictionaries, assigns to words a probabilistic affinity for a particular emotion. For example, to the term "accident" we can assign the 80% of probability that indicates a negative affect, as in "car accident", "hurt by accident".

In this case, the lexical affinity, operating at word-level, can easily be tricked by negative sentences like "He avoided an accident" and others word like "She met her boyfriend by accident".

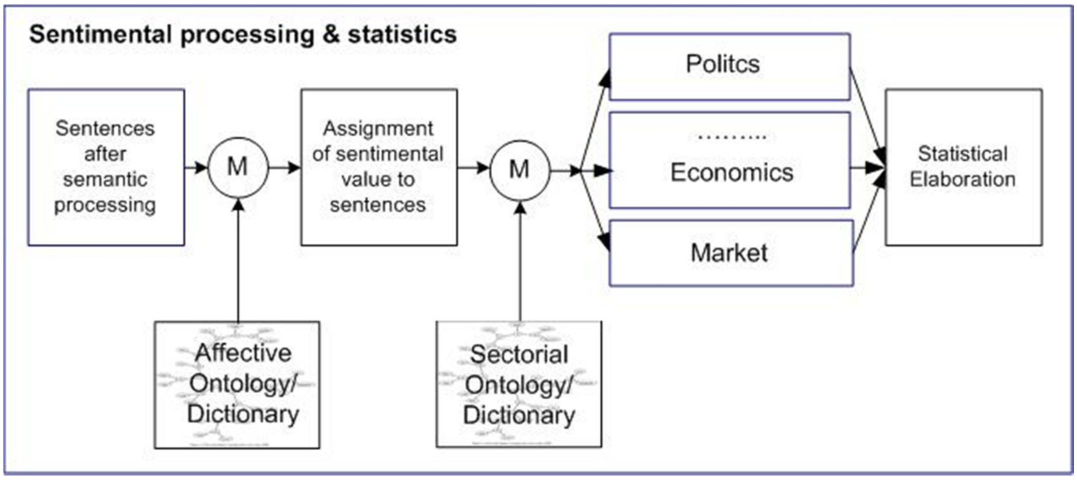
In the Statistical Method (Jpennebaker et al., 2001) that feeds a machine learning algorithm with a large affective training corpus, it is possible, for the system, to learn the affective valence of affective keywords (keyword spotting) and take in account the valence of other arbitrary keywords (lexical affinity). We can do an analysis on the lists containing a lot of emotional adjectives and after, with appropriate statistical techniques, reduce these.

Esuli and Sebastiani (2006) have created SentiWordNet, a lexical resource for opinion mining, where they assign to each synset (set of synonyms) of WordNet a sentiment scores: positivity, negativity and objectivity (neutrality). The opinion is positive if the positivity of its terms is higher than negative and objective scores. WordNet Affect (Strapparava and Valitutti, 2004) is a linguistic resource for a lexical representation of affective knowledge. In WordNet Affect each synset of WordNet is labeled by one or more affective-labels, representing the affective meaning of the synset. Examples of affective-labels are emotion, mood, trait, cognitive and physical state, etc…

Consoli et al. (2009b) have developed an original algorithm for sentiment analysis that mainly focuses on six Ekman emotional indexes: anger, disgust, fear, happiness, sadness, and surprise (Ekman, 2007).

Our model of sentimental processing, that exploits these Ekman emotional indexes, is shown in the Figure 8.

Figure 8: Sentimental processing and statistics



The sentences, which we have processed by the semantic module, macth with an ontology that contains affective terms like good, bad, fear, anger, etc.... In this way, we can assign to sentences a different affectivity that depends mainly from the frequency of the affective terms inside each sentence. The sentences can be related to specific topics of politics, biology, geography, literature, economics, market, etc.. An additional matching allows to separate the sentences for topics. For example if, in the sentences, there are a lot of terms contained in a Economic Ontology then we can consider this sentence in Economic topic.

At the end, we can apply a statistic processing to individuate if the positive opinions prevail or are egual to negative and viceversa. In this way for each topic we can know if the sentiment is positive, negative or neutral. In statistical processing we can use different algorithms of data and text mining (Mastrogiannis et al., 2009)(Bolasco et al., 2005) to classify the opinions. At the moment the precision of the model is about 76% and we studied a manner, with the improving of affective algorithms, to increase it.

8. Conclusions

On the web, especially with the advent of social web (web 2.0), there is a lot of information and in particular a lot of opinions on a specific product/service, politician, teacher, law or regional rule. It is interesting to gather and process this information by semantic technologies of the web 3.0 that allow communications among humans and machines.

This technology enables companies to know how consumers are moving in the buying process and use this feedbacks to improve their product/services. Semantic technology provides an immediate perception of signals that coming from the market and from the voice of the customers. Processing this information manually it becomes very stressful and therefore it is necessary an automatic model. In this paper we propose a model to extract the sentiment of opinions from the web using semantic technology. The model present different stages of opinions' processing. In addition to semantic layer, it is necessary to consider a sentiment layer for assigning an emotional weight to opinions of individual sentences and for evaluating the positivity or negativity of the opinion. At present the model presents a discrete precision that, in the future, we would improve, optimizing the algorithm.

This model of sentiment analysis is very interesting to interpret the opinions expressed by people about a product/service, a public administrator or a law. In this way, the companies, also public administrations can improve their products/services in a context of the customer/citizens satisfaction.

References

Boley H., Tabet S. and Wagner G. (2001). "Design Rationale of RuleML: A Markup Language for Semantic Web Rules". In Proc. Semantic Web Working Symposium (SWWS'01), Stanford University, July/August 2001, pp. 381-401

Boucouvalas A. and Zhe X. (2002). "Text-to emotion engine for real time internet communication", in Proceedings of International Symposium on CSNDSP. UK: Staffordshire University, 2002, pp. 164-168

Brickley D., Guha R.V. (2003). "RDF Vocabulary Description Language 1.0 : RDF Schema",W3C Working Draft 23 January 2003 available online at http://www.w3.org/TR/rdf-schema/

Consoli D., Diamantini C. and Potena D. (2008). "Improving the customer intelligence with Customer enterprise Customer model". In Proc. of the 10th International Conference on Enterprise Information Systems, Software Agents and Internet Computing (SAIC), Barcelona, Spain, June 12-16 2008, pp. 323- 326.

Consoli D., Diamantini C., Potena D. (2009a). "Knowledge Mining from web customer opinions to improve enterprise product". In Proceedings 2th International Conference on Mathematics and Informatics, Bacau, September 8-10 2009, Vol. 19 (2009), n. 2, pp. 163-178

Consoli D., Diamantini C., Potena D. (2009b). "Affective algorithm to polarize customer opinions" In Proc. of the 11th International Conference on Enterprise Information Systems, Human-Computer Interaction (HCI), Milan, Italy, May 6-10 2009, pp. 157-160.

Dean M. and Schreiber G. (2003). "Web Ontology Language (OWL) Reference Version 1.0", W3C Working Draft, 21 Febbraio 2003 available online at http://www.w3.org/TR/owl-ref/

Ekman P. (2007).“Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life” NY: OWL Books, 2007.

Eiter T., Ianni G., Krennwallner T. and Axel Polleres (2008). “Rules and Ontologies for the Semantic Web”. In Reasoning Web, Cristina Baroglio, Piero A. Bonatti, Jan Maski, Massimo Marchiori, Axel Polleres, and Sebastian Schaffert (Eds.). Lecture Notes In Computer Science, Vol. 5224. Springer-Verlag, Berlin, Heidelberg, pp. 1-53

Elliott C. (1992). “Knowledge acquisition within the affective reasoner” in Notes for the AAAI Spring Symposium: Cognitive Aspects of Knoweledge Acquisition, 1992, pp. 64-72

Esuli A. and Sebastiani F. (2006). “SentiWordNet: A publicly available lexical resource for opinion mining” in Proceedings of LREC-06, 5th Conference on Language Resources and Evaluation, Genova, IT, 2006, pp. 417-422

Flouris G., Manakanatas D., Kondylakis H., Plexousakis D. and Antoniou G. (2008). “Ontology change: Classification and survey”. Knowl. Eng. Rev. 23, 2 (June 2008), pp. 117-152

Gerber A. J., Barnard A. and van der Merwe A. J.(2007). “Towards a semantic web layered architecture”. In Proceedings of the 25th conference on IASTED International Multi-Conference: Software Engineering (SE'07), W. Hasselbring (Ed.). ACTA Press, Anaheim, CA, USA, pp. 353-362

Gerber A., van der Merwe A. and Barnard A. (2008). “A functional semantic web architecture”. In Proceedings of the 5th European semantic web conference on The semantic web: research and applications (ESWC'08), Sean Bechhofer, Manfred Hauswirth, Hoffmann J., and Manolis Koubarakis (Eds.). Springer-Verlag, Berlin, Heidelberg, pp. 273-287.

Gomez-Perez A. and Corcho O. (2002). “Ontology Specification Languages for the Semantic Web”. IEEE Intelligent Systems 17, 1 (January 2002), pp. 54-60

Herman I (2010), “Short introduction to the Semantic Web”, avaiable online at: http://www.w3.org/People/Ivan/CorePresentations/IntroThroughExample/Slides.pdf

Khong Chua W.W. and Soong Goh A.E. (2010). “Techniques for discovering correspondences between ontologies”. Int. J. Web Grid Serv. 6, 3 (September 2010), pp. 213-243

Kuhlins S. and Korthaus A. (2003). “A multithreaded java framework for information extraction in the context of enterprise application integration” in ISICT ’03: Proceedings of the 1st international symposium on Information and communication technologies. Trinity College Dublin, 2003, pp. 518–523.

Lyman P., Varian H.R., Charles P., Good N., Jordan L.L. and Pal J. (2003) “How much information?”, available online at http://www2.sims.berkeley.edu/research/projects/how-much-info-2003

Mastrogiannis N., Boutsinas B. and Giannikos I. (2009). “A method for improving the accuracy of data mining classification algorithms”, Comput. Oper. Res. 36, 10 (Oct. 2009), pp. 2829-2839

McGuinness D.L. and van Harmelen F. (2003) "Web Ontology Language (OWL): Overview", W3C Working Draft, 10 Febbraio available online at http://www.w3.org/TR/owl-features/

Ortony C.-G. and A. Collins (1990) “The cognitive structure of emotions”, New York: Cambridge University Press, 1990

Pennebaker J. W., M. E. Francis, and R. J. Booth (2001). “Linguistic inquiry and word count: Liwc” in Computer Software, Mahwah,NJ, 2001.

S. Bolasco, A. Canzonetti, F. M. Capo, F. della Ratta-Rinaldi and B. K. Singh (2005). “Understanding text mining: A pragmatic approach” in Knowledge Mining, ser. Studies in Fuzziness and Soft Computing, S. Sirmakessis ed., Springer Verlag, 2005, vol. 185, pp. 31–50.

Strapparava C. and Valitutti A. (2004). “Wordnet-affect: An affective extension of wordnet” in The 4th International Conference On Language Resources and Evaluation, 2004, pp. 1083-1086

Strobbe M., Van Laere O., Dauwe S., Dhoedt B., De Turck F., Demeester P., van Nimwegen C., and Vanattenhoven J. (2010). "Interest based selection of user generated content for rich communication services". J. Netw. Comput. Appl. 33, 2 (Mar. 2010), pp. 84-97

Sumi K. (2008). "Anime de Blog: animation CGM for content distribution". In Proceedings of the 2008 international Conference on Advances in Computer Entertainment Technology (Yokohama, Japan, December 03 - 05, 2008). ACE '08, vol. 352. ACM, New York, NY, 187-190.

Tanawongsuwan P. (2010). "Part-of-Speech Approach to Evaluation of Textbook Reviews", in Proceedings of the 2010 Second international Conference on Computer and Network Technology (April 23 - 25, 2010). ICCNT. IEEE Computer Society, Washington, DC, pp. 352-356

Tim Berners-Lee, James Hendler and Ora Lassila (2001). "The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities", Scientific American Magazine, May 17, 2001

Valitutti A., Strapparava C., and Stock O. (2004). "Developing affective lexical resources" vol. 2, n. 1, pp. 61-83

